TRUST ON THE LINE

A philosophical exploration of trust in the networked era.

ONLINE VERTROUWEN ONDER DRUK
Een filosofisch onderzoek naar vertrouwen in het netwerktijdperk.

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.
The public defense shall be held on

Thursday, 21 April 2016 at 13.30 hours
by
Esther Lieve Omer Keymolen
born in Sint-Niklaas, Belgium.

Erasmus University Rotterdam

Doctoral Committee

Promotors:

Prof.dr. J. de Mul
Prof.dr. V.A.J. Frissen

Other members:

Prof.dr. M. Hildebrandt
Prof.dr. S. van der Hof
Prof.dr. P.P.C.C. Verbeek

TRUST ON THE LINE

A philosophical exploration of trust in the networked era.

This research was supported by TNO

Printing CPI Koninklijke Wöhrmann
Design Tiemen Seinstra

Dankwoord

Dit proefschrift heeft niet alleen mij, maar ook een heleboel andere mensen bezig gehouden. Mijn dank gaat uit naar de oud-collega's van de WRR die mijn wens om te doctoreren altijd hebben gesteund. In het bijzonder dank ik Anne-Greet voor de aanmoedigingen en Corien die niet alleen de praktische proefschriftmolen in gang zette, maar mij ook inspireert als vrouw in de academie. Dennis, die de categorie van oud-collega al lang is ontstegen, dank ik voor de vriendschap, de gesprekken aan de keukentafel, niet te vergeten samen met Linnet –dank Linnet!-, en voor het vakkundig schrappen van de helft van een hoofdstuk.
Ik dank TNO voor het mogelijk maken van dit proefschrift en de faculteit Wijsbegeerte, in het bijzonder de vakgroep Mens en Cultuur waarbinnen ik de ruimte en tijd kreeg om mijn eigen onderzoek vorm te geven. Ik dank al mijn Rotterdamse collega's voor hun interesse en hulp en zeker ook de promovendi-groep waarin ik mijn werk mocht presenteren. In het bijzonder dank ik mijn kamergenoot Rolf voor de immer inspirerende gesprekken én voor het verzinnen van de titel van dit boek! Zo ook dank ik Awee voor de filosofische gesprekken die mijn denken hebben gescherpt en om mij de kans te geven de liefde voor de techniekfilosofie te delen met studenten.

Dat mijn twee promotoren Jos en Valerie een proefschrift over vertrouwen relevant genoeg vonden is enigszins verwonderlijk aangezien zij daarover–zo heb ik mogen ondervinden- duidelijk alles al weten! Dank jullie wel voor het onwrikbaar vertrouwen in mij, de vrijheid die ik heb gekregen om mijn ideeën vorm te geven en de begeleiding daar vervolgens bij. Dank voor jullie ontspannenheid die mijn innerlijke controle-freak enigszins wist te sussen (niet volledig, natuurlijk) en voor de gezamenlijke congresbezoeken en projecten. Ik kijk al uit naar nieuwe samenwerkingsverbanden in de toekomst!

Van het ene warme bad in het andere! Wat een geluk. Ik dank al mijn eLaw collega's aan de Universiteit Leiden voor het hartelijke welkom anderhalf jaar geleden. Met jullie samenwerken is enorm inspirerend en het was een ontzettend goede stimulans om er bij het schrijven de vaart in te houden. Ik dank in het bijzonder Simone voor de

kans om samen het Advanced master programma verder te ontwikkelen en voor de vrijheid om mijn eigen academische weg te vinden.

Dit boek was bijna op losse A4'tjes in Comic Sans bij elkaar geniet tot jullie gekomen, ware het niet dat Lisa daar een stokje voor stak en de lay-out van de tekst voor haar rekening nam. Dank! De prachtige cover en dito boeklegger dank ik aan Tiemen. Het oog wil tenslotte ook wat. Shazade Jameson dank ik voor het nalopen van het Engels.

'Oude' vriendinnen die zich niet lieten afschrikken door mijn filosofische aspiraties: Mieke, Jurith, Jona en Sonja, dank jullie wel voor jullie vriendschap, steun, koffie, wijntjes, chocolade en taart op zijn tijd. Als ik weer eens mezelf aan het voorbij rennen was, kon ik me bij jullie altijd terugvinden.
Hoewel het proefschrift een heel groot deel van mijn aandacht heeft opgeslokt de afgelopen jaren, bleken er toch mensen te zijn die zich daar niet door lieten weerhouden en nu deel uitmaken van mijn leven. Hulde! Ik ben ontzettend blij en vereerd jullie tot mijn vrienden te mogen rekenen. Anna, Petra en Jacqueline; zonder jullie was het een stuk minder leuk geweest.

Het stelt me enorm gerust dat ik tijdens de verdediging hen aan mijn zijde heb als paranimfen: Sem en Bibi! Sems kritisch meedenken –al dan niet onder de geneugten van een hapje en een drankje- en humor bleken een noodzakelijke voorwaarde voor het slagen van dit boek. Bibi is 'buiten categorie' en aan haar zou een heel dankwoord gewijd kunnen worden. Vriendin, collega en zo vertrouwd als familie. Ik kijk uit naar alles wat nog op ons pad komt ('en hunnie zijn @#$%' ;-).

Ik dank mijn familie die –over de landsgrenzen heen- altijd in mij geloofd heeft: mijn moeder Karin en zus Ellen. Dichterbij huis, de familie Hemmes en familie Naaborg en dan zeker opa Fred die trouw door weer en wind kleinkinderen ophaalt en natuurlijk ook tante Jacqueline voor haar sprankelende zelf.

Tenslotte draag ik dit proefschrift op aan de liefste mensen ter wereld: Rogier, Hannah en Nina. Niemand kent mij beter dan Rogier. Ik heb het boek dan wel geschreven, zijn liefde en vertrouwen hebben het mogelijk gemaakt. Hannah en Nina, met jullie ontdek ik de wereld helemaal opnieuw. Daar kan geen boek tegenop.

Contents

1 Introduction: Philosophy begins in wonder. 10

1.1 Research on trust: Pandora's box 13

1.2 Guido Möllering. Trust: reason, routine, reflexivity 15

1.3 e-Trust 19

1.4 The Internet 23

1.5 Outline 26

2 The concept of trust. 30

2.1 Introducing Niklas Luhmann 31

2.2 Luhmann's theory of trust 33

2.3 Trust: bridging the hiatus 45

2.4 Helmuth Plessner and philosophical anthropology 48

2.5 Three times in a row: Positionality 51

2.6 An anthropological perspective on trust 58

3 System trust in late modernity. 61

3.1 Introducing: system trust 63

3.2 System trust and late modernity 66

3.3 Trust: The interaction between system trust, interpersonal trust, and familiarity in late modernity 74

3.4 Conclusion 85

4 The Internet: A familiar world? 87

4.1 A short history of the Internet 89
4.2 The Internet as a layered infrastructure 95
4.3 A new online reality 101
4.4 Conclusion 115

5 Trust in context: a theory of mediation. 117
5.1 Human beings, technology, and environment 119
5.2 The networked era 124
5.3 Artificial by nature in the networked era 131
5.4 Artificial by nature in the networked era: some challenges for trust 142
5.5 Mediated immediacy in the networked era 144
5.6 Mediated immediacy in the networked era: some challenges for trust 152
5.7 Utopian standpoint 155
5.8 Utopian standpoint: Challenges for trust in the networked era 158
5.9 Conclusion: challenges for trust 162

6 Open Sesame. When your phone becomes your key. 166
6.1 The hotel key and a cumbersome but handy key chain 168
6.2 When a key becomes a card 174
6.3 When a keycard becomes a smartphone 178
6.4 Conclusion 195

7 Interpersonal system trust in Airbnb. 197
7.1 What is collaborative consumption? 200
7.2 Airbnb 205
7.3 Context 207
7.4 Construction 211

7.5 Curation 214
7.6 Codification 218
7.7 Conclusion: Why the sharing economy is not just about you and me 224

8 A too familiar world. 226
8.1 Construction 229
8.2 Curation 232
8.3 Codification 234
8.4 Context 238
8.5 A familiar world 242
8.6 Profiling 244
8.7 A too familiar world online 246
8.8 Challenges for trust 247

9 Epilogue 252

10 Bibliography 257

11 Samenvatting 287

12 About the author 291

1

Introduction:

Philosophy begins in wonder.

You wake up and for another 5 minutes you stay in the warmth of your bed to check the latest news on your smartphone. You see the headline about the Snowden revelations and read that the NSA –the U.S.A.'s National Security Agency- has compromised the backbone of the Internet, constructing secret back doors to enable them to eavesdrop on literally everybody. Although the scope of the surveillance seems to be much bigger than most experts had expected, you cannot say you are surprised –this is what secret intelligence services do, right? - You leave the news app to look at today's weather forecast.

In the tube, on your way to work, you check your Facebook page and like a few of the messages some of your 436 friends have posted. Next, you go to your Linkedin-app and accept the pending requests to connect. Some of these people you only know vaguely, but they can become an interesting connection, you never know.

At work, you log in on the company's network. Before starting your busy day in the office, you quickly go to the website of your bank. You want to wire the money for a theatre ticket your friend has bought for you in advance. You log in with your username and password and verify if the green icon is visible in your browser to ascertain the trustworthiness of the connection. To be honest, you feel pretty good about doing that. You will not be fooled by any smart-ass cybercriminal pretending he –they are often male- is your bank, filching your log-in credentials. Also, the e-mails from the Nigerian prince who regularly contacts you to ask for your help with his enormous inheritance, you delete without a second thought.

During your lunch break you check some travel sites, as you are still not sure

where to go to for the holidays. You do have some promising visions about white beaches, cocktails, and a lot of doing nothing. These daydreams are reflected in the advertisements following you around on the Net. Where you normally are quite able to ignore these commercial distractions, now you cannot resist clicking on them. Perhaps the golden destination is just one click away.

Ping. There is a new message in your Whatsapp-theatre friends group. Someone has posted a link to a review of the play you will go to next week. The critic is ruthless. A funny Whatsapp-conversation arises about possible –and hilarious- ways to leave the theatre if the play is as awful as the reviewer claims.

Just some highlights of a regular day that you encounter - but I could just as easily have written *you and me*.

We have learned to navigate the World Wide Web, post messages, and check e-mail. We know how to wire money online, order our groceries, books, and other stuff in web shops. We chat, leave comments, friend and de-friend as we feel like it. Using the Internet and all the artefacts, applications, and services powered by that Internet seems to come naturally to us. It is not that we are not aware of the possible dangers attached to our use of all these handy artefacts. We read the news on the NSA, we know of cyber-attacks, and we are aware that these personalized advertisements following us around must be based on some personal information we have provided, willingly or not. We know all that; but in everyday life it simply does not seem to affect us.

If philosophy begins in wonder, as the proverb goes, then my wonder is this: how can it be that we have so ostensibly easily adopted a technology, while we obviously are aware of the fact that this makes us vulnerable to all sorts of jeopardy? Do we trust the NSA to honour our privacy? Probably not. Do we believe that all the information platforms we use take care of our data in a responsible way? Perhaps not. Are we really sure that when the green icon pops up in our browser, we are in the clear? Maybe not. This uncertainty does not seem to hold us back, however. To put my wonder differently:

How come we trust even when we know that there are clear reasons not to do so?

When I first started to think about this, I was very excited to have come across this paradox of trust. Philosophers, in general, love paradoxes. A paradox mostly unites two opposing statements, which on logical grounds cannot both be true, but after examination do seem to be true.

What I found out during my research is that this paradox is in fact the *main explanatory force* of what trust is. It is the core of what trust is about. Trust is acting *as if* we are sure about our affairs while we at the same time are also aware that *we will never be sure* about our affairs. This connection between [the need to] *trust* and *the contingency of life* is crucial. If we could be sure of how things would turn out, trust would be redundant. We would simply *know* what the future looks like; no need for trust there. We would not need to act *as if* we know, we would just know.

Trust is, however, *a strategy to cope with the uncertainties* inherent in human life. Trust does not take these uncertainties away. Trust makes them bearable in the sense that in the end "uncertainty need not be problematic in practice" (Möllering 2006: 6).

These uncertainties are derived on the one hand from our awareness of an ever-changing future and on the other hand from the fundamentally unpredictable behaviour of our fellow human beings. These uncertainties can be a burden or a chance to accomplish great things, depending on your take on life. Notwithstanding an optimistic or pessimistic view, on a daily basis every one of us has to deal with the fact that things could have been different and that we are never completely sure about what the person in front of us is thinking or how he or she is going to act.

Of course, we can make some well-estimated guesses. My agenda can give you a pretty accurate idea of what I will be up to the next couple of days and based on our past experience with others we can also more or less predict how they will behave. However, there will always remain some odds and ends we have to deal with and that is the moment where trust comes in. Without some basic trust in the continuity of everyday life and in the self-evident aspects of our interactions, we would not even be able to get up in the morning. All the possible futures lying ahead of us would overwhelm us. We are just not equipped to think everything through or to find evidence for every aspect of our actions. In that respect, trust is always, more or less, *blind trust*.

1.1 Research on trust: Pandora's box

My initial wonder about our eagerness to make use of the Internet and all the services provided through that Internet, lead me to dive into the scholarly research on trust. It was as if I opened Pandora's box.

Seppanen et al. (2007: *cited in* Lyon et al. 2012) claim that in their review they found more than 70 different definitions of trust. And this was 'just' a review of the empirical research on inter-organizational trust. "As pervasive trust appears as a phenomenon, as elusive it seems as a concept", Judith Simon (2013: 1) rightly concludes.

The only thing scholars seem to agree upon is that trust is a *fuzzy concept*, meaning that depending on the domain in which one is working and the basic beliefs one holds about the defining features of human beings, the definition will differ.

Trust is being researched in a large variety of domains, from psychology, economy, and sociology to philosophy. For psychologists, trust is one of the key aspects of our daily interactions (Desteno 2014). It has to do with the child expanding its view of the world by learning from others (Harris 2012), developing strong ties with primary caretakers by building what has been called "basic trust" (Erikson 1950). This basic trust is a necessary condition to build up trust with other people and institutions further away from the safety of the home (also see Giddens 1991; Giddens 1990).

For economists, trust has to do with calculation and risk-assessment, averting negative outcomes. The main idea is that "…trust is reasonable when the trustee is trustworthy, which […] simply means unlikely to act opportunistically" (Möllering 2006: 24). Based on this rational calculation actors decide whether or not to trust. Some key authors who conceptualize trust in the rational choice tradition include Coleman (1982; 1994), Bradach and Eccles (1989), and Axelrod (1984).

For *sociologists*, trust is the cement or social capital of society (Fukuyama 1995; Putnam 2000; for social capital on social networking sites: Grabner-Kräuter 2009) and an important enabler of social interaction (Goffman 1959, 1990; Misztal 2001).

For *philosophers*, trust is an important aspect of the *human condition*; the way human beings are in the world. Trust is closely connected to autonomy (O'Neill 2002a) and it is been acknowledged as a necessary condition for a healthy democratic

state (O'Neill 2002b), built on trustworthiness and institutional integrity (Meijboom 2008). Trust has also been seen as a strategy to deal with vulnerability and uncertain relations with the environment and other people (Baier 1986).

As this quick overview illustrates, the ways in which the concept of trust is applied strongly differs from one context to the other. Rather than deciding which uses of the concept are "proper" or "improper", or trying to come up with one unifying theory on trust, it would be better to aim at developing a perspective on trust which encompasses the complexity as well as the richness of the phenomenon (Simon 2013: 2).

While it is true that developing a complete, all-encompassing theory on trust is rather impossible, we are able to identify a minimal number of key concepts that have to be taken into account when speaking of trust (Möllering 2006: 7-9). These are concepts that reappear in different research domains and therefore can be regarded as central to our understanding of trust.

One of the first things is that we have to be able to identify an *actor* –often referred to in the literature on trust as the *trustor*- who has expectations of the intentions or behaviour of someone else (the *trustee*). We have to be able to distinguish a trustor and a trustee in order to speak of trust (Möllering 2006: 7). The *expectations* the trustor has of the trustee have to be positive and favourable. These expectations do *not have to be conscious* expectations. It is often the case that we have put trust in someone and only after the trust has been breached (e.g. 'I should never have lent her my scarf. She forgot it on the train and now I have lost it') we become aware of this act of trust.

The relevance of trust is due to the principal *vulnerability and uncertainty* of the trustor towards the trustee. The trustor does not know for sure how the trustee will act. He or she can *harm* the trustor (idem: 8). Depending on the room to act giving by the trustor to the trustee, the trustor might be harmed more or less by the trustee. Therefore, the actions of the trustor and the trustee are *interdependent (idem: 8)*. It has to be said that in order to speak of trust, the "social vulnerability and uncertainty have to be *irreducible* (emphasis added)*"* (idem: 8). It is not merely that the trustor does not have all information about the intentions of the trustee, but also that both are actors with a certain amount of autonomy; they have *agency*.

This also means that trust *cannot be forced or guaranteed*. This willingness of

the trustor to be vulnerable, based on the expectation that the trustor will perform a specific action which is important to the trustee, does not imply some "masochistic desire" (idem: 9) to get hurt. On the contrary, this willingness to get hurt is actually the "highly optimistic expectation" that vulnerability is not a problem and no harm will be done (idem: 9). Here, trust differs crucially from other social processes that are about avoiding or diminishing vulnerability, rather than positively accepting it.

Finally, another important aspect to recognize is that the trustor and the trustee are embedded in a *social context* which influences how exactly they can define themselves as actors and enact their agency (idem). Trust is, therefore, not something which is merely a part of an isolated interaction of the trustor and the trustee. The social context, history, and other actors also play a role in the establishing of trust.

1.2 Guido Möllering. Trust: reason, routine, reflexivity

In trying to develop such a rich account of trust, the research done by Guido Möllering (2001, 2006; 2005) –who we already met in the previous section- gave me a head start. Studying his work was of key importance to this book for at least two reasons. First, notwithstanding our different academic backgrounds –he is a professor of organisation and management, I am a starting philosopher of technology interested in philosophical anthropology- both our perspectives on trust show some important points of resemblance. Consequently, his work helped me to sharpen my own conceptual framework. Second, his focus on *the leap of faith* as the core of trust became the starting point for my theoretical analysis of trust in the next chapter.

As to the first, the paradox I identified lying at the heart of the concept of trust is by and large in line with Möllering's approach. He states that "trust is ambivalent" –I would say *paradoxical*- "because it solves a basic problem of social relations without *eliminating* the problem" (Möllering 2006: 6). He also asserts the 'as-if' character of trust. By acting as-if, by creating fictions, uncertainty and vulnerability are not removed – rather, they are suspended. Trust may be based on a fiction, but it simultaneously is also productive by enabling a reality to take shape.

In his 2006 book, '*Trust: Reason, routine, reflexivity*', Möllering convincingly explains that:

> "Trust is an ongoing process of building on reason, routine and reflexivity, suspending irreducible social vulnerability and uncertainty as if they were favourably resolved, and maintaining thereby a state of favourable expectation towards the actions and intentions of more or less specific others" (Möllering 2006: 111).

With *reason, routine, and reflexivity*, Möllering refers to the three major approaches to trust he discerned in the scholarly domain. The first refers to the perspective that trust is foremost a *rational choice*, the second sees trust as a *routine behaviour*, and the last approaches trust as a *reflexive reinforcement*.

As to the first, in the rational choice approach trust has been dominantly defined as a decision of the trustor based on an estimation of whether or not it is likely for the trustee to act in an expected manner and honour the trust that is placed in him (see for example Hardin 2001, 2002, 2006; Coleman 1994; Bacharach and Gambetta 2001). Following the rational choice paradigm, actors are self-interested and trust is the rational outcome of the imperfect estimations of a trustor's perspective on the trustworthiness of a trustee.

With the second approach, trustors, instead of making difficult rational choices in a complex and often puzzling world, base their actions on the things that are given and relatively stable. We all act following certain rules and adopting certain roles which make our actions much more predictable and stable. Trust then has a "taken-for-granted" character.

The third approach is based on the idea that trust can be built up in interaction; it is a process in which actors learn to trust each other (Nooteboom 1997, 2002; Nooteboom and Six 2003). Even when actors do not 'really' trust each other but just take a shot at it and start cooperating, trust may emerge out of this interaction and become 'real' trust. Step by step, trust can be established (Axelrod 1984).

In his book, Möllering (2006: 105-106) makes the argument that all three perspectives highlight important and meaningful grounds for trust, but by classifying trust as merely one of these perspectives, "the concept is stripped of its *unique* explanatory power".

Moreover, all three approaches seem to 'explain away' what trust is rather than to really look for its core, because in the end when do we have enough certainty to act,

how do we deal with the possibility of people acting 'out-of-character' and how do we take that first step in the process of building trust? To Möllering, this *leap of faith* that all three grounds of trust prepare is the essence of trust. Reason, routine, and reflexivity provide us with the grounds from which we are able to take this leap. *Suspending* our quests for more certainty, we can then whole-heartedly take the chance of being rewarded for our trust, or of getting hurt. In that sense, trust is always a "risky business" (Luhmann 1979) and is strongly connected to vulnerability (Baier 1986). *With trust, there is always something at stake.*

Furthermore, Möllering argues that although concepts such as *suspension* and *the leap of faith* are at the centre of understanding trust, they are underdeveloped in the scholarly research. Next to his own work on bringing these concepts to the fore, he also refers to some other scholars who, often implicitly, have built on these notions in developing their account of trust.

One of these scholars is the German sociologist, Nikolas Luhmann (1979). On several occasions in his book, Möllering refers to Luhmann as a scholar who developed "key initial ideas" for his own approach. Möllering (2006: 116) states:

> "…Luhmann argues that trust involves 'a type of system-internal "suspension" (Aufhebung) (Luhmann, 1979, p.79). When actors achieve suspension they treat uncertainty and vulnerability as unproblematic, even if it could turn out that they are problematic. Luhmann (1979) describes trust as 'a movement towards indifference: by introducing trust, certain possibilities of development can be excluded from consideration. Certain dangers which cannot be removed but which should not disrupt action are neutralized' (p.25)."

By analysing Luhmann's work on trust, which focuses on its function to *reduce complexity*, I will take up Möllering's challenge to further explore the leap of faith he identifies as being essential to our understanding of trust (Chapter 2). I will bring together Luhmann's work on trust and the work of the German philosopher Helmuth Plessner (1975) to deepen our understanding of the leap of faith, which I will refer to as the *bridging of a hiatus*. From a philosophical anthropological perspective, I will relate *the leap of faith* or *the bridging of the hiatus* to the ontological distance inherent in human life. Because human beings can, as it were, *from a distance* look at themselves, at others, and at the world around them, they do not fully coincide with

themselves. They always have to bridge this three-fold distance, resulting in the previously mentioned uncertainties, which are inherent in human life. Referring to this external position human beings can take, Plessner (1975) speaks of an "eccentric positionality". I will show how trust is a product of this eccentric positionality and that the 'as if' character of trust should be understood in the light of the ontological homelessness all human beings have to bear.

A caveat. Do not expect to find a clear-cut, final definition of trust in this book. As Simon (2013) already wrote, there is not one all-encompassing definition that one can use to analyse this phenomenon without simultaneously losing the richness of it.

In line with Wittgenstein (2009 [1953]: 67), I have therefore chosen to look for "family resemblances", instead of trying –and inevitably failing- to come up with what is essential to trust. I have focused on a *family of related concepts* that I found chiefly in the work of Luhmann (1988, 1979), which in their interrelatedness shed a light on what trust might come down to. This trust-family consists of the following members: a *familiar world, interpersonal trust*, *confidence* or *system trust*, and the *reducing of complexity*.

Trust as a way to *reduce complexity inherent in human life* can only take place in a *familiar world*. Although trust is a strong strategy to deal with uncertainty, it cannot function in a world, which is not to a certain extent already 'familiar'. A world is familiar when it is based on shared values and norms. People inhabiting this familiar world presume that other people perceive the world in a more or less similar way and take certain aspects of life for granted.

Taking this family of trust concepts to examine the networked era leads to interesting questions such as: how is the familiar world constituted online? Is it still possible to speak of system trust when smart artefacts and Internet services increasingly display pro-active and personalized functionalities? How do human beings deal with the complexity inherent in human life in the networked era?

Next to these related concepts to analyse trust, looking into the work of Luhmann brought me another advantage. Because Luhmann not only considers interpersonal trust, but also studies *system trust* to understand how people can have confidence in large, opaque technical systems, he clears the way for an analysis of trust in cyberspace. As interpersonal trust in a globalized and technologized world is no longer sufficient to reduce complexities, a new kind of trust is being developed,

based on the belief in the complexity-reducing mechanisms of systems such as political systems, air traffic systems, banking systems etc. In the networked era, the Internet can be seen as one of the most dominant systems providing an infrastructure that impacts almost every aspect of daily life.

1.3 e-Trust

Every scholar has his or her final ground, the fundamental prepositions on which his or her analysis is built. Let me be upfront about mine. I basically assume two things: first, if you want to try to understand human beings you have to look at them *in relation to their environment*. Second, all our interactions are *mediated interactions*. The way we look at ourselves, others, and the world around us is influenced by the technologies –or other artificial means- that mediate these relations. Taking into account my background in the philosophy of technology and keen interest in theories of mediation (Plessner 1975; Ihde 1990, 1993; Verbeek 2000; Verbeek 2011b), these two beliefs may not come as a surprise.

In my aim to understand my puzzlement about trust in a world mediated by the Internet, I will therefore be focussing on the way in which *trust and the Internet interact*. Or, to be more precise, *how human beings establish trust through smart artefacts*. Investigating the way in which the relation between man and technology takes shape is actually the focal point of the philosophy of technology (see for a full definition of the philosophy of technology Kaplan 2009). How do Internet-connected devices co-shape our experience of the world and influence the way in which we establish trust? And how do we as trust-giving beings attach meaning to these devices?

From this specific theoretical angle, when reviewing the research on trust and e-trust more specifically, I encountered two difficulties.

First, the dominant paradigm in trust research sees trust as first and foremost a phenomenon that manifests itself on the *interpersonal level* (McLeod 2014; Good 1988) or at least as something that *starts* at the interpersonal level (Kohn 2008). Trust chiefly exists in the interaction between persons (trustor and trustee) and often these persons are seen as rational agents (remember the 'trust as reason' approach) (see for example in the e-trust domain Taddeo 2010a; or McGeer 2004: who

investigates trust between online rational agents who developed 'friendship trust'). As a consequence, the focus on the context in which this interaction takes place or on the technologies mediating this interaction is rather low.

Second, in the literature on *e-trust*, where the technology obviously is part of the analysis, the conceptualisation of the technology remains rather broad. It often depicts the Internet as an 'online world' or 'environment' without really specifying which platform or service one is referring to (see for example: Pettit 2004). Also the devices through which the Internet is being accessed in general fall out of the scope of these analyses.

In line with the *empirical turn* in the philosophy of technology (Achterhuis 2001), I will argue in chapter 4 and 5 that in order to understand the ways in which trust is established in Internet-mediated contexts, it can be more fruitful to look at specific devices and practices than to investigate "The Technology" (capital T intended).

There are of course also exceptions. For example, Kiran and Verbeek (2010) explore a post-phenomenological account of trust and illustrate this approach by looking into empirical cases such as the role of telemonitoring in e-health and the use of prosthetics. Also Coeckelbergh (2012), in his analysis of trust in robots, develops a phenomenological account of trust. In his philosophical analyses of online trust, De Laat (2005, 2012) pays a lot of attention to the different environments online, consequently bringing to the attention the importance of the specific, empirical aspects of the online world for our understanding of trust. And even in research that is predominantly based on a rational approach, the contextual aspects of trust do trickle down. For example, O'Hara (2012), although taking a rational approach to trust, simultaneously emphasizes the messiness of the human world and the increasing integration of people and machines, leading to a 'mediated reality', which poses a challenge to modelling trust in systems.

While the post-phenomenological, mediated perspective is still underdeveloped in the research on trust and the Internet, there are several seminal works in the domain of e-trust, notwithstanding their differing theoretical bases, of which the findings can help to develop such a contextual account of Internet-mediated trust. Unfortunately, it is not in the scope of this research to provide a complete overview of all these works

(For a more comprehensive overview I like to refer to Simon 2013; Ess 2010; Taddeo 2009), however, due to their direct influence on this book, I do explicitly want to list the following works.

One of the earliest philosophical explorations of e-trust was undertaken by Hellen Nissenbaum (2001; 2004). Nissenbaum convincingly argues that trust cannot be replaced by *security online*. While it is true that security is an important precondition for trust to be established, it cannot be simplified to security, because such a shift would damage the creativity, the freewheeling, the political… facilitated by trust (Nissenbaum 2004: 179). She warns that although security is important and even crucial to online activities such as banking and e-commerce, we should be cautious to let this security paradigm take over the nature of cyberspace[1].

Over the years, a couple of special issues have been published also devoted to the subject of e-trust. In 2004, *Analyse und Kritik* published a special issue titled "Trust and community on the Internet. Opportunities and restrictions for online cooperation". Dominantly from an analytic point of view, the contributions focus on the question of if and how cooperation online is possible. Trust is valued as an important precondition for any cooperation. In this special issue "… the potential, the pre-conditions and the limits of the Internet for the emergence of trust and community building are discussed" (Lahno and Matzat 2004: 1). Also in this issue, Pettit (2004) retakes his influential article "The cunning of trust" (Pettit 1995) to argue that a rich account of trust on the Internet is impossible due to the absence of three key forms of evidence: the evidence of face, the evidence of frame, and the evidence of a file of the shared history. According to Pettit, online you miss bodily cues (face) you do not see how people interact with others (frame), and you cannot keep a record of past behaviour (file), consequently, trust between actors merely known to each other in an online context is impossible. In the same issue, also Hardin (2004) shows himself skeptical about the possibility of trust online. In line with

[1] In their article "The case of online trust" Turilli et al. (2010) seem to suggest that Nissenbaum rejects the possibility of trust online. However, in my reading Nissenbaum only rejects the possibility of obtaining *trust through security online*. The obstacles to trust online Nissenbaum identifies are not unsolvable nor that fundamental to dismiss trust and replace it for security online altogether. They do invite the rethinking of the balance between security and freedom (the latter which is nurtured by trust).

Pettit, he finds the relationships online to be too thin to foster cooperation.

In 2010, *Knowledge, Technology & Policy* published an issue called "Trust in Technology", guest edited by Mariarosaria Taddeo (2010b) revolving around the claimed problematic nature of trust in technology invoked by the widespread development of ICTs. From a diverse disciplinary background, the scholars contributing to this issue addressed –amongst others- the social responsibility of Internet Service Providers and hosting companies (Cohen-Almagor 2010) and the importance of value-sensitive design for trust online (Vermaas et al. 2010). This issue also contains a thorough overview of the literature on e-trust (from an ethical perspective) (Ess 2010) and the previously-mentioned phenomenological account of trust by Kiran and Verbeek (2010).

In 2011, Taddeo, now together with Luciano Floridi (2011), edited an issue of *Ethics and Information Technology* called "The case for e-trust: a new ethical challenge". The issue promotes a *hybrid approach*, combining conceptual analysis with empirical data. It especially draws on models developed in the domain of artificial intelligence (AI). I would also like to highlight the contribution of Grodzinsky et al. (2011) who provide a thorough review of the research on e-trust and specifically focus on the way in which the *human actors* design, introduce, and use the artificial agents. Also the article of Pieters (2011) who focuses on the role of the explanation provided to the users for assessing the trustworthiness of systems is highly relevant. In chapter 5 of this book, the importance of users being explained how systems function returns.

Also in 2011, a book was published called "Trust and virtual worlds. Contemporary perspectives" edited by Charles Ess and May Thorseth (2011). One of the starting points of this book is that in Internet research –and therefore in e-trust research as well- one has to look beyond the online-offline divide, a point I will elaborate in the next chapter. Consequently, the online or virtual environment is of utmost importance for analyzing trust[2].

[2] I especially want to mention the interesting academic discussion that unfolded between Taddeo (2011) and John Weckert (2011) in this volume concerning the question of whether Artificial Agents (AAs) have some sort of moral agency and to what level human beings are defined by this moral agency which AAs may or may not possess. Where Taddeo provides a rationalistic, Kantian account of trust, Weckert suggests that trust is more like a hermeneutic framework influencing how we act and perceive

1.4 The Internet

Next to coming to grips with what trust entails, also the Internet itself needed to be conceptualized. The Internet is not a one-dimensional technological innovation. Although it is a 'global network' and it has entered virtually every domain in life, it cannot be approached as a heteronomous system. One cannot say anything meaningful about trust in or on 'the' Internet, because what the Internet comes down to simply is not clear-cut.

The conceptualization of the Internet in this book is based on the layered design of the Internet itself, which will be discussed in chapter 4. In order to take into account those aspects of the Internet, which dominantly influence the formation of trust, I have discerned four cornerstones of the current Internet: *construction, curation, codification* and *context*.[3]

These four Cs I deem to be crucial to understand trust mediated by current smart artefacts and Internet platforms. I explicitly say 'current' as the Internet is still a system 'under construction' and it can be questioned if it will ever leave this state. These four Cs, therefore, may still change over time. However, for now they constitute the conceptual lens I will use to analyse trust in the specific cases presented at the end of this book.

---

the world.

[3] I do not pretend to have come up with a thorough definition of what the identity of the Internet then in fact is. Based on the conceptual analysis of trust that I develop in chapter 2 and 3, I have mainly brought to the fore those aspects of the Internet, which are key to understanding trust in the networked era. Without any doubt, I therefore have left out fundamental aspects of the Internet, such as the physical tubes of the network (Blum 2012) or its cultural imbeddedness (Deibert 2008; Deibert et al. 2012a; Deibert et al. 2010). These aspects of the Internet would definitely deserve their place in a more general definition of the Internet. They are, however, left out here because they are of less importance to understand the way in which users trust or have confidence in and through their devices.

1.4.1 Context

Context refers to *users experiencing the world* through their smartphones, *interacting with others* on social network sites, making use of services of *information intermediaries* such as Google or Facebook. As we have seen in the short overview on the academic literature on trust, it is generally accepted that trust arises on this interpersonal, micro-level. Based on insights deriving from theories of mediations, I will show how smart artefacts invite users to act and interact in certain ways and how this may enable or challenge trust. Because of their easy-to-use design and proactive and personalized services, smart artefacts generally persuade users to have confidence in the functioning of the artefact. Consequently, users may tend to "forget" the mediating workings of the smart artefact, assessing their interactions as direct instead of mediated. This may lead to the utopian belief that interactions can be based on merely interpersonal trust, without taking into account the system component of the interaction.

Although the context level generally can be seen as the starting point for the analysis of the way in which users are experiencing trust, it should never be the sole focal point. Due to its networked ontology, Internet technology is far more complex than what merely becomes present or visible in the phenomenological experience on the micro-level. The other Cs should therefore also be taken into account when analysing trust mediated by smart artefacts, platforms, or online services.

1.4.2 Curation

Curation stands for the actors who *govern* the Internet. This can be, amongst others, governments, international organizations, private parties, and/or civil society organizations. These actors, on the one hand, contribute to the familiar world by maintaining the infrastructure, developing standards and protocols, designing user-friendly interfaces and providing users with personalized services. On the other hand, these curators may endanger the familiar world when they make use of the Internet for their own interest. Remember the NSA altering the Internet infrastructure to enforce backdoors in its technical backbone. Or, think of information intermediaries such as Google or Facebook collecting all sorts of data from their users to sell to third parties. Due to the networked ontology of smart artefacts, a new relation of mediation occurs between curators and users. Users become increasingly *visible* to curators in

an *invisible* manner due to the data that is collected by smart artefacts and Internet services. Although these relations of users and curators almost never enter the phenomenological experience of users, they may have an impact on trust in the services provided by the curator and in the smart device as such.

1.4.3 Codification

With codification, I explicitly refer to the rules and regulations put forth by the curators. Through their terms and conditions, but also by the way they design the device and set up the protocols regulating their platforms and services (Lessig 2006), they pre-sort certain behaviour in their users. As we have seen, trust is only possible in a familiar world. A legal framework, technical protocols, corporate rules and regulations, and implicit norms and values contribute to this familiar world. However, when these rules and regulations are susceptible to change, losing their self-evident character, it becomes more difficult to trust as the complexity (that has to be reduced) rises.

1.4.4 Construction

With construction, I refer to the design of the artefact. Smart artefacts generally have hardware as well as software elements. The often-slick design of smart devices invite users to assess them in a one-dimensional way: do they or do they not work.

The ways in which algorithms perform the collecting and mining of data, the way these data are interpreted and used to influence the user's experience are difficult to ascertain. *Devices are designed to be used, not to be understood*. Therefore, it becomes increasingly difficult for the average user to come to grips with the impact of his or her interaction with the smart device or online service. Even the programmers themselves do not even always understand how their coding ends up delivering certain outcomes (Aupers 2002).

1.4.5 A two-fold use

The analytic framework - context, curators, codification, and construction- can be put to use in two different manners. First of all it is a *descriptive* tool. By taking into account these four layers and not just limiting ourselves to the *interpersonal level* or

context level, I can do justice to the influence of the *networked character* of the smart artefacts and services.

Second, this conceptual lens can also be part of more *evaluative practices*. Analysing specific cases of trust mediated by internet-based artefacts through the proposed conceptual lens of the 4Cs, can help ethicist and policymakers to determine whether or not trust is being given in a justified manner. Moreover, the 4Cs enable them to pinpoint on which level measures might be taken to strengthen trust.

1.5 Outline

This is how this book is composed. Overall, there are three parts. The first part contains chapters 2 and 3, which focus on trust. The second part, chapters 4 and 5, focusses on the Internet and how trust is established there, and the third part, chapters 6,7, and 8, consists of two cases and an analysis of a more encompassing trend to illustrate in a more detailed manner how trust is shaped in cyberspace. The book ends with a short epilogue.

1.5.1 Chapter 2

In the following chapter, I will first study trust on the ontological level, meaning that I will dive into the question of what trust means to human beings and how it relates to the way human beings are in the world. Central to this analysis will be the work of Niklas Luhmann who focused on the functionality of trust: a strategy to reduce complexity. By replenishing the work of Luhmann with insights from the work of Helmuth Plessner, I will show that although all living things have to deal with complexity, the complexity human beings have to cope with is even more radical due to their eccentric positionality. I will also pay attention to the "three anthropological laws" Plessner has formulated. These laws will return as an instrument to investigate trust on the micro-level (context) in chapter 5.

1.5.2 Chapter 3

In chapter 3, I will look at trust from a sociological/historical perspective and argue that in modernity and late modernity, interpersonal trust increasingly had to make

way for system trust due to the arrival and dominant presence of large and opaque systems in society. Elaborating on the work of Luhmann and Giddens, I will focus on system trust and confidence and how these new forms of trust relate to interpersonal trust. At the end of this chapter, the whole family of trust notions will have been discussed.

1.5.3 Chapter 4

Central to this chapter is the Internet as infrastructure and it revolves around the question if and how the Internet can function as a familiar world. I first retell the history of the Internet by focusing on the role trust plays in the collaboration between the founding fathers of the Internet. Next, I look at the layered structure of the Internet and show how this layered structure has been a fruitful starting point for conceptualizing the Internet not only for me but for other scholars as well. I will argue that curators such as governments, private parties, and non-profit organizations on the one hand contribute to the Internet as a familiar world by developing rules and regulations, protocols and by maintaining the infrastructure. All these actions help to make the Internet a reliable infrastructure and environment where trust can thrive. On the other hand, these curators also use the Internet as a tool to set and reach their own goals. This may conflict with the earlier goal of creating a reliable and steady Internet infrastructure and environment. As a way of conclusion, I reflect on the transition of the Internet as an *open* environment to the Internet as a more controlled and *closed* environment.

1.5.4 Chapter 5

Where in chapter 4, I take a macro-perspective on the Internet, I now move to the micro-level, where the experience of users is central to the analysis. In this chapter the context level of the Internet is the main point of departure for my study of trust. By replenishing the three anthropological laws of Plessner with insights deriving from the philosophy of technology and constructionism, I analyse the current networked era and the way in which trust is being established.

I think of this chapter as being the heart of this book. It not only shows how interpersonal trust starting at the context level is intrinsically connected to the other

layers –curators, codification, and construction- consequently transforming into interpersonal system trust. It also is where I develop, based on the building stones crafted in the earlier chapters, my take on the specific challenges posed by the networked era on the building and maintaining of trust. These findings will be further substantiated by two cases –in chapter 6 and 7- and an exploration of online personalization as a dominant and recurring theme in the networked era – in chapter 8-.

1.5.5 Chapter 6

The first case starts off with the often-cited example of the not-returned hotel key, by Bruno Latour. In this example, Latour investigates how by adding an extra weight to the key, the imperative "bring back your hotel key when you leave the hotel' is being inscribed in the artefact. This new 'actant' alters the interaction and trust relation of the actors involved. This chapter shows how due to technological changes the hotel key transforms, simultaneously also transforming the trust interaction of the actors involved.

1.5.6 Chapter 7

The second case will focus on the online platform AirBnB. The aim of this platform is to connect people who want to rent their house or a room in their house and people who are looking for a place to stay while traveling. AirBnB is a prime example of the shared economy movement. This movement wants to change the economic system, which is based on *ownership,* to a system that is based on *access*. In short, it should no longer be important to own things but to have access to them. By cutting out the middleman, in this case the hotel owner, old forms of trust based on reputation and reciprocity can be reinvented. In other words, trust is solely something that is part of the context level. I will, however, argue that this view is based on a utopian belief that technology can restore the indirectness of the ontological distance that is at the centre of every interaction. I will show how not only the construction of the platform is shaping the building of trust, but that also the interests of the curator – AirBnB, a privately owned company - and the codification of AirBnB and several state actors should be taken into account when analysing trust between users of AirBnB.

1.5.7 Chapter 8

This chapter not so much focuses on a specific case but on a dominant tendency that affects a variety of practices in the networked era: personalization. Increasingly, online services are tailored to a specific profile of a user based on collected data that are mined to discover patterns of behaviour. In this chapter, I will argue that although this may lead to a *personalized world* very familiar to the users themselves as it seemingly fits their individual preferences, it is not a familiar world where trust can easily thrive. A world dominantly based on personalized preferences does not necessarily also function as a familiar *background of shared norms and values*. Instead it confirms a person in her initial beliefs and convictions, raising the bar for projecting herself into the position of someone else with different beliefs. The distance inherent to every interaction becomes more difficult to bear if this situation is not reflected- even contradicted- in the personalized environment.

2

The concept of trust.

Niklas Luhmann meets Helmuth Plessner

To understand the meaning of trust in human life, we have to ask the preliminary, transcendental question: "what makes it even possible for human beings to experience such a thing as trust?" This chapter will focus on what is (often implicitly) presupposed when we talk about trust but which in fact should be acknowledged as its core explanatory force: *bridging the hiatus* inherent in human life. This *hiatus* or *distance* lies at the core of human existence. The uncertainty about, amongst others, the *action of others* and the way in which *the future will unfold*, results in a hiatus between human beings and the world they inhabit. Fundamentally, trust is about dealing with this hiatus. The hiatus cannot be resolved or taken away. The positive expectations we hold towards others enable us *to act as if the future is certain*, as if the hiatus is not there. Trust cannot erase this ontological distance, but neutralizes the anxiety, as it were, for a potential bad outcome.

In the introduction, it was already stated that Möllering (2006) argues that the majority of research being done on trust chiefly focuses on the *grounds* for trust – which he clusters around three themes: reason, routine and reflexivity- instead of analysing how these grounds *interact* with the bridging of this hiatus –which he refers to as "a leap of faith"- that occurs when people act on trust. He argues that more effort should be put into understanding why and how this *leap of faith* or the *bridging of the hiatus* takes place and how, in fact, the act of bridging is essential to reinforce and consolidate the grounds on which trust is made possible in the first place.

In this chapter, I will take on this challenge by conceptualizing trust as the bridging of a hiatus, consequently, connecting it to the ontology of human beings. The work of Niklas Luhmann (1979) on trust will be our starting point. Not only is his theory, which first and foremost focuses on *the function of trust*, very influential in different research domains, it also already incorporates several elements to help us understand the connection between trust and the bridging of the hiatus[4]. Although more than informative, Luhmann's theory on trust is not sufficient to grasp the consequences of the hiatus at the heart of human existence.

In the second part of this chapter, I will therefore make use of some key insights deriving from philosophical anthropology, and more specifically from the work of German philosopher Helmuth Plessner, to replenish Luhmann's account of trust. It has to be noted that although Luhmann himself positively refers to Plessner in some footnotes in *Trust,* he later in his career explicitly distanced himself from philosophical anthropology. It therefore is questionable if Luhmann himself would appreciate this connection between his work and that of Plessner. Nevertheless, I will show that bringing together Luhmann's trust account and Plessner's philosophy gives us a better understanding of trust as a way of dealing with complexity caused by the hiatus inherent in human life.

2.1 Introducing Niklas Luhmann

Niklas Luhmann (1927-1998) was an influential German sociologist who published more than 50 books and 300 articles on a variety of topics between 1964 and 1997. In 1997, he published his main work called *Theory of Society* which according to Arnoldi (2001: 2), can be seen as "a synthesis of his sociological work and his general system theory", the latter being the accomplishment he is most famous for. However, an earlier work called *Social Systems* (1995) should also be mentioned because of the

[4] Möllering (2001) argues that in fact some key notions of Luhmann's theory can and should be traced back to the work of Georg Simmel. Also Misztal (1996) has acknowledged the importance of Simmel's work on trust. However, since the work of Simmel is hardly ever directly recognized by other trust scholars and in general is only known through the work of Luhmann, I will mainly focus on Luhmann's theory.

introduction and thorough analysis of some of his key concepts (Paul 2001).

Unfortunately, it would lead us too far to give a detailed description of his whole oeuvre, especially because our main interest concerns just one book called *Vertrauen* published in Germany in 1968, translated into English and bundled together with another work called *Power* in 1979[5]. *Trust* belongs to the early period of Luhmann's work and stands rather apart from his later work. It is, next to one article published in 1988, his only work that is completely dedicated to the analysis of trust.

In this early stage, Luhmann was mainly interested in expounding and justifying the concept of the *reduction of complexity* (Poggi 1979). His chief assumption is that there are *empirical systems* that have to establish their place in an environment far more complex than their own internal structure (Paul 2001). Systems try to reduce this external complexity by building internal or ordered complexity. Trust, next to other options such as a legal framework and the use of contracts, is a way to reduce this complexity in social life.

Although key notions in Luhmann's thinking such as *autopoesis* and *double contingency* are only elaborated later on in his career, his most important thoughts on how society should be conceptualized and analysed are already set in place in *Trust*. Where useful to understand the specifics of his trust account, I will try to embed these theoretical notions.

Possibly, the reason for the popularity of Luhmann's work on trust in the social sciences and beyond can partly be found in the preface to his book *Trust*. In the one-page introduction, he states that far too often sociology makes use of concepts coming from common usage or from other disciplines such as ethics. However, before starting the dialogue with other disciplines, sociology should strive at formulating a theory of its own. He sees it as his task, difficult but not impossible, to bridge the gulf between theory and empirical work. His effort to thoroughly think through the concept of trust and relate it to a network of grounding ideas has been an important source of inspiration for many trust-scholars.

As the first sociologist to provide a conceptual framework for understanding

[5] Although the official title of the English translation is *Trust and Power*, I will refer to it as *Trust*, simply because Power is another work that stands on its own.

trust (Misztal 1996), Luhmann's theory can be indicated as "the starting point for the modern approach to trust and its cognate concepts" (Taddeo 2009: 3). He provided trust-scholars "with what is no doubt the richest set of insights and understandings of trust currently available" (Seligman 1997: 18).

Because of his functionalistic approach, leaving out normative connotations, Luhmann's conceptualization of trust is open to different realizations, which might be the main reason why he has "inspired trust researchers across a broad range of disciplines" (Möllering 2006: 5). Despite the abstract nature of his theoretical framework and his rather inaccessible style of writing, the level of abstractness in fact turned out to make his theory applicable and useful in a variety of research domains. Luhmann himself has illustrated this by using his theoretical concepts as a basis to write about a wide range of topics such as love, risk, and mass media.

2.2 Luhmann's theory of trust

In his book *Trust*, Luhmann launches the hypothesis that trust is a manner to reduce the complexity of the world. Simply put, the environment of every living system contains more possibilities than the system itself can actualize. Therefore, every system has to make selections in order to persevere (Bednarz 1984). Although all systems have to deal with the complexity of their environment, only human beings are conscious about the *world's contingency*. This awareness poses upon human beings the burden of choice. They have to make selections, because they cannot accept all the possibilities the world inhibits. Because these selections cannot be made based on sufficient evidence –in the end, we are not sure about what tomorrow brings nor can we foresee all the actions of our fellow-human beings – trust is "a blending of knowledge and ignorance" (Luhmann 1979: 25). It is to act *as if* the future is certain.

In the following paragraphs, I will give a compact outline of Luhmann's theory on trust (1979), based on the key concepts: *complexity, risk, familiarity, roles, confidence,* and *system theory*. Another important concept in Luhmann's theory, namely *system trust* will be elaborated in the following chapter. For now, I will mainly focus on *interpersonal trust*.

Furthermore, I will pay attention to the phenomenological influence on his

conceptualization and the connection this conceptualization might have with philosophical anthropology (Poggi 1979). Principally in his early work, it is clear that Luhmann is influenced by philosophical anthropology as well as by phenomenology, specifically by Husserl (Paul 2001).

I will argue that the connection between his approach on trust and some basic ideas originating from philosophical anthropology is very fruitful to grasp the core aspect of trust. It has to be stressed that Luhmann himself, especially in his later work, never explicitly made this connection. Even more so, *he categorically distances himself* from the philosophical anthropological discipline. Therefore, in the second part of this chapter, by taking Luhmann's concept of trust and relating it to insights deriving from philosophical anthropology, we move beyond what Luhmann aimed at in his work on trust. I will especially look at the work of Helmuth Plessner, one of the founding fathers of philosophical anthropology, to explore more in depth the hiatus central to trust. Finally, in line with Möllering (2001, 2006) I will then make the argument that up and foremost trust is defined by the suspension of uncertainty and by the attempting to bridge the gap.

2.2.1 Complexity

Trust is a *basic fact of social life*. Without some sense of trust, it would be impossible to get up in the morning. We would be overwhelmed by the idea of all the possible turns fate could take when we leave the warm safety of our bed. Without the conviction that others will act *in character* and that the ways of the world will not change overnight, we would be paralysed. It would become impossible to do the things we want to do, to believe in our own ability to act. Our lives would stagnate, as if time had frozen and we had gotten stuck in the bubble of an everlasting present.

Trust, as Luhmann defines it, is a way to *reduce complexity* that is inherent in the world we inhabit. To human beings the world is open and has no boundaries. It excludes no possibilities and, therefore, it is always more complex than the systems which are living in it. While it is true that all systems, whether made of stones, animals, or human beings, live in a selectively constituted *environment*, human beings are the only ones who "are conscious of the world's complexity and therefore of the possibility of selecting their environment" (Luhmann 1979: 6). The world as a

whole is the "universal horizon of all human experience" (idem: 5) and, from a more Husserlian perspective, human beings have to focus their attention against this background of all other possibilities over and over again (Arnoldi 2001: 5). Consequently, human beings are aware of the fact that things could have been different. This awareness of contingency implicates uncertainty about "a future characterized by more or less indeterminate complexity" (Luhmann 1979: 15). Trust is necessary to reduce this complexity and to tolerate the risks and uncertainties which accompany it.

This complexity enters the human world by means of two elements: the *other* and *time*, revealing a social and a temporal level in the complexity of the world.

On a social level, this complexity is connected to the "subjective –I-ness" of other human beings (Luhmann 1979: 6). Complexity comes into the world because of the possibility of unanticipated actions by other human beings, constituting a source of insecurity. For Luhmann, other *egos* are in fact black boxes. After all, other people are, to a certain degree, free to see things differently, to have their own perspective and understanding of the world. We do not have direct access to the others we interact with, so, in consequence, they can act in unforeseen ways.

On a temporal level, human beings are aware of the discrepancy between possible futures and the one future that will become reality. In the present they have to cope with an over-complex future. Therefore, trust has also to do with anticipating the future. Trust is "to behave as though the future were certain" (idem: 10).

Because of this future-orientatedness, trust is closely tied to *expectations*. Possible fulfilment, if it happens at all, only appears after the action has taken place, while commitment has to be there beforehand. "This problem of time is bridged by trust, paid ahead of time as an advance on success for a certain time" (idem: 25). To trust is to have *positive expectations* concerning the future actions of other actors. It is not about having control over events, rather it is *a move to indifference* (idem: 25). It is an attitude of becoming indifferent to the many, very different ways the future can unfold. With the act of trust, certain options can be set aside and dangers which cannot be removed are neutralized. The act of trust enables one possible future to stand out, blending all other options into the background. This ability to trust is not given and has to be learned. Its process of learning already starts in the earliest stages of life in interaction with family members and steadily expands to other actors and

systems in the broader social world.

2.2.2 Risk

Trust is a *risky investment* (Luhmann 1979: 42) because there is, by definition, always something at stake for the trustor. The goal the trustor wants to achieve cannot be reached without the interference of a trustee. To trust someone means to be vulnerable and dependent on the action of a trustee who in his turn can take advantage of this situation of vulnerability and betray the trustor. Luhmann takes it even a step further and explicates that:

> "Trust therefore always bears upon a critical alternative, in which the harm resulting from a breach of trust may be greater than the benefit to be gained from the trust proving warranted." (Luhmann 1979: 24).

Luhmann's emphasis on the fact that trust has to make a difference in a decision – otherwise we have merely hope - does not entail that trust also has to be rational. Trust can happen thoughtlessly and carelessly, almost completely based on routines. Rationality for Luhmann does not refer to the decision-process as such, as is the case in rational choice theory, but to the fact that trust is functional for the system to reduce complexity (Möllering 2006). Therefore, Luhmann should not be categorized as a disciple of rational choice theory or any other rational approach to trust as some scholars seem to argue (Coeckelbergh 2012; Taddeo 2009). Luhmann (1979: 88) himself clearly states:

> "Trust is not a means that can be chosen for particular ends, much less an end/means structure capable of being optimized [...] Trust is, however, something other than a reasonable assumption on which to decide correctly, and for this reason models for calculating correct decisions miss the point of the question of trust".

That being said, Luhmann nonetheless recognizes the importance of reasons to support an act of trust. However, he interprets reasoning rather as a way of upholding self-respect and justifying oneself socially than as building a sufficient basis for trust (idem: 26). With Lewis and Weigert (1985: 976), who are influenced by Luhmann, we can conclude that "trust begins where prediction ends".

2.2.3 Familiarity

While it is true that the complexity of the environment can be overwhelming, it is equally true that we live in a *familiar world*. We take the presence of the world, our fellow human beings and the objects we encounter for granted. In everyday life, we do not doubt their existence. Moreover, we expect to see and experience the world in a similar way as our fellow human beings do. They are, so to speak, "presupposed and co-experienced" (Luhmann 1979: 18).

Trust can only take place in a familiar world in which existence is already structured in a pre-reflexive way. Our experience of the world automatically entails the intersubjective constitution of meaning. "There is no differentiation in the operation of constituting meaning and world, which brings everybody together in a diffuse consensus" (Luhmann 1979: 18).

As long as our fellow human beings do not shatter this shared worldview and are only perceived as objects inhabiting this familiar world, trust is redundant. However, when another actor appears in the trustor's consciousness, Luhmann speaks in line with Husserl of an "alter ego", where she becomes, due to her freedom to act, a source of complexity. It is through this alter ego's mediation of the world that "man's environment becomes man's own world" (idem: 7).[6] Simultaneously, by presenting us with other perspectives of the world, the alter ego makes us aware of the world's horizon of infinite possibilities. As we will see later, this familiar world resembles the *lifeworld,* a central notion in *phenomenology*.

2.2.4 Roles

In addition to this familiar world, which functions as a precondition for trust, Luhmann also pays attention to *role-taking* and *self-presentation* as essential in building trust. To earn trust, actors have to take part in social life and be able to absorb the expectations of others into their own self-presentation (Luhmann 1979: 62).

[6] This is the first time Luhmann (1979) refers to Helmuth Plessner in the notes.

He frequently refers to symbolic-interactionist sociologists such as Herbert Mead (1934; 1938), Harold Garfinkel (1963), and Erving Goffman (1959) to explain that trust rests on the assumption that people *act in character*. In an interaction, actors are signalling and detecting important behavioural cues such as the definition of the situation, social status, and intentions (Vanderstraeten 2002). However, for Luhmann trust does not just come down to a trustor who is expecting the trustee to conform to her role-repertoire. On the contrary, trust entails that eventually these general expectations are replaced with expectations tailored to the specific capabilities of the trustee.

Role-predictability differs from the predictability that is part of rationalist theories such as rational choice theory and standard theories of economics. Where in the latter theories predictability is understood as a calculative indicator of trustworthiness, in the interactionistic view this predictability is a stepping-stone for trust, shaped by unquestioned routines and procedures (see also Möllering 2006).

Trust itself is for Luhmann an act of self-presentation. "People and social systems strive to draw a consistent picture of themselves and make it socially accepted" (Luhmann 1979: 81). To trust is to assume that a certain trait of behaviour will fit meaningfully with one's own expectations and patterns of life (idem 1979: 71). Or in the words of Henslin (1967: cited in Möllering 2006: 68):

> "When an actor has offered a definition of himself and the audience is willing to interact with the actor on the basis of that definition, we are saying that trust exists."

This emphasis of Luhmann on self-presentation to build trust entails that first and foremost trust is a matter of representation. It has not so much to do with the actual characteristics of the trust relationship rather than with the beliefs people attach to it. Or more precisely, with the beliefs people hold about other people's beliefs on the matter (Misztal 1996). It therefore becomes apparent that trust is a 'risky business' not only for the trustor but also for the trustee.

> "Anyone who has been around for some time is known, has trusted and enjoys trust, is thus entangled with his self-presentation in a web or [sic]

> norms which he himself has helped to create, and from which he cannot withdraw without leaving parts of himself behind" (Luhmann 1979: 63).

Because in every action people disclose more information about themselves than they intended or even are aware of, every appearance presupposes a minimum amount of trust (Luhmann 1979: 40). Every actor has some basic trust that the other will not misinterpret the performance and that she will fill in the informational gaps and inconsistencies all communications entail.

For Luhmann, trust is built, step-by-step, in the process of interaction. Trust is a learning process which he also refers to as the "principle of gradualness" (Luhmann 1979: 41). In the starting phase, the stakes will not be high: people help each other with small tasks or display trustworthy behaviour for example by returning a forgotten scarf. Only after a basic form of trust has been established, it can be tested more thoroughly.

Luhmann describes several elements, which could stimulate the deepening of trust. I will focus on the two most important ones: *freedom of action* and the presence of a *risky investment*.

First, it must be possible to attribute actions to a person in order to judge her trustworthiness. A person's trustworthiness cannot be judged in a situation where she is *forced* to act in a certain manner. What counts as a 'free action' is often determined by *social expectations*. Luhmann illustrates this with an example of an *employee* whose actions are a result of a direct order of a supervisor. These actions do not really deepen trust because they are not perceived as being 'free'. Therefore, if the employee wants to show herself as trustworthy, she has to act *beyond* what is generally expected of her. On the other hand, the *supervisor* is generally perceived as an actor who has the freedom to make her own decisions –even though in reality these decisions are often pre-sorted by corporal structures. Consequently, her actions are more likely to be regarded as tokens of trustworthiness.

We will see in the second part of this book that attributing an action to a person might become problematic when technologies are involved. Because, who is responsible for a plane crash? Who is responsible for the leaking of data? Or as Luhmann puts it:

> "…(t)he outcome of any complex technological process […] appears to be relatively impersonal. The greater the combination of recognizable causes, the more difficult it becomes to isolate who originated the action" (Luhmann 1979: 41).

Second, Luhmann stresses that especially in situations where there is a significant risky investment involved and it is possible and even desirable for the trustee to abuse the trust invested, one can really test and, eventually, build trust. When the stakes become high and the interests of the trustor do not align with those of the trustee, trustworthiness might really be put to the test. Therefore, especially "supererogatory performances", which are performances that move beyond mere duty and rule-following, increase the possibility to deepen trust (Luhmann 1979: 43).

Thus, where Luhmann on the one hand stresses that trust can only take place in a *familiar world* made possible by –amongst others- a social structure with clear roles and rules of interaction in order to temper the radical complexity human beings encounter, he on the other hand also emphasizes the importance of *risk-taking* and the "deliberate imprudence and deviance on the part of the actors" (Möllering 2006: 88) to deepen trust.

It, however, has to be noted that this *risk-taking perspective* of Luhmann is generally considered rather extreme. Many authors argue that mere positive experiences, preferable consistently recurring, are sufficient as a foundation for trust (idem: 88). Nonetheless, the basic ideas that trust is often *process-based* (Nooteboom 1997, 2002; Zand 1972: cited in Möllering 2006: 85-87) and always involves some sort of *risk-taking* (Sztompka 1999; Hardin 2006) are widely shared amongst trust-scholars.

2.2.5 Confidence

In an article in 1988, Luhmann elaborates the distinction between *confidence* and trust in relation to familiarity (also see: Jalava 2003). Confidence is what I would refer to as the *default setting*. You are confident that your expectations will be confirmed and that things turn out as anticipated. You do not call in to question the security of the bridge you cross, the honesty of your husband, or the value of your money in the bank. You ignore the possibility of disappointment because the

alternative would be "to live in a state of permanent uncertainty and to withdraw expectations without having anything with which to replace them" (Luhmann 1988: 97).

Considering the fact that familiarity, confidence, and trust can be seen as "different modes of asserting expectations" (idem: 97) confidence is situated *between familiarity and trust*. Where familiarity structures the world by ignoring its contingency, confidence and trust both have to do with expectations that can turn out to be disappointments. However, where trust requires previous engagement and a certain *element of choice*, confidence is characterized by *not considering alternatives*. In the case of confidence, disappointment will lead to *external attribution* (it is the fault of the government the bridge collapsed because there was not enough money reserved to maintain it). In the case of a breach of trust, there will be *internal attribution* (I should never have trusted her taking care of my child). Where trust has to do with considering other options and is linked to a situation of risk, confidence is about putting aside possible alternatives.

2.2.6 Systems theory

As stated in the introduction of this chapter, Luhmann is known for his work on systems theory. Also in *Trust*, he conceptualizes trust with the aid of rather abstract terms deriving from systems theory.

> "The objective world is more complex than any system: it comprises more possibilities than the system itself provides and can realize. In this sense the system exhibits a greater degree of order (fewer possibilities, less variety), than the world. This discrepancy in the degree of order, as already indicated, is offset through the system developing a 'subjective' image of the world. That is, the system interprets the world selectively, overdrawing on the information which it possesses, reduces the world's extreme complexity to an amount of complexity to which it can meaningfully orient itself, and so structures the possibilities of its own experience and action" (Luhmann 1979: 32).

Although it is perfectly feasible to understand Luhmann's trust account without making use of the language of general systems theory, this quote imbedded in the

systems theory approach is illuminating for three important reasons. First, this quote discloses the *dialectic manoeuvre* that occurs in the act of trust. Second, it can be interpreted in such a way that it reveals the common ground shared with *philosophical anthropology*. Third, this quote illustrates the influence of *phenomenology* on Luhmann's conceptualization of trust.

As to the first point, a system does not eliminate trust but reduces it "to an amount of complexity to which it can meaningfully orient itself" (Luhmann 1979: 32). Complexity is made manageable and at the same time is preserved by it. Consequently, in the act of trust, as a manner to reduce and manage this complexity, a dialectic manoeuvre occurs. Complexity is taken in, absorbed by trust and internally transformed into an "ordered complexity", maintaining uncertainty in a bearable form. The ontological structure of a system is, as it were, defined by complexity.

As to the second point, an essential element in philosophical anthropology is the innate relation between organism and environment. In the quote above, Luhmann defines this relation in terms of the *difference in degree of complexity* between system and environment. The world always contains more possibilities than a system can realize. The main function of a system is to reduce this complexity. Because a system is open in the sense that to survive it has to interact with its environment, the primary act to fulfil this function is "…to establish a border which filters the environment for the system" (Paul 2001: 381). The reduction of complexity takes place "through the stabilization of an inner/outer difference" (Bednarz 1984: 58). The outer complexity of the world forces the system to make selections and establish an inner-simplified complexity. Or as Poggi (1979: 4) formulates it: "The trick, at any rate for living and social systems, is to manage complexity without being overwhelmed by it or entirely sacrificing it". This ontological necessity of a border and border traffic for a system to be able to exist in an overly complex environment is one of the central themes in the work of Helmuth Plessner and it will be of great importance in the further conceptualization of trust.

Finally, in this quote hints the influence of phenomenology. *Phenomenology* or the *science of phenomenons* is a 20th century philosophical discipline, often traced back to philosopher Edmund Husserl whose rallying cry was "to the things themselves".

Apart from scientific knowledge and its strict apparatus, he and other phenomenologists are concerned with the description of how things, other human beings, and the world show themselves from *a first-order perspective* in everyday life.

The focal point in the phenomenological description is a so-called *intentional consciousness*. *Intentional* refers to the fact that consciousness is always *conscious of something*. In the *act* of intentionality, objects do not 'just' appear but they show themselves under a certain perspective. In the intentional act of consciousness, meaning is constituted. This constitution of meaning happens against the background of what is called a *lifeworld*. A lifeworld refers to an often-unreflected background consisting of meanings and beliefs that ground everyday interaction. In addition, Husserl also claims that it is part of the structure of intentional consciousness to assume that other human beings more or less perceive the world in a similar manner. Garfinkel (1963 in Möllering 2006: 56) speaks of a "common-sense world" and Schütz (1967 [1932]) emphasizes the necessity of this taken-for-grantedness in all social interaction.

Taking into account this -rather brief- description of phenomenology, it nevertheless becomes clear that the manner in which Luhmann describes systems as actors that interpret the world by drawing subjective images to orient and structure their actions, resembles Husserl's concept of intentionality. For Luhmann, social systems are systems of communication that 'make sense' of their environment. This element of 'sense-making' he shares with the phenomenological tradition. Moreover, the earlier mentioned familiarity resembles Husserl's lifeworld. For Luhmann trust can only take place in a familiar world. Although other human beings are a source of complexity, this cannot be as absolute as to doubt the existence of some minimally shared worldview. This kind of complexity would be a paralysing form of complexity, one that cannot be tamed.

It is important to understand that although Luhmann is influenced by phenomenology, -even in his dissertation *Functionen und Folgen Formaler Organisationen* (Luhmann 1964), there are clear traces of philosophical anthropology to be found (Fischer 2006)- he is also critical about it. Especially in his later work, Luhmann distances himself from fundamental phenomenological concepts such as *intersubjectivity* or *first-order-experience*. In his analysis, he takes

on a third persons perspective (Arnoldi 2001) and argues that there is no such thing as intersubjectivity. He speaks of persons as being "black boxes" to each other (Arnoldi 2001: 6). Communication between persons does not happen in a direct manner, because of what he calls *double contingency*. Focussing on communication instead of perception, as did Husserl, he conceptualized communication as an apart, closed system. Communication is therefore not a direct transmission of meaning between persons but always contains a third system (idem).

However, in the early stages of his work this critical attitude towards phenomenology, and philosophical anthropology for that matter, is not that outspoken and clear-cut as is the case in his later work. In the beginning of his career, Luhmann did not deny being influenced by philosophical anthropology and philosophers like Gehlen, Plessner and Husserl. In 1968 he claimed:

> "Uberhaupt trifft die hier skizzierte Theorie sozialer Systeme sich in wesentlichen Punkten mit einer anthropologischen Soziologie, welche die "Weltoffenheit" und die entsprechende Verunsicherung des Menschen zum Bezugspunkt von (letztlich funktionalen) Analysen macht. Siehe auch Helmuth Plessner, Conditio Humana, Pfullingen 1964" (Luhmann 1970: 116).[7]

However, twenty years later he declares "…Philosophical Anthropology, I have never liked it…" (Luhmann in Hahn 2004: 285). Although Luhmann casts aside the connection between his social systems theory and philosophical anthropology and in his later work no longer mentions philosophical anthropology, Habermas already confronted him with this link to philosophical anthropology in a debate in 1972 (Habermas and Luhmann 1972). Hahn (2004) explains this change of heart by the fact that after 1968 it was no longer fashionable to develop a sociology that touched on philosophical anthropology. Moreover, Fischer (2006) argues that Luhmann's

[7] English translation: "At important points, the above described social systems theory comes together with the anthropological sociology which takes as its starting point for an, in the end, functional analysis of man concepts such as 'openness to the world' and the uncertainty that comes along with that. Also see Helmuth Plessner, Conditio Humana, Pfullingen 1964."

sudden dislike of philosophical anthropology had to do with his need for conceptual freedom to elaborate his own theory. Moreover, Luhmann himself speaks of a paradigm shift in system theory. All in all, it becomes clear that the role of man and hence of philosophical anthropology in Luhmann's theory changed fundamentally over time. Nevertheless, unlike Luhmann, I will, through the work of Plessner, incorporate insights deriving from philosophical anthropology in my trust account.

2.3 Trust: bridging the hiatus

Now that we have gained a basic understanding of what Luhmann's theory on trust is about, I would like to draw your attention to one specific element which logically follows from his perspective on trust, but is only implicitly present in his theory (also see Möllering 2006). The difference in complexity between an actor and her environment and the fact that human beings are *aware* of this difference in complexity results in a *hiatus*, a void or gap between system and environment (Keymolen 2008).

"Trust rests on an illusion" Luhmann (1979: 32) states. There is never enough information to give assurance and let complexity dissolve. Trust *reduces* complexity; it *does not take it away*. As a consequence, trust always entails a kind of gap that has to be bridged, a hole in the road that cannot be filled with evidence of a certain and clear-cut future. In his conceptualization of trust, Luhmann also indirectly indicates trust's somewhat *transcendental nature* (Möllering 2001: 409). By speaking of *reducing* complexity instead of *eliminating* it, by characterizing trust as an *illusion* and emphasizing its *as if* nature, he implies that the act of trust is oriented to deal with something rather than to erase it. Trust for Luhmann is "functionally rational" but simultaneously "epistemologically and ontologically transcendental" (idem).

All in all, this *hiatus* is a *grounding aspect* of trust and needs to be included in our analysis to fully grasp the meaning of trust; it is so to say its *essential explanatory force*. After all, if we know for sure how others are going to act and things are evolving, trust would be redundant. It is the unsolvable uncertainty brought forth by this distance between ourselves and the world around us that brings trust in our lives.

My focus on this hiatus reminds of Möllering's analysis of trust as a *leap of faith* that

entails *suspension* and *bracketing*. According to Möllering (2006: 115), actors bracket out irreducible social vulnerability and uncertainty as if these issues were resolved. They suspend looking for evidence or certainty and act as if they are in control. My analysis of trust can be situated in the domain of trust research, set on the agenda by Möllering (2001, 2006), who takes the suspension of vulnerability and uncertainty (the leap of faith) to be the core elements of conceptualizing trust.

Although Luhmann's theory on trust presupposes the hiatus and, therefore, is a fruitful ground for a preliminary exploration of this grounding concept, it does not elaborate nor draws a connection between the bridging of the hiatus and the other related concepts, such as familiarity, risk, and roles.

Although he is very explicit on the fact that only human beings are aware of the world's contingency and the overwhelming complexity this brings along, he does not face the question as to where this special position derives from. While he admits employing a transcendental-phenomenological account, he constrains himself by referring to human beings as *systems that interpret and give meaning to the world*. He does not concern himself for example with bodily perception, intersubjectivity or the question how it is even possible that human beings are open systems that in interaction with their environment have to uphold their boundaries in order to exist. Or as Poggi (1979: xi) formulates it:

> "Consistently with his functionalist viewpoint, Luhmann concerns himself not so much with what makes meaning possible, as with what meaning makes possible- that is, a peculiarly effective complexity-reducing strategy."

Luhmann does not question but simply issues the thesis that in experiencing and giving meaning to the world certain beliefs and expectations are constituted, which reduce complexity and enable action. For Luhmann, experiencing the world is first and foremost about beliefs, about deciding what is in and out, in short about *communication*.

All in all, to explicate the presupposed hiatus, we need to look further than Luhmann's account of trust. I will, therefore, turn to the domain of philosophical anthropology, a branch of philosophy, which revolves around the fundamental

question: “what is man?”.

The basic assumption in philosophical anthropology is that in our understanding of human life, we already, often implicitly, have a perception of what human beings are as a whole. It is these presuppositions the philosophical anthropologist focuses on when clarifying the structures or categories that are already set in place when people think, argue, love, go online or…trust. Because of this objective to explicate suppositions and to map the distinctive features or ‘make up’ of man, philosophical anthropology can also be characterized as a *transcendental discipline* (Corbey 1986: 51). It is the aim of the philosophical anthropologist to understand human beings as part of their “concrete social, historical, every day and natural world” (Borsari 2009: 119). Consequently, a philosophical anthropologist aspires to explicate the structures that ground scientific knowledge as well as human experience as a whole (de Mul 1994: 5). This holistic approach should not be understood as leading to a full-scale ontology but rather it is a *methodological position* or *model* that enables the philosophical anthropologist to analyse different aspects of human nature (Thies 2009: 38). Turning to philosophical anthropology seems fitting to analyse why trust is essential to our way of coping with complexity and how the hiatus plays a crucial role in establishing trust.

In one of the only passages in Trust where Luhmann does refer in a (more or less) explicit manner to the hiatus, he grounds his analysis on the thoughts of Helmuth Plessner (1978), one of the founding fathers of philosophical anthropology:

> “The complexity of its inherent possibilities [of the world] does nevertheless make itself felt in particular as a break, a schism, between the familiar and the unfamiliar, the strange, the uncanny, something which has to be either fought against or treated as mysterious” (Luhmann 1979: 19).

In the notes, we can read that Luhmann positions himself in line with Plessner, who he thinks “rightly sees a fundamental difference between the familiar world of the close-at-hand for humans and the environment of animals” (idem: 22). While the connection between Luhmann’s thinking and that of another protagonist of philosophical anthropology, namely Gehlen, is well-known (Poggi 1979; Paul 2001), his connection with Helmuth Plessner is often overlooked.

In the second part of this chapter, I will bridge the gulf between the work on trust of Luhmann and the philosophical anthropology of Helmuth Plessner; certainly not because this is something Luhmann would encourage or like –I think not- but because I regard it as the most fruitful strategy to conceptualize the hiatus, key to our understanding of trust. Making use of Plessner's theoretical framework that he developed in his magnus opus *Die Stufen des Organischen und der* Mensch (Plessner 1975), I will come up with a philosophical anthropological account of trust in which trust as a way to bridge the hiatus is elaborated. This account of trust will function as the basis for our further investigation of trust in relation to the Internet in the following chapters.

2.4 Helmuth Plessner and philosophical anthropology

Considering Luhmann's emphasis on a radical complexity inherent in human life, it only seems logical to investigate if there are any points of support to be found in the building scheme of human beings that might ground such an account. The second part of this chapter will therefore be dedicated to explaining how this complexity is brought forth by the specific, *eccentric positionality* of human beings, making use of Plessner's theory developed in *Die Stufen des Organischen und der Mensch*. The aim is to develop an account of trust that honours the hiatus as grounding explanatory force and is based on a multi-layered understanding of the way human beings are in the world.

Plessner, as a true philosophical anthropologist, poses the very fundamental and ambitious question: "what is the nature of the preconditions that make human life possible?" Or, in other words, "what is the human *a priori*?" If we adapt these questions to our quest, the rallying query of this section becomes: "what are the preconditions for trust to be possible?"

One of Plessner's basic assumptions is that to map man's ontological blueprint, he has to move beyond the Cartesian divide. Cartesian dualism discerns two fundamentally ontological poles, namely: *res cogitans* and *res extensa; mind* and *body*. How these two poles relate or which pole should be leading in the analysis of

human life has been up to debate since Descartes formulated the issue in the 17th century.[8] Also in current debates the divide can be discerned and the tendency to make one pole outweigh the other. For example:

> "...evolutionary biology claims now to explain not only life but the sociocultural world as a whole and, vice versa, culturalism, by means of the linguistic turn, explains natural science and the evolutionary pattern as a mere cultural interpretation-scheme of special historicity" (Fischer 2014: 43-44).

Of course the human mind, reasoning, and subjectivity are all important aspects of the human make-up. Nonetheless, in order to use this human mind, one above all has to *live*, and this existence always implies a material ground. Choosing one perspective above the other would do no right to the fact that man in fact is both. A human being is mind as well as body, inner as well as outer, subject as well as object. The self-experience of human beings is open to both perspectives; the domain of the mind and the domain of the body are continuously intersecting each other. In fact, it is this ontological *conflict of existence* that is the ground of man's existential structure, which Plessner wants to grasp (Plessner 1975: 32). Plessner aims at describing man as a *psychophysical indifferent unity*[9], a *lived body*, taking into account both sides of the equation. In doing so, he begins his analysis with the fundamental category that grounds both mind and matter: *living nature*.

2.4.1 Living nature

For decades, *living nature* has been the domain of biologists and empirical researchers. Although Plessner acknowledges the importance of their work, he also blames them for mistakenly using categories where they in fact are speaking of concepts (Plessner 1975: 116). This mix-up makes it almost impossible to determine what is mere empirical, a-posteriori or a-priori knowledge in their work. For Plessner,

[8] The mind-body problem in different terms also occurs with the pre-Aristotelian philosophers.

[9] I have translated all German concepts and quotes of Helmuth Plessner myself, unless noted differently.

who wants to unfold the building scheme of human beings, the so-called *existential structures* are *categories* of an *a-priori* character, which cannot be exclusively analysed in an empirical way. Empirical findings such as 'metabolism' and 'genetics' are very interesting concepts that most likely point to a-priori structures, but they cannot be aligned with them. It is therefore up to philosophical anthropology to take on this *transcendental* assignment (Weiland 1999).

From a phenomenological perspective, Plessner starts off by describing how we intuitively perceive a difference between non-living and living objects.[10] Where both kinds of objects fill in an objective place in space and time, there is a difference in the way they have this place. A living thing does not only *has* a place -as do all objects whether or not they are alive- it also *takes* its place. There is some kind of *activity* in the way they maintain their space. A living thing does not only have a boundary that separates it from the outer world and other objects, it also *upholds its own boundary*. This results in a two-way relation: "both directed into the body and away from the body" (Grene 1966: 261). "[I]t is the *way an organism bounds itself* that is essential. It is a question not only of a *Grenze*, but of *Begrenzung"* (idem: 255). Plessner (1975: 129) proposes to refer to this characteristic way of boundary-upholding, when a living thing both has its body and is it as *positionality*.

Living things, by upholding their own boundaries, are defined by an inner and outer junction. Plessner (1975: 128) speaks of *bi-aspectivity*. Living creatures, whether they be plants, animals, or human beings, have a relation towards their environment and towards themselves. Or to put it differently, *"they have a relationship to both sides of their constituting boundary, both to the inner and the outer side"* (de Mul 2003: 252). As a result, there is a cut, an in-between, a distinction, a *hiatus* (here it is!) between living things and their environment. A living thing has its boundary as part of itself and to interact with the environment, it has to cross this boundary. The way in which this *boundary traffic* -the bridging of the hiatus- takes place, ontologically

[10] Plessner, who studied with Husserl, was certainly influenced by phenomenology but as Marjorie Grene eloquently puts it: "…without the heavy emphasis on the new 'method' and its new certainty which makes much phenomenological philosophy so difficult for the outsider to penetrate (Grene 1966: 250)."

defines the positionality of living nature.

Unlike living things, non-living things have *contour*s instead of boundaries. They simply stop where the environment or *medium* as Plessner calls it begins. There is no distinction between a non-living thing and its contours; it coincides with it. Therefore, in non-living things there is not at all something like boundary traffic or positionality. Non-living things are characterized by *passivity*.

It is important to keep in mind that the perceived difference between non-living and living things is not based on an empirical but an intuitive, phenomenological observation. All the attributes we associate with liveliness such as *movement, irregularity,* and *plasticity* are only indicators of life and cannot be aligned with life as such.

2.5 Three times in a row: Positionality

With positionality, Plessner refers to the typical way in which living things uphold their own boundaries, defining their place in and towards the environment. Analysing the manner in which this positionality is organized, three *ideal-types* are disclosed, namely: *plant, animal,* and *human being*. From one stage to the other, Plessner gradually builds up the different positionalities of living nature, respectively characterized as being *open, centric (or closed),* and *eccentric*.

2.5.1 Plants

Plants, according to Plessner (idem: 219-220), are characterized by an *open positionality*. In the totality of their existence, plants are directed towards their environment. Although the boundaries of a plant are entirely part of it, there is nothing behind the boundaries. There is no *centre* that steers organs to gather food or initiates action. The movement some plants display is not mediated by a centre but consists out of impulses that arise in the interaction between plant and environment. For example, the flower opens and closes its calyx in reaction to day and night-time. Because of the plant's open positionality, the bi-aspectivity, which characterizes all living nature, is not standing out as much as it will be the case at the following stage that is reserved for the animal.

2.5.2 Animals

Animals are defined by a *closed positionality*. Unlike plants, they are driven by a *centre* that is often represented by a central nervous system or a more primitive equivalent (Grene 1966). It is through the *mediation* of this centre they can, to a certain extent, control their body. They *are* both a body and *are in* this body; they have a body (*Körper*) and a lived body (*Leib*).

Animals are aware of their environment, which Plessner refers to as *Umwelt*, and of their body, but this awareness does not correspond to what is often referred to as *self-consciousness*. First, animals live in the present, in what Plessner calls the *here and now* (*Hier-Jetzt*); they do not experience a past or future. Their centre of experience is "absorbed without residue into the here and now" (Grene 1966: 273).

Second, although they are grounded by the double position of *being and having a body*, they are not aware of this double position. This ontological organization resulting in their closed positionality is kept from them. They carry it, but aren't familiar with it. The animal is just out there, ascending in the *here and now* (Plessner 1975: 239-240). "The animal lives out from its centre, into its centre, but he does not live as a centre" (idem: 288).

As far as an animal is aware of the outside world and his own body, he can spontaneously respond to the stimuli coming from his environment. Plessner speaks of *frontality* to describe this direct interaction of animals with their environment (idem: 241). The animal "takes not only a place, but a stand" (Grene 1966: 271).

However, animals are 'captured' in a *Funktionkreis,* a building scheme that limits this spontaneity. Animals are aware of their environment as far as their *Funktionkreis* permits them. In line with Von Uexküll, Plessner (1975) speaks of an *Umwelt*, a species-specific environment animals inhabit. The information animals receive from their *Umwelt* can only be of use in a specific situation, for example when they perceive an enemy close by and have to choose between fleeing, fighting, and freezing. These limited inclinations always fit in a fixed and pre-existing knowledge frame, related to the here-and-now.

Consequently, animals cannot reflect upon their choices. They cannot break out of the actual situation, sit down and wonder how to bring their strategies to perfection (Keymolen 2014a). The animal world, therefore, may not be a reflected world; it is, nonetheless, a *familiar world*. Generally, the way the world shows itself to the animal is in line with the repertoire of actions the animal itself is able to

produce.

2.5.3 Human beings

The third stage Plessner discerns is that of human beings, who are defined by their *eccentric positionality*. Human beings still live out from and into their centre, but unlike animals they also live *as a centre*. Plessner (1975: 293) writes:

> "Positionality there is a threefold situation: the living thing is body, is in its body (as inner life…) and outside the body as the point of view from which it is both (body and inner life)" (translation by Grene 1966: 274).

Without cutting across the animal centricity, the life of human beings is also placed outside themselves, eccentric. This detachment enables awareness and reflexivity. "Man not only exists, and experiences his existing, but he also experiences the experience of his existence" (Plessner 1975: 364). For this reflexivity to take place, a distance has to be created -a new, *second hiatus* if you want- between me, as an 'I', and my centre. To enable this detachment, the centre of experience has to split and this division can only happen if the realm of the here and now is forced open and a past and future are fed in (Plessner 1975: 289).

What differentiates human beings from animals is the fact that their positionality is based on this *detachment*. To be human means to be shattered, to be broken. Human beings are defined by a hiatus, which makes it possible to take a position *outside of the centre* and subsequently also *have a relation* towards this centre. Although human beings, just like animals, have a natural place and live in the here and now, they are not fully merged into it. Unlike animals living in a familiar world, human beings are aware of the *world's contingency*, consequently, the human world is not at hand but has to be built. It is only by means of men's creative powers that, the world they inhabit can become a familiar world. With an existence that is literally "based on nothing" (translation by Grene 1966: 274) they can be everywhere and nowhere. Human beings are *homeless* by constitution. Or as Plessner writes (idem: 291):

> "There he stands, on both sides of the hiatus, bound to its body, bound to its soul but at the same time, nowhere, homeless except for the ties of time and space. And like that, he is man."

2.5.4 Three worlds

The bi-aspectivity –the fact that by upholding their own boundaries living things have an outer and inner side- becomes *radical* in human beings because, unlike animals, they are *aware* of this ontological distinction. As a consequence, they find themselves in a world, which, depending on the position they take can be defined as an *outer world*, an *inner world*, or a *shared world of culture*. Grounded in bi-aspectivity, these three worlds are characterized by a double perspective: both the relation *from* as *towards* the boundary is taken into account.

First, in the *outer world*, human beings are aware of their body as a lived body on the one hand and as an object amongst other objects on the other. Plessner (1975: 294) refers to this specific effect of bi-aspectivity as lived body (*Leib)* and body (*Körper)*.

Second, in the *inner world*, human beings know themselves as centres of experience and action and at the same time grasp that they are at the mercy of their feelings and emotions. Plessner (1975: 295) speaks of experience (*Erlebnis)* and soul (*Seele)*.

Third, the *world of culture* (*Mitwelt*) is the world that is reserved for human beings only and in which their eccentricity manifests. It is the world in which you, the other, and me as an "I", a person, are inextricably intertwined. It is where a "we" appears (Plessner 1975: 308). Human beings are grounded in and supported by this world of culture and simultaneously it is up to those human beings themselves to create and shape it. The necessity of shaping, of making and building one's own life directly derives from human being's eccentric positionality. Based on nothing, they can only lead the life they have constructed first (Plessner 1975). The world of culture, characterized by "protection and familiarity" (*Geborgenheit und Vertrautheit*) is the only place where human beings can seek refuge (Plessner 2003: 185). One's country, mother tongue, family, and rituals but also institutions and technologies are all used to create a home for this ontologically homeless creature. Consequently, "technology and culture are not only- and not even in the first place- instruments of survival but

an ontic necessity" (de Mul 2003: 254).

The inner and outer worlds rest on this latter world of culture because it is only in interaction with other human beings, an individual can objectify the outside world and himself. Man's ability to take the point of view from which he is both body and soul rests on this world of culture. Consequently, it is in the world of culture the bridging of the hiatus takes place. Or as De Mul describes:

> "The world of technology and culture is the expression of the desire of human beings to bridge the distance that separates them from the world, their fellow humans and themselves" (de Mul 2003: 254).

Sometimes, the world of culture has been judged as comparable to the *Umwelt* animals inhabit (Plessner 1975: 307). However, this is not completely accurate. Although at first sight both the Umwelt and the world of culture are closed environments, the world of culture can, contrary to the Umwelt, only be built and understood against the background of an open world. The world of culture only functions as a *filter* between human beings and the open world they live in, enabling them to develop a fragile and frequently disturbed balance that lies at the root of their daily life. Culture can be seen as man's second nature and because it is a world that has been made and not been given, trust is a necessary condition for it to become reality (Grene 1995).

Animals cannot leave their *Umwelt*. For them, there is nothing beyond the world they inhabit. They interact in a direct manner with their environment. Intuitions guide them through life; no reflection is possible or needed. Human beings often seem to forget that their world of norms and values actually is made of brittle compromises. They gladly embrace the comforting idea of a closed and steady world in which their daily life rests on a fixed order of values and norms. While it is true that human beings heavily lean on the way society, with its predictable roles and expectations, is organised, the open world with all its unforeseen and disturbing stimuli is continuously shining through, unsettling the delicate balance human beings obtained by furnishing their world with culture (Plessner 2003: 186). This balance therefore can only be temporary. It never resolves the ambivalence of human existence.

2.5.5 Three anthropological principles

In the last chapter of the *Stufen,* Plessner elaborates his notion of eccentric positionality by identifying *three anthropological principles*. He claims that man is *artificial by nature*, is defined by *mediated immediacy* and has the desire to employ a *utopian standpoint*. All three principles have in common that they identify a specific aspect of the continuous and never completely fulfilled endeavour of human beings bridging the hiatus lying at the heart of their existence. Moreover, to underline that this conflict of existence cannot be resolved, Plessner connects for every anthropological principle two -at first glance- contradicting concepts, which together illustrate the ambivalent character of human life.

His observation that man is *artificial by nature* is a direct consequence of the necessity for human beings to build the environment they lack on biological grounds. To compensate for the ambivalent character of their eccentric form of life, human beings have to bring forth things that have enough weight to anchor their own existence. For a creature with an existence that is *based on nothing*, it is *natural* to build an *artificial* environment to live in (Plessner 1975: 310). Or as Plessner (idem: 320) says: "By means of production man only wants to provide himself with that what nature owes him".

As a result, it is impossible to make a clear distinction between natural and artificial adaptations in human life (Plessner 2003: 183). All non-natural features human beings need to exist, are in fact natural because of the inescapable demands their eccentric positionality imposes on them. Nevertheless, Plessner emphasizes that by producing artefacts only a temporary equilibrium can be reached. Artefacts, when they enter the domain of culture, gain their own momentum, they have a kind of heaviness that stands apart from the people who created them. Or as Plessner (1975: 321) writes:

> "Equally essential for the technical artefact is its inner weight, its objectivity that discloses the aspect of technology that only can be found or discovered, but never made. Everything that enters the sphere of culture shows its dependence on human creation. But at the same time (and to the same extent) it is independent from man" (translation by de Mul 2003: 261).

The second anthropological principle of *mediated immediacy* points out the way in which human beings relate to the things around them. Because human beings have a relation towards their centre and do not completely coincide with it, this centre has a mediating function between "I" and the environment. As a result, the way human beings perceive the world and interact in the world is of a broken, indirect nature. Man needs a detour along artefacts, language, and other human beings to establish a meaningful relation with his environment. They form, as it were, the *link* between man and its environment. They make direct what is indirect by nature. Consequently, notwithstanding the indirect relation human beings have towards the world, they still experience the world -just like other animals do- in a direct manner (Plessner 1975: 325). The principle *mediated immediacy* may be contradictive on logical grounds, it nevertheless is perfectly feasible when applied to the relation of man and its environment (Plessner 1975: 324). Both indirect and direct at the same time, the nature of this relationship reflects man's eccentric positionality.

The third anthropological principle is that of the *utopian standpoint*. It refers to man's awareness of the triviality (*Nichtigkeit*) or contingency of his existence and that of the world. Being aware of his existence as just a chance coincidence (*Zufall*), he does not get to know what his place in the world is (Plessner 1975: 342). His desire to find a final ground for his existence leads him to the domain of religion. Throughout history, religion has had different forms and names, but it always appeals to that what man lacks and at the same time defines him, a *definitivum*. As Plessner (1975: 342) describes:

> "A final connection and order, a place for life and death, protection, a reconciliation with faith, an explanation of reality, a home; it can only be bestowed by religion."

Because of man's eccentric positionality, he can imagine a god as a final ground and capture the idea of the Absolute. However, at the same time it is through his eccentric positionality that he is also open to doubt the presence of a god. *"Giving up this idea means giving up the idea of one unifying world. It is easier said than done, being an atheist" (idem: 346)*. On the one hand man builds himself a reality in the inner world, the outer world and the world of culture. On the other hand he makes way for the awareness of his own contingency. Only religion can bring the final order to man.

Those who believe will always come home (Plessner 1975: 346). However, those who choose to stay in the realm of the mind will never return (Weiland 1999: 114).

2.6 An anthropological perspective on trust

The comprehensive theory Plessner developed in the *Stufen,* which I have briefly summarized above, is of great value to our understanding of trust. Where Luhmann mainly focuses on the *functionality* of trust, we are now able, by applying the work of Plessner, to replenish his account of trust with a solid *anthropological basis*. As a result, we can now make explicit that trust, as a strategy to reduce complexity in fact is the bridging of a hiatus that lies at the heart of human existence. Taking into account the eccentric positionality of human beings and the three anthropological laws deriving from that positionality, we can substantiate the "as if" character of trust and the "bracketing" and "suspension" of uncertainty.

First, we can explain why the high level of complexity that Luhmann identified is inherent in human life. Because of their eccentric positionality – caused by the ontological distance between themselves and their centre- human beings are aware of the fact that they are both outer and inner, body and soul, matter and mind.

Moreover, it is due to this eccentric positionality that the world's contingency shines through, that the freedom of action of other human beings is undeniably present, and that the inner world can be a settling as well as a disturbing experience. While it is true that a fragile balance can be reached in the world of culture, this always concerns a temporary balance that continuously is being disturbed by the open and unpredictable world on which it is based. The complexity human beings have to endure is *radical* because it is indissolubly attached to their 'humanness'. Consequently, trust can never resolve the uncertainty deriving from this complexity; it can only reduce it to a bearable level. Although we live in a familiar world, where uncertainties can decrease because of role-predictability and self-presentation, this world remains fragile and susceptible to change. For Luhmann, trust, therefore, is always *blind trust*. It has a fictive aspect. Trust is to act as if the future is certain and as if all complexity has vanished.

Second, where trust is first and foremost a strategy to deal with complexity in interactions situated in the world of culture, with Plessner we see that this is only one

specific aspect of trust. Just like other animals, human beings live in an outer and inner world, the former defined as the realm of the body, the latter defined as the realm of the mind. As a result, in analysing trust, one should not only look at the interpersonal level, but also take into account bodily aspects (reliance on the outside world) and mental aspects (emotions, self-confidence, ontological security).

Third, in the next chapter we will see that Luhmann and other authors argue that in modern times institutions and technologies increasingly mediate interactions, resulting in a growing demand for trust. With Plessner we can already substantiate this claim. Because human beings, unlike animals, do not have a direct relation with their environment, but need some kind of mediation to restore this direct experience, artefacts are made and set in place to bridge this gap. However, these artefacts shaped by human beings are at the same time standing on their own, bringing forth new complexity. In other words, in and through the use of artefacts trust can be placed and undermined; it can be established and questioned. To understand how trust takes shape, we have to include the workings of the artefacts that mediate our interactions. As we will see in the following chapters, this mediation becomes particularly conspicuous in modern society.

Fourth, the hiatus we have already indirectly identified in the work of Luhmann becomes central to Plessner's analysis of eccentric positionality. In fact, one could say that this hiatus is twofold. The *first hiatus* grounds the existence of all living things. By upholding their own boundaries, living things have an inner and outer side, which makes that there is a hiatus between the living thing and its environment. This hiatus brings forth the complexity all living nature has to process. Or in Luhmanns terms: this hiatus is the reason why all *systems* have to develop ways to absorb and transform complexity. The *second hiatus,* however, is reserved for human beings who not only experience a distance between themselves and their environment, but also between themselves and their centre of experience. It is this second hiatus, which is unique for human beings and grounds their eccentric positionality that brings in the radical complexity trust has to reduce. Consequently, to speak meaningfully about trust as a way to reduce complexity, one has to take into account the effects brought forth by this second hiatus.

Finally, the three anthropological laws identified by Plessner illuminate how trust can only reduce and never completely take away complexity. The distance human beings experience in themselves, between each other and their environment

can only be bridged temporally. The paradoxical character of the three anthropological laws reminds us of the fundamental openness of human beings to the world. Being *artificial by nature* refers to the active shaping of a world of culture, which brings human beings a familiar world but simultaneously confronts them with new complexity. The principle *mediated immediacy* refers to the double-faced character of their interactions, which are experienced as direct and therefore unambiguous, but which are in fact broken and open-ended. Human beings can never fully grasp the intentions and motivations of others; therefore, a sense of trust is part of every interaction. The *utopian standpoint*, being the last anthropological law, illustrates how human beings, notwithstanding the paradoxical nature of their existence, strive for stability and certainty. *Trust is good, certainty is better*, as the proverb goes. Although the hiatus lies at the heart of human life, there is always this deep felt urge to dismiss its presence.

3

System trust in late modernity.

In the previous chapter, we took Luhmann's influential theory of trust as a starting point to map the fundamental aspects of the concept of trust. By defining trust as a way to reduce complexity, Luhmann takes a functionalistic approach, focusing on the specific way human beings as social systems uphold themselves in a radically complex environment. This complexity is brought forth by the human awareness that others can act in unforeseen ways and that the future can unfold in many different directions. Trust is a strategy to deal with this complexity. Trust neutralizes the dangers, which cannot be taken away; consequently, it enables people to act as if the future is certain.

Although Luhmann never elaborates it, his theory logically entails that the act of trust merely reduces and does not take away complexity; there always remains some kind of informational gap or hiatus, some uncertainty that needs to be dealt with in interactions. It is the bridging of this ontological hiatus, which is essential for our understanding of trust. By taking into account both Luhmann's theory and the work of Plessner, it becomes possible to substantiate this approach and show that not only is the bridging of the hiatus fundamental to human life, but also that in everyday life the act of trust is an essential strategy to cope with this hiatus.

The previous chapter, in fact, lays the ontological foundation of the concept of trust. It answers the question: "what makes trust an indissoluble and fundamental aspect of human life?" by showing how trust is necessary to bridge the distance human beings experience towards themselves, others, and the world around them.

As Möllering (2006: 192) convincingly argues, most trust research, at best, presupposes this hiatus, but is not focused on explaining how the leap is made.

Rather, the emphasis lies on how to avoid, reduce or eliminate the gap. However, to fully grasp the idea of trust as the bridging of the hiatus, we cannot suffice with a purely theoretical approach as laid down in the previous chapter. We have to put flesh on the bones of these concepts in order to understand how trust is placed and shaped in everyday life.

In this chapter, I will take a first step by analyzing how socio-historical developments affect the meaning and workings of trust. More specifically, this chapter will explain how changes related to the arrival of modern and late modern society have reshaped trust and other closely related concepts belonging to the trust family such as familiarity and confidence.

Since modernity, we have witnessed the arrival of what can be called system-interactions. To go about everyday life, we increasingly have become to depend on systems such as the financial system (although the general confidence people have in this system is rather shaky nowadays due to the financial crisis), democratic institutions, technologies and corporations, (see Giddens 1991; Seligman 1997; Luhmann 1979; Misztal 1996). This has lead to a new form of trust, referred to as system trust or trust in abstract systems. System trust holds the promise of reducing societal complexity, enabling global instead of merely local interactions. However, authors such as Giddens, Beck, and Seligman argue that together with the arrival of this system trust, the earlier, self-evident character of interpersonal trust has changed. In what has been called the period of late modernity, systems are constantly being shaped and reshaped, bringing forth new complexity instead of primarily reducing it. This new complexity penetrating everyday life, influences the way trust takes shape on the interpersonal level. In other words, where system trust is on the one hand a solution to deal with the complexity of everyday life in late modernity, it is, on the other hand, also the source of new complexity, which human beings, consequently have to deal with.

Although trust is first and foremost something that is established between persons, the context in which an interaction takes place and, more specifically, the systems and artefacts that are used to initiate and support everyday life have an impact on the way in which trust is built. This may not come as a surprise. With Plessner, we have already seen that due to the second hiatus defining human life (see 2.6), human

beings are simultaneously privileged and deemed to make use of artificial means (ranging from language to technology) to engage with each other and the world around them. After all, they are artificial by nature. This mediation or necessary detour is never neutral. A telephone call differs from a face–to-face interaction, even when the words that are spoken are exactly the same. Artefacts gain, as Plessner says, their own weight. They make a difference and therefore influence the way trust is being established.

In late modernity, due to technological developments, this mediation becomes particular conspicuous. Interactions are no longer confined to the local society to which one belongs, but increasingly bear on the interplay with layered and often opaque systems such as the healthcare system, the air traffic system, and –central topic in the next chapter - the Internet. Under these circumstances, interpersonal trust is no longer sufficient to support such interactions. Trust, then, has to be put into and found within the system itself.

We will begin with a short recap of the conceptual trust-cluster "familiarity-confidence-interpersonal trust" and see where system trust fits in. Next we will look into the arrival of system trust in late modernity, based predominantly on the work of Giddens and Luhmann.

Finally, we will look deeper into the interplay between system trust, familiarity, and interpersonal trust in late modern society. Some empirically-informed examples will be used to illustrate the most important aspects of system trust and the way in which new complexity arises due to these systems.

At the end of this chapter, the basic ideas that were outlined in the previous chapter will have been placed in a socio-historical setting and the preconditions for system trust will have been explained and analyzed. In the following chapter, this will enable us to further explore the bridging of the hiatus in one of the most influential realms in contemporary life: cyberspace.

3.1 Introducing: system trust

Trust is first and foremost interpersonal, something which takes place in the interaction between human beings. Without trust, the uncertainty caused by not knowing for sure how others will behave and act, bringing along an unknown future,

would result in a situation of total paralysis. Luhmann sees trust therefore as:

> "[t]he generalized expectation that the other will handle his freedom, his disturbing potential for diverse action, in keeping with his personality…" (Luhmann 1979: 39).

Notwithstanding the fact that trust is interpersonal, when social reality becomes too complex, interpersonal trust no longer suffices to temper this complexity, therefore, it has to be extended to other domains of human life (Luhmann 1988, 1979; Giddens 1991; Seligman 1997). In contrast to the traditional world where control, socialization, and familiarity were adequate to establish trust, in the modern world with its wide variety of contingent risks, trust based on solely interpersonal interactions is no longer adequate (Jalava 2003: 174).

Luhmann introduces the concept of system trust to describe the way in which human beings have become used to putting trust in abstract systems such as the political system or the banking system to reduce complexity. This generalized trust replaces the enormous amount of personal interactions that would be necessary to ensure a stable and trustworthy interaction (Luhmann 1979: 51).

3.1.1 Familiarity, confidence and trust revised

Before looking deeper into the connection between the arrival of (late) modernity and system trust, and how this influences the relations of some of the members of the conceptual trust family, let us first recapture some key notions surrounding trust in general. In the previous chapter, we have seen that trust is closely connected and depending on the concepts of familiarity (2.2.3) and confidence (2.2.5). Where, in this cluster of concepts, can system trust be placed?

Familiarity, confidence, and trust can all be seen as "different modes of asserting expectations" (Luhmann 1988: 97). Where familiarity structures the world by ignoring its contingency, confidence and trust both have to do with expectations that can turn out to be disappointments. Trust requires previous engagement; there is a certain element of choice involved and acting on that choice makes you vulnerable to significant risks. Confidence, however, is characterized by not considering alternatives. It is the default setting. You assume the bridge will not collapse, that

your lawyer is competent, and the food you are eating is safe. In the case of confidence, disappointment will lead to external attribution (it is the fault of the government I have eaten poisoned bread because they did not sufficiently control the production process). In the case of a breach of trust, there will be an internal attribution (I should never have trusted him to mail that letter in time). Where trust has to do with considering other options and is linked to a situation of uncertainty and doubt, confidence is about putting aside possible alternatives.

Unlike confidence and trust, familiarity structures the world by ignoring its contingent character. We take the presence of the world, our fellow human beings, and the objects we encounter at face value. In everyday life, we do not bother to question their existence. Moreover, we even expect to see and experience the world in a way similar to how our fellow human beings do. They are also, so to speak, "presupposed and co-experienced" (Luhmann 1979: 18). Trust can only take place in a familiar world in which existence is already structured in a pre-reflexive way.

So, where does system trust fit in, in this cluster of concepts? System trust, just like interpersonal trust, always entails the risk of disappointment. However, where a breach of trust will lead to internal attribution, the attribution in the case of misplaced system trust will be, similar to the case of confidence, external (everyone had an account of this Icelandic bank!). And although system trust in general has a latent character, it fundamentally differs from the ignorance that accompanies familiarity, because with system trust one is aware that:

> "everything that is accomplished is a *product*, that each action has been *decided* on after comparison with other possibilities. System trust counts on *explicit* processes for the reduction of complexity, i.e. on people, not nature (emphases in original) " (Luhmann 1979: 58).

All in all, we can conclude that system trust resembles in fact confidence and that it therefore can be situated between familiarity and interpersonal trust. In his later work, Luhmann indeed seems to use system trust and confidence interchangeably (Luhmann 1988; see Möllering 2006; Jalava 2003: 184; Seligman 1997: 19).

3.2 System trust and late modernity

Now that we have recaptured the relations between familiarity, confidence and trust, and specifically, the way in which confidence and system trust relate, we will now turn to the connection between modernity and system trust.

Luhmann (but also other authors such as: Giddens 1991; Beck 1994; Seligman 1997; Möllering 2006) observes that interpersonal trust as a basis for interactions is eroding in modern societies, blurring the boundaries between the familiar and the unfamiliar.[11] He (1988: 96) characterizes the arrival of modernity as a shift from "*cosmology to technology*", marking the transformation from a world in which unfamiliar and unforeseen events were seen as "*an expression of the hidden meanings of nature or the hidden intentions of God*" towards a world where unexpected events may be the simple effect of our own actions and behavior.[12]

Although the arrival of modernity is linked to Luhmann's focus on system trust, in his work *Trust* he does not elaborate this connection. We, therefore, turn to the sociologist Anthony Giddens (1991), definitely influenced by the work of Luhmann, who analyses new conditions for trust brought forth by a transition from traditional to modern society.

Giddens (1991: 14-15) uses the term modernity in a very general way, referring

[11] Of course, also in pre-modern time people have put their trust in systems. However, these systems were mostly grounded on religious assumptions or natural law. There were only few alternative grounds systems were built on. The human disposition was presupposed and fully encapsulated in the system. There was no critical distance between persons and the systems they were living in. The complexity inherent in human life was therefore assumed to be already reduced by the ordering of the system. Whenever there was the need to explain certain ordering principles, one turned to *authorities,* which functioned as *third party trust,* such as gods, priests or other "wise men".

[12] In line with Luhmann, the sociologist Seligman observes a similar shift. In his analysis of Christian thought, he argues that in traditional society, the presence of God provides a foundation for personal relations. The rules set in place by a shared faith in an all-overseeing God structured the interactions of fellow believers. "Thus, in the transcendent otherness of God and of 'amore Dei', people found not only their own individuality, but the very model for relations with the mundane other" (Seligman 1997: 48). It is with the death of God, Seligman claims, man with his unpredictable behavior no longer mediated by a transcendent entity, became "a problem for human knowledge" (Seligman 1997: 50).

to institutions and patterns of behavior which were first established in post-feudal Europe, but which have become dominant in a "world-historical" sense in the twentieth century.

He identifies three closely interconnected elements that define *modernity*, which he roughly refers to as the "industrial civilization" or the "modern society" (Giddens and Pierson 1998: 95): the *separation of time and space*, the presence of *disembedding mechanisms*, and *institutional reflexivity*.[13]

These three elements are important to our understanding of trust because they form the preconditions for the development of *system trust*, which is not merely the intensification of the initial *interpersonal trust* or the *eroding of familiarity*, but in fact becomes a different and specific form of trust in modern and late-modern society.

First, *the separation of time and space* refers to the fact that in late modernity, time and space no longer are connected through a fixed place. For Giddens, our interactions increasingly become disembedded.

> "[L]arger and larger numbers of people live in circumstances in which disembedded institutions, linking local practices with globalised social relations, organise major aspects of day-to-day life" (Giddens 1990: 79).[14]

Social relations can be established across wide spans of time and space, even globally. As we will see in the next chapter, this specific aspect of late modernity is an

---

[13] When Giddens specifically focuses his analysis on the *separation of time and space*, the presence of *disembedding mechanisms*, and *institutional reflexivity*, he in fact refers to the "post-traditional order of modernity", also called *high or late modernity*. For the sake of clarity, I will refer to it as *late modernity*.

[14] Seligman (1997) claims that especially in modern society, characterized by its differentiation in roles, the room to negotiate these roles becomes more substantial. Not only do people have more roles than in premodern society, they also encounter more 'relevant others' and in their different roles have to deal with a wide range of interactions (Seligman speaks of role-sets). There is much less overlap between these different roles which brings along a greater potential for conflict and contradiction. Or in other words, "(t)he greater indeterminacy and the greater negotiability of role expectations lead to the greater possibility for the development of trust as a form of social relations" (Seligman 1997: 39).

important characteristic of Internet technology as well.

Second, the separation of time and space also makes possible the use of *disembedding mechanisms*. Giddens distinguishes two types of disembedding mechanisms: *symbolic tokens* (such as money) and *expert systems* (healthcare system, telecoms system, science system), which he refers to both as *abstract systems*.[15]

These mechanisms make it possible to separate or "lift out" activities from local contexts. Events happening on the other side of the world are increasingly shaping our local everyday lives. Also economic exchanges are taking place on a global level resulting in labor practices being "lifting out" of the local community and which are "recombined across time and space" (Giddens and Pierson 1998: 98).

It has to be noted that unlike other sociologists such as Luhmann and Seligman, Giddens deliberately chooses not to use the term *differentiation* here. Differentiation in general refers to the separation of roles and functions bringing forth a specialized and precise society, where Giddens (1991: 18) values the *detachment between action and context* and how these two become *re-embedded* as more fundamental to modern society. Perhaps, Meyrowitz's (2005: 25) concept of the *glocality* describes best what Giddens has in mind: the fact that in today's consciousness "the local and the global co-exist in the glocality". What we for example learn through media about other places, foreign politics and cultures is as much of importance in the shaping of everyday life, as is the influence of our local environment. Modernity is just as much about fragmentation as it is about unification, Giddens (1991: 27) claims.[16]

---

[15] Luhmann speaks just of "systems" instead of "abstract systems" as Giddens does, but in fact they both refer to the same phenomenon. Therefore, I will use them interchangeably throughout my dissertation.

[16] Although Giddens specifically focuses on the influence of global events and developments on everyday life, he does not deny that the physical place or locality is still an important aspect of our daily interactions. While less explicitly, with this statement he affirms Meyrowitz's (1985) principal argument that although the physical context and societal context become separated, the local context still is an important setting for everyday life.

The third fundamental aspect of late modernity is *institutional reflexivity,* which touches on the use of knowledge in all domains of social life. Knowledge gained about a certain aspect of life - education, health - flows back to that domain, simultaneously constituting it. Giddens' concept of reflexivity –Adams (2004) speaks of "heightened reflexivity"- should not be confused with the *eccentric positionality* we discussed in the previous chapter. The latter refers to the ontological distance of human beings which entails that "man not only exists, and experiences his existing, but that he also experiences the experience of his existence" (Plessner 1975: 364). The former, on the other hand, refers to the flows of information about possible ways to organize society (macro-level) as well as everyday life (micro-level). This knowledge is not bound to the borders of institutions and therefore can always become a question of debate, by experts as well as by laymen. Living in a modern world means living in a world of change and, even more importantly, in a world of radical doubt. Leaving religion and dogmas behind, the Enlightenment falsely promised that reason would bring us certainty, however:

> "[N]o matter how cherished, and apparently well established, a given scientific tenet might be, it is open to revision – or might have to be discarded altogether – in the light of new ideas or findings. The integral relation between modernity and radical doubt is an issue which, once exposed to view, is not only disturbing to philosophers but is *existentially troubling* (emphasis in original) for ordinary individuals" (Giddens 1991: 21).

3.2.1 Risk in late modernity

Perceiving and dealing with risks and uncertainty becomes inevitable in a society where all knowledge is questionable and traditions are increasingly becoming eroded.

In general, *risk* refers to an active and explicit engagement with future threats. When we talk about risks, we talk about the *chance* or *probability* that a certain– often undesirable - event will occur. When we refer to *uncertainty*, on the other hand, we face possible *unpredictable outcomes*. For example, when a new technology is being introduced it is difficult to predict possible side effects because there is no previous experience on which one can fall back. In such a situation of high

uncertainty, it becomes increasingly difficult to clearly identify risks.

Nevertheless, in contrast to what this rather strict distinction between risk and uncertainty might seem to imply - eloquently described by Knight (1921) in his influential study *Risk, Uncertainty and Profit* -, we have seen in the previous chapter that risk and uncertainty are indissolubly intertwined both on the *ontological* and the *epistemological* level (also see: WRR 2008). Not only does the world confront us with uncertainty because of the variability and indeterminacy of social processes, we are also aware that our knowledge of risk-determinacy, impact, and causal effect are limited. To put it differently, even when we talk about risks there is uncertainty because we might have doubts about our *risk perception*. From some risks we may be more convinced than from others. Some authors, therefore, see uncertainty as an *attribute* of risk (Asselt 2000).

The connection Giddens makes between modernity and risk is in line with and partly based on the work of Ulrich Beck (1994, 1992b) on the *risk society*. What makes *risk* – nuclear powers, biotechnology - different than *danger* - natural disasters, epidemics - which were already present in pre-modern society is that the former is the consequence of a techno-economic decision, where the latter is nothing more and nothing less than "'strokes of fate' raining down on mankind from 'outside' and attributable to an 'other' – gods, demons or Nature" (Beck 1992a: 98).

However, these techno-economic decisions that bring forth risks cannot be easily attributed to individuals nor is the average citizen in general actively involved in such decision-making processes. As we will see in the next chapters, with the arrival of ICTs, people seem more and more to be responsible for what happens in the world, while simultaneously it becomes increasingly more difficult to point directly at the specific sources of responsibility (Floridi 2015a: 21).

Also Luhmann (1990) sees a difference in the appreciation of risk between decision makers and other people. The former believe they put technologies into practice based on rationally calculated risks, whereas the later feel they are simply exposed to dangerous technologies. People who are not part of the decision-making process are less willing to accept such dangers. Apparently, a double standard of evaluation is being used, based on whether or not someone is in control of the situation (idem: 226). The distant locus of decision-making in relation to the significant impact these decisions may have on everyday life means that "risk and

danger are part of our daily lives" and that simultaneously these risks are "out of our control and there is nobody who could be held accountable" (Misztal 1996: 93).

Beck (1992a: 99) speaks of the "mathematical ethics of the technological age". This is a type of ethics without morality, where insurance and liability laws replace personal accountability. Nowadays, we develop actions based on prevention, compensation, and precautionary principles, making events that not yet have occurred our object of interest and concern. All these actions are aimed at providing us with security in the face of an uncertain future (Beck 1992a: 100).

3.2.2 Trust in abstract systems

Notwithstanding the presence of risk in late modern society, Giddens does not presuppose a lack of trust, rather he sees the erosion of local and traditional order as an instigator for the reconstruction of new traditions and structures on the global level (also see Misztal 1996: 89). For Giddens we do not abandon or move beyond modernity, but we continue it with different means. As we have seen, trust is closely connected to risk because it can function as a way to cope with it.

> "Trust is also about the binding of time and space, because trust means giving commitment to a person, group or system across future time" (Giddens and Pierson 1998: 101).

Trusting a teacher or the healthcare system might turn out to be a very effective way of setting potential bad outcomes aside and –as we have seen with Luhmann - to be able to act *as if* the future is certain. For Giddens trust, consequently, is a characteristic of modernity *par excellence* because it is essentially about *organizing and confronting an open future*.

Giddens differentiates between two types of trust: trust in persons (*facework commitments*) and trust in abstract systems (*faceless commitments*). He claims that trust in abstract systems – which by and large overlaps with Luhmann's concept of system trust - becomes dominant in modernity.

Abstract systems, consisting of *symbolic tokens* such as money and *expert systems* such as healthcare, enable people not only to temper the uncertainty about an open future, but also to cope with the reflexivity of knowledge that shapes the

organization of society. The trust we have in abstract systems provides us with a sense of security, which is necessary to lead our everyday life. Every time someone uses an ATM machine, drinks tap water, uses aspirin or brings her child to school, she puts – often implicitly - trust in the abstract systems that enable these actions. We all suppose that the money we have in our pocket will keep its value overnight, we trust that there are competent people who are checking the water running out of the tap for bacteria and other harmful substances, we trust the science that underlies the workings of medicines, and we trust the schooling system to teach children adequately.

Characteristic for late modern society is that it is rather difficult to opt out from these abstract systems (Giddens 1990). The overall presence of "low-probability high consequence risks" in late modernity, people cannot bear individually. Not only is it impossible to withdraw from the risks involved in some abstract systems (even when you refuse to use nuclear power, a meltdown of an installation will nevertheless affect you), - in order to function in society you need to trust the expert systems, which enable you to fulfill your *role* –referring to Seligman (1997) - in everyday life. Abstract systems, and especially expert systems, not only provide a sense of security, they also produce the world with all its chances and possibilities we are nowadays living in (Giddens 1990: 84).

3.2.3 Giddens versus Luhmann

For Luhmann, anyone who puts her trust in a system basically counts on its functioning well more than she believes the people who make that functioning possible are all trustworthy actors in a personal and intimate setting. Although the personal sphere can still be an element of importance in an interaction (for example, your doctor can also be a friend of the family) for the continuity of the interaction, it becomes more decisive that you as a patient trust the *medical system* in which the doctor is educated and trained and that you trust the *safety valves* that are set in place in the system: ranging from a second opinion to evidence-based treatments. It would be impossible to be acquainted with all the employees of the hospital, the bank or any other large institution for that matter. You trust them as representatives of an abstract system. The trust you invest in them is not longer necessarily based on personal and intimate experiences. Their presence merely reassures you that the

system, which technicalities you cannot fully understand, is properly functioning. Luhmann (1979: 52) states:

> "In other words, he has to be able to depend and to rely on the processing of information by other people. He knows, that is, others who know how the engine of his car works, how his gastritis can best be treated; he might mistrust the newspapers but still assumes that their news is at least news; he relies on the fact that the representatives of his insurance firm give him factually correct information on insurance matters. In a highly complex environment this type of trust can not longer take the form of person trust…"

Although Giddens' analysis of trust and modernity is by and large in line with the analysis of Luhmann, there are some differences of opinion in the details. Where they both see an important role for *system trust* or *trust in abstract systems*, Giddens, contrary to Luhmann, values the *personal* and the *impersonal* –or the *facework commitments* and *faceless commitments*- differently. More than Luhmann, Giddens emphasizes the continuous state of trust by showing how the impersonal and the personal are *intertwined* in modernity. He pays more attention to the different ways in which people place their trust in abstract systems (Möllering 2006: 73) and how system trust instigates a transformation of interpersonal trust. More than Luhmann, who sees it to be the function of experts to control the system, Giddens argues that people working as representatives for a system have an important role in transforming impersonal interactions at the "access points" (Möllering 2006: 74). This different approach leads Giddens (1990: 33) to state in the introduction of his book *The Consequence of Modernity* that he will "conceptualise trust and its attendant notions differently" than Luhmann has done.

While it might be true that Luhmann does not analyse in detail the role of interpersonal interactions in establishing system trust, I believe Giddens' reading of Luhmann's demarcation between interpersonal trust and system trust (or simply trust and confidence) might also be too stringent.

The difference Luhmann makes between trust and confidence is not primarily based on the question of whether or not "individuals consciously contemplate

alternative courses of action" as Giddens (1990: 32) claims. As we have seen in the previous chapter (see 2.2.2*),* trust, for Luhmann, can be placed in an almost careless way and certainly is not the outcome of a rational calculation of the most optimal action (Luhmann 1979: 88). In discerning trust from confidence, it is more important "whether or not the possibility of disappointment depends on your own previous behaviour" (Luhmann 1988: 98). Consequently, this makes differentiating between trust and confidence much more *subjective* –and therefore perhaps also more diffuse- than Giddens presumes. It could therefore well be that only after an interaction turned out wrong an actor would become aware of the fact that she had placed trust in that person, because, looking back, she would hold herself responsible for getting involved.

On the other hand, disappointment can also lead to external attribution, in which case we would speak of confidence. Consequently, also Luhmann acknowledges that this analytic distinction in fact can become complicated because a relation based on trust can quickly turn into confidence and vice versa. However, unlike Giddens, he merely confirms this interlocking without explicating it.

3.3 Trust: The interaction between system trust, interpersonal trust, and familiarity in late modernity

The final part of this chapter will be dedicated to the interaction between system trust, interpersonal trust, and familiarity in late modernity. First, we will look into the ways in which interpersonal interactions are still part of overall system trust. Next, we will address the role of familiarity for trust in late modernity by analysing basic trust as an important pillar for familiarity and by discerning the overlap and differences between familiarity and the *Umwelt*.

3.3.1 Facework commitments in abstract systems: the air traffic case

Interpersonal, trustworthy interactions form a necessary condition for stable system trust to occur. Where Luhmann merely focuses on the workings of the abstract system, Giddens argues that *facework commitments* –the interaction with the operators- at the *access points* of abstract systems are an important aspect of system trust. Access points are the places were the laypeople or users meet the

representatives –not necessarily the experts- of the abstract systems. They form "the meeting ground of facework and faceless commitments" (Giddens 1990: 83). Although, not all abstract systems presuppose interactions between laypeople and representatives or operators of the system, most of the abstract systems do involve at a certain moment in time an interaction between both parties. In addition to new knowledge that can be spread by media and other sources, trust towards abstract systems is, therefore, strongly influenced by the experiences people have at access points. These facework commitments at access points are the interpersonal interactions that co-shape trust in systems.

System trust and interpersonal trust are not part of a zero-sum game. It is not the case that more system trust necessarily entails less interpersonal trust or the other way around (Luhmann 1988: 99). On the contrary, a social evolution bringing forth increasingly complex societies requires not only more confidence in systems but, in order to seize opportunities and chances, more trust in other partners as well (idem).

> "So it is not to be expected that scientific and technological development will bring events under control, substituting mastery over things for trust as a social mechanism and thus making it unnecessary. Instead, one should expect trust to be increasingly in demand as a means of enduring the complexity of the future which technology will generate" (Luhmann 1979: 15-16).

As Luhmann foresees, the arrival of increasingly complex technological systems in late modernity, not only provides human beings with a new repertoire, new possibilities to interact, communicate and create their artificial world they lack on natural grounds, these systems also produce new complexities, new uncertainties human beings have to relate to.

To illustrate the importance of trustworthy facework interactions at access points, I will fall back on a personal experience I had when I took an airplane to fly from The Netherlands to Canada, attending a conference there. One should know that I am afraid of flying, and it is especially on these occasions, where existential uncertainties are apparent and the act of giving trust is no longer latent but transforms into an intentional and conscious act, one increasingly becomes aware of the trust we all put into these expert systems on a daily basis.

First, there is some general knowledge I try to bear in mind when boarding for the flight. I know that statistically, flying is much safer than the car trip that brought me to the airport[17]. Nevertheless, the anxiety I am feeling now and that was absent in the car tries to convince me otherwise.

Second, I know that the company I am flying with is a trustworthy company. KLM has an excellent track record of successful flights executed by professional pilots with well-kept planes.

Finally, I know that a lot of other, well-educated and rational people whom I trust travel by plane without giving it a second thought. So, why shouldn't I? Trust in abstract systems is often based on the fact that others trust and make use of that abstract system too. Because, as a layperson, I am not able to check all the detailed workings of the plane, I have to depend on the experts and control mechanisms of the system. Their trustworthiness, however, I also deduce to a certain level from the way in which others value their functioning.

Despite all these reassuring thoughts, I still hesitate to get on the plane. I then find myself carefully observing the flight attendants. How do they behave? Do they look nervous? What are they talking about? I see that they are laughing and are helping passengers to find their seat. They look competent in their blue uniforms. Nothing seems to be out of the ordinary.

> "At access points", Giddens (1990: 85) explains, "the facework commitments which tie lay actors into trust relations ordinarily involve displays of manifest trustworthiness and integrity, coupled with an attitude of 'business-as-usual', or unflappability."

Encouraging, a flight attendant smiles at me and welcomes me on board. I let myself be willingly directed to my seat, and then put on the seat belt and wait for the safety instructions. When all is done and the plane takes off, I still keep my eyes on the crew. When they retrieve to their seats, they shut the curtains between their staff cabin and the passengers' section, blocking my view. I feel the anxiety returning. Is there something I should not see? Something that might prove that my trust was in

[17] At least if one takes into account the statistics on death per kilometre. The death per journey statistic is in favor of the car!

fact unjustified?

Elaborating on Goffman's (1959) notions of *frontstage* (the passenger's section) and *backstage* (the staff cabin), Giddens (1990: 86-87) explains that controlling these two areas is part of the essence of professionalism.

First, the work experts do might require great mental concentration. Not being disturbed can therefore be essential to the success of their actions.

Second, even experts can get things wrong. Showing this to laypeople may lower their trust, even if the mistake of the expert does not negatively influence the outcome.

Finally, contingency always remains a relevant aspect of the functioning of abstract systems. Elements of hazard and luck can enter the performance of the representative of the system. In general, however, a representative or expert prefers to conceal how these elements might come into play, because, being in control is part of a trustworthy performance.

All in all, it becomes clear that interpersonal trust and system trust are connected. Although, I was aware that the real storage of trust was located in the air traffic system that is behind the cabin crew, rather than in the cabin crew itself (if the plane malfunctioned or the pilot became unwell, they probably would not have been able to stop the inevitable, tragic event of a crash), they nonetheless are an essential part of the overall trust that I put into the system when I fastened my seatbelt.

Statistical knowledge lacks the reassuring smile I needed to get on board of that plane. Trust in abstract systems provides the security we need to cope with a radically open future, but it cannot offer us the intimacy we experience in personal trust relations. Giddens (1990: 115) claims that:

> "This is one of the main reasons why individuals at access points normally go to great pains to show themselves trustworthy: they provide the link between personal and system trust".

System trust, just as interpersonal trust, has to be learned. It is established in cycles of positive experiences with the system at hand and partly rests on the perception that others trust the system as well. The impossibility of backing out of abstract systems makes that it almost never becomes a subject of public debate. "One can only feel unhappy and complain about it" (Luhmann 1988: 103). Its latent character even helps maintain its integrity because laypeople are not in the position to control the

detailed functioning of a system anyway (Luhmann 1979: 57). They depend on the expert whose job it is to control the system.

For Luhmann, control mechanisms should be built into the system and made explicit that "trust in the ability of systems to function includes trust in the ability of their internal controls to function" (Luhmann 1979: 57-58). This focus on internal controls may partly explain why Luhmann does not pay much attention to the facework interactions at the access points of these systems, while with Giddens we, nonetheless, might add that these controls can also be partly external when it is the function of the representative to perform certain safety measures.

For example, it is the task of the cabin crew to check if the doors are properly closed and all the hand luggage is safely put away. By showing these actions to the passengers –or the audience in Goffman's terms- they carry out the message that safety is an important aspect of flying. Next to well-functioning internal control mechanisms, this external and visible control might add to the trust put in abstract systems by laymen.

3.3.2 Basic trust and familiarity

Trust –whether it takes the form of interpersonal trust or system trust- can only take place in a *familiar world*. This familiar world –also referred to as the "lifeworld"- is the un-reflected background of beliefs and meanings against which the world is perceived. We can only get to know and deal with the unfamiliar or the unexpected in a familiar way (Luhmann 1988). The unfamiliar therefore does not necessarily impose a problem upon trust as long as actors are able to engage in a *process of familiarization (*Möllering 2006).

Also, in late modernity, familiarity remains a necessary condition for trust, interpersonal trust and system trust alike. In general, human beings do not question their own identity or that of the others around them. They presuppose the continuity of their environment and they believe that others experience this environment in a more or less similar way. Even philosophers whose core business revolves around rather uncomfortable questions such as "do I really exist?", "what is identity?" or "what is it like to be human (or a bat)?" do not seem to doubt their daily actions. This basic 'taken-for-granted' character of the world is a necessary ground for trust to occur (see 2.2.6).

Giddens (1991: 36) speaks of *ontological security* to denote in fact the same familiarity, which he, more specifically, relates to the "bracketing" inherent in the natural attitude in everyday life. Ontological security, therefore, also refers to the deep-rooted belief people generally have in the continuity of their self-identity and the consistency of the world around them (Giddens 1990: 92). This ontological security or familiarity is grounded on the acquiring of *basic trust* or *elementary trust* in early-childhood.[18]

Next to getting familiar with the way the world works (every morning the sun rises, things fall down and not up, when the ball rolls under the table it has not vanished but is merely out of sight), as a child we also develop a *basic trust* by interacting with parents, siblings and close family (Harris 2012). Infants learn they can rely upon others. In the nourishing relation with their parents, they form a sense of self, which will hopefully become the stable basis from which they can interact and build relations with others outside their family.

The developing of a self or an identity is closely connected to the *distance* infants increasingly experience in the interactions with their parents. They have to learn that the absence of their parents, for example when their parents leave them at the daycare facility, does not mean that they have left them for good or that the love between them has vanished. Trust, again, means bridging the gap with an unknown future. And in this early stage in human life, the bridging of this gap is mainly focused on coping with the increasing distance between parents and children; it is in fact about the blocking off the

> "existential anxieties which, if they were allowed to concretise, might become a source of continuing emotional and behavioural anguish throughout life" (Giddens 1990: 97).

[18] Next to basic trust, Giddens also speaks of elementary trust. Misztal (1996: 91) remarks that although he uses the two concepts interchangeably, based on the context in which both concepts are adopted, basic trust seems to be connected to ontological security and elementary trust is associated with the predictability of everyday life. I am not going into the underlying nuances and will connect both concepts to the covering idea of familiarity.

That a disruption in the development of basic trust may eventually lead to people living in an unfamiliar world instead of a familiar world becomes clear when we, for example, look at the devastating case of Savannah. This Dutch toddler was severely abused by her mother, Sonja de J., causing her death in September 2004. This tragic case led to a wide public outcry and stricter rules in the youth care system (Keymolen and Prins 2011; Prins and Keymolen 2011; Keymolen and Broeders 2013).

In a book chapter, appropriately called *Tragic Parenthood*, De Mul (2014a) argues that the responsibility we often presume all parents feel towards their children –a responsibility belonging to the ontological structure of man as it were- in fact needs to " 'be activated' in relation to certain experiences" (de Mul 2014a: 188). External stimuli are needed "to develop this inborn capacity" (idem). Likewise, we can claim that *ontological security*, or in a broader sense *familiarity*, which presence is a necessary precondition for human beings to thrive, can only be activated by developing basic trust in early life.

During the trial, it turned out that Sonja de J., the mother of Savannah, had gone through a very traumatizing childhood herself, most probably destroying her ability to feel a deep-rooted sense of responsibility towards Savannah. We might also say that it deprived her of the possibility to develop basic trust. In the violent relation with her parents, there was little room for security and feelings of reciprocity. The absence of basic trust results in a world which does not bracket out existential fears or risks but, on the contrary, reinforces them. Sonja de J.'s statements throughout the trial outlined a world in which the normal state of affairs is not one of setting aside uncertainties and doubts, but of presupposing animosity and distrust as the fundamental ground of every interaction. Her world is not a shared or common sense world. It is not a place of continuity but of disruption. The world she was familiar with was unfortunately not a familiar world reigned by normality, but one reigned by abnormality.

It would go too far to completely ascribe Sonja de J.'s deviant behavior –as it was referred to in the media- to her traumatic childhood inducing this defective basic trust. Fortunately, not all abused children grow up to kill their children. However, the case of Savannah does illustrate that developing basic trust is an important condition to activate ontological security and to, consequently, live in a familiar world. Moreover, a sense of ontological security is necessary to be able to place trust in persons and systems later on in life. Anticipating the next chapter on trust in and

through digital technologies:

> "Trust in the reliability of nonhuman objects, it follows from this analysis, is based upon a more primitive faith in the reliability and nurturance of human individuals" (Giddens 1990: 97).

3.3.3 Distinction and overlap between the familiar world and the Umwelt

Interestingly, Giddens (1991: 126-133) –with a reference to Goffman (1967)- speaks of an Umwelt as a protective cocoon that encloses human beings with an environment they perceive as normal and uneventful. As we have seen in the previous chapter (see 2.5.2), the Umwelt is associated with the environment animals inhabit. Animals perceive their environment always in line with their Funktionkreis, their ability to react on the gathered information. This balance between information and action makes that the animal world indeed can be seen as a world reigned by normality, more or less similar to the familiar world of human beings. In both worlds, actions are mostly un-reflected, based more on routines and successful behavioral patterns than on conscious decision-making. Without such a sense of normality, it would become more difficult to have confidence in the continuity of our self-identity and of the world around us. In addition, the ability to plan for the future would diminish, weakening overall trust (Misztal 2001).

However tempting it is to equate the Umwelt with the familiar world, the difference between both environments is nevertheless so fundamental that using them interchangeably rather obscures the meaning of the familiar world than elucidate it.

First, the familiar world of human beings includes more than the mere physical immediate environment, as is the case in the Umwelt. Giddens, therefore, alters the original idea of the Umwelt by adding that for human beings the Umwelt also includes the perception of high-consequence risks over "indefinite spans of time and space" (Giddens 1991: 127). The information animals derive from their environment is always in line with their range of action. All this information, therefore, is only displayed in and important to the present, to the 'here-and-now'. Animals cannot reflect upon the possible consequences of their actions in the future. This makes that they do not know trust, at least, not in the manner as we have

described it: to act as if the future is certain. For animals there is no need to trust, simply because they are not consciously aware of the future and the uncertainties it may bring forth. This might make the animal Umwelt a familiar world but also a non-reflected one (see chapter 8).

In the familiar world of human beings, however, the uncertainties of the future always shine through, even more so in late modern society where high consequence risks are part of everyday life. The most fundamental difference between the animal Umwelt and the familiar world of human beings, therefore, is that the Umwelt cannot be shattered, -there is no world beyond the familiar world, so to say- where the human familiar world, on the opposite, always remains a world 'under-construction'.

The Umwelt is the only world animals live in. They are only aware of the 'here-and-now'. "Animals do not give meaning to the world but bear a meaning" (Lijmbach 2002: 106 with reference to Buytendijk 1939). Human beings tend to forget that the routines they follow and the rules that are set in society in fact are social constructs and not natural laws. From this perspective, the familiar world of human beings is also an un-reflected world. The third anthropological law of Plessner captures this human urge to live as if the world is a well-ordered place, by speaking of a utopian standpoint people strive for. There is, however, always the possibility of questioning routines and changing them. A life-changing event –getting married or giving birth to a child- can shake someone's familiar world to its foundations. By reducing the familiar world to an Umwelt, one runs the risk of losing sight of its man-made character. It always remains in close connection to the complex world, which it filters in order to neutralise fundamental uncertainties. It is in fact the interplay between the familiar world and the open and complex world, which is always shining through that brings the need for trust into human life.

3.3.4 Familiarity contested in late modern society

Not only a problematic development of basic trust might negatively influence the robustness of the familiar world, but also on the macro level there are changes related to modernity, which weaken familiarity.

The conditions for familiarity have fundamentally changed since the arrival of the printing press (Giddens 1991; Seligman 1997) and in its wake recent, digital technologies (Castells 1999, 1996; Turkle 2011, 1995, 1984). Our current, complex

society faces the problem that we are now able to collect and store more information than any person would ever be capable of coming to know. This brings us in the awkward position of simultaneously having a lot of knowledge and knowing that there is so much more information with which we will always be unfamiliar as well. While others may find themselves in the position to utilize the information we do not grasp, we might have access to information others do not have. Everyone is an expert as well as a layperson, because diversification has made it virtually impossible to become an expert in all domains. In modern society, everyone has to deal with a multitude of systems and of most of them we can only grasp in a superficial way their technicalities.

More knowledge present in society, therefore, does not necessarily entail more *shared knowledge*. While in general, it is still widely accepted that we all perceive – phenomenological speaking- the world more or less in a similar way, we increasingly become aware of the different *epistemic contexts* in which these perceptions are interpreted. This reflexivity, as Giddens calls it, puts pressure on familiarity because it questions the "taken-for-granted" aspect of the familiar world.

Also the earlier mentioned *separation of time and space,* characteristic for modern society may cause difficulties for the firmness of the familiar world. Although all human interaction is in a way mediated, the mediating workings of for example the television or newspaper may generate a so-called "reality inversion": the real object, then, seems less 'real' than the representation of that same object (Giddens 1991: 27). Events happening at a distance may well enter daily life in a very "real" manner, thereby affecting the familiar world.[19] Especially media, such as television and the Internet, make us aware that the place we are living is not the one and only community, but merely just one of the many possible communities we could be living in. These other "localities" become the "generalized everywhere" serving as mirrors to view and value our own everyday life (Meyrowitz 2005). Place and familiarity are no

[19] It has to be noted that although it is true that global processes –such as the capitalist market-, structure our local life, it is too big of a claim to state that in everyday life we are constantly aware of this global influences- instigating what Thomlinson (1994) refers to as a "phenomenology of the global". Rather it is the continuous interaction with distance events and actions, which transform the familiar world.

longer indissolubly attached to each other. This does not necessarily lead to alienation from the local context, but it is more closely related to the "integration within globalised 'communities' of shared experience" (Giddens 1991: 141).

Finally, because of the arrival of system trust, the nature of interpersonal trust – which is in fact the main level for familiarity to develop - has changed. It is not so much that system trust has replaced interpersonal trust; rather, it has transformed it. What we experience as being personal is indissolubly intertwined with the abstract systems on which modern society relies.

Where in pre-modern times basic trust between persons was integrated in relations in the community, family relations, and friendships, in late modernity through the disembedding and re-combining working of systems, people are able to keep in touch with others all over the world. There is, therefore, no use to differentiate between the detached character of systems and the intimate character of interpersonal relations. Trust is no longer self-evidently based on personal ties in one's local community. Rather, interpersonal trust becomes "a project, to be 'worked at' by the parties involved, and demands *the opening out of the individual to the other*" (Giddens 1990: 121). Trust no longer is pre-given, but involves a "mutual process of self-disclosure" (idem).

As a consequence from the transformation of self-evident to *active trust*[20] (also see Möllering 2006), familiarity loses its more given nature and becomes something that has to be worked at as well. It has to be noted that interpersonal trust always is "active" (also see Adams 2004). As we have seen in the previous chapter, trust can be seen as a leap of faith in which actors set aside uncertainties and act as if the future is certain. This always implies an effort, from both trustor and trustee. Active trust, as Giddens describes it, therefore, should be more valued as an intensification of the

[20] Active trust refers back to the concept of suspension and the bridging of the hiatus, which was discussed in chapter 1. While these elements are always closely related to trust, we could say that in modern society, where familiarity and ontological security are increasingly difficult to attain, trust between persons becomes more active in order to be able to bridge the hiatus and suspend the insecurity brought forth by high-consequence risks. Möllering (2006), building on –amongst others- Giddens, elaborates active trust to make it one of the keystones of his conceptual framework.

active aspect of interpersonal trust than a radical different form of interpersonal trust.

3.4 Conclusion

Where in the previous chapter we have focused on the preconditions for trust and the manner in which trust is related to the human condition, in this chapter we have added a layer to this ontological perspective by taking into account the socio-historical context in which trust occurs. More specifically, I have shown how trust and related concepts such as familiarity and confidence have altered in late modernity.

First and foremost, there is the dominant presence of system trust. Increasingly, it has become necessary to put trust in systems in order to go about everyday life. As we have seen, complexity is inherent in human life. The uncertainty brought forth by the fact that human beings are aware that the future is capricious and that others have the freedom to act in unanticipated ways brings trust as a strategy to cope with this complexity in the realm of human life. The arrival of late modernity characterized by globalization, the decline of religious beliefs and fate, and the presence of low probability/high consequence risks, has conspicuously shaped this complexity. Where in traditional society, interpersonal trust and familiarity sufficed to cope with the uncertainties at hand, system trust becomes a necessary filter to deal with the complexity of late modern society. Systems, which we cannot fully understand or control, from money, politics, to the educational system and healthcare, are set in place to neutralize the complexity inherent in human life in the modern era.

Interestingly, the development of system trust did not in any way force out interpersonal trust. On the one hand there is the development of facework commitments at the access points of systems to smoothen the trust put in systems, on the other hand the active characteristic of trust – which was always there, because trust always entails an act of suspension - becomes more radical and transforms trust into something that has to be invested in and worked on. Trust between persons becomes “a project” (see Giddens 1991; Giddens 1990; Giddens and Pierson 1998).

Also the familiar world, which we have seen is a necessary precondition for any form of trust to thrive, keeps being important in modern society, although, its pre-given character is slightly eroding. The ontological security, which is one of the pillars

of a familiar world, might be pressured by the presence of a multitude of frames of interpretation in modern society, diminishing the ground for a shared perspective on everyday life. Additionally, the more active character of interpersonal trust, another important point of support for the familiar world, reduces its taken-for-granted character.

4

The Internet: A familiar world?

All interactions are mediated interactions. In late modernity, these forms of mediations are increasingly being shaped by large, technical, and often opaque systems. In the previous chapter, a few of these dominant systems – such as the political system and the financial system - passed in review. We also looked deeper into one specific system –the air traffic case- to illustrate how *system trust* comes about.

However, we have not yet dived into the specific workings of probably the most dominant technological system of our time: *the Internet*. The Internet as a system will be the main focus of this chapter.

There were several reasons for postponing this endeavour. First, although the Internet surely shares some key characteristics with the earlier-mentioned systems in late modernity, it also differs from these systems in such fundamental ways that it would be fictitious to approach it in a similar fashion.

The arrival of the Internet has been judged to have a disruptive effect on a wide range of domains in everyday life so that we could truly speak of a *revolution*, bringing about the need for new narratives and concepts to capture this fundamental upheaval. Some speak of the *Network Society* (Castells 1996; van Dijk 2012), the *Information Age* (Castells 1999), *Hyperhistory* (Floridi 2015b), or the *Networked Era* (van den Berg and Keymolen 2013) to mark the fundamental shift in the way human life is being organized nowadays.

The Internet is considered to be a *critical infrastructure* just as water and electricity are, vital to society as a whole and to everyday life (Lewis 2006). Moreover, the Internet has become the enabler of all sorts of transactions: from financial

transactions such as online banking, to physical transactions like controlling engineering plants and flood-control dams. It has been adapted in a wide range of processes, knotting different formerly separate infrastructures together (Luiijf et al. 2003.) by "what are known as 'supervisory control and data acquisition' or SCADA systems" (Singer and Friedman 2014: 15), fundamentally changing their *modus operandi*. All in all, dealing with the Internet as just another system in late modernity would be misleading. Such an analysis would not do justice to the impact it has on everyday life and on the way it shapes trust.

A second reason for not immediately diving into cyberspace is that the intrinsic elusiveness of trust has lead me to look for 'family resemblances' of trust-related concepts first, instead of grounding the analysis on an already existing, all-encompassing definition of what trust might be. This strategy provided me with a family of connected concepts of trust: *interpersonal trust, system trust* or *confidence*, *familiar world*, and *the reduction of complexity*. We can now employ these concepts to investigate trust in relation to the Internet. As trust can only thrive in a familiar world, the central question of this chapter will be:

'Can the Internet function as a familiar world?'

In other words, is the Internet designed and organized in such a way that it facilitates shared expectations, that it enables a common view on the online reality? In order for trust to be a fruitful way of dealing with complexity, there already has to be set in place an environment in which the most rudimentary forms of complexity have been reduced. Trust is a strong force to deal with complexity, but it cannot carry it all. Therefore, the Internet has to function in such a way that it to some extent is predictable and reliable. If users have to cope with too much fundamental insecurity, for example if they have to make sure that protocols are still set in place, if certificates are genuine, or if they doubt that their e-mail is been compromised, then trust becomes pressured and more difficult to give.

As a way of analysing the Internet and its role as a familiar world, I have – inspired by the layered design of the Internet itself which will be discussed in section 4.2- chosen to focus on four cornerstones of the current Internet, which are crucial to our understanding of trust mediated by Internet technology: *context, construction,*

curation, and *codification*. Of these, *curation, codification,* and *construction* will be central to this chapter[21]. *Curation* refers to the key actors who govern and steer the Internet. *Codification* refers to the rules and regulations put forth by these curators. *Construction*, in this chapter, refers to the design of the Internet as a system, an infrastructure on which applications can be built. In the next chapter, where the focus will shift from the macro to the micro level, *construction* will primarily concern the *design of smart artefacts and services*. *Context* -the phenomenological experience of users on the micro-level- will then be the main point of departure.

In this chapter we will start with a short history of the Internet. In that section, I will focus on the role *trust* has played in the *collaboration of the developers* of the Internet. Their trust-based approach conspicuously contributed to some of the essential characteristics of the construction of the Internet. Next, I will look into the functioning of the Internet and, more specifically, I will concentrate on its *layered structure*. I will show how this technical structure has influenced Internet scholars in the social and political sciences to conceptualize the Internet in a layered manner as well. Finally, I will retake the history of the Internet, now focussing on *the shift from an open to a controlled Internet* and the way in which curators such as governments and private companies contribute to the familiar world online as well as jeopardize trust online because of their conflicting interests.

4.1 A short history of the Internet

In the introduction of this chapter, I conveniently spoke about the Internet as if it is a conspicuous, clearly defined object. However, the opposite is in fact the case. In general, the average user –including myself- is well able to operate and navigate the Internet, but hardly ever do we comprehend its technical functioning.

In this section, we will therefore first take a short dive into the history of the Internet and analyse some key design concepts of the Internet: layering, decentralized design, TCP/IP, and HTTP. Moreover, we will differentiate between the

[21] *Context* -the contextual, phenomenological experience of users- will be the main focus of the next chapter.

Internet as an *infrastructure* and the World Wide Web as an *application,* focussing on the way in which the trust-based approach to develop both the Internet and the World Wide Web has had a significant impact on their workings.

Although we nowadays make use of the Internet as if it has always been there, just as we have become used to having tap water and the ubiquitous availability of electricity, it in fact is a fairly new invention. Leonard Kleinrock (2010), one of the founding fathers of packet-switched networks –which, as we will see later, became the design backbone of the Internet- discerned "two threads" emerging in the late 1950s and coming together in the early 1960s that led to the early development of the Internet.

On the one hand, there was the *academic research thread*, mainly localised at Massachusetts Institute of Technology (MIT) where Kleinrock wrote his PhD on a *mathematical theory of packet switching for dynamic resource sharing* and subsequently at the university of California at Los Angeles (UCLA) where he started working after finishing his PhD research.

On the other hand, there was the creation of the defence organization called Advanced Research Projects Agency (ARPA)[22]. J.C.R. Licklider, the first director of the Information Processing Techniques Office (IPTO) of ARPA, envisioned how "networking computers could support social interaction, and provide networked access to programs and data" (Kleinrock 2010: 28). ARPA wanted to connect its defence investigators to the few large and very expensive computers that were spread around the country, enabling them to "…share each other's hardware, software and applications in a cost-effective fashion" (idem: 29). Packet-switch network theory as developed by Kleinrock laid the technical foundation for this endeavour. Independently, around that same time, other researchers such as Paul Baran at Rand (U.S.A.) and Donald Davies and Roger Scantlebury from the UK also worked on related subjects (see Abbate 1999).

The ARPAnet kicked off with, as its first node, the computer of UCLA where Kleinrock was in charge. Soon, three other institutes and partners of the ARPAnet

[22] The agency has changed its name to Defence Advanced Research Projects Agency (DARPA) in 1971, then back to ARPA in 1993, and back to DARPA in 1996. I refer to it as ARPA, its original name.

followed: the University of California at Santa Barbara, the Stanford Research Institute, and the University of Utah. In her book *Inventing The Internet*, Abbet (1999) describes how especially two approaches, one about the *content*, the other about the *process*, were perceived to be essential and eventually successful aspects of the ARPANET, namely: *layering* and an *informal* and *decentralized management style*.

First, *layering* is a way of computer programming where the programme consists of separately functioning components. Interaction between layers has to follow certain rules; consequently, a designer of one layer does not have to fully understand the workings of the other layers, only their interaction. Moreover, layers can be designed and modified separately, as long as the designers agree to use shared interfaces. Consequently, layering is a programming strategy which enables to work with systems that, when taken as a whole, are rather complex. Abbet (1999: 51) concludes:

> "Thus, layering has both technical and social implications: it makes the technical complexity of the system more manageable, and it allows the system to be designed and built in a decentralized way."

Second, the participants in the ARPANET project worked together on an *informal basis*. They trusted each other to do their best and this management style was praised not only by the contractors but by an external consultant who evaluated the project as well. Abbet (1999: 55-56) quoted the report, which stated that the ARPANET had:

> " '... been handled in a rather informal fashion with a great deal of autonomy and an indefinite division of responsibilities among the organizations that address the various elements of this function'. The report continued: 'Personal contacts, telephone conversations, and understandings are relied upon for day to day operation. This environment is a natural outcome of the progressive R&D atmosphere that was necessary for the development and implementation of the network concept.'"

This trust-based approach, valuing the autonomy and expertise of the scientists

involved, and the layering approach together have had a significant impact on the way in which the ARPANET and, subsequently, the Internet has evolved.

For instance, it made it possible for the participants to *incorporate their own values* into the design of the network. Some found it important that the network ensured a quick response as if one was interacting with a local device and not with a computer on the other side of the country. Others, such as Kleinrock, insisted that there would be measurement software installed to monitor the performance of the network (Abbate 1999: 56). Although there were clear conflicts of interest between the participants, the "dominant paradigm remained one of collaboration" (idem: 72).

After the connection between the first four nodes had been established in 1969, the ARPANET project subsequently focused on the linking of other networks. The 'internetworking' project led by Vincent Cerf and Robert Kahn took off in the 1970s (Naughton 2012: 45). Together with their colleagues coming from both academia and government, they wrote the *Transmission Control Protocol/Internet Protocol* (TCP/IP), often simply referred to as the *Internet protocol*: a universal language for computer networks.

> "It would allow applications to run over an internetwork while hiding the differences between network protocols by using a uniform internetwork protocol" (Kleinrock 2010: 34).

All kinds of devices and networks –from personal computers to the network of governmental departments- would be able to connect and share information. The challenge of connecting all these different devices –even taking into account the possibility of connecting devices that still had to be invented!- directed the founders to develop an internet that was *open, minimalist,* and *neutral* between applications (Goldsmith and Wu 2008: 23).

The Internet design is *open* because it accepts almost any device that wants to join. It is *minimalist* because the requirements to join are low. No internal changes to the network have to be made in order to connect. Basically, if a device or network can run the Internet protocol it can join. The internet is *neutral* between applications because it does not matter whether one wants to send e-mail, movies, or any other kind of application through the network, all will be treated alike.

On January 1 1983, the ARPANET made the transition to TCP/IP. This

changeover was carefully prepared and went very smoothly, resulting in a distribution of buttons saying "I survived the TCP/IP transition" (Leiner et al. 2009: 7). Moreover, the transition to TCP/IP enabled the ARPANET to be split into a "MILNET supporting operational requirements and an ARPANET supporting research needs" (idem).

All in all, this transition marks the start of the Internet as we know it (Naughton 2012: 45). Within a couple of years after the transition, the Internet was a widely used technological system for researchers and developers and, moreover, it started to become used for computer communications by other communities as well (Leiner et al. 2009: 8).

4.1.1 The World Wide Web

It was then Tim Berners-Lee, who came up with what later turned out to be the most successful application of the Internet, the World Wide Web, entered the stage. Berners-Lee was triggered by the problem of sharing and creating multimedia content. Up until then, it was only possible to share texts, however, the rise of *personal computers*, which were largely image-oriented, brought along the question of sharing data other than merely texts.

Again, a trust-based working environment facilitated Tim Berners-Lee's invention. He worked at CERN, the European high-energy physics laboratory in Geneva. That one of the most important Internet applications ever has been developed at this institute that actually focuses on physics rather than on computer science has everything to do with the freedom given to Berners-Lee and his team. It was not so much that CERN put a lot of support or means into the Web project, but they did not stopped it from growing either. His boss Mike Sendall wrote on the web project proposal "*vague but exciting*" and gave it the space to develop (Naughton 2012: 53). It was this trust in the abilities and cleverness of the scientists which facilitated the rather quick development. Within a couple of months, Berners-Lee had a working version of the Web. Three criteria lay at the foundation of the documenting system Berners-Lee developed (idem: 52).

First, the system should be *decentralized*. "In Berners-Lee's vision, the Web would create 'a pool of human knowledge' that would be easy to access" (Abbate 1999: 215). Just as the ARPANET of his colleagues at ARPA, the system should

accommodate diverse computer technologies. Berners-Lee built his new application on top of the TCP/IP protocol, which also ran on the computer systems of CERN. Together with his team, he designed the hypertext transfer protocol (HTTP) to enable the transfer of information between Web browsers and Web servers.

Second, central to the system was the idea of *hypertext*. Hypertext basically means that documents must be able to hold internal links to other documents. With hypertext it is no longer necessary to present information in a linear way. This idea of hypertext was central to the hacker counterculture of the 1960s and 70s. Berners-Lee added to this idea of hypertext the use of multimedia, in order to build a 'world wide web' of information (Abbate 1999: 214).

The third design requirement was that it would have to become possible for this system to link and connect documents across the worldwide Internet. This global aspect of Berners-Lee's system made it particularly important to also develop a uniform way of identifying the information one wanted to access. He therefore created the *Uniform Resource Locator* (URL), which is a standard address format to specify both the type of application protocol and the address of the computer that has the requested information (idem: 215). Important to note is that, again in line with the design philosophy of the ARPANET, the URL could refer to a variety of protocols, not just HTTP. As a consequence, it became possible not only to refer to existing content residing on older Internet services, but also to enable connections to content running on new protocols.

On August 1st, 1991, Berners-Lee and his team released all their info concerning the World Wide Web onto the Internet and subsequently in 1993, they made sure that CERN provided a certification that the computer technology and program code was in the public domain, for anyone to use, alter, and improve. As Lessig (2001) pointed out, openness of code is an important feature to keep a platform neutral. It enables tinkering and makes sure that "users are not held hostage" (idem: 54). Just as the decentralized design of the Internet itself reflected trust in the users to find their own solutions, so does open code. All in all,

> "[a]n open code platform keeps a platform honest. And honest, neutral platforms build trust in developers" (Lessig 2001: 54).

It was the Web which made the Internet popular, supported by on the one hand the *wide-spread access* to the internet –provided by privatization- and on the other by the *technical means* for individual users to run the web software –provided by personal computers (Abbate 1999: 215).

> "[The Web] solidified the Internet's traditions of decentralization, open architecture, and active user participation, putting in place a radically decentralized system of information sharing… The Web's exciting multimedia format and the seemingly endless stream of new features offered by entrepreneurial companies put the Web at the center of public attention in the late 1990s. By which time 'the Internet' and 'the Web' had become synonymous to many people" (Abbate 1999: 217-218).

4.2 The Internet as a layered infrastructure

The layered design approach of the scientists involved in developing the ARPANET – which eventually evolved into the Internet- turned out to be influential beyond its own domain. In the following section, I will briefly describe the layered and decentralized design of the Internet and look into the way in which this particular set-up has an important impact on the way in which the Internet is being *conceptualized* in other research domains. Moreover, I will give a concise overview of some influential adaptations of the layered model and take into account some of its weak spots. I will show how these layered models have also served as a point of reference for the development of my own conceptual lens consisting of the "4 Cs": context, construction, codification, and curation.

The multi-layered construction of the Internet has found its way in a wide range of academic disciplines which focus on the Internet, such as the philosophy of technology, law and ICT, Internet Governance Studies, etc.,… (for an overview see: Broeders 2014: 18). Depending on the research focus, the number of layers and level of detail in the analysis of the layers may differ and vary. However, most scholars agree that the Internet has a *physical layer* (idem). Sea cables, servers, and modems are all necessary, physical components for the basic existence of the Internet (for an analysis of this physical layer of the internet see: Blum 2012). Next to this *physical*

layer (infrastructure), at least two more layers have to be added: the *layer of transport and operations* (software code) and the *layer of application services* (content) (van Dijk 2012: 51). These three layers are central to all computer networks. Subsequently, these three basic layers can be subdivided in seven more-detailed layers, following the *Open System Interconnection Model*, which most network engineers refer to as the *OSI Reference Model* (See Table 1).

The OSI model is the standard model for conceptualizing computer networks. It describes how "…different applications and protocols interact on network-aware devices" (Briscoe 2000: 13). When a message from an application running on device A has to be sent to an application running on device B, this message has to descend from the application layer, all the way down to the physical layer and go up again to be delivered to the receiving application. The application layer, where the process starts, is the only part of the process a user actually sees and it in fact is only a small portion of what the application does to prepare the message before it can be sent over the network (Gralla 2007).

The Internet is, partly thanks to the work of Leonard Kleinrock, a 'packet-switched network'. On the Internet, there is no straight, unbroken connection between the sender and the receiver, in contrast to, for example, a traditional telephone line, which is dedicated to just one contact after a connection has been made. " [I]nstead, when information is sent, it is broken into small packets, sent over many different routes at the same time, and then reassembled at the receiving end" (Gralla 2007: 13).

The TCP/IP protocol, the universal language for computer networks, which is central to the functioning of the Internet, can also be found in the OSI Model, on layer four and three respectively. Essentially, the function of TCP is to break up "every piece of information and message into pieces called *packets*, deliver those packets to the proper destinations, and then reassemble the packets into their original form", where IP "is responsible for ensuring the packets are sent to the right destination" (Gralla 2007: 19).

At the physical layer, the packets are encoded into the medium that will carry them and that sends the package to that medium. Finally, at the receiving node, "the layered process that sent the message on its way is reversed" (Gralla 2007: 15).

Table 1 Seven technical O.S.I. network layers enabling network use (adapted from van Dijk 2012: 52)

Nature	Network Layer	Function
Content	7. Application	Enables the use of the content of applications (telephone conversation, Internet exchange, broadcasting, etc.) determining the identity and availability of communication partners and synchronizing communication
Software code	6. Presentation	Formats and encrypts data of applications using a different data language in order to be readable across networks.
	5. Session	Controls sessions ('dialogues') between different computers or hosts.
	4. Transport	Reliable transmission of data between end-users. Control of data streams. E.g. TCP (Transmission Control Protocol).
	3. Network	Path determination and addressing of data (packages) between different networks. E.g. Internet Protocol (IP4, IP6, IPsec).
Infrastructure	2. Data link	Physical addressing between multiple devices and a transmission medium.
	1. Physical	Electrical and physical specifications for devices and for the connection between devices and transmission media.

As the OSI Model is the standard way of conceptualizing computer networks, it became -often in a slimmed-down tripartite form of *physical infrastructure*, *code*, and *content*- the starting point for other conceptualizations of the Internet. Because the OSI Model and other derived conceptualizations primarily focus on what happens

behind the interface, scholars in the broad domain of social sciences interested in the interactions of actors with and on the Internet have added different layers to this model, making it more fitting for their research.

For example, Deibert et al. (2012b) in their study on cyber warfare in the 2008 Russia-Georgia War, differentiate between *cyberspace* and *Internet*. In line with the current definition of the US Department of Defence, they define cyberspace as:

> "a global domain within the information environment consisting of the interdependent network of information technology infrastructures, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers" (Deibert et al. 2012b: 5).

Two aspects are important: first, this definition confirms the *close connection between the physical and informational domain*. Second, it presents cyberspace as the covering realm of all information structures, including but not limited to the Internet. Next to the physical infrastructure level and the code level, which we know from the OSI model, they have added two other levels: the *regulatory level* and the *level of ideas*. The regulatory level includes the norms, rules, laws and principles, which govern cyberspace. The level of ideas is the sphere "through which videos, images, sounds, and text circulate". In the context of cyber warfare, Deibert et al (2012b: 6) refer to the standing practice of governments to "generate information effects" on this level. But more generally one could say that in the interaction between users and content, all sorts of new, meaningful information can emerge. The level of ideas, therefore, is not restricted to the domain of cyber warfare.

DeNardis (2012) takes the technical infrastructure of the Internet as her starting point for analysing how Internet governance is not just about maintaining the technical backbone of the Internet but also revolves around *arrangements of power*. Though the "complex technical architecture beneath the layer of applications and content" may be out of view for the average user, its "design and administration internalize the political and economic values that ultimately influence the extent of online freedom and innovation" (Denardis 2012: 721). By looking further than the visible layers of the Internet, Denardis uncovers the power struggles taking place on the technical layers below and to analyse how they have an impact on the Internet experience of the average user.

And finally, Schermer and Lodder (2014: 3-6) not only present a layered

conceptualization of the Internet as a communication system[23] but also develop a layered perspective on the *services* that are offered through the Internet. They show how actors on different levels are involved in governing and facilitating the Internet.

4.2.1 Critique

It has to be noted that there has also been some critique on the use of the OSI Model and related technical models to analyse the Internet. For example, Jonathan Zittrain (2008) convincingly argues that an analysis of the Internet which primarily focuses on the literal network itself does not sufficiently take into account the effect of the endpoints–this can be a laptop, smartphone, tablet or any other internet-based device- and, consequently, on the experience of the users. As we will see in the next chapter, the way in which endpoints or applications are designed –as open or closed/tethered devices- may even fundamentally shape the organization of the network itself. It has to be stressed that the endpoints are not included in the OSI Model. Although Zittrain (2008: 8) does not deny that an analysis of the network itself is important, a too stringent focus on the network may obscure "the reality that people's experiences with the Internet are shaped at least as much by the devices they use to access it"[24].

In a fashion similar to the scholars above and taking into account the critique of Zittrain, the conceptualization of the Internet in this book is based on a layered approach. Instead of perceiving the Internet as a homogeneous system and consequently missing out on the different actors (material, individual, and organizational alike), it focuses on four levels of the Internet, in order to take into account those aspects of the Internet which dominantly influence the formation of

[23] Schermer and Lodder make use of the TCP/IP model as described in the RFC1122, which is related to the OSI model. However, the former model focuses more than the OSI model on the transport layer. The transport level roughly corresponds with layer 4 of the OSI model.

[24] Preceding the next chapter, Zittrain's emphasis on the importance of including in the analysis the devices which are used to access the Internet is in line with the focus of Science and Technology Studies (STS) and the current Philosophy of Technology on the role of artefacts in the way users interact in and perceive the online world.

trust. As already presented earlier in this book, I conveniently called them 'the 4 Cs': *construction, context, curation,* and *codification*. It becomes clear now that some of these levels overlap with some of the levels or layers as described and used by other scholars.

For example, the *regulatory level* as articulated by Deibert shows some resemblance with the *level of codification* as they both refer to the norms, laws, and regulations steering the Internet and interaction online. The *level of curation* is more or less in line with Denardis' idea of *arrangements of power* because both take into account the influence of different stakeholders in cyberspace.

However, in some aspects, levels also differ. For example, in this chapter, which focuses on the Internet as a system, *construction* chiefly refers to the layered design of the Internet as described above. However, in the next chapter that revolves around the user's experience of the Internet through smart artefacts, the focus will shift to the micro-level. Then, I will mostly take into account the design of the artefacts themselves.

By focussing on this micro-level in the next chapter, taking into account the phenomenological experience of users or, to put it differently, the *context* in which the interaction with and through their devices takes place, a new layer is being added to the conceptual models described above. By including *context* as one of the important cornerstones of the conceptualization of the Internet and taking into account the construction of the smart devices and services mediating these interactions, Zittrain's objections that the individual experience of the Internet is too important to be left out of the analysis are met.

All in all, looking at the way in which the Internet as a technical infrastructure has been built, trust turned out to be one of its key characteristics. In its early days, *the Internet was a familiar world*. The trust-based collaboration of the developers of the Internet is reflected in the design of the system itself. *The open and decentralized architecture of the Internet expresses confidence* in the ability of users to solve problems and innovate. Moreover, it is *an acknowledgement of the unexpected*, of the creativity inherent in human life. *The founding fathers of the Internet knew that they could not know* what the possible purposes of the Internet would be. Therefore, instead of aiming at control, they chose trust as their dominant strategy to deal with complexity. As a result, they ensured the freedom to create and innovate (also see

Lessig 2001).

4.3 A new online reality

As a way of ending this chapter and simultaneously introducing the following one, we will recapture the evolution of the Internet, focussing now on how the dominant *Internet ideologies* have changed over the years and how the Internet itself has lost some of its innocence along the way.

Where the early Internet (phase of the Open Commons 1960-2000) was grounded on trust and the design was set up in an open and decentralized manner bringing forth a familiar world, this strategy for handling the complexity accompanying the Internet increasingly became *contested* with the *commercialization of the Internet* in the 1990s (Deibert et al. 2012a).

Instigated by the widespread use of personal computers, the Internet entered the home and opened up a new space for governments, citizens, companies, and customers all over the world. New companies and services found their way online and also governments entered the scene. Although the latter were rather late to catch up on the important developments that were taking place online, from 2000 on, governments regained their place and asserted a more dominant position in cyberspace (also see: van Eeten and Mueller 2013: 722)[25].

This interplay of new actors online brought forth *three fundamental changes* to the early Internet that have had a major impact on the current establishment of trust online and the idea of the Internet as a familiar world.

First, the initial tech community no longer is in charge of the Internet

[25] Security-expert Bruce Schneier (in Gasser et al. 2013: 11) suggests that this delayed picking up of the Internet by the government is related to the functioning of technology in general. He claims "technology magnifies power in general, but the rates of adoption are different." Those who are not organized and distributed (early-adapters, geeks, hackers, criminals, etc.) can make use of new technologies faster. However, when institutionalized powers come to grips with the new technology, they can make use of it more effectively and, therefore, establish and even expand their influence through the new medium. To put it differently, it takes more time for states to adapt to new technologies, but when they succeed, their initial position of power is consolidated or even increased.

infrastructure. It is now up to a heterogeneous group of actors –states, companies, NGOs, civil society- to work together and maintain the Internet as a familiar world, a background against which trustworthy action can develop.

Second, these actors not only maintain the Internet, they also steer and mould it to cater their own needs. Looking after their own interests may conflict with their function to ensure the familiarity and stability of the infrastructure of the Internet.

Third, online *information intermediaries* such as Google and Facebook increasingly mediate the online experience of users. This mediation pre-sorts the actions and interactions of users in ways opaque for the users. The role of these intermediaries will be addressed in more depth in the next chapter.

These three key changes have a fundamental impact on the way in which the Internet currently *functions as a familiar world* and on the way in which *users experience the Internet* and build trust (context level).

Where in the initial stage of the Open Commons, *interpersonal trust* was the fundament for building and interacting online, in the subsequent stages of increasing commercialisation and regulation the open character and trust-based interactions become more and more pressured. How this shift from an *open* to a *controlled* Internet influences the experience and actions of Internet users (context level) will be the focal point of the next chapter.

4.3.1 The phase of the open commons

As we saw in section 4.1, the designers of the ARPANET and later the Internet, collaborated on a personal, rather informal basis. There was no detailed plan laid out and they tackled problems as they came. If there would be no consensus on how to solve a problem or on which direction to take, they mostly just discussed the issue up until one of them was able to convince most of the other parties involved. In the words of one of the designers of the early Internet, Dave Clark: “We reject: kings, presidents, and voting. We believe in: rough consensus and running code” (Goldsmith and Wu 2008: 24). This open way of working had as a consequence that the designers were able to translate their own values into the design. In other words, the design of the Internet reflected their motives. The legal scholar, Jonathan Zittrain (2008: 28) noted that the creators of the Internet:

> "...had little concern for controlling the network or its users' behavior. The network's design was publicly available and freely shared from the earliest moments of its development. [....] Energy spent running the network was seen as a burden rather than a boon. Keeping options open for later network use and growth was seen as sensible, and abuse of the network by those joining it without an explicit approval process was of little worry since the people using it were the very people designing it."

This design approach presupposes two fundamental assumptions, which Zittrain (2008: 31) refers to as the *procrastination principle* and the *trust-your-neighbour approach*. The former refers to the belief that other users can address most problems occurring in a network. In other words, the designers did not aim at foreseeing and fixing all the problems that might occur when using the network. This would presuppose a vast amount of control which not only would cost them a lot of time and money, it would, even more importantly, conflict with their aim of developing an open network to which artefacts, not yet even invented, could connect without a great deal of effort. After all, too much control would hinder innovation and the generative quality of the network itself. They therefore trusted that other users could come up with working solutions as well. Consequently, the robustness of the network partly depended on the actors using the network.

This brings us to the latter assumption, the *trust-your-neighbour approach*, which refers to the belief that the users were competent enough and with good intentions not to deliberately hinder the functioning of the network. *Trust between the developers* as well as *trust in and between the other users*, all belonging to the same tech community, was essential to the functioning of the Internet in its early days.

From a trust perspective we can say that the designers chose *trust over control* as the dominant strategy to deal with complexity online. They had no intention to monitor exactly what the users online were doing or what kind of content travelled over the network. Striving for simplicity, by no means easy to attain, allowed them to *prioritize connectivity over security*. Instead of installing a variety of safety measures that would probably strengthen the security of the network but also conflict with their aim to create an open network, the designers chose to let the safety issues be taken care of at the end points, so on the level of the users, and not in the network itself. Building their network on trust made it possible to really focus on the functionality of

it. It allowed them to develop a network that was open to anyone who wanted to join and any device that wanted to connect, to treat all data in the same manner and to be sent from anyone to anyone (Zittrain 2008: 32).

Goldsmith and Wu (2008) in their explanation of the Internet design made a connection to the zeitgeist of the 1960s and 1970s. Although the creators of the Internet were not explicitly engaged in political activism, their design did reflect the growing belief that there could be governance "liberated from national or physical identity" (idem: 16). Goldsmith and Wu (2008: 23) found that the designers "built strains of American libertarianism, and even 1960s idealism, into this universal language of the Internet." They developed a global network, which reflected distrust for "centralized control" (idem). Consequently, the Internet is probably the first information-related innovation that resulted in a technology almost everybody can access, making use of a multitude of devices on a neutral net (Zittrain 2008; Wu 2011). All in all, there was the general conviction that the arrival of the Internet would enable a shift from national, governmental power to a more bottom-up, self-regulating domain –often referred to as "cyberspace"- standing apart from the physical world, which was also called "meatspace".

When at the end of the 1980s and even more in the first part of the 1990s the Internet increasingly became populated by individual users (users not aligned with a research or governmental defence organisation), these values of openness, decentralization, and self-regulation put into the system by the creators, still remained the basic assumptions for interaction online. Deibert et al (2012a) refer to this phase as "The Open Commons"[26] [27] by which they emphasize the separate status of the Internet as a domain where people were able to govern themselves. The Internet was addressed as

[26] For Deibert et al (2012) the Open Commons phase starts already in the 1960s and ends in 2000. They do not make a distinction between the early design phases where ARPA was developed and subsequently the phase where the Internet became known and used also by non-academic actors. Because the values central to the Open Commons perspective are by and large in line and, even more so, depending on the values of the designers of the Internet, it is for the purpose of this chapter not necessary to make a strict distinction on the matter.

[27] The Open Commons is also referred to as the *Digital Commons* or the *Open Internet Perspective*.

a “fresh start” for democracy, free from traditional governmental intrusion. The possibility for collective action was valued as its main democratizing force. By making use of blogs, online communities, and cheap technologies, it became possible to support a real *global civil society*. The Internet was a place where people could experiment with their identity (Turkle 1984), find like-minded people (Rheingold 1993), and hope to form the “first truly liberated communities in human history” (Goldsmith and Wu 2008: 16). It was also the time that *Electronic Frontier Foundation* (EFF) started off, probably up until today one of the most influential Internet NGOs. This non-profit organization, founded in 1990, sees its goal as to defend civil liberties in the digital world, such as free speech online, user privacy and innovation. It defends the values held dear by the designers of the Internet and its early-adopters. John Perry Barlow (1996), one of the founders of the EFF, posted his “Declaration of the Independence of Cyberspace” online in which he declared:

> “Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather. [...]
>
> You claim there are problems among us that you need to solve. You use this claim as an excuse to invade our precincts. Many of these problems don't exist. Where there are real conflicts, where there are wrongs, we will identify them and address them by our means. We are forming our own Social Contract. This governance will arise according to the conditions of our world, not yours. Our world is different.
>
> Cyberspace consists of transactions, relationships, and thought itself, arrayed like a standing wave in the web of our communications. Ours is a world that is both everywhere and nowhere, but it is not where bodies live.
>
> We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth.
>
> We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence

> or conformity. [...] "

Looking at the Declaration, it appears that there was a *utopian belief* that in cyberspace it would be possible to interact without a top-down government imposing rules and regulations. Instead, the inhabitants of cyberspace would, from the bottom-up, built their own society, solve their own issues and install –if at all necessary - their own rules and regulations. The trust-based approach the designers of the Internet had chosen for their collaboration and which consequently allowed for the open design of the network, was seemingly adopted by most members of the online community.

Where in the previous chapter, we analysed the arrival of *system trust* or *confidence* in late modernity, enabling people to interact with abstract systems and strangers, in the open commons phase of the Internet, the dominant view is that cyberspace is grounded on *interpersonal trust*. This belief in a self-regulating cyberspace presupposes a return to the local, pre-nation state ideal where interaction is dominantly based on interpersonal fundaments such as reputation, shared norms and values, third party trust, and strong ties. It is revealing that probably the largest technological system of our time brings along the existential longing for a pre-modern manner of interaction and society building. Botsman and Rogers[28] (2010: xiii-xiv), two contemporary advocates of the 'online commons' formulate it as follows*:*

> "Online exchanges mimic the close ties once formed through face-to-face exchanges in villages, but on a much larger and unconfined scale. In other words, technology is reinventing old forms of trust".

In chapter seven, on the role of trust on the platform Airbnb, I will argue that this belief turns out to be mistakenly utopian nowadays. Although interpersonal trust, beyond any doubt, is still an important strategy to reduce complexity in the online environment, it is also fundamentally mediated. The technology itself, but also key actors involved such as the company Airbnb and different regulators, have an impact on the way in which trust is being established. Interpersonal trust online is not just

[28] Also see chapter 7.

about you and me (context), but about you, me and the other three Cs involved (code, curation, and codification).

4.3.2 Denied, controlled, contested[29]

The *open commons* perspective is definitely still an important and influential *normative perspective*. Key authors, such as Tapscott (2006), Chesbrough (2006), Benkler (2006, 2011), and Bauwens (2012) write extensively on collective action, online community building, open innovation, and the digital commons[30]. Also on a micro-level, concepts underpinning the *open commons* perspective are still dominant. The average user experiences online interactions as being merely facilitated and not presorted by the technology involved. The Internet, as for example the movement of collaborative consumption proclaims, enables users to gather and collaborate on interpersonal grounds without being steered by governments or companies.

However, although the open commons perspective is still alive and kicking as a normative perspective, as a *descriptive perspective* of how the Internet –including all the layers behind the visible application layer- functions, it has lost its strength (Deibert et al. 2012a). The arrival of major actors such as governments, companies, and non-profit organisations has fundamentally changed the character of the Internet. The initial idea of the founders of the Internet, that self-regulation and trust in the users of the network would be sufficient to develop a robust Internet, had to make way for *a balancing act* of different state and non state-actors that are governing the Internet, which has been referred to as the *multi-stakeholder model* (also see: Schermer and Lodder 2014: 16). International institutions, government agencies, and governments, but also ISPs, search engines, social media platforms, and web hosting companies all regulate a substantial part of the Internet. An often-cited definition of this Internet governance, coming from the report of the Working

[29] Deibert et al (2012) make a stricter distinction between the phases of Access Denied, Access Controlled, and Access Contested.

[30] In chapter 7, we will focus on trust in the domain of Collaborative Consumption, a new way of doing business, facilitated by the Internet. Collaborative Consumption can be seen as a practical application of the Open Commons.

Group (2005: 4) on Internet Governance states:

> "Internet governance is the development and application by governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet."

Instead of the initial, rather homogeneous tech community who developed the Internet, now all these different curators have become responsible for developing standards, protocols, and Internet policy, maintaining a *stable Internet environment* (for an overview of the actors in internet governance also see: van Eeten and Mueller 2013; Mueller 2010; Simonelis 2005). To function as a trustworthy technology, the Internet has to have a 'taken-for-granted' character to its users.

For example, the Internet should not transform from a packet-switched network to a centralized network overnight, nor should users suddenly have to type all the digits of the IP address in order to go to a website. Such uncertainties could have devastating consequences for the overall confidence people have in the functioning of the Internet (or in the Web, for that matter). If users had to decide every time they use an online service or make use of an application on their smartphone, whether or not to trust the underpinning infrastructure of the Internet, the costs would simply become too high. People would be overwhelmed by the uncertainty of such a complex and unstable environment and probably reject it as a valuable means of interaction. For trust to take place on the interpersonal level –or between a user and an organisation-, the environment, whether or not online, should be familiar first, that is, stable and predictable.

In general, the Internet user is not consciously aware of the presence or functioning of these stakeholders. It is often only when the Internet connection is down that a user thinks of the Internet Service Provider. Or only when one needs more space for a personal website, the webhosting company comes into mind.

The so-called content intermediaries (Denardis 2014: 153-172) however, are the exception to the rule. They offer the platforms where users can share their photos (Instagram), post messages (Twitter), interact with friends (Facebook) and upload and watch movies (YouTube). In general, users are more personally involved with these companies than they are with their ISP or web-hosting company. Everyone recognizes the typical white webpage with the search bar in the middle as being the

start page of Google. And the bluish "thumbs up" of Facebook is hard to avoid when going online. We know how these companies look like and we use their services on a daily basis. All large content intermediaries (Google, Facebook, Twitter, Youtube, etc.) are in the top ten of most-visited websites. Google heads this list with 1, 100 000 000 unique visitors a month[31].

Although these content intermediaries, just like the other technical actors working in the background, contribute to a familiar world online, they do this on a different level. Where ISPs and hosting companies mostly add to the *accessibility and technical stability* of the familiar world, the platforms, on the other hand, *mediate our experience of the familiar world online*. The way information intermediaries choose to design their interfaces and enable functionalities, co-shape the actions and interactions of users online. This will be the focal point of the next chapter.

It is not in the scope of this book to judge the sustainability or the success of this stakeholder model. Where these Internet governance actors succeed to develop protocols and procedures, ensuring the stable functioning of the Internet infrastructure, they contribute to the familiar world online, a necessary condition for trust to be established.

The way these actors govern the Internet partially moulds the online environment in which trust is being established and influences the way users put trust in the Internet as such. In other words, to understand how trust is developed between users online, between users and companies online, and between users and governments online, one has to take into account how these actors shape the familiar world of the Internet. In line with political philosopher Langdon Winner's (1980)[32] "artefacts have politics", we can say that there is a *politics of the Internet* as well. These actors –although out of sight for the average user- have become *active players* in the Internet environment. Or as DeNardis (2014: 7) puts it:

> "The complex institutional and technical scaffolding of Internet Governance is somewhat behind the scenes and not visible to users in the

[31] http://www.ebizmba.com/articles/most-popular-websites, accessed on Sept 22 2014.

[32] The ideas of Langdon Winner will be discussed more in depth in the following chapter.

> same way applications and content are visible. Although these technologies lie beneath content, they nevertheless instantiate political and cultural tensions… Bringing infrastructures of Internet to the foreground reveals the politics of this architecture".

For most stakeholders their role online does not end with the governance of the Internet. These different actors, besides their task to maintain the Internet infrastructure, also make use of the Internet to pursue their own interests, which certainly does not always neatly align with their responsibility for a stable Internet environment and, therefore, may even put the familiarity of the online world in jeopardy.

Increasingly, a shift from "governance of the internet to governance using the internet" (Broeders 2015: 4) occurs which has an impact on the Internet's functioning as a familiar world. These curators not only *maintain* the Internet but also *steer* it in certain directions, based on their political or business motives and ideologies. Their choices bear *values* and are influenced by economic and political forces. They are crucial for the way the Internet develops and hence for the way trust is established in and through Internet technology.

The 2009 revolution in Iran, often referred to as the *Twitter Revolution* due to the use and coverage of the events through Twitter, is a prime example thereof. Where at first, it seemed that Twitter and other information intermediaries chiefly played an important role in organizing the resistance against President Ahmadinejad, these intermediaries turned out to be also excellent tools in the hands of the Iranian authorities to search and find the activists and, in their view, enemies of the state (Morozov 2011: 1-5).

Another example: March 2014, with the elections around the corner, the Turkish government simply blocked Twitter, and shortly after wiretapped recordings, damaging the government's reputation, were leaked on the medium. Amnesty International condemned this action as: "a blunt attack on Turkey's citizens' right to share and receive information"[33]. Previous to the blocking of Twitter, Erdogan was clear on the matter:

[33] https://www.amnesty.org/en/articles/news/2014/03/turkey-pre-election-twitter-shutdown-brings-internet-freedom-new-low/. Accessed 15 May 2015.

> "We are determined on the issue, regardless of what the world may say"… "We won't allow the people to be devoured by YouTube, Facebook or others. Whatever steps need to be taken we will take them without wavering" (quoted in Rawlinson 2014).

What these two examples illustrate is that while the average user may experience the Internet as an open and neutral infrastructure, it actually is being strictly monitored and controlled.

In his book *Black Code: Surveillance, privacy, and the dark side of the internet,* Deibert (2013) points out how we all increasingly rely on technology of which we actually know little about. That is where the "black" in the title of the book stands for. It refers to "that which is hidden, obscured from the view of the average user" (idem: 6).

Deibert shows that not only are cybercrime manoeuvres beyond the average user's awareness, but also private companies and governments operate under the radar. In the name of security and effectiveness the latter increasingly aim and succeed at controlling citizens, while steering clear of democratic control. The Snowden revelations on the functioning of the National Security Agency (NSA) in 2013 demonstrated that the *modus operandi* of this American governmental agency is on bad terms with the core aspects of the rule of law. The law, for example, does not seem to restrain government officials working for government agencies like the NSA to ensure that values like freedom and autonomy central to a democratic state are safeguarded. Sufficient control on the functioning of such security agencies is lacking. It becomes difficult to speak of a liberal democracy when it is permitted to "indiscriminately listen in on, watch, or otherwise collect everything we do and say online" (Deibert 2013: xiv).

Moreover, governments do not operate alone; often they delegate their surveillance activities to the *information intermediaries*. Google, Twitter, Facebook, Apple, Microsoft and others have all been pressured to block or remove content. A vast amount of these government requests are not accompanied by a court order and the receiving companies are generally not allowed to go into details about it, again side-stepping the checks and balances central to the rule of law.

It has to be noted that measures taken by governments and information intermediaries to control the online world can stem from the best intentions. A *controlled* environment is often also a *more secure* environment. One could therefore advocate that controlling the Internet and intervening online makes this online world more predictable, therefore, less complex. As a result, it might even become easier for the average user to establish trust, even when this means that the initial idea of the founders of the Internet that the infrastructure should remain open and neutral has to be abandoned.

However, one has to remember that for trust to thrive a delicate balance needs to be established between a stable and predictable environment on the one hand and an environment which allows for freedom to act and developing new thoughts and initiatives on the other. If we have 100% security –which is unlikely ever to happen, but perhaps 80% is just as significant- trust will become redundant, as we would know how things would end up. Insecurity must be brought back to a bearable level in order for trust to be established, but when this leads to a world that is not just familiar but completely ruled and controlled, trust loses its meaning (see chapter 8).

Also Nissenbaum (2004) warns against the seemingly self-evident move to strive for more security as it may endanger trust and the worthwhile practices it facilitates. Nissenbaum discerns three security mechanisms: *access control, transparency of identity,* and *surveillance*.

The first refers to passwords, firewalls, and other measures to ensure that only those actors who are allowed to enter –clients, citizens, members-, do enter and those who are not allowed –hackers, spies, criminals- are blocked.

The second mechanism is about making actors more identifiable. Where in the early days of the Internet, if you were technically able to connect you could go online, nowadays it has become increasingly necessary to give at least some of your credentials in order to go online or make use of a service. Even if a user does not willingly provide personal data, all sorts of cryptographic and profiling techniques are used to authenticate users. The basic idea is that if your identity is known you will think twice before acting malicious, as you can be held accountable for your deeds.

The third mechanism is based on the idea that monitoring actions and behaviour online can prevent bad things from happening or at least could help to quickly and easily find the wrongdoers.

Nissenbaum is clear on the fact that for specific actions online such as banking

or e-commerce, high security is necessary. However, this should not be a permit to strive for overall security in a way that it narrows down freedom, which is nurtured by trust to develop –amongst others- innovative, creative, or political practices.

Preliminary to the next chapter, it is not only the way in which governments together with private companies are policing the Internet which is threatening the open character of the Internet, but also the way in which users in general access that Internet is increasingly becoming sorted. In 2010, Chris Anderson and Michael Wolff (2010) wrote the influential article "The Web is Dead. Long Live the Internet" in tech-magazine *Wired*. In that article, they describe how users turn their back on the *open* World Wide Web –the Internet's most important application- in exchange for sleeker, but also more *controlled services* providing us with personalized information, tailored to our needs, sometimes even before we have become aware of those needs[34].

People do not search for the latest news on the Web, they just open their personalized news app. They do not look for like-minded people on the Web, but go directly to the walled garden called Facebook. And even when Internet users do end up on the World Wide Web, the top 10 Web sites account for the vast majority of pageviews.[35] Moreover, while users are under the impression that they are anonymously surfing the Web and that nobody is really interested in their online activities (Benoist 2008: 168), almost 80% of the most frequently-visited websites use tracking technology to gather information on their visitors (Angwin 2010).

On top of that, when entering the World Wide Web, almost all users make use of what Wu (2011) coined: "the master switch", better known as Google. To find information and to connect with other people we dominantly make use of Google's search engine. As a consequence, Google has a very important say in what we believe to be important information and what is not. Or as Wu (2011:281) puts it: "whatever shows up on the first page of a Google search is what matters in forming our sense of any reality; the rest doesn't."

[34] Chapter 8 on personalization will focus on the influence of online personalization on trust and the familiar world online.

[35] According to Compete, a web analytics company, cited in Anderson and Wolff (2010) the top 10 Web sites accounted for 75 percent of the pageviews in 2010.

According to Anderson and Wolff (2010) this gradual transition from "the wide-open Web" to an Internet colonized by "semiclosed platforms" does not mean that users reject the idea of the Web, rather it is just that these semiclosed platforms work better. In addition, semiclosed platforms in general do not only perform better and are more secure, they are also *more easily controlled* and therefore *more easily monetized*. All in all, semiclosed platforms make an excellent basis for companies to develop their services and consequently they steer innovation in the direction of these platforms.

The arrival of semiclosed platforms is closely related to the devices we use to access the Internet. For a long time, the personal computer was the one and only way to access the Internet. However, increasingly we make use of smartphones, smartwatches, tablets, and other devices to go online. These artefacts and more specifically their interfaces lead us to connect to the Internet in a different way than we were used to. Browsing the Internet on an iPhone is much less convenient than directly going to the designated app. Google maps on your phone is much more handy than on your desktop when trying to find your way back to the hotel. Moreover, artefacts increasingly are online even *without our active intervention*. The Internet of Things, which will be discussed more at length in the next chapter, represents the trend to connect all artefacts to the Internet, from refrigerators and cars to coffee machines. Through their connections these artefacts do not just remain updated, but when they are also so-called 'smart' they can learn from their interactions with users and pre-sort the interaction to cater the needs and wishes of their users. This is all fully automated.

All in all, artefacts running on Internet technology are just as the earlier-mentioned semiclosed platforms designed in a way that chooses control over openness. In his book *The Future of the Internet and how to stop it*, Zittrain (2008) warns for the advance of what he calls "*tethered devices*", devices which bundle hardware and software and which are controlled by the companies that sell them. In contrast to the values of openness, creativity and trust in the users to come up with their own ideas and solutions, users are becoming more and more dependent on the companies from whom they buy their devices. These devices are already completely programmed, the company remotely updates them, and users cannot break them open without losing their guarantee.

Although users pay a lot for their iPhones and other devices, *these artefacts never truly become theirs,* because even after they have left the Apple Store or other retailer, the devices remain connected and under remote control of the manufacturers who have the power to change the workings of the device (see also Deibert 2013: 229). The freedom central to the early Internet has to make way for user-friendliness and fashionable designs. In the next chapter we will look deeper into this *new ontology of Internet-mediated artefacts*, but for now it is enough to see that the devices we use to connect to the Internet are not neutral artefacts but important physical points of control which may endanger the openness of the Internet.

4.4 Conclusion

What all these studies on control on the Internet show, whether they focus on governments policing the Internet, companies involved in surveillance practices, or tethered devices enabling control over their users, is that the *commons* are not built in a vacuum but in an online environment in which free access to the online world becomes more and more *contested* (Deibert et al. 2012; Morozov 2011). The open commons perspective is no longer an adequate description of the state of the Internet today; rather, its principles became something in need of protection. In the words of Deibert et al. (2012a: 8):

> "The core elements of an open commons have now become the touchstones for a set of constitutive principles to be shored up and defended, as opposed to assumed away as invincible. Perhaps ironically, what were once assumed to be the immutable laws of a powerful technological environment are now potentially fragile species in a threatened ecosystem."

From a trust perspective, we can frame this shift to a controlled Internet as *a familiar world coming under pressure*. Trust can only flourish in a familiar world. A familiar world is the shared background of un-explicated norms and values against which we all interact. It consists of uncontested basic beliefs held by all actors, for example, that in general all actors perceive the world in a similar way and that they will act in

line with their social roles.

Dominant actors governing the Internet -governments, companies, and NGOs- contribute to the familiar world online by providing stability through political, societal as well as technical measures. However, at the same time they also pursue their individual interests, striving for more control of the Internet and its users. This double-sidedness should be acknowledged as an important aspect of the analysis of trust mediated by Internet technologies.

In this quest for more control, the curators of the Internet may destabilize the familiar world online. Interactions become sorted in a way users are not aware of nor have consented to, which –when they do find out- can have a devastating effect on the trust placed in the Internet as such, as well as in the actors who operate on the Internet.

With the arrival of a variety of new actors online, all with their own –often conflicting- interests, the familiar world online is no longer a self-evident, stable background. This does not mean that on an interpersonal level people cannot cooperate based on trust, but that online trust can easily be shattered by *external* influences.

This brings us to the core challenge of understanding and analysing trust mediated by Internet technologies. How to relate the power of the Internet to connect people and enable them to develop interactions based on interpersonal trust with the unstable familiar background shaped by the often conflicting interests of major players such as governments and companies? In other words, how do interpersonal trust and system trust come together in an environment mediated by Internet technologies?

5

Trust in context: a theory of mediation.

At the heart of this book lies my wonder about the ostensibly effortless way in which people interact online. They do not seem to doubt if the people they are interacting with are genuine nor do they seem to be bothered by the fact that the screen they are looking at and projecting themselves into is in fact an artificial world, consisting of bits and bytes. Even negative reports on the actions of governments and companies online do not really seem to hinder their online activity (yet)[36]. People using the Internet do not seem to be disturbed by the fact that an online interaction is a truly *mediated interaction*, enabled by networked and smart artefacts such as smartphones and tablets, steering and pre-sorting their experience. A *paradox of trust* seems to occur: although we are aware to a certain extent that there are risks when we make use of these artefacts, we act as if there are none.

Where we normally are tempted to try to solve paradoxes, in the second chapter, revolving around the ontological question if and why trust is a necessary aspect of human life, we saw that this paradox in fact lays at the centre of the concept of trust itself. To trust is to act as if the future is certain, as if uncertainties do not matter for the outcome of our interaction. It is the function of trust to enable us to act despite our uncertainty about the future and the way in which others might act. Trust is not about resolving uncertainties, but about accepting them. Trust provides us with

[36] It has to be noted that the Snowden-revelations concerning the practices of the NSA certainly did have (and still have) an impact on for example the US information technology industry, for example: "…foreign countries both react to protect their citizens' privacy and use the trust outage as a means to advance local competitors" (Richards and King 2014: 415). However, judging by their behavior, 'average users' seem to be less disturbed.

the fiction we need to face reality.

In the third chapter, we looked at trust through a broader, sociological lens and took into account its historical context. In modernity, the arrival of large systems in society such as the banking system, political systems, and global corporate systems altered the character of trust. Trust is no longer predominantly a part of interpersonal interactions, but increasingly also becomes a strategy to deal with uncertainties in interactions between persons and systems. In times where interactions increasingly gain a global and impersonal character, interpersonal trust had to make way for confidence or system trust.

In the fourth chapter, we delved into the Internet, the most dominant technological system of our time, and analysed the way in which the Internet was designed. Trust was not only a key factor in the cooperation between the founding fathers of the Internet, but it also found its way into the design itself. The construction of the Internet leant on openness and trust in the ability of the users to deal with problems themselves rather than aiming at anticipating all possible problems by securing and closing the network.

Due to the ambivalent role of the current curators of the Internet, this familiar world online is under pressure. Governments, companies, and actors from civil society on the one hand add to the stability of the familiar world online by producing and maintaining the protocols and technical standards, but on the other hand, when they pursue their own interest, they may endanger the stable and shared background against which trust can thrive.

Each chapter has provided important building blocks for understanding how trust online is established. In the second chapter, I determined the function of trust on the ontological level and discerned a family of concepts, which help us to understand the fuzzy concept trust turns out to be: interpersonal trust, confidence, system trust, a familiar world, and the reduction of complexity. I took into account the more socio-historical developments by looking at *system trust* in the third chapter. In the fourth chapter, I analysed the workings of the Internet itself and the challenges it might pose for the familiar world online.

In contrast to the fourth chapter, which approached the Internet as a technical system (construction), in this chapter, I will descend to the micro-level and focus on

the personal experience of users (context) mediated by networked artefacts.

I will take a *contextual approach* by taking up the thread of Plessner's three anthropological laws, based on the eccentric positionality of human beings. Central to Plessner's philosophical anthropology is that human beings, in their most fundamental form as living nature, should always be understood as being in interaction with their environment. Consequently, I will analyse how this interaction in the current *networked era* takes shape and which challenges this poses to trust.

5.1 Human beings, technology, and environment

As we have seen in the second chapter, one of the fundamental principles for Plessner's philosophical anthropology is the *interaction* between living nature and its environment. Whether one looks at plants, animals or human beings, their existence can only be understood if one takes into account their positionality; the way in which they actively uphold their boundaries and regulate the boundary traffic between them and their environment.

Focussing on human beings, characterized by their eccentric positionality, I discerned a first and second hiatus (in the second chapter of this book). The first hiatus all living nature has in common. By upholding their own boundaries, living things have an inner and outer side. Consequently, there is a hiatus between them and the environment. This hiatus or double aspectivity as Plessner calls it, brings forth the complexity that all living nature, including human beings, has to process. The *second hiatus,* however, is reserved for human beings who not only experience a distance between themselves and their environment, but also between themselves and their centre of experience. It is this second hiatus, which is unique for human beings and grounds their eccentric positionality that brings forth the radical complexity trust has to reduce.

Notwithstanding this second hiatus and the complexity it brings forth, human beings nevertheless remain, like animals, central creatures, living in the here and now, residing more often in a state of action than in a state of reflection, longing for the wholeness of the central position, which –as we will see later in this chapter- guides their interaction with technologies.

Plessner captures the eccentric positionality in three anthropological laws or principles. First, human beings are *artificial by nature*. This points at the ontological

necessity to replenish and anchor their lives by creating artefacts (ranging from language to smart artefacts). Human beings can only live the life they create first. Technology is therefore an *artificial answer to a natural need*.

Second, there is the principle of *mediated immediacy*. It refers to the way in which human beings experience the world. Because they need artefacts –and culture more generally - to shape their lives and assign themselves a place in the world, all experience is in fact mediated through these artefacts. However, although these experiences are always mediated, they are nevertheless experienced as being direct.

Third, there is the principle of the *utopian standpoint*. It refers to man's awareness of the triviality (*Nichtigkeit*) or contingency of his existence and that of the world around him. Although he is not able to find a final ground, a certainty, which will undo his eccentric positionality, there is always this driving force present to reach a wilful balance between himself and the environment.

5.1.1 Mediating technology

If trust is a positive attitude towards the future, a strategy to enable human beings to deal with the complexity of human life brought forth by the double hiatus lying at the heart of their existence, similarly, *technology is a material strategy* to cope with the same complexity. The interaction between human beings and their environment is strongly influenced by the technologies or artefacts human beings make use of. With the first anthropological law of being "artificial by nature", Plessner (1975) underlines this indissoluble intertwining of human beings and their artefacts. The way in which artefacts currently 'mediate' these interactions will be the focal point of this chapter.

Human beings need technologies to handle the ontological distance they experience in their interaction with others, the world around them, and in themselves. We need clothes to shield our sensitive skin, we need houses and cities to find shelter against the power of the elements, we need culture to meaningfully shape our lives, we need language and books to share our thoughts and nowadays, we need Internet technology to mould every piece of our daily life. Weibel (1992) stated that in the end all technologies are in fact *tele-technologies*. Technologies all aim at bridging a gap, a distance, and this distance lays between you and me, us and the world, or, more closely, in our relation towards ourselves.

One important aspect of this need to develop artefacts to bridge the hiatus is that although human beings *create* these artefacts, they do not completely *control* them. Artefacts gain their *own weight*, evoking events that were not foreseen nor intended. So although we need technologies to build our home, our so-called *familiar world*, the stability we seek can only be of a temporary nature because the same technologies also bring forth new complexity we need to reduce.

5.1.2 The relation of human beings and technology: instrumentalism, determinism, and constructionism.

The observation that human beings need some sort of technology to exist is rather uncontroversial. From the palaeontologist who confronted with possible ancient human remains goes looking for some kind of artefact in the vicinity to determine the origin, to the economist who sees technological innovation as the dominant driving force for human development, all acknowledge the importance of technology for human life, even in its most rudimentary form.

Although the connection between humanity and technology is rather uncontested, the way in which this dependency emerges certainly is open to debate. Currently, three dominant positions on the man-technology relation are generally discerned: *instrumentalism*, *determinism,* and *constructionism*[37].

In short, *instrumentalism* is the stance that technology is merely a neutral vehicle for the aims and intentions of their human users. Technology is nothing more but also nothing less than *applied science*. It is a value-neutral instrument; as a consequence, it is up to the user to decide how it is being employed. A current example of such an instrumental view on technology can be found with the USA-

[37] Another very informative way of discerning the dominant perspectives on technology has been described by De Mul (2002). He contrasts instrumentalism with a substantial perspective on technology. Instrumentalism refers to the perspective where technology is a neutral instrument, the starting points of the substantial perspective are that technology steers and pre-sorts our actions. This substantial perspective knows two sub-positions: on the one hand technological determinism, which refers to the autonomous force of technology, on the other hand constructivism, which presupposes that actors inscribe meaning in technology after which this technology steers our actions.

based National Rifle Association (NRA). Their slogan "guns don't kill people, people do" emphasizes that in the act of shooting the main actor is the human shooter. The gun solely functions as the enabler of this shooting behaviour. The focus therefore lies on the human being and not on the technology. It is people who act –or shoot-, and the device by which this action is made possible is only of a secondary importance (also see: Latour 1994).

A second important position is called *technological determinism* or *determinism* in brief. This perspective on technology assigns technology as the driving force for human life and development. Technology is the instigator of all human action. Who human beings are is shaped partly by the technologies they use and which surround them. In the determinist perspective, technology and human beings oppose each other. Technology is then the dominant, autonomous force steering human life.

Determinism comes in two flavours: an optimistic and a pessimistic one. People who adhere to the former, such as the technologist Kevin Kelly (2010), see the leading role for technology as a good thing. In his book "*What Technology Wants*" Kelly argues that history has shown that technological development always brings along more good than evil and that by giving way to technology –by listening to 'what technology wants'- a prosperous future lays ahead of us.

More pessimistic perspectives on the impact of technology on human life can be found with the classical philosophers of technology, such as Martin Heidegger (2010), Herbert Marcuse (2009 [1964]), and Jacques Ellul (1990). They all worried about the technological rationality imposed on human beings by technology, pushing aside human autonomy and dignity. Not only did technology determine the organization of society and the way we live and work, but also on an even more fundamental, ontological level it pre-sorted the way human beings experienced the world and themselves. For Heidegger, one of the biggest threats human beings face because of the dominant force of technology is the occurrence of a restriction in our way of thinking; a restriction which means that we can only understand our being in technological terms (Dreyfus 2009: 27). Marcuse (2009 [1964]), from a more political perspective, warns that this dominant position of technology will lead to *one-dimensional men* who will completely conform to prevailing technological demands.

What both instrumentalism and determinism have in common is that they

presuppose a clear and impenetrable distinction between object and subject, between human beings and the outer world (also see: Verbeek 2011b). In the instrumentalist perspective, the outer world and the material objects within that world are at the disposal and under the control of human beings. In the determinist perspective, human beings are steered and controlled by the outer world, constantly threatening their freedom (negative reading) or coming to bloom (positive reading)[38].

Since the 1980's, especially as a reaction to the deterministic stance of classical philosophers of technology such as Heidegger and Ellul, a third perspective on technology was developed called *constructionism*.

Constructionism refers to a number of schools such as Social Construction of Technology (SCOT), 'mutual shaping' approaches, Actor Network Theory (ANT), and the Social Shaping of Technologies (SST) (also see: van den Berg 2009: 29-30).[39] SCOT is one of the most dominant approaches and can be described best as *strong social constructivism* as it sees social actors and social practices as the core explanatory force of technology (see for example: Bijker 1995; Bijker and Law 1992; Bijker 2001). I will make use of some of the key concepts of SCOT to analyse how being *artificial by nature* takes shape in the networked era (section 5.3).

---

[38] Plessner (1975) aims at thinking beyond this modernist subject-object dualism. In the second chapter of the *Stages* he goes to great lengths to show that the division between *res cogitans* and *res extensa,* famously proclaimed by Descartes, is not that fundamental as it is generally put forward. Rather than falling for one of the two possible poles –the material world or the world of the mind- he shows that human beings, being first and foremost *living nature*, have both and should only *be understood* as being both. He speaks of human beings as *psychophysical indifferent unions*, with an inner world and an outer world. Instead of a schism, lying at the roots of its existence as with Descartes, Plessner speaks of a hiatus that constantly is being bridged in the experience. The inner world cannot be understood without a necessary detour along the outside world. Likewise, the outside world comes to human beings mediated by the inner world. Human beings should always be understood as being in interaction with their environment, which they shape and are simultaneously shaped by.

[39] It has to be said that although all these different schools share the proposition that technology always gains its meaning in a specific socio-historical setting and simultaneously also shapes the social actions in which it is used, they also differ in fundamental ways (methodology, focus, theoretical background). Unfortunately, it is not in the scope of this thesis to address these differences.

Social shaping approaches are a form of *mild social constructivism* as they are –in contrast to SCOT- "willing to attribute properties and effects to technology" (Brey 2009: 101). Instead of merely looking at the social actors, they take into account the interdependency of social actors and artefacts. Nevertheless, in the end social shaping studies explain these technological properties still by referring to social practices.

ANT goes a step further than the social shaping approach, abolishing the divide between social and technological actors all together. In this perspective, artefacts can have agency and therefore invoke potentially unforeseen consequences. De Mul (2002) refers to this kind of approach, where there is effectively room for taking into account the influence of both social and technological actor, as *technological interactionism*. ANT will be the starting point for a case study into the innovation of hotel keys in the next chapter.

Together and in close connection with these schools, the philosophy of technology also took an empirical turn towards "a more practical, contextual interpretation of artefacts and machines" (Kaplan 2009: 1). This *contextual approach* becomes particularly conspicuous in the postphenomenological school and other related empirical-based *theories of mediation* where the central presupposition is that human beings *shape* their environment and, simultaneously, *are being shaped* by their environment.

This focus on the interplay or *mediation* of human beings and technology, is also the most related to Plessner's fundamental assumption of the interrelatedness of human beings and their environment. Because of his attention to the way in which living nature and its environment are closely connected to each other and how, for human beings, culture, language, and especially technology play an important *mediating role* in these interactions, Plessner actually anticipated the current theories of mediation (Kockelkoren 2014: 327).

5.2 The networked era

Although it is always difficult to interpret one's own epoch, there is some compelling evidence suggesting that we currently live in a period of time that is fundamentally different from earlier times (van den Berg and Keymolen 2013).

In our article (van den Berg and Keymolen 2013), we referred to the current timeframe as the *networked era* because of the arrival of ICTs, and especially of the Internet as the network of networks which have fundamentally changed –and are still changing- the prevailing way we live, interact, and communicate. We chiefly focused on the new, *radical nearness of technology*, which arises due to the technological developments of our current time.

Where in the end of the 19th and the beginning of the 20th century, technology was being perceived as consisting of large, opaque, and dominant systems –think of the highway system, airports, and electricity infrastructure-, this perspective shifted due to the arrival of consumer electronics after the Second World War. Technologies were not just something happening 'out there', but increasingly took up a central place in the home and in the office –think of washing machines, vacuum cleaners, and PCs-.

With the arrival of ICTs, this tendency of technology to become an integral part of everyday life, to become *radical near*, by not merely entering the intimate sphere of the home but of our body as well, becomes very conspicuous.

On a similar note, Floridi (2015b) speaks of *hyperhistory*[40] -instead of the networked era- to discern a society that fundamentally rests on ICTs and data processing powers. One of the most important characteristics he discerns is that in this ICT-empowered society, it is no longer *us* who is processing data, ICTs are doing it *for us*. He writes (idem: 52):

> "in hyperhistory, there are ICTs, they record, transmit and, above all, process data, increasingly autonomously, and human societies become vitally dependent on them and on information as a fundamental resource."

Our current western society is in a state of hyperhistory, not because of the mere presence of ICT, but because our facilities and organization of society crucially hinge

[40] For Floridi (2015: 52), hyperhistory does not refer to a specific period in time. It does not say anything about when and where people live but how people live. Hyperhistory refers to a dominant way of organizing society.

on the working of these ICTs. Or, as Floridi[41] puts it: "[w]ho lives by the digit, dies by the digit". Only a society that relies on ICTs for its most fundamental functioning can become vulnerable through these same ICTs.

In their collaborative research, The Onlife Initiative[42] (2015) speaks then of a "Hyperconnected Era" which, to a large extent, pinpoints at the same changes Van Den Berg and I listed. These scholars, chaired by Floridi, conclude (2015: 44-45) that the rapid and fundamental changes brought forth by the ubiquitous presence of ICTs instigates the need to rethink or, even better, "reengineer" our conceptual toolbox.

They discern four major transformations (Floridi 2015b: 2) which cause the need for such a reengineering. First, there is the *blurring of the distinction between reality and virtuality*. Second, they depict also a *blurring of distinction between human, machine and nature*. Third and fourth, The Onlife Initiative also foresees the impact of *the transformation of information scarcity to information abundance* and of *the shift from stand-alone things to the functioning in processes and networks* (idem).

In the following paragraphs, I will illustrate these transformations leading to the networked era by briefly looking at some key technological developments, revolving around smart artefacts, ambient intelligence, Big Data, and the Internet of Things.

5.2.1 Smart artefacts

An important change instigated by the Internet is the arrival of *networked artefacts*. Increasingly, the Internet becomes part of the *ontological structure* of a wide range of artefacts we use in everyday life, surpassing its function of a mere infrastructure, as analysed in the previous chapter.

The innovation of networking artefacts is often referred to as the *Internet of Things*. The Internet of Things is characterized by the fact that artefacts become

[41] https://www.youtube.com/watch?v=riT-ew7n7RU, accessed 8 July 2015.

[42] The Onlife Initiative consists of a group of influential scholars coming from diverse backgrounds such as philosophy, law, computer science, and ethics.

embedded in networks of information often *invisible* to the user. Mere physical objects become what is been framed as *smart* by adding a computational component to them, bridging the gap between the physical world and the online world (Kopetz 2011: 308).

A prominent example of this evolution is the smartphone, formerly known as the telephone. Because of the added possibility to connect to the Internet, the usage and meaning of the smartphone differs radically from its offline predecessors. Calling someone is no longer the primary function of the smartphone. Because of the services that can be offered through the Internet connection, people use their smartphone to chat, book a place for the holidays (see chapter 7), open their hotel door (see chapter 6), post messages on their timeline, check the weather, etc. That the dissemination of the smartphone is by no means merely an instrumental change, but also has an impact on our self-perception, social interaction, and society as a whole has been extensively researched (see for example: Pariser 2011; Turkle 2011; boyd 2014).

The introduction of smart artefacts, such as the smartphone, has partly instigated the merging of the online and the offline realm. Increasingly, devices mediate the way in which we perceive the world around us. We literally find our way in the city by making use of handy navigator apps. We also use apps to add layers of information to our environment, creating an *augmented reality*. Friendships and relations –in bygone days located in the offline world- now also thrive online. On social media platforms, users post pictures, report on their daily activities, and engage in discussions, adding a new online layer of activity to their relations. As described previously, the separate online sphere where people would be freed of the offline meddling of governments, companies, and other unlike-minded people seems to be definitely finished in the networked era. We truly have become 'onlife' creatures (The-Online-Initiative 2015), living in a world, which is both online and offline at the same time[43].

[43] Following the colonization of the Internet by governments and companies, the introduction of these smart objects seems to have been the death blow for the independent cyberspace that was once so forcefully defended by Barlow and other online utopians. Meat space and cyberspace have merged.

5.2.2 New kinds of services

The 'smartness' of these networked artefacts does not solely depend on their ability to offer a wider range of services. Also *the kind of services* they offer is conspicuously different from their non-networked counterparts and can be characterized as dominantly *personalized*, *pro-active, and persuasive*.

Personalization of services is the aim to offer the user not a general service but to tailor it to his or her specific needs. To enable personalization, companies, governments, and other service providers make use of algorithms to automatically mine databases loaded with all kinds of data to look for correlations that may indicate something about the preferences of a user.

A simple example of personalization is the weather app on your phone, which, based on your geo-location data, provides you with weather info of your current location and not just with general predictions for the whole country. A more excessive example of personalization is online price-differentiation based on the devices you use to access the Internet[44].

In chapter 8, we will delve deeper into personalization by looking at the impact of the personalized interface on trust. For now it is enough to understand that personalization brings along fundamental questions about for example the ontological status of an artefact, but also has an impact on the familiar world which is necessary for trust to thrive.

Closely connected to personalization is the *pro-active character* of the networked artefacts. It is not just the aim to deliver services in a personalized manner; these services should also *precede* the request of the user. Based on all this collected and mined information, it becomes possible to predict what a user needs and as a result present him or her with these services, even before the request has been explicitly made.

An example of such a pro-active service, that at least all users of the search engine Google are acquainted with, is the famous Google search bar. When you type

[44] http://www.wsj.com/articles/SB10001424052702304458604577488822667325882, accessed 10 March 2015.

in a search query, the search engine automatically –based on your search history combined with that of millions of others- tries to complete your words in order to provide you as fast as possible with the relevant results.

5.2.3 Ambient intelligence

Also in 'the real world', networked artefacts are increasingly being designed to display pro-active behaviour. A hotel room that projects artwork by your favourite artist on the wall and puts on your favourite heavy metal record, the refrigerator which automatically orders your groceries when you need them or the thermostat which adjusts the temperature to the preferences of the person who enters - these are all examples of what has been referred to as the *ambient intelligence* vision. In this vision, technology becomes

> "invisible, embedded in our natural surroundings, present whenever we need it, enabled by simple and effortless interactions, attuned to all our senses, adaptive to users and context and autonomously acting" (Lindwer et al. 2003:1 cited in van den Berg 2009: 59).

In line with the observation that objects are becoming pro-active, Van den Berg speaks of *anticipative* artefacts (van den Berg 2009: 70-71). Next to being embedded, context-aware, personalized and adaptive, these networked artefacts *anticipate the need of the user (idem)*. This means that, as van den Berg (2009: 71) argues, "…systems will be given a large responsibility in managing and maintaining a user's information sphere". It will be up to the pro-active networked artefacts to:

> "…decide what information is relevant, useful and even meaningful for the user in his current situation; the responsibility of finding, filtering and processing this information is removed from the user and placed squarely on the shoulders of the technology" (van den Berg 2009: 71).

Because smart technologies are able to personalize their services to the profile of the user and even anticipate his or her needs, smart technologies can also become *persuasive*. Persuasive technologies are able to "explicitly influence the behaviour of users in specific directions, effectively persuading people to behave differently"

(Verbeek 2011b: 19).

An 'old school' example of persuasive technology is given by Bruno Latour (1990) who shows that hotel keys, intentionally made heavier by a weighty keychain, persuade the guest in a material, non-verbal way to return them to the hotel reception. Instead of the hotel owner explicitly stating that the keys have to be returned, this request is partly delegated to the keys themselves. The design of the keys indirectly persuades the guests to display the desired behaviour. In the next chapter, this influential notion by Latour will be updated by looking at the current version of the hotel key, namely the *rfid key card* and the *digital key*.

In the networked era, persuasive technology can particularly be found in the domains of healthcare and personal development. Health coaches in the form of apps on the phone or tablet help users to act –and persist in acting- in a healthier way. For example, the *7 Minute Workout App*[45] not only shows you which fitness exercises you have to do to get in shape, it also sends you a notification if you haven't done your daily exercise yet. You can support friends who also use the app to keep them motivated. Another app like *RoomForThought*[46] sends a push message once a day to let you take a break and take a picture for which you only have three seconds in order to capture your life. It is an app that persuades you to stop and relax for a moment, escaping the hecticness of everyday life.

5.2.4 Big Data

All these devices and apps can only function if they are fed enough data. Only by the automated analysis of sets of data, often referred to as *Big Data*, looking for correlations based on which predictions can be made, can smart devices deliver their services (Mayer-Schönberger and Cukier 2013). A wide range of different data, ranging from information on people, sensors, to (online) behaviour, -are mined by algorithms for insight, leading to what has been called *data-driven decision making*. Because data is often re-used and combined, it becomes increasingly difficult for the average user to understand how the collection and mining of data takes place

[45] https://7minuteworkout.jnj.com/, accessed 07 December 2015.

[46] http://www.roomforthought.nl, accessed 25 February 2015.

(Richards and King 2013; Hildebrandt 2011b). Although the innovations in the domain of Big Data are promising, Richards and King (2014) argue that the secondary use of information we deem to be confidential may damage the trust we have in the institutions we share this information with (also see Keymolen 2014b).

5.2.5 Intimate technology

Technology is not only being woven into the very fabric of our homes, offices, and public spaces (*ambient technology*) and the objects we use daily (*smart artefacts*), but technology also becomes an integral part of our *bodies*; think of neuro-enhancers and implants such as bionic ears, pacemakers and new steel joints. In the near future it should become possible to directly connect brains to the Internet or to see in infrared. Important steps taken in the domain of molecular medical science will lead to new 'bio-sensors' curing –and even preventing- diseases and handicaps such as sickle cell anaemia and deafness. Because technology increasingly becomes part of our bodies and integrated in our environment, it has been characterized as *intimate technology* (van Est et al. 2014).

This new 'nearness' of the technology is not merely conspicuous for the relation of man and technology but also for the technology itself. The above-mentioned innovations are partly spurred by the converging of formerly separate technologies such as nanotechnology, biotechnology, information technology, and cognitive science.

In their book *Life As A Construction Box,* Swierstra et al. (2009: 10) explain how the converging of these technologies lead to a whole new holistic self-perception because of the control we gain over the "building blocks" of living as well as non-living nature. Technologies are used to enhance the body not by adding components to it on the outside –such as a pair of glasses or a wheel chair- but by transforming the body from within; consequently making it increasingly more difficult to determine where the technology stops and the body begins.

5.3 Artificial by nature in the networked era

As a result of this *radical nearness of technology* in the networked era, the principle of being "artificial by nature" becomes increasingly topical. There is not just the

general interplay of human beings and their artefacts and how they are mutually shaped in this interaction. In some instances, for example when people receive implants to help them see or when they take certain medication to enhance their moral behaviour (Specker et al. 2014), even an actual *merging of the human and technology* takes place resulting in a *new enhanced entity*.

Peter-Paul Verbeek (2011b: 144) introduces the *cyborg relation* to emphasize how the presence of such a new entity "physically alters the human" because of this kind of absorbing association with technology. Although de Mul (2003: 254) is rightly claiming that "human beings always have been cyborgs" because culture and technology are not mere "instruments of survival" but form an "ontic necessity", this cyborg-intertwining becomes particularly real in the networked era where the division between human beings and technology is no longer clear and the body increasingly consists of organic and non-organic components.

Through the lens of Plessner's first anthropological law, being 'artificial by nature', I will now first look into the openness or 'multistability' of artefacts and the way in which *humans construct the meaning of artefacts*. Subsequently, I will concentrate on the other side of the equation by focusing on the way in which *artefacts invite human beings to use them in a certain fashion* and *how artefacts steer interactions*. I will discern the 'own weight' of artefacts in the networked era.

It has to be noted that in fact human and artefact are simultaneously constituted in their interaction. The meaning of an artefact comes about in the way it is taken on by a human being; the identity of a human being is shaped by the artefacts he or she uses. It is certainly not the case that one precedes the other, which would lead us back to the stance of instrumentalism or determinism. It is therefore only to ensure a clear analysis that both sides are addressed separately here.

5.3.1 The openness of artefacts in the networked era

For Plessner (1975), who sees artefacts as an ontic necessity, the openness of artefacts is crucial for the way in which human beings live in their environment. The world of human beings is not given but has to be built first; as a result, the contingency and subsequent openness of their artefacts is a fundamental aspect of their 'being-in-the-world'. It is typical for human beings that they can use artefacts in different contexts,

shaping their environment over and over again. The American philosopher of technology, Don Ihde (1990)[47], refers to this openness of artefacts with the concept of "multistability". Although human beings experience their world often as stable and governed by rules as strict as if they were natural laws, contingency, nevertheless, always shines through. The openness of artefacts can be negatively framed as *a burden*, causing instability in human life. However, it can just as well be defined as *a change,* instigating new and fruitful practices or, in other words, innovation. The openness of artefacts is closely connected to the kind of creativity only humans seem to demonstrate.

For example, the elderly man who likes to sit comfortably when reading his newspaper on a sunny afternoon can experience armrests on the benches in the park as an extra luxury. For the homeless man wandering around at night, looking for a place to sleep, these same armrests hinder him from using the bench as a resting place, because they divide it in separate places to sit, making it impossible for him to stretch out and get some rest.

The way we perceive artefacts can also change over time. In the early days of mobile telephones, my mother hid away in the car when she wanted to use her cell phone. Back then, calling in public was perceived as improper. This attitude towards calling in public spaces has completely changed over the years, notwithstanding the fact that it still can be quite annoying when done too loudly in public transport.

Taking it even a step a further, where the mobile phone was introduced as 'a mobile way of calling', users unexpectedly started to use their phone to text rather than to call. Texting was initially added to the phone as a funny gimmick. However, it turned out to be a game-changer in social life. Besides that it solved the earlier-mentioned problem of inappropriate calling in public all together, it also brought forth new ways of interaction, had an impact on language, and simultaneously introduced a new problem as texting dragged people into conversations elsewhere, creating an "absent presence" (Gergen 1991).

---

[47] Don Ihde stands in the postphenomenological tradition, which I have assigned to the general domain of constructionism because also central to this stance in philosophy is the fundamental idea of the mutual shaping relation of human beings and artefacts.

All in all, users may "*domesticate*" (Silverstone and Hirsch 1992; Frissen 2004) their artefacts, meaning that technologies are 'tamed' or appropriated to fit the context and experience of the user. As we have seen with the example of texting, the *meaning* of the artefacts changed (the mobile phone is no longer primarily a 'calling device', but also a 'texting device'), as well as the *materiality* of the phone. Because people increasingly used the phone to text, telephone companies changed their design of the device, interpreting the mobile phone in a new manner.

The fundamental openness of artefacts and the possibility to read different meanings into them can of course only be understood by looking at *who* is attaching certain meanings to the artefact.

In the social construction of technology approach (SCOT), which also falls into the domain of constructionism[48], *relevant social groups* are the starting point for analysing the workings of technology in society (Bijker 2001: 26). By looking, as it were, through the eyes of different social groups, for example, users, developers, policy-makers etc., it becomes possible to map the "*interpretative flexibility*" of artefacts. Interpretative flexibility refers to the malleability or "social dimension" of the design of artefacts (Bijker 1995: 76). When the meaning attached to a certain artefact is stabilized, consequently guiding the interaction with the artefact in a homogenous way, the identity of an artefact can become more obdurate and fixed. When the interpretative flexibility of an artefact diminishes, Bijker speaks of *stabilization* and *closure*. One can speak of closure when "[c]onsensus among the different relevant social groups about the dominant meaning of an artefact *emerges* and the 'pluralism of artefacts' decreases" (Bijker 1995: 86). Stabilization refers to a similar development towards a fixed identity of an artefact but then *within* a social group (idem: 87). When a technology is stabilized, it becomes a black box and its "properties come to determine the way that the technology functions in society" (Brey 2009: 101).

[48] Brey has shown that the approaches in constructionism -he speaks of constructivism- may have interesting perspectives to add to the philosophy of technology. Their focus on technological change and more specifically on the development phase of technology may "…provide a potentially fruitful basis for normative and evaluative philosophical analyses of technology and its impacts" (Brey 2009: 108).

Fuglslang (2001), influenced by the concept of the product life cycle, discerns three phases in the innovation of a technology to ground this development from openness to closure, and sometimes, back again.

The first is *the phase of flexibility,* which refers to the "initial stage of technological innovation" (idem: 44). In this phase the interpretative flexibility often is extensive. The second phase is *the phase of momentum* (idem). In this phase, the public generally has accepted the artefact. They have invested time and money in it and the crucial decisions have been made. Now it becomes more a question of fine-tuning the artefact. The final phase is *the phase of diffusion* (idem: 45). In this phase the artefact is finalized and "is diffused to consumer industries". Users domesticated it, making it fit their user context. There is a matter of closure and stability. At that moment, however, a "reversed product life cycle" may also occur (idem: 45). Artefacts integrated in user practices may lead to new innovations. Remember the example of the mobile phone; when people increasingly interpreted their mobile phone as a texting device rather than a calling device, this led to new materializations of the artefact. The design of the phone was adapted by adding easy-to-use keyboards, mobile dictionaries, and automated typing suggestions.

When we now apply this perspective of *openness* to smart artefacts and Internet applications in general, some interesting observations can be made, relevant to our analysis of trust.

First, although the initial openness of the Internet is still available for anyone who is able to code and develop his or her own applications, when it comes to regular consumer products, *the room for users to shape and adapt their smart artefacts is rather limited.*

It is true that users domesticate artefacts to give them a proper place in social interaction and social interaction changes to absorb new artefacts. It is also true that often some superficial changes can be made in the set-up of the interface, changing some preferences in the programme or tweaking the design of the artefact by adding gadgets such as colourful sleeves. This, however, does not affect the more fundamental working of the smart artefact itself. For example, one can personalize one's Facebook and Twitter profile with a background picture, one can share movies on Youtube and pictures on Instagram but one cannot influence the settings for sharing data beyond what is provided by the curators of the platform. Or as Deibert

(2013: 229) puts it:

> "the experimentation that is encouraged actually operates on these shallow planes. On deeper, more fundamental levels, it is strictly controlled".

It is not exceptional for artefacts in the diffusion stage to have gained a certain stabilized identity and use, bringing forth a "take it or leave it" option for users (also see Bijker 2001: 29). A bike, to re-use the example of Bijker's (1995) often-cited research in the development of the bicycle, is difficult to imagine in a way different than what we are used to it now. Although users may still alter and domesticate their bicycles, this often only happens in rudimentary ways; for example, by adding components such as a speedometer, flags, and stickers. It seldom touches upon the 'essence' of the bicycle. Borrowing the concepts of SCOT, when it comes to the development of the bicycle, there is *closure*. Amongst the different social groups there is a shared and defined image of what a bicycle is for, how it should be used, and what it should look like.

However, when we look at smart artefacts this *closure does not seem to arrive*. Because smart artefacts have a networked ontology, *they never leave the phase of flexibility*. Or, to put it more precisely, *the phase of flexibility and the phase of diffusion converge*. By definition, smart artefacts reside in a never-ending beta-stage (also see de Mul 2002: 38). The stage of developing and designing smart artefacts and applications does not end with the user taking the smart device home from the shop or with the downloading and installing of a new application. The smart artefacts and applications persist in the sphere of influence of the curators, remaining therefore open to their interpretation. In other words, while for the average users the interpretative flexibility of the smart artefact is rather small, for the curator behind the artefact the interpretative flexibility remains extensive. The data curators can collect through the smart artefacts is a source for endless *innovation* or –more pessimistically- for an endless *function creep*. This continuous reinterpreting of the artefact may disturb the relation users have with their smart devices and services and consequently impact the trust vested in the artefact.

There are legions of examples of such disturbances. Facebook changing their design of the timeline caused public outcry amongst users. Or Whatsapp adding

check marks behind messages, the blue colour indicating a message has been read, also set off a storm of complaints. But because in the interaction between user and smart artefact the interpretative flexibility, the openness to develop alternative uses, is rather small, this disturbance in the established stability of the artefact leaves the user with no other option than to choose between "exit, voice, or loyalty" (Hirschmann 1970). Where sometimes opting for "voice" -letting those in power know of your discontent- may help to alter the situation (e.g. Whatsapp added the possibility to disable the blue check marks after public outcry), users are generally left to choose between "exit" –quitting the service- or "loyalty" –sticking to the service. However, because some information intermediaries such as Google, Facebook, and Linkedin provide services that increasingly become a necessary condition to function in Western society, the option to "exit" may come at a very high price.

To avoid any suspicion, I am not advocating a technological determinist perspective here, denying that users have a say in the way in which artefacts are embedded in a user context or are ascribed meaning in social life. However, I do contest the idea that the inherent and undeniable openness of artefacts is accessible to or can be played with by all actors involved. Moreover, looking at the ideal-typical appearance of smart artefacts, it is not far-fetched to conclude that when it comes to the smart artefacts and services targeting large consumer populations, there is a tendency to locate more flexibility in the interaction between curator and smart artefact than in the interaction between user and smart artefact[49].

5.3.2 Artefacts' own weight in the networked era

While in the previous section I looked into the way in which human beings ascribe meaning to an artefact and adapt it to their personal user context, I will now approach the concept of being 'artificial by nature' from the other side of the

[49] It has to be noted, however, that there of course will be counter examples of devices and applications in the networked era where this difference in interpretative flexibility between the two relations are different or even not existent at all. One always has to take into account the actual use and setup of an interaction to establish the openness and hardness of the artefact. We will therefore be looking at some in-depth cases in the following chapters.

spectrum and look at the way in which artefacts steer users in certain directions. Although human beings are the creators of artefacts, as these are an ontic necessity for the existence of human beings, humans do not completely *control* these artefacts. Producing artefacts is only half of the job (Plessner 1975: 321).

Artefacts also bring forth unintended consequences. Artefacts gain, as Plessner describes, their *own weight*. They are not mere neutral instruments, the materialization of the creative urge of human beings, but they can steer the interaction of human beings with and within their environment in unforeseen ways. Human beings have to discover this weight in order to relate to their artefacts in a meaningful way.

Don Ihde (1990) speaks of "technological intentionality" to refer to the way in which artefacts steer our actions and interactions. Although it is not impossible to use artefacts differently, they do *invite* users to adopt them in specific ways. He writes: "Technologies, by providing a framework for action, do form intentionalities and inclinations within which use-patterns take dominant shape" (Ihde 1990: 143).

Verbeek (2011b: 107) makes an informative distinction between different forms of invitations or "forms of mediation". Artefacts can *force* users to act in certain ways. For example, you need to agree to the terms and conditions of a website in order to make use of the service. Then the room to manoeuvre for users is rather low or non-existent. Artefacts can also *persuade* users to act in certain ways. For example, a smart meter providing feedback on energy use stimulates users in a transparent way to alter their behaviour. Or artefacts can *seduce* users to act or refrain from acting. For instance, the Facebook app on a smartphone can be used to maintain intimate relations with the people one loves. It is, however, designed and set up in such a way that it rather stimulates and pre-sorts users to share information *beyond* this group of family and close friends. So, although it is not impossible to use it for the former goal, the way the artefact –in this case the Facebook app- is designed to stimulate alternative uses.

Madeline Akrich introduced the concept of "script" to refer to the way in which the –sometimes implicit- presuppositions of developers and designers concerning the envisioned users and user-context, would find its way in the design of the artefact. She writes:

> "Designers [...] define actors with specific tastes, competences, motives, aspirations, political prejudices, and the rest, and they assume that morality, technology, science, and economy will evolve in particular ways. A large part of the work of innovators is that of 'inscribing' this vision (of prediction about) the world in the technical content of a new object" (Akrich 1992: 208).

Although it is difficult to say something meaningful about technological intentionality in general, when looking at the networked era and especially taking into account the above-sketched trends of data gathering and data analytics (Big Data), it can be argued that from a user's perspective, a vast amount of the smart artefacts used today *persuade and even force users to produce and share data*. Even when it is not the primary aim of an artefact, gathering data often becomes an important by-product. Where data gathering is first a means to an end, it often turns into an end, causing a *data-function creep*.

For example, it is not the primary goal of a smart energy meter at home to produce data but to regulate as efficiently as possible your energy use. However in order to perform well, it does need to crunch a lot of data. Next to the data of other users, it needs to process data on the behaviour and preferences of the people living in the house in order to adapt, in this case, the central heating. Moreover, the usual business model of companies behind these kinds of technologies is not merely to enable efficient energy use and to add to a sustainable environment. More often, the gathering of data is also used for other goals such as improving the user experience and the services provided. Sometimes, these data are also shared with or sold to other companies, for example to provide personalized advertisements.

5.3.3 The political weight of artefacts

Whether or not deliberately, the values and presuppositions of designers find their way into the design, consequently steering the use of the artefacts in certain directions. The political philosopher Langdon Winner (1980) argues that artefacts may also evoke *political consequences*, bringing forth certain power relations. He makes use of several examples of practices mediated by technological artefacts to analyse such political ramifications.

One of Winners' particular starting points is that artefacts are by no means

neutral but the carriers of political power. Having the capability to inscribe political views into artefacts is, in a society where technology is fundamental for its functioning, a very powerful skill.

In his article, Winner carefully moves from examples of *intended* political working of artefacts to *unintended* political working of artefacts. The most-cited example of an intended political artefact he analyses is that of the 'bridges of Moses'. In this example he describes how Moses, the architect of some bridges to Jones Beach in New York, designed the bridges in such a way that only cars could make use of them. Because of the low-hanging overpasses, busses were kept away from the beach and consequently, the poor, Afro-American citizens predominantly using this transport were too. Put differently, the architect deliberately designed the bridges in a way that became the carrier of his racist vision, encompassing "purposes far beyond their immediate use" (Winner 1980: 125).

This example aroused some heated debate amongst scholars (Joerges 1999a, 1999b; Woolgar and Cooper 1999), even refuting the analysis that these bridges were the only way to get to Jones Beach or arguing that Moses was no more racist than his contemporaries (also see: Verbeek 2011b: 44).

However, the own weight of artefacts does not limit itself to the way in which artefacts, deliberately inscribed by their developers, invite users to engage with them in a specific practice or, in the case of the bridges of Moses, how they exclude people. Winner also shows how, unintentionally, differently-abled people are excluded from public spaces because these are designed with able-people in mind. On a similar note, he analyses the introduction of an automated tomato harvester that caused a disadvantage for small farmer companies, completely rearranging this agriculture sector as a result. Without a preconceived opinion, designers can invent an artefact that, when introduced in society, provokes changes that were not foreseen nor desired.

Subsequently, in his article Winners focuses on *inherently political technologies*. These are technological artefacts, which also bring forth *unintended political consequences*, just as the tomato harvester did, but with that crucial difference that for these technologies there are not that many *feasible design alternatives*. A bridge does not have to have low-hanging overpasses to function as a bridge nor do public spaces stop being public spaces when they are designed in a way that makes them

accessible to disabled people. It is not hard to imagine a bridge or a public space, designed differently, which still carries out its tasks but without the above-mentioned unwanted consequences. Other technological systems, such as a nuclear plant, or a laboratory in which people work with dangerous chemicals, lack this flexibility. To choose such a technology means "to choose unalterably a particular form of political life" (Winner 1980: 6).

Winner makes a distinction between a *strong* and a *weak* version of this claim concerning inherently political artefacts. In the strong version, "the adoption of a given technical system actually requires the creation and maintenance of a particular set of social conditions as the operating environment of that system" (Winner 1980: 7). The weak version holds that "a given kind of technology is strongly compatible with, but does not strictly require, social and political relationships of a particular stripe" (idem: 7). Both versions can be relevant to an artefact, meaning that it may happen that some aspects of an artefact or a technological system are rather impossible to change if one wants to have a functioning device or system while other aspects, although they strongly lean towards a certain setting, can nonetheless be altered.

Winner notes that what counts as a "practical necessity" and, therefore, is impossible to adapt is not merely an empirical question. He claims that often the argument that 'it cannot be done in any other way' unjustly overrules other arguments. He writes (Winner 1980: 9):

> "to say that some technologies are inherently political is to say that certain widely accepted reasons of practical necessity – especially the need to maintain crucial technological systems as smoothly working entities – have tended to eclipse other sorts of moral and political reasoning."

It certainly is not the case that Winner here suddenly takes a deterministic approach, as the classical philosophers of technology do. Actually, he opposes their view, because such a deterministic perspective hinders the possibility of holding people accountable for the political consequences they bring forth through their artefacts. Rather, he shows that the openness or "multistability" as Ihde refers to it, is not the same for all artefacts. Some artefacts lend themselves better to being mould by users or society than others. As a result, the decision of whether or not to introduce such a technology is crucial as the possibilities to adjust it when it is in use are limited. With

his analysis, Ihde therefore also criticizes an all too strong belief in the social power to shape the meaning and use of artefacts, which can sometimes be found in the SCOT studies. Not all artefacts are sensitive to the influence of social actors and not all social actors are able to influence the way in which artefacts find their way in daily life. As Winner (2001: 15) writes:

> "The smart people are those able to 're-engineer' their organizations and careers by liquidating older roles, relationships, and institutions in response to technical and economical necessities that loom ahead. The less proactive to these conditions are doomed to suffer as the new technical order crashes in on them."

All in all, with Winner's analysis of the political consequences of artefacts it becomes clear that artefacts' own weight can take different forms. Artefacts not only steer users in a certain direction because of the presuppositions inscribed into the artefacts by the designers, intentionally or otherwise. Winner (1980: 10) adds also the option that "intractable properties of certain kinds of technology are strongly, perhaps unavoidably, linked to particular institutionalized patterns of power and authority."

5.4 Artificial by nature in the networked era: some challenges for trust

In the section on being artificial by nature in the networked era, I analysed the way in which both the *openness* and *own weight* of smart artefacts generally take shape. While on the one hand it is clear that smart artefacts are characterized by openness, even conspicuously so, because their networked construction makes it possible to easily adapt and mould them to fit different contexts, they on the other hand also leave their undeniable mark on the social context. Artefacts' own weight does not only become visible in the way in which they are the carrier of intentions of designers and companies but also in the relations of power and control they inherently bring forth.

We can now already perceive that being artificial by nature in the networked era brings along some specific challenges for trust. In theory, the networked character of smart artefacts pre-eminently makes it possible for users to adapt them to fit their own situation and to cater to their personal needs, making the artefacts easier to control and therefore lowering the barrier to trust. However, because in practice this

openness is reserved for curators of the artefacts, the predictability of how these artefacts are functioning is under scrutiny. Trust, a way to act as if the future is certain, becomes more difficult when no closure occurs and the phase of flexibility and diffusion collide.

It has to be noted that when these adaptations of smart artefacts take place beyond the user's awareness, for example privacy settings being changed without a clear notification, a breach of trust doesn't necessarily need to occur immediately. However, when a user does become aware of it and the imposed change does not fit the meaning that the user initially had attached to the artefact, *the discrepancy between what the artefact was and what it has become* may cause new complexity that trust cannot neutralize. It therefore can be argued that because an interruption in the interaction with the artefact may lead to a user questioning the use of the artefact or may even lead to the user's decision to quit the service all together, curators will generally go to great lengths to ensure that changes in the artefact or service are unobtrusive in nature.

As we have seen, personalization is one of these changes that often occur with smart artefacts. Anticipating chapter 8 on the personalized bubble caused by online profiling, one of the intriguing questions nowadays is how trust might be impacted because of the personalization of the interface, often taking place beyond the user's awareness. Trust can only thrive in a familiar world structured by shared norms and values. These shared norms and values are not absolute. They fluctuate, change, and evolve in and through our interactions. What if this background online is no longer a shared but an individualized one, no longer reflecting shared norms and values but first and foremost personalized ones? How well will people be able to deal with other norms and values if the possibility to fruitfully confront their own values with those of others diminishes online? What if, in a never-ending feedback loop, they are presented with the affirmation of their own beliefs?

On the other hand, also chances for developing trust may rise. If a user does perceive the changes made by a curator as being in line with his or her own interests, this may add to trust vested in the company or designer as well as in the artefact itself. As Luhmann already wrote, trust becomes especially apparent when it is put to the test. It is in times of increasing complexity that trust not only becomes evidently important but that one can also determine how robust the invested trust actually is.

5.5 Mediated immediacy in the networked era

Where in the previous section we looked at the way in which artefacts and human beings constitute each other, we now, with the focus on the second anthropological law "*mediated immediacy*", dive into the way in which human beings *experience* artefacts and the world mediated through those artefacts.

The eccentric positionality of human beings, characterized by the double hiatus, brings forth that human beings do not have direct access to the world around them. Their world is never given, but always has to be built first. Artefacts mediate the interaction between human beings and their environment. An artefact is, so to say, the missing link between human beings and their environment. It restores the imbalance human beings experience through their positionality; it makes direct what is indirect by nature. Human beings are, because of their eccentricity, aware of the fact that their knowledge, their language, and their tools occupy a fundamental place between them and their environment. Because human beings can, from a distance as it were, relate to themselves and understand that it is *them* who are cognizant of and interacting in the world, they have to cope with the indirect and mediated character of their interactions. In other words, *human beings are aware of the fact that their world is always a mediated world*.

Although human beings are *eccentric*, they simultaneously are also *centric*. They are both animal and human beings. This shows itself in the fact that although all interactions are mediated and therefore *indirect*, human beings experience these interactions often as *direct*, just as animals do.[50] Plessner "borrows Husserl's idea of the intentionality of consciousness" (Kockelkoren 2014: 324) to understand the

[50] On logical grounds, this might seem implausible, because: how can a relation be simultaneously immediate and mediated? Although this apparent contradiction might lead to the belief that a human being can relate in two different ways to its environment –directly and indirectly- this is a misrepresentation. Because human beings are both animal and human, both centric and eccentric, their interaction is simultaneously direct and indirect and cannot be understood otherwise (see Plessner 1975: 325-326).

directedness of *all living nature* towards their environment. What human beings in their interaction perceive comes to them as being direct and real. Artefacts are being incorporated in the perceptions and interactions of human beings and move to the background of their attention when they interact with the world around them, with others, or relate to themselves.

With this perspective on the mediated relation between human beings and their environment, Plessner is a predecessor of the current stance in the philosophy of technology called postphenomenology (Ihde 1990; Verbeek 2000; Verbeek 2011b, 2011a). Central to this perspective is the idea of *relationality* between human beings and their environment (Ihde 1990). Human beings are always directed towards their world, they shape it and give it meaning. Simultaneously, the world 'carries' and 'supports' human beings, making them who they are. All in all, human beings and their world constitute each other. This continuous interaction makes human beings and their world inherently "*interrelated*" (Verbeek 2000: 279). The postphenomenological framework conceptualizes the mediating role of artefacts. Borrowing some of the central concepts of this framework, I will now look into the way in which the mediated immediate experience in the networked era unfolds.

5.5.1 Ready-to-hand smart artefacts

Although mediated, human beings experience their interactions generally as direct. Heidegger has conceptualized in a very informative manner this 'forgetfulness' by describing technologies or more specific *tools* as being "ready-to-hand" (*Zuhandenheit*). Artefacts that are "ready-to-hand" are not the *object* of experience but the *means* of experience[51]. As such they "withdraw" as it were from the attention

[51] It has to be noted that Ihde (1993) in his conceptualization of the relations of mediation puts Heidegger's "ready-to hand" on the same level as his embodiment relation. Although the similarities are clear –in both cases the technology withdraws from the attention of the user- opening up the world in a certain way to human beings, I approach, in line with Plessner's anthropological law of mediated immediacy Heidegger's "ready-to hand" in a broader manner. Experiencing an interaction as direct, even when it is in fact mediated, does not limit itself to embodiment relations, but can also occur in other mediation relations. To be absorbed in an interaction, 'forgetting' the artificiality of it, is inherent in the eccentric positionality of human beings.

of human beings in order for them to function as they are supposed to. Artefacts direct our attention and action towards a certain practice, which Ihde (1993) refers to as the *technological intentionality* of artefacts. Artefacts come with a certain purpose, a said functionality, which invites users to adopt them in a certain way. However, it is not just that when artefacts work smoothly, they facilitate a certain practice. They also "shape what it means to be a human being, by opening new ways of being-in-the-world" (Kiran and Verbeek 2010: 422). Artefacts are *context-dependent*, meaning that their meaning is derived from the situation in which they are put to work. Simultaneously, artefacts-in-use also bring forth a kind of coherence in the environment. The objects surrounding the interaction with the artefact gain a certain implicit meaning because of the interaction with the artefact. A familiar world is constituted around the artefact.

For example, when I am working on my computer –let us presume I am drafting my doctoral thesis-, I become absorbed by it; forgetting the computer as such. I am not really consciously aware of the screen or the typing I undertake; I read myself as it were into the text; I am present with the text and the computer withdraws. This all changes instantly, when the computer crashes and the screen turns black. Heidegger claims that especially when an artefact malfunctions, instead of being mediated our attention becomes directed towards the device itself. We suddenly are aware of its presence. The artefact is then "present-at-hand" (*Vorhandenheit*). Moreover, all the other objects that were included in my actions while typing the doctoral thesis suddenly seem to lose their self-evident presence as a result of the malfunctioning of the computer. The coffee I was drinking, the radio that was playing in the background, the chocolates I was eating - all these actions were connected to working on the computer and now come to stand on their own. The *familiar world* attached to working on the computer has been shattered. The confidence put in the computer has not necessarily vanished, but its implicit character has. Once more, the computer has to be rebooted and fit in to the user-context to regain my confidence.

All in all, we can conclude that if *the intentionality of the user aligns with the intentionality of the artefact*, the mediated and therefore indirect character of the interaction moves to the background. Users then experience their interaction as direct. As long as the computer functions, I perceive the text of my doctoral thesis on the screen in front of me as direct.

For an artefact to become the *object of our attention*, instead of *the means of our experience*, it, luckily, does not necessarily need to break down first. Taking into account the fact that smart artefacts never leave the realm of design, therefore, staying malleable, shifts in identity of smart artefacts may also instigate such a detached position, directing the attention to the device as an object. Updates affecting the design of the interface or the usage of programs may initially put me in a relation to the computer and the programs itself, rather than that they mediate my interactions. Changes in the working of the computer may disturb the familiar world carrying my interactions and the confidence I have vested in the device. This may be a chance to renew my confidence in the device, but it may also be a threat as changes to the device may hinder my interaction with it and makes the complexity, which I wanted to make bearable by an act of trust, sensible. In the end, it may turn out that my trust in the device was unjustified. This may lead to withdrawing myself from the device entirely.

5.5.2 The design of ready-to-hand smart artefacts

For an artefact to be 'forgotten' or to be "ready-to-hand", it helps if it is designed in such a way that it is easy to use. Smart artefacts such as mobile phones, smart energy meters, and tablets in general are developed with much attention to user-friendliness. Interfaces are designed in such a way that complex scripts and algorithms do their work out of sight of the user. The device itself can be easily controlled by straightforwardly clicking through a menu of presented options. Often, this last step is not even necessary as the default setting of the artefact enables direct use.

Also the hardware remains out of reach for the user. The products of Apple, characterized by their slick and clean black or white design, are a prime example of this unburdening design philosophy. Smart artefacts are first and foremost designed to deliver personalized and sometimes even pro-active services and certainly not to burden their users with difficult usage questions. The whole idea behind personalized and pro-active services is exactly that the needs of users are catered in such a fluent and inconspicuous way that users are not even aware of the fact that these actions are taking place. Therefore, beyond actually interacting with the smart device or service, not much other input is expected from the user.

The flexibility for users to interpret these devices differently than intended by the designers is rather limited as the software as well as the hardware is difficult to penetrate. Generally, smart artefacts are designed in such a way that they are *easy to use and difficult to tweak*. As a result, the design of smart artefacts *invites* human beings to approach these devices in a mere functionalistic, dual way: "Do or do they not function properly?"

This lack of possibilities to relate oneself to artefacts is criticized by the philosopher of technology Albert Borgmann (1987, 2000, 2009), who sees such unburdening and sealed artefacts as hindering a more profound and engaging relation with artefacts and, consequently, with the world these artefacts open up for human beings. Such devices instigate "paradigmatic consumption" which "attenuates human engagement with material reality" (Borgmann 2000: 419). As we will see more extensively later, such a functionalistic attitude towards artefacts may have significant consequences for trust because it limits its *scope*.

5.5.3 Relations of mediations in the networked era

The way in which smart artefacts mediate interactions or open up the world for human beings can take different forms. At first sight, when we think about going online on our computer or making use of an app on our smartphone, it may seem as if we find ourselves in what Ihde has coined a *hermeneutic relation*. To illustrate what he means with a *hermeneutic relation*, Ihde (1990: 84-85) refers to the use of a thermometer[52]. When I am inside, I can read the thermometer and know how cold it is outside. I do not have a direct sensory perception, but I can *read* myself into the situation of feeling the temperature outside by looking at the display of the thermometer. The thermometer does not so much open up our view on the world, as that it *represents* reality. The thermometer is not transparent but *opaque* and the user has to master certain skills (in this case reading and comprehending the Fahrenheit scale) in order to interact with the device. It is not *through* but *by* (Verbeek 2000: 142) the thermometer's ability to make clear in a sensible way what the temperature is, the world becomes meaningful to the user. The thermometer is

[52] In chapter 8, I will look deeper into the different relations Ihde discerns.

my "*object of perception*", however, simultaneously it is "referring beyond itself to what is not immediately seen" (Ihde 1990: 82).

Likewise, when I check my Facebook timeline on my computer, the computer is my "*object of perception*", however, simultaneously it is "referring beyond itself to what is not immediately seen" (idem). The interface of the computer represents the online world to me. It represents in a sensible way the bits and bytes which I cannot perceive with the naked eye. Through the interface of the computer, I am able to read myself into the online world without actually being there.

However, reading the temperature on the thermometer and consequently having access to a specific aspect of the world is something fundamentally different from me looking at the online interface of my computer, which is not referring to a specific aspect of the world, but it is referring to a *completely different reality*! The online world includes bits and bytes (technical level), references to virtual contexts (my Facebook timeline), and the offline world (the pictures of my friends represent real persons, the information I find on Google shapes my experience of the 'offline' world). The interface, which enables me to read myself into the online context, is simultaneously an integral part of this context.

In contrast to the thermometer that represents the world and therefore is under the influence of a specific aspect of that world (if it gets warmer, the temperature I read on the display will be higher), the interface representing the online world is influenced by a myriad of actors. The developers and companies behind the virtual contexts, the self-learning algorithms, other users, and, of course me as a user are constantly shaping and reshaping the online context.[53] *The networked ontology of smart artefacts increasingly involves actors beyond the mere users.*

The online environment in which I read myself into through the interface has an immersive character (also see Verbeek 2015: 219). I get deeply engaged with the

[53] It has to be emphasized that the influence of the interface on my experience of the world is *not limited to the online context*. In the networked era, the idea of a separated *cyberspace* and *meat world* is completely redundant. The way in which smart devices deliver services –from the way in which I retrieve information online, keep in touch with friend and colleagues, read the news and do my groceries- inherently shapes my interactions in and experiences of the world.

online service or the interaction with others through an online platform while smart algorithms constantly monitor my actions and –for example by profiling activities- simultaneously adapt and personalize my interface.

Taking into account these differences between a "traditional" hermeneutic relation and the characteristics of an interaction with a smart device, we need to broaden Ihde' framework. When analysing these interactions with smart artefacts and services, we encounter *hermeneutic aspects* as well as *immersive aspect*s, leading to what I refer to as an *immersive hermeneutic relation*. Through the interface we read ourselves into the online context, while simultaneously this context is pro-actively engaging with us in visible, but often also in invisible ways.

5.5.4 Immersive hermeneutic relations

It is generally accepted that with technological mediation some aspects of reality are highlighted while others move to the background. This in fact is essentially what mediation is about. *Mediation always entails a transformation*. If there would be no difference in the experience of the world with or without a mediating artefact, the artefact would be futile.

To make it more concrete, in its mediating activity the thermometer directs my attention to the temperature outside, it represents the world in a specific way, namely as a world where it is cold. This representation of the world steers my behaviour in certain ways because now I will wear a winter coat, a scarf, and a pair of gloves when I leave the house. The thermometer does not say anything about the time of day, whether or not it is snowing or if there are a lot of people on the street. Because they focus on specific aspects of the world, leaving out others, mediating technologies direct our attention and help to shape what we think of as being real. They mould reality.

Just as in the interaction with 'traditional artefacts' like the thermometer, with smart artefacts, some aspects of the world are emphasized while others are hidden from sight. However, the transformation smart artefacts bring forth has a fundamentally different character.

Smart artefacts and online services draw heavily on personalization. As we have seen, in the networked era, services are tailored to cater to the specific needs of

individuals. The mediation by smart artefacts, therefore, is increasingly becoming personalized. Consequently, what is shown and hidden from sight for one person can be completely different for somebody else.

While it is true that people may act differently and in unique ways upon the thermometer being below freezing point –some will want to go outside and enjoy the cold, others will directly walk to the canals to check the ice, others (like me) will turn on the central heating and stay inside- the thermometer itself will display the same temperature to all people. However, online the temperature fluctuates depending on who is looking and who has enough technical savvy. If smart artefacts mediate our interactions and therefore help us to shape reality, *reality increasingly becomes individualized*.

A final important aspect of the mediation of smart artefacts is that they not only mediate the experience of users and the –augmented- world but also of *the curators perceiving the users*. Through the use of smart devices and online services, users are increasingly becoming *visible* to the curators of the smart artefacts. Without looking at specific artefacts or services, it can be expected that smart artefacts of which at least partly the business model depends on monetizing data, represent, through data mining, the user for the curators behind the artefact. This visibility is often translated in profiles that are used to optimize and adapt the functioning of the device. Data may, under certain conditions, also be shared with or sold to third parties.

Although users become *visible* to the curators, this relation of mediation is often *invisible* to the users themselves. Although users are, to a certain extent, aware of the connection they have through their smart devices with these curators, they do not have a clear idea of the profoundness and scope of it. The directedness human beings experience makes that they 'forget' the mediation of their interaction; a mediation of which, in the case of smart artefacts, curators are indissolubly taking part.

With a reference to the anthropological laws of Plessner, I identify this relation of users and curators as *the relation of invisible visibility*. Users become increasingly visible in for them an invisible way through their use of smart artefacts. This extra relation of mediation enabled by smart artefacts obviously brings forth questions of privacy and accountability, but as we will see in the following paragraphs, also challenges for trust.

5.6 Mediated immediacy in the networked era: some challenges for trust

In the previous section on *mediated immediacy*, I focussed on the way in which human beings experience the world around them mediated through the artefacts they use. Artefacts, when they function properly and perform in a way that is aligned with the expectations of the user, withdraw from attention. Through their artefacts, human beings are directed to the world; *human beings are always present in their world*.

5.6.1 Designed for confidence

From a perspective of trust, we could say that when people experience their environment in a direct manner, "forgetting" as it were the mediating workings of their smart devices, they are in a state of *confidence*, of *system trust*. Users presume that the smart artefacts work as they were intended to. As users in general do not have in-depth knowledge about the functioning of the device, they put confidence in the expertise of others and rely on them developing well-thought-out devices. Confidence often thrives on a sense of "everyone does it, so why can't I?". Without confidence working in the background, it would be impossible to get around in the networked era. The information systems and smart devices, increasingly taking up a central role in everyday life, would lose all their attractiveness if users had to completely understand them or had to consciously take into account all the possible complexity these devices bring forth. Interpersonal trust could never reduce the complexity inherent in smart devices. In general when using smart artefacts, users are therefore in a state of confidence.

The design of smart devices adds to this state of confidence. Smart devices are generally developed in such a way that they are easy to use and do not ask much technical understanding of the user. The user knowing how to handle the device does not necessarily imply that he or she also knows exactly how the device works.

As we have seen in the previous section, if the device itself becomes the object of attention and is being evaluated, this often happens in a dual manner: "does or does it not work?"

When it functions properly, the artefact or service is taken for granted and is not questioned. As a result, confidence in the artefact again facilitates a smooth interaction. If the device unexpectedly does not function, as it should –and confidence, therefore, seems misplaced- the reason for the device to malfunction is ascribed to external causes. If the problem can be solved, the reasons for malfunctioning are accepted by the user, and possible damages are low, then confidence will be restored.

Although the dual "does it or does it not work" approach is not wrong –a well-functioning device is one of the most important conditions for system trust to be established- it sometimes is too narrow an approach. It may hinder questions about the *design* of the artefact and how that design relates to the *interest of users*. Merely focussing on its functionality, we may lose sight of *the mediating qualities of the device*; the way in which the device opens up the (online) world.

5.6.2 A personalized familiar world

The personalized way of mediating the world by smart devices may result in a world, which is very *familiar to the individual but lacks the common ground*, the implicitly shared norms and values that are needed to create a world where trust can thrive. The actions of others are one of the main sources of complexity human beings have to deal with. If human beings increasingly perceive the world as reflecting their initial beliefs and are not confronted with different perspectives, personalized mediation may lead to a *moral bubble* instead of a familiar world. If everyone lives in his or her *personalized familiar world* instead of in a *shared familiar world*, this may hinder fluent interactions with others, as trust becomes more demanding to give. If a shared familiar world declines, the complexity which trust has to reduce becomes larger. The hiatus inherent to human life therefore widens.

A second challenge for trust is the new relation of mediation between curators and users as a consequence of the networked ontology of smart artefacts. The directedness of human beings towards their environment together with the design of smart artefacts focussing on user-friendliness and efficiency makes that for users the presence of curators is not self-evident. The influence and power of curators goes

beyond the actual experience mediated by the smart artefact. Data collected through the smart artefact is not only used to adapt and optimize the functioning of the device or service itself, but may also serve other goals such as targeted advertisements on other websites and services, pro-active services and price-differentiation.

The invisibility of these actions makes it rather difficult to escape the state of confidence and actively engage with the question if curators are trustworthy and if the actions they undertake are in line with the interests of the users. As we have seen in chapter 3, confidence or system trust is first and foremost about *not* personally looking for certainty but about relying on the experts who do know the functioning of the system and are able to let it work as it is supposed to. However, in the networked era, the expertise of curators is not merely active in a certain, limited domain or system like it was the case in the air traffic example. The influence of curators reaches beyond the mere artefact they develop and maintain. Consequently, it is unclear what this confidence users put in curators is worth, as users often do not oversee the curator's actions that move beyond the mere curation of the artefact.

First, to have confidence or trust in something or someone, one has to be aware that one is in a *relation of uncertainty* and in a *dependent position*. One has to know that there is *something at stake*. These conditions do not seem to be met in the relation of mediation between curators and users due to the invisibility of their relation. Second, not only is the relation as such invisible to the user, the consequences of the relation, which have an impact beyond the mere mediated experience by the smart artefact, are also difficult to perceive. How can a user come to know that a price is higher for her because she uses a Macintosh computer instead of a Dell device? How can a user oversee the impact on her life because data brokers have collected and sold her data to an insurance company? In these situations, Dewandere (2015: 212-213) speaks of a risk of "reality theft". Personalization threatens the idea that reality is a shared reality where some fundamental aspects are generally the same for everyone.

In line with Heidegger, an answer could be: it only becomes obvious when it has been done in a completely wrong way; when personalization malfunctions. When I do not want to book a plane ticket but I receive wrongly targeted advertisements about cheap tickets wherever I go online, at such moments, I may become aware of the fact that I am being tracked. However, who can I hold accountable? Which curator or curators are influencing my perception of the world beyond the mediation

of the smart artefact they control? Moreover, where such badly targeted advertisements definitely will annoy me and perhaps will make me a bit suspicious, they may also wrongly hush me because if they function this improperly, what harm can they do? Unfortunately, badly working personalization does not say much about algorithms that do work excellently, as we are not aware of the latter. *The mal-functioning services are not representative of those who do work properly*.

All in all, to be in a trust relation, system trust or interpersonal trust alike, there has to be some understanding of the relation one is in; one has to know there is something at stake. For the relation of mediation of curators and users, the awareness of users seems to be lacking due to the invisible character of the relation. Consequently, when users put confidence in a smart device, this confidence should be understood as confidence in the mediating function of the smart device and not necessarily as confidence in the actions of the curators going beyond this mediated experience.

5.7 Utopian standpoint

The final anthropological law Plessner describes is the law of the "utopian standpoint". The contingency of human life and the awareness human beings have of it leaves them with a feeling of triviality. They have to bear the thought that their choices could have been different and, therefore, that their lives could have been different. Because of their eccentricity, human beings are *homeless by nature*. They can only live the life they make for themselves. While they on the one hand experience their life through and through as their own, a life no one else can lead for them, on the other hand they see the smallness of it.

The law of the utopian standpoint actually refers to the inner drive all human beings have to strive for that utopian goal of a home, a native soil, a place where their interactions with others, the world around them, and themselves is no longer broken but gains a direct character. The utopian standpoint refers to the desire of human beings *not to bridge but to overcome* the three-fold distance in them, between them, and between them and the world.

A dominant manner of coping with this contingency inherent in human life, Plessner localizes in *religion*. Although, the face of religion changes over time and can take on a different shape in different cultural settings, its core, namely, to provide a

certainty, a final ground which human beings lack for ontological reasons, remains essential. To be able to hold on to the belief that there is a meaning to life which is pre-given and not has to be made first provides a certainty which nature does not provide.

When Luhmann analysed trust, he came up with three related concepts: confidence, trust, and faith. Where the first two both were an affirmation of contingency –confidence in an implicit way and trust in an explicit way- faith is a *denial of contingency*. It is not to act *as if* the future is certain but *to believe that it is certain*. The utopian standpoint in its most extreme variant, as a fundamentalist conviction, with no room for doubt or wonder, excludes trust and in fact is in contradiction with the eccentric organisation of human life.

Similar to the previous anthropological laws, Plessner however also emphasizes the *paradoxical character* of this utopian standpoint. Naturally deprived of a final ground, human beings turn to a God to ensure their self-made home of a solid fundament. Simultaneously, however, their eccentric positionality also leads them to doubt the existence of such a divine creature. While it is true that the ability to believe in a higher power may reduce complexity inherent in human life, the fundament provided by religion remains shaky due to the eccentric positionality of human beings. Again, complexity can be reduced but not diminished. On a similar note, De Mul (2014b: 459) interprets this paradox as the "tragic" nature of human beings. The coming together of "necessity and freedom, brute contingency and significance," (idem) is simultaneously the burden and the splendid chance human beings have to relate to.

5.7.1 A utopian standpoint in the networked era

It can be argued that especially in late-modern Western society the place of God has increasingly been occupied by technology (de Mul 2003). Every new technological innovation is accompanied by promises to help human beings overcome the three-fold ontological distance they experience, provide them with complete mastery over things, and solve a myriad of other societal problems.

Also in the networked era, such high expectations are abundant. Big Data applications in the health domain should make it possible to personalize and therefore optimize treatments, pro-actively functioning smart artefacts and smart

services should solve problems even before users are aware of them, ambient environments will bend to cater to the needs of their inhabitants. All in all, the friction that could be experienced because of the mediated character of the interaction may vanish because of the introduction of smooth-operating and pro-active smart and networked technologies. The interplay between technology and human beings in the networked area is increasingly gaining a *fluent*, almost *natural character*. The *invisibility* of their functioning feeds the longing of people to live in an environment enabled by smart technologies without the interference of those technologies. Human beings actually want *the transformation without the mediation*. Ihde (1993: 75) speaks revealingly about a "doubled desire" of human beings, which:

> "on one side, is a wish for total transparency, total embodiment, for the technology to truly "become me." Were this possible, it would be equivalent to there being no technology, for total transparency would be my body and senses; I desire the face-to-face that I would experience without the technology. But that is only one side of the desire. The other side is the desire to have the power, the transformation that the technology makes available" (Ihde 1993: 75).

However, as we have now repeatedly seen, technology is never neutral. It always co-shapes the situation. Consequently, human beings may long for the outcome provided by technology; the mediating workings of that same technology cannot be erased.

Also the Internet itself has been acknowledged to hold such utopian promises (see chapter 4). The potential to connect people whole over the world, to make the world "flat" (Friedman 2005), to ensure self-government (Rheingold 1993), and total access to information are just a few of the aspirations surrounding the Internet.

The developers and early adapters of the Internet upheld similar beliefs. They claimed that the Internet would make nation state-based governments redundant, enable self-regulation, and restore freedom and autonomy. Moreover, online people would no longer be bound to the material world with its restricting laws –natural and social alike- and the limits of their body. On the Internet, people could truly become themselves, experimenting with their identity because "all they see are your words" (Turkle 1995: 184).

Also the availability of information through the Internet instigates the belief of godlike *omnipotence* and *omniscience*. Complexity inherent in human life suddenly seems to become something solvable if we would just be able to collect, connect, and use all the information to predict, and if necessary, prevent a certain state in the future from happening. Although, with Luhmann, we have already seen that technology used to reduce complexity always brings forth new complexity, Dewandere (2015: 198) rightfully observes:

> "In scientific terms, contingency is just another name for 'epistemic failure', a not-yet-known. By denoting contingency with the term uncertainty, i.e., as a negative, certainty is made the norm or the ideal".

5.8 Utopian standpoint: Challenges for trust in the networked era

At the heart of the anthropological law of the utopian standpoint lies the confrontation human beings have with their own contingency. Although human beings experience that they are at the steering wheel of their life, simultaneously they are also confronted with the triviality of it. Their ontological homelessness cannot be shaken off; they always take it with them wherever they go.

In religion, Plessner sees a fundamental strategy human beings apply to deal with this confrontation. In late-modern western society, religion increasingly has to make way for technology. The high expectations surrounding technological innovations, as for example advocated by the open Internet movement, reflect the utopian desire to transform the world through technology without actually taking into account this technology.

5.8.1 Faith in technology

A first challenge the utopian standpoint imposes on trust is its inclination to *move towards faith*. If the utopian standpoint loses its paradoxical character and, consequently, the openness inherent in the eccentric positionality diminishes, trust or confidence is no longer possible. To trust or to have confidence is to accept contingency and the possibility that things may turn out differently than expected.

However, faith is a denial of contingency. To believe is to be sure about what the world is and how it should and will look like.

Similarly, if one is completely dedicated to technology, owning this technology as if it is an integral part of one's being, one may lose sight of the unintended and sometimes unwanted side effects it may cause. A utopian belief in technology is a longing for the transformations it brings forth without taking into account the new complexity that also arises with every device or service that is used. The changes may be small and even invisible to the naked eye, but all artefacts influence the context in which they are put to work. *To surrender to technology is to give up the critical stance towards technology*, as well as to others and ourselves. This critical stance human beings are able to take because of their eccentric positionality, may be sometimes experienced as a burden because it deprives human beings from a direct and uncomplicated interaction. However, it is also the starting point of new ideas, creativity, and innovation. A strong utopian belief in technology may at first sight seem attractive, however, it comes at a high price.

5.8.2 Interpersonal system trust

A second challenge, which is a direct consequence of an overly-embracing faith in technology, is the belief that Internet technology will create a global community where 'traditional values' such as reciprocity, reputation, interpersonal trust, and thick social ties will, just as in pre-modern times, be leading in the interactions of users. As a result, top-down regulation, such as provided by states in legislation and by commercial actors in contracts, would increasingly become redundant. In the networked era, especially *information intermediaries* such as Facebook, AirBnB, Twitter, and Uber create platforms, which enable interactions between persons by pro-actively facilitating easy-connection and information exchange. They set up an environment where people can present themselves and get to know each other based on social cues such as pictures, shared history, and known reputation brought together in so-called profiles. In these environments, system trust as discussed in chapter 3 might seem to be no longer necessary as people, just as in pre-modern times, can act based on interpersonal trust.

While it is true that through the Internet, the possibility has been created to develop relations that are dominantly based on interpersonal trust, the context in

which this interaction takes place cannot be left out of the equation. As we have seen, information intermediaries, and curators more generally, may influence and steer the interaction. Moreover, the technology itself may invite users to display certain behaviour. It is not so much that it is impossible to develop interactions, which are based on interpersonal trust, through the Internet, however, this interpersonal trust has obtained a *strongly mediated character*. Trust developed for example between users of an online platform like AirBnB cannot therefore account for interpersonal trust. However, it is not system trust either. Where system trust is about trusting the experts behind the system and the employees as contact points of those systems, the system in the networked era increasingly moves to the background of the user's experience evoking the idea of seamless and fluently natural interactions. *While the system is definitely present in its consequences, it is absent in the phenomenological experience of users*. This development may mistakenly lead users to believe that their interaction is interpersonal where it is in fact truly part of a system process. I have defined this kind of trust-based interaction as: *interpersonal system trust* (Keymolen 2013). People experience their interaction as interpersonal, while there is –often unnoticeable in the interaction as such- a mediating system involved.

5.8.3 Solving the problem of contingency

A final challenge to be faced is the strong belief in *predictability* in the networked era. The collecting and mining of data on a big scale has set in motion a new way of approaching reality. The basic belief is that if enough data can be gathered and correlations can be found, it will become possible to predict the future almost completely. As a consequence, risks do not have to be mitigated anymore, but can be prevented from happening at all. The ubiquitous call for Big Data analytics in almost all domains of life is the most conspicuous example of this current overall focus.

If it were true that the future can become something foreseeable, that the unknown could be disposed of, then trust would be redundant as there would not be any complexity to reduce. The two main sources of complexity, the awareness of human beings that a myriad of possible states could become reality in the future and that human beings can never completely predict the behaviour of others, would simply run dry. Where faith was a denial of contingency, the belief in a calculable world even goes a step further. Believing in a completely predictive world is not

denying contingency, it is *solving* the 'problem' contingency causes.

Scholars in Big Data analytics have pointed out that a 100% solid prediction is an unfeasible goal. Therefore, contingency can never be completely resolved and human beings will always need strategies to deal with the complexity caused by their understanding of this contingent world.

However, in general, human beings are very poor in assessing chances and probabilities (Kahneman 2011). What in fact is only likely will quickly become sure and proven, especially when the solution comes rolling out of a computer, lacking a clear and understandable explanation of how the result was reached. Even if people are dealing with probabilities, they often tend to approach the outcome in a dual "yes or no" way. This problem is accurately and painfully funnily illustrated in the sitcom Little Britain. The character of Carol Beer works in different customer service contexts and replies, after typing in the information on the computer, on almost every request with the words "computer says no". Jeroen van den Hoven speaks of "artificial authority" to explain the dependence of users on the functioning of machines. They can often only read the results produced by the artefact (van den Hoven 1998).

Nissenbaum is wary when we move to a world where safety and certainty are the two main goals worth striving. She states:

> "In a world that is complex and rich, the price of safety and certainty is limitation. Online as off, (…) the cost of surety –certainty and security- is freedom and wide-ranging opportunity" (Nissenbaum 2004: 173-174)

Also Nicole Dewandere (2015) raises fundamental questions concerning this longing for certainty. In her argument against an "omniscience-and-omnipotence" utopia, where the main goal is to gain sufficient knowledge and control, Dewandere (2015: 206), building on the work of Hannah Arendt, warns for a society where "relations create no surprise". Too strong a focus on predictability and control, she argues, may hinder the "societal intelligence and resilience" inherent in human life to thrive. It leads to approaching people in a mere instrumental way, at the costly price of losing the central values of "natality and plurality", both key concepts in Arendt's work. *Natality* refers to the ability of human beings to create and to initiate new beginnings. So, instead of trying to smoothen all possible problems lying ahead, one should have confidence in the ability of human beings to deal with difficulties and

unforeseen situations. *Plurality* is the threefold idea that the human condition is defined by equality (people experiencing others as other selves), specificity (the uniqueness of human beings), and the reflective nature of identity (in Plessner's terms: human beings' membership of the *Mitwelt*). By reducing human beings to a data set, focusing on correlation instead of meaning, approaching them in a merely functionalistic way, plurality is pressured.

> "Indeed if, together with Arendt, we believe that the purpose of politics is freedom, it is high time to endorse and make sense of the world we are living in…It is high time for plurality to substitute, or at least complete, the other metaphors underlying policy-making, i.e. the invisible hand (which encourages the pursuit of one's own interest, decoupled from all forms of empathy towards other selves) or the competitive race (which considers others as competitors to be defeated)" (Dewandere 2015: 215).

5.9 Conclusion: challenges for trust

In this chapter, I looked at the contextual layer of trust interactions. Making use of the three anthropological laws of Plessner, I analysed the way in which trust and smart artefacts are intertwined on the micro-level. An important starting point for this analysis is that human beings and their environment are indissolubly connected. What human beings are is established in their interaction with the environment. Simultaneously, the environment is shaped and gets meaning through the interaction of human beings. Currently, its pro-active, persuasive, and connective character shapes the environment in the networked era. Artefacts become smart by adding a computational component to them, bridging the gap between the material and the virtual.

The three anthropological laws refer to different aspects of the intertwinement of human beings and their environment, when adopting them to analyse the current interaction of human beings and their environment, they therefore highlight different but nonetheless closely connected challenges for trust. The most important ones are listed below.

5.9.1 Artificial by nature

-The openness of artefacts to change and adaptation is generally preserved for curators and not so much for users. These changes to artefacts may happen with or without the awareness of users which –when it does come to the attention- may instigate issues of trust in the artefact as well as in the curators behind the artefact.

-As a result of the difficulty for users to adapt their artefacts, their assessment of the artefacts remains in the realm of functionality. Does or does it not function properly? This perspective may hinder more fundamental questions about the trustworthiness of the services provided and the intentions of the curators behind the artefacts.

-Artefacts may impose codifications, for example through their terms and conditions. This may result in certain distributions of power and control, which can help to reduce complexity, because it makes the interaction with an artefact more predictable. However, it may also destabilize the interaction, when these rules change on a regular basis, are set up in incomprehensible ways or do not take into account the interests of the user.

5.9.2 Mediated immediacy

-Through smart artefacts a new relation of mediation becomes possible: namely, a relation between curators and users. Facilitated by the networked ontology of smart artefacts, the actions of users are made visible to the curators in an often-invisible way for the users themselves. Again, this may lead to users mistakenly assessing only a part of their interaction with smart devices, clouding important questions concerning the interests of curators.

-The tendency to personalize the services delivered by a smart artefact may lead to an i*ndividualized* familiar world, where, for trust to be possible, there is the need for a *shared* familiar world.

5.9.3 Utopian standpoint

-The utopian belief in the innovations brought forth by smart artefacts may lead to a denial of contingency. As a result, the mediating workings of the devices are left out of the equation. This may result in the inclination that interactions are

interpersonal where online platforms and smart artefacts in fact mediate them.

I have intentionally spoken about 'challenges for trust' and not about 'problems for trust', because the premise that human beings are "artificial by nature" means that in the end the way in which artefacts are shaped and designed is not based on natural laws which are given and unquestionable. The anthropological laws with their paradoxical character, uniting contradicting poles (artificial-natural, mediated-direct, utopian-grounded) beautifully illustrate the constantly changing and ambivalent character of the interaction of human beings and their environment. Human beings are not free in the sense that they can shape their lives without artefacts and beyond the influence these artefacts have on their lives, but human beings do have the freedom to shape their lives *in relation to these powers (see: Verbeek 2011b: 73)*.

Verbeek and Kiran (2010) suggest that there are two specific conditions that have to be met in order for human beings to "trust themselves to technology".

> "First, the technology in question needs to leave room to develop an explicit relation to its mediating role, rather than being dominating and overpowering. And second, human beings need to have the ability to 'read' the mediating roles of the technology, and the skills to 'appropriate' it in specific ways" (Kiran and Verbeek 2010: 424).

What I aimed at showing in this chapter is that, taking into account these conditions; trust meets some challenges in the networked era. First, the openness necessary to shape such a free relation to technology is not the same for all actors. In general curators of smart artefacts have more influence in shaping the interaction of users and their environment than these users are aware of. Moreover, curators may have conflicting interests when their business model is built on monetizing their users' data. So, although smart artefacts do not necessarily need to be developed in such a way that they, for example, leak data to third parties in order to function –just as the bridges of Long Island did not need to be designed in a way inaccessible to busses- the incentive to do so nevertheless is very strong as it is currently the key strategy to make money.

Second, the invisible character of the mediations brought forth by the smart artefacts hinders a thorough questioning of our interactions with these artefacts. How to style an interaction if you are not even aware that you are in such an interaction?

In other words, how to have trust when you are not aware that you are in a dependent situation where something is at stake?

Third, the closed design of the smart artefacts hinders users to appropriate them in a trustworthy manner. While there is always the possibility to use an artefact in a way that was not foreseen by designers or to put it in a completely different user context than was first intended (an image that comes to mind is a TV advertisement where an older man uses his iPad as a chopping board to prepare dinner), due to the low interpretative flexibility for users, the options are limited.

Finally, the promise of smart services and smart artefacts to cater to our every need, smoothening the paradoxical character of human life may wrongfully persuade us to believe that we can completely control and even resolve the complexity inherent in human life.

6

Open Sesame. When your phone becomes your key.

Central to this chapter is a rather mundane artefact: a key. Generally, it brings to mind a small metal device, used to open and lock doors. However, as small and at first sight insignificant as a key may look, it plays an important role in everyday life. Through the use of a key one can decide and control who is allowed to enter a place and who isn't. Where walls are merely meant to keep people out –or 'in', depending on your perspective- keys offer choice. Keys bring along flexibility. As long as you have a matching key, you can go in and out of a room as you please. You can grant people access by giving them a key. You can deny them access by taking the key away. By locking the door, you distance yourself from the 'outside world' because only people with the proper key can enter. Keys therefore are strongly connected to the private sphere; they enable the creation of a private domain in which the control to access is delegated to the key and its owners.

A branch of industry where the key is central to everyday business is, of course, the *hospitality sector* and more specifically the *hotel sector*[54]. After checking in at the hotel desk, providing the hotel owner with some necessary personal information such as name, address, a copy of an ID-card or driver's license and credit card credentials, a hotel visitor receives the key to his or her room in the hotel. This key is the central artefact in the interaction between hotel owner and hotel guest. For the guest, the key

[54] However, it has to be noted that large global hotel chains no longer define the hospitality sector. We will see in the next chapter that increasingly also individual homeowners step in to the world of hospitality, providing places to stay by renting out their own houses.

is a necessary condition to obtain a private domain in a rather public environment. By offering the key, the hotel owner *transfers* a small part of her 'ownership' to the guest who now –within the boundaries of the hotel policy and common sense- 'owns' the hotel room for a certain period of time.

Mutual expectations come along with such a transfer. The hotel visitor expects the hotel owner and the hotel staff to honour her privacy by not entering the hotel room unannounced. The owner wants the guest to keep the room intact and bring in the keys when leaving the hotel. This latter expectation, however, has been a challenge for visitors to live up to. In the rush of the moment, returning the keys may easily be forgotten.

The problem of not-returned hotel keys is elaborated by sociologist and philosopher Bruno Latour (Latour 1992: 104; Latour 1990) who shows how a key attached to a weight may persuade the hotel guest to return the key to the hotel desk. By the association of different actants –hotel owner, key, a spoken request, an information board, a weight connected to the key- hotel visitors are prompted to change their *action programme*.

I will recapture Latour's analysis of the hotel key by retelling it in terms of trust. How does the chain of actants –human and nonhuman- change the action programmes and therefore the trust between hotel owner and visitor?

While the weight attached to the hotel key may have had a significant impact on hotel practices, the innovations in the hotel sector did not end there. In recent years, there has been *a* shift from *keys* to *keycards* that make use of a magnetic stripe, smart chip technology, or RFID technology. The introduction of these keycards enables a different kind of transfer than the 'old-fashioned' hotel keys as Latour described them. Not only has the problem of forgotten keys become less important –a new keycard can easily be printed, making the old keycard merely 'a card' as it will no longer be able to open the door of the hotel room- it also opens up the possibility of adding new functionality to the key. Where the 'old-fashioned' key had a rather limited repertoire of functions, the keycard can be programmed to do much more. For example, it can be used to monitor the use of hotel facilities linked to the keycard or it can track the presence or absence of the hotel guest in the hotel. I will show how the keycard co-shapes a specific kind of trust relation between the hotel owner and the hotel guest, differing from the relation mediated by the hotel key.

Finally, I will look into the newest, state-of-the-art hotel key, which actually no

longer is a key or a card but a smartphone. In 2014, the famous high-end hotel chain Hilton has invested 500 million dollars in the development of a digital environment brought together in an app, which not only makes it possible for a customer to select a specific room in the hotel and pre-order extra services, but which also turns a telephone into a digital key. By waving a smartphone in front of the lock, the door opens. With this innovation, a hotel guest no longer has to wait at the hotel desk to check-in and the hotel owner can assign tasks to his employees other than checking in guests.

By focusing on the 4 Cs: context, construction, curation, and codification, I will show how trust is shaped through this new digital key and how it changes the character of the relation between the hotel owner and her guests.

6.1 The hotel key and a cumbersome but handy key chain

Who has not, at least once, forgotten his keys? I do not know if it was intended, but it seems that Latour could not have picked a better artefact than a plain and often forgotten key to illustrate how artefacts which at first glance only seem 'neutral' instruments in the hands of their users and, therefore, often forgotten in social analysis, do nonetheless matter and make a difference. In different writings, Latour therefore makes a convincing plea not to forget the nonhumans when analysing social interactions (Latour 1992; Latour 1993).

To understand the importance Latour attaches to this inclusion of artefacts and other nonhuman actants, we first briefly have to go back to the previous chapter. There, we saw that in reaction to on the one hand the instrumental view of technology as a neutral instrument and on the other hand the deterministic –and overall pessimistic- perspective on the relation of human beings and technology proclaimed by the 'classical' philosophers of technology, more empirically-based, contextual approaches were developed. These different schools, which can be gathered under the umbrella of *constructionism*, consist of amongst others the Social Construction of Technology (SCOT), the Social Shaping of Technology (SST), and Actor Network Theory (ANT) developed by Latour together with John Law (van den Berg 2009: 29-30). Likewise, also the philosophy of technology itself underwent changes and became, while interacting with these constructionist schools, more empirically informed and focussed on specific practices (Kaplan 2009).

While it is true that all these disciplines have different starting points and a specific methodology, they are united in their critique on the instrumental and determinist stance. Where the *instrumental perspective* does not take into account the way in which technologies bear values and pre-sort the actions of their users, the *deterministic perspective* has a rather one-dimensional view of the influence of technology, only taking into account the influence of technology on society. The constructionist disciplines offer an alternative view, which is based on a *mutual shaping* approach (Frissen 2004, 1994, 1997). In the interaction, both the users and the artefact are shaped and form an identity.

6.1.1 Radical thinker

Latour is probably the most radical thinker on this matter. He not only criticizes the deterministic and instrumental perspective but also other constructionist disciplines for grounding their position on the subject-object dichotomy and not really overcoming this divide, despite of their claimed intentions. Social constructivists, such as the adherents of the SCOT school, are too focussed on social factors in their analyses. Doing their best to avoid the pitfall of technological determinism, they fall in the pitfall of social determinism (van den Berg 2009: 32). Phenomenological accounts, on the other hand, with their emphasis on the intentionality of human beings towards their environment only assent to the subject-object dichotomy, Latour claims (Latour 1993; also see Verbeek 2000: 180-188).

Latour develops therefore a theory, or better, a set of concepts that could replace the "technology/society" divide by instead focusing on *technical mediation*. The basic idea is that humans and nonhumans can only be understood through the networks that connect them. It is in their relation with other actants that humans and nonhumans are shaped. An artefact only gets meaning in the interaction with humans, and humans become who they are in their interaction with artefacts. Following Latour, to understand the way power relations work in society, we therefore also have to take into account the nonhuman actants and the way in which they persuade and mobilize other actants to display certain behaviour in social links.

> "I argue that in order to understand domination we have to turn away from an exclusive concern with social relations and weave them into a

> fabric that includes non-human actants, actants that offer the possibility of holding society together as a durable whole" (Latour 1990: 103).

In other words, if we want to understand power relations and moral behaviour we should include the nonhuman actants in our analysis and think through exactly how this technological mediation takes place. To illustrate how nonhumans are part of power relations and how all networked actants influence each other, Latour comes up with the example of the hotel owner who seeks a way to persuade his guests to bring back their hotel keys (Latour 1990; Akrich and Latour 1992).

6.1.2 The problem of missing keys

The history of modern hotels allegedly started in 1862 with the opening of the marvellous Le Grand Hotel in Paris (Ambrosino 2014). With 800 rooms and beautiful architecture, the Grand Hotel set a new standard in the hotel sector. Also the Grand Hotel's key policy was a prime example of the way in which modern hotel business should be run. Metal keys were "attached to a big key-ring, which was hung on a board at the concierge office" (Ambrosino 2014: 3). Consequently, guests had to visit the concierge first in order to obtain their key or turn it in, as it was not allowed to have keys outside the hotel[55].

The hotel owner kindly requesting to leave the key at the front desk, did not seem to have much effect on the elegant French guests and also a sign with the explicit inscription "please leave your room key at the front desk before you go out" did not result in the behaviour demanded by the hotel owner (Latour 1990: 103). It is only when an "innovator" comes to the rescue and "displaces the inscription by introducing a large metal weight, the hotel manager no longer has to rely on his customer's sense of moral obligation" (idem: 103).

[55] Latour never explicitly refers to an actual existing hotel or period of time in which he situates the problem of the missing hotel keys, because it obviously is more a thought experiment than it is an actual empirical case. However, I like to think of it as taking place in the early days of Le Grand Hotel. The board with the keys hanging in the room of the concierge still had to be invented and the hotel owner was desperately looking for a way to persuade the guests to turn in their keys.

> "Where the sign, the inscription, the imperative, discipline, or moral obligation all failed, the hotel manager, the innovator, and the metal weight succeeded. And yet, obtaining such discipline has a price: the hotel manager had to ally himself with an innovator, and the innovator had to ally herself with various metal weights and their manufacturing processes" (Latour 1990: 104).

What happens here cannot easily be understood by upholding a rigorous distinction between humans and nonhumans. For Latour, social interactions do not merely consist of human agents but include human and nonhuman actants alike. To understand how a bulky keychain changes the behaviour or 'program of action' of both human (i.e.hotel owner, guests) and nonhuman (i.e. the key, the weight) actants, one has to see how "the original program of action is thus translated or transformed in the technical mediation into a new one" (Verbeek 2000: 173).

The program of action of the hotel manager is 'I want the hotel guests to bring back their keys' which may be in conflict with the program of action of the guests which are more focused on 'having a nice holiday'. The latter is an anti-program because it does not align with the intentions of the hotel owner. The hotel owner can now try to connect with other actants to fortify her message. She can add an oral message to her wish: "please Miss. Anderson, return the key when you leave the hotel"; she can put up a sign with the same message; and she can attach a cumbersome weight to the key.

Every time, the hotel owner includes a new actant in the chain of mediation, she tries to persuade the hotel guests to adapt their program of action. These associations with other actants are always a balancing act. If the hotel manager would not tolerate a single missing key, she would have to align with guards at each door to ensure that all guests give back their key. Although this might solve the problem, it would probably also lead to new problems, such as having no customers at all. And of course, there are always stubborn clients who may try to remove the weight from the key or new actants entering the scene like dogs traveling together with their bosses who see the key weight as something fun to play with. To become a predictable, stable action, not necessarily all but most anti-programs have to be countered. It then becomes something which people just do, without really thinking about it or questioning the request. "The customers obey the order, with only a few exceptions, and the hotel manager accepts the loss of a few keys" (Latour 1990: 105).

The initial message 'return the keys when you leave the hotel' is no longer the same because of the associations taken upon by the hotel manager. It has been *translated*. By displacing the message in to the weight added to the key it has transformed. The key together with the weight attached to it substitutes the hotel manager's demand for returning the keys. The design of the key weight helps the hotel guest to return the key to the front desk. It is no longer something the hotel guest has to do by herself; it is partly *delegated* to the bulky key chain. In these associations of humans and nonhumans, changes occur. Because of their connectedness they are no longer the same entities. As Latour describes:

> "Customers no longer leave their room keys: instead, they get rid of an unwieldy object that deforms their pockets. If they conform to the manager's wishes, it is not because they read the sign, nor because they are particularly well-mannered. It is because they cannot do otherwise. They don't even think about it. The statement is no longer the same, the customers are no longer the same, the key is no longer the same –even the hotel is no longer quite exactly the same" (Latour 1990: 105).

6.1.3 Trust between the hotel owner and the hotel guest

Now that we understand the way in which Latour analyses the technical mediation taken place in the case of the forgotten hotel keys, it becomes possible to retell this story now focussing on the issue of trust. The hotel owner (*trustor*) has to deal with the complexity of not knowing for sure if her guests (*trustees*) will return the key to the front desk as they are supposed to. In the transaction of giving the key to the guest, there is *something at stake*. If the hotel owner wants to have a flourishing hotel business, she is bound to providing the guest with a key, running the risk of losing the key if the guest does not return it. Of course, there are some checks and balances in place. The guest has handed over her personal information to the hotel owner, making it possible to identify and trace her if something might go wrong. Moreover, the guest has signed a contract, agreeing to act according to the rules set in the hotel policy. Still, the hotel guest has the freedom to act (*agency*) in a way that is in conflict with the *expectations* of the hotel owner, making the hotel owner *vulnerable* nonetheless. Trust can never be forced or guaranteed. In that sense, trust is always

blind trust. It entails the suspension of looking for more evidence, more certainty, and accepting the uncertainty inherent in every social interaction. Trust is a fiction necessary to face reality. The hotel owner providing the hotel guest with the key acts *as if* she is sure about the way in which the guest will behave, while in fact she is not.

Unfortunately for the hotel owner, this trust is often shattered, because of the absent-mindedness of careless hotel guests. The hotel owner finds it increasingly difficult to trust the hotel guests with the key and therefore tries to influence the situation by trying to steer the behaviour of the guests in the right direction. She calls upon the guests and puts up a sign, and although these actions help to remind the guests of the fact that they are in a situation where trustworthy action is expected from them, it does not significantly seem to impact them.

In an effort to turn the tide, the hotel manager calls in the help from an innovator who comes up with the plan to add a weight to the key in order to make it physically less attractive for guests to take the key with them when leaving the hotel. Also in this new relation of the hotel owner and the innovator trust issues arise. The hotel owner has to trust the innovator to come up with a successful plan to persuade the guests. Next, when this invention is adopted in the interaction of hotel owner and guests, it also becomes an object of trust - of system trust, more precisely. The hotel manager then not only has to trust the guests, but also has to have confidence in the way in which the keychain functions.

Finally, when adding the heavy keychain to the key, most of the guests adapt their behaviour and bring back the key. Although the interaction has changed because of the introduction of the weight attached to the key, it remains an interaction where trust is present. Trust is now *distributed trust*, as the trust first uniquely vested in the hotel guest is now shared between the hotel guest and the key with weight; or even more specifically, it is invested in the new *association* of hotel guest + key + weight.

Although guests display more trustworthy behaviour because of the bulky keychain, they can still breach the trust of the hotel owner and take the key with them when they leave the hotel. In the altered relation, guests still have agency, which is a precondition for trust. If it could be possible to completely control the returning of the keys, trust would be redundant. Thus, although their actions are more predictable now they interact with the bulky keychain, hotel guests can still act differently than expected and hoped for by the hotel owner. Moreover, as already mentioned by Luhmann (1979), where technology is used to reduce complexity, it may solve the

problem while simultaneously creating new ones. For example, the new key chain persuades guests to return the keys but it also attracts new actants such as dogs who perceive the keychain as a toy that should definitely be played with. These actants may induce the hotel manager to rethink her policy for animals in the hotel and change the trust vested in hotel guests who travel with their dogs.

6.2 When a key becomes a card

In the example of Latour, the cumbersome keychain is a material strategy to persuade hotel guests to display trustworthy behaviour. Another strategy, however, could be not to try to change the behaviour of the hotel guests, but to see if it might be possible to rearrange the situation in such a way that there is *less at stake* for the hotel owner. In other words, if it is possible to provide the guest with a key which, when not returned, does not impose too much of a burden on the hotel owner. As we have seen, trust always is a risky business. If there is less to be risked, there is less need for trust as well.

The hotel keycard seems to tick all the boxes. This plastic card, generally having the looks and size of a credit card- can be programmed to open a specific door for a certain period of time. If it is returned to the front desk, it can –depending on the type of card- often be reused by overriding the initial data and putting new data on it. If it is not returned, the costs to replace it by a new card are rather low. As a result, in this new situation there is less at stake for the hotel owner.

While it is clear that the keycard is a successful strategy to deal with the uncertain behaviour of distracted hotel guests, it was not the prime reason to introduce the card in the hotel business. It was actually a 1976 lawsuit by famous singer Connie Francis that was the wake-up call for hotel owners to abandon metal keys and to look for alternatives. The lawsuit Garzilli vs. Howard Johnson's Mother Lodges Inc led to $1,5 million in damages awarded to Connie Francis and her husband. In 1974, she was raped after an intruder opened her apparently locked sliding hotel door and entered her room. The Court found that the hotel had not fulfilled its duty of "reasonable care" and listed several reasons to substantiate its verdict. Amongst others: the doors were easily opened although they appeared locked, burglars had already entered the hotel four times through these sliding doors, and the safer locks that were ordered were still not installed (Sherry 1993: 355). As a

result of this verdict, hotels, together with their insurance companies, "learned that entrance and exit of their rented rooms related to their own liability and guests' safety." As a result, "[t]hey became more willing to pay for good security" (Giordano 1997: 1).

The keycard seemed to be the solution. A Norwegian inventor named Tor Sornes, who heard about the tragic event that happened to Connie Francis –one of his favourite singers-, invented the card in 1975. It was a plastic card with 32 holes, which made it possible to compose a unique code for every new guest (Ambrosino 2014). Sornes calculated that there were over 4 billion possibilities, enough to provide the whole population of the Earth at that time with their own personal hotel card.

With this card the privacy of the guests was assured more than in the situation with the 'normal' key. If previously a key was not returned, the only way to fully ensure the privacy of the room was to change the lock and buy a new key. This obviously is a time consuming and costly solution. More likely, the key was therefore merely replaced by a copy. Consequently, the hotel guest who still possessed the key, could access the room long after he or she was allowed to do so or sell the key on the street. Indeed, in the 1960s and '70s these keys were sold on the black market for $500 (Sherry 1993: 355). Because in general the name of the hotel and number of the door was printed on the key chain, it was quite easy to make use of such an orphaned key. In the case of the keycard, however, when a card was lost or not returned, a new card could easily be printed. The privacy of the room would be less compromised as it was only in the time between losing the card and replacing it that unauthorized individuals could open the door, and then if and only if they would be able to trace the matching room number, which was not printed on the card.

6.2.1 How do hotel keycards work?

Where the first hotel keycard Sornes invented was a mechanical card –called the VingCard- of which the punched holes in the keycard had to match the template card put in the lock (Sornes 1979), he kept on working to improve the security of the card to finally introduce the electronic keycard, powered by LEDs in the beginning of the 1980s. It was the predecessor of another type of keycard that became widely adopted: the card with a *magnetic stripe*. The magnet stripe has to be run over a sensor, which can then read the information on the stripe. There are several ways to encode the

information on the card.

The first method that can be used is to "encode the check-out date of the hotel guest and the lock information. This tells the electronic lock that the key is supposed to open the door until the specified date and time. The hotel keycard is also issued the lock information. Every lock has an individual code"[56]. When the card is then inserted into the lock, the lock compares the information on the card with the information locally stored in the lock.

The most common way to embed these magnetic keycards, however, is to wire every lock to a server. This gives the hotel staff more control over the keycards and makes it also easier to replace a lost card. Where with the first method the lock has to be manually reset when a keycard has gotten lost, with the second method this can be done from behind the front desk. Wiring the lock to the server can be done by making use of hardwired connections to a central computer or by making use of different sorts of radio waves.[57]

The most recent keycard is the *RFID* (Radio Frequency Identification) keycard. These cards are provided with a radio sensor chip. When they are held close to a corresponding reader, the doors can be unlocked. Because these cards contain both microchips and radio technology, they are considered to be the most secure ones. Also with the RFID card, locks are connected to central computers. This can be done *wirelessly*, by installing routers and gateways that connect the locks to the LAN (Local Area Network) of the hotel. Or it can be done *through wires*, directly connecting the locks to the LAN of the hotel. Next to the lock itself, RFID lock suppliers offer hotels also different monitoring options, which I will discuss in section 6.3.2.

Up until now, I have talked about RFID *keycards*, but actually, as RFID facilitates *contactless* interaction, the shape of the key becomes less absolute. Not only can cards be used, but keyfobs and wristbands for example as well. Moreover, as the state-of-the art lock systems are compatible with NFC (Near Field Communication)

[56] http://www.plastic-card-services.co.uk/information/hotelkeycardsinfo.html, Accessed 10 June 2015.

[57] http://www.magnetickeycards.com/, Accessed 10 June 2015.

and BLE (Bluetooth Low Energy) technology, it also becomes possible to enable contactless communication between the locks and smartphones equipped with NFC or BLE[58].

6.2.2 Revised: Trust between the hotel owner and the hotel guest

Just as the trust relations changed because of the introduction of the bulky keychain, the trust relations changed again with the introduction of the keycard. The keycard mediates in a different way the interaction between the hotel owner and the hotel guests, resulting in an altered trust relation between the hotel owner and the hotel guests. As we have seen, with the metal key, the complexity the hotel owner had to reduce by trusting his guests to return the key was rather large.

With the introduction of the keycard, however, the hotel owner (trustor) becomes less vulnerable to breaches of trust. If the guest (trustee) does not return the keycard, this is much less of a problem than it was the case with the metal key. The owner –or the hotel staff- can just print a new one when a key has got lost.

Moreover, the introduction of the keycard also in another way influenced the trust relation of hotel owner and hotel guest. In comparison with the traditional key, the keycard is more secure[59]. As a result, it becomes less likely that an intruder can enter the room. The positive expectations of the hotel guest, that the hotel owner provides her with a save place to stay will come to pass. As we have seen, trust always implies vulnerability, uncertainty; there has to be something at stake. By replacing the metal key with the keycard, this vulnerability is been substantially reduced.

Is trust then no longer existent in the interaction of hotel owner and guest?

Trust did not become redundant, as the interaction of hotel owner and hotel guest is certainly not completely defined by the transaction of the key or keycard. The hotel

---

[58] http://www.assaabloyhospitality.com/en/aah/com/press-room/product-documentation/, p8, Accessed 09 June 2015.

[59] Of course, also keycards can be stolen or forged. Security is never 100%. For example, magnetic stripe keycards can be cloned and there are several tutorials online on 'how to hack your hotel keycard'.

owner also still expects the guest not to cause any annoyance for the other guests and not to ruin the room. And the guests still expect the hotel owner and the staff to honour the privacy of their hotel room. Moreover, *technology's own weight* brings along new complexity both hotel owner and hotel guest have to cope with. The hotel owner has to trust her suppliers to provide her with reliable cards and systems, the hotel guest has to become familiar with the way the cards work (how to put them in the card reader, how to make sure they do not get damaged).

All in all, the introduction of the keycard did alter the trust relation, not because it led to a new distribution of trust –as was the case with the adding of the key chain- but because it reduced the vulnerability of the hotel owner and helped to establish a familiar world. While the introduction of the keycard adds up to a familiar world, necessary for trust to be established, it also introduces new complexity the actors involved have to relate to.

6.3 When a keycard becomes a smartphone

On 28 of July 2014, the prominent hotel company Hilton announced that as part of extending their customized digital services to hotel guests, they would make it possible for guests to use their phones as a key to open the lock of their hotel door. These new services would, amongst others, enable customers to choose their own room, pre-order services, skip the check-in at the front desk, and let them go straight to their room. Hilton declared to invest $500 million dollar to start this operation[60]. Also other hotel chains such as Starwood Hotels & Resorts (SPG) and Hyatt Hotels and Resorts started testing the possibility of using a smartphone to open doors (White 2014). Moreover, in brochures of suppliers, the possibility to convert current lock systems into systems that facilitate NFC and BLE which makes it possible to let smartphones and locks connect is being promoted[61].

We started this quest with a mundane metal hotel key that went missing more often than the hotel owner was willing to accept. Now, we find ourselves engaged with

[60] http://news.hiltonworldwide.com/index.cfm/newsroom/detail/27192. Accessed 09 June 2015.

[61] http://www.assaabloyhospitality.com/en/aah/com/press-room/product-documentation/. Accessed 09 June 2015.

the hotel key of the networked era, no longer a key but a phone. It seems as if with the future arrival of this keyless key, a new milestone in the history of hotel keys will be reached, bringing together the advantages of RFID enabled key systems, the managing software systems connected to the locks, and the functionality of the smartphone. By integrating the functionality of a key –opening doors- into the workings of a smartphone, the keyless key is a prime example of the *connectedness, personalization* and *pro-activity* we saw to be crucial to the current networked era.

Then again, the fact that this technical change is not neutral, but brings along new complexity, new vulnerabilities, and therefore new trust interactions, does probably not come as a surprise. To include the way in which the networked ontology of both key and lock, both phone and lock, mediate and co-shape the trust of hotel owners and hotel guests, I will analyse them through the conceptual lens of the 4 Cs framework I discerned to analyse Internet technology: *context, construction, curation,* and *codification*. As there is not yet much known about the specific, technical workings of the door-opening smartphone, I will base this part of my analysis on the digital key apps that are already online, the press releases of the hotels themselves, their current privacy policies, the coverage in the media, and the brochures of suppliers of the lock systems and accompanying software to manage these systems in order to make the network of actants as complete as possible.

6.3.1 Context

How do hotel guests interact with the hotel and its staff when the digital key becomes an integral part of their stay at the hotel? How do they perceive their hotel stay when it becomes co-shaped by the app?

First, guests have to download the designated app. This can be the app issued by the hotel itself –for example Hilton and Starwood both have their own branded applications- but there are also companies developing 'third-party apps'[62]. Whereas the hotel-issued apps are designed to fit the specific services and brand of the hotel, the third-party app is more generic and compatible with all hotels that have installed the proper locks and interfaces.

[62] https://wefunder.me/leapindigitalkeys, accessed 10 December 2015.

Generally, all apps are designed in such a way that the interface is easy to use and functions intuitively. The average user should be able to just open the app, tap through the menu and find his or her way around the digital hotel environment without having to depend on extra instructions or assistance.

Next, when the app has been installed and a reservation has been made, the guest can opt in to request a digital key. If the hotel already supports the digital key, the guest will receive a push notification –after checking in and confirming payment in the app- with the room number and the digital key –some kind of encrypted code- on the day of arrival. The guest can then go straight to the room, avoiding the check-in at the hotel's front desk. By waving the phone close to the lock on the door, the door can be opened. When guests want to check out, they can do that making use of the RFID tags around the hotel or in the app itself. After checking out, the digital key is cancelled and the phone can no longer open the hotel door.

In addition to the digital key, the app also allows guests to personalize their stay. They can choose –making use of digital floor plans- which particular room they want and they can pre-order all kinds of services: all to make their stay as pleasant and seamlessly as possible.

Already in 2010, a test was conducted in the Clarion Hotel in Stockholm to collect feedback of guests on the use of digital keys (see Pesonen and Horster 2012: 14-15). The hotel provided 30 guests with an NFC-enabled phone they could use to make reservations and to receive a digital key. The results of the survey that was part of the trial showed that: participants appreciated not having to check in and out, they all saved time, almost all participants would use digital keys again if NFC compatible phones where available, and a majority of the guests also declared that "the service made their hotel stay more pleasant" (Brown 2011).

Before starting their new digital services, also the Hilton chain took a survey to become aware of the wishes of their clients. They found that 84% were in favour of choosing their room and two out of three wanted more control over the room where they stayed[63].

Next to this first experiences reported by users, others who have tried out

[63] http://www.hoteliermiddleeast.com/22579-exclusive-qa-hilton-worldwides-digital-check-in/1/print/, Accessed 07 June 2015.

similar digital keys, found it superior to the magnetic swipe cards because:

> "First, it's much harder to lose a smartphone. Second, your smartphone can't be demagnetized by other things in your pocket. Third, you can skip check-in completely and go straight to your hotel room – and you can skip the check-out, too" (Anthony 2014).

Neuhofer et al. (2015: 7) developed an overview of some of the key experiences guests may have during their stay at the hotel and mapped the way in which smart technologies change those experiences. First, in the old situation –without smart technology- the settings for the *comfort of the room* are uniform. With smart technologies they can be personalized based on preferences known prior to the arrival of the guest, the settings can be dynamically updated during the stay, and also employees can update their observations through smart technologies.

Second, without smart technologies the *welcome moment* at the desk is standardized and rather impersonal and general. With smart technologies this interaction can become personalized because the staff members not only already know the name of the guest and his or her preferences, the guest also already knows the staff members because they were already presented to the guest through the application.

Third, without smart technologies, visits to the restaurant or other services also remain standardized. With smart technologies involved, the greeting and welcome can become personalized as the staff can already know the preferences of the guest. Updates of preferences can dynamically take place. All in all, the introduction of the digital key and the extra digital services provided by the app has to lead to a personalized and comfortable stay for the guest.

Notwithstanding all these positive expectations concerning the digital key and the supporting smartphone application, from a user's perspective some uncertainties also arise concerning the use of these digital hotel services. Ambrosino (2014) for example wonders what the impact of the keyless key may be on the guest's interactions with hotel staff. Where Neuhofer et al. (2015) chiefly focus on the personalized experience made possible by smart technology, Ambrosino (2014) wonders if digital keys not just lead to more impersonal hotel experiences. If guests skip the front desk and can order

everything in the app, will there still be face-to-face interaction? For that reason, more traditional hotels such as the Ritz in London hang on to the metal key, in order to preserve the personal interaction with hotel guests. If the digital key will stimulate or hinder the interaction with hotel staff remains to be seen, but that the interaction will change is a given.

Another question concerns the security of the application. As we will see when we look at the construction layer, the key and lock system is generally being judged as the most secure option. However, this judgement is primarily based on possible threads of malicious intruders or hackers coming from outside and does not so much take into account the interests of the 'insiders' or the curators (the hotel owner) and the way data is collected, stored and used through the use of the app.

In addition, this focus on threats coming from the outside is also rather blind to the ways in which hotel guests themselves may misuse the application. Whereas in the situation of the old-fashioned metal key problems arrived because of the negligence of the guests, now problems may arise because guests create easy-to-break passwords, not being fully aware of the consequences it has when their phone suddenly becomes more than just their phone, but a way to gain access to their personal domain. The guest's phone becomes more valuable and therefore also more attractive to steal or hack.

6.3.2 Construction

Where the guests interact with the interface of the app, all the technical processing is conveniently tucked away behind the sleek and intuitive design of the interface. As a result, the hotel guests are not directly confronted with the technical workings of the app or with the values that by means of the technology are embedded in the app. This makes their assessment of the digital key often superficial, as the most dominant options become: does or does it not work? Or, do or do I not use it?

To take into account the technology of the digital key, we have to look –in Latour's terms- at the network of nonhumans and humans that make it function. Consequently, it is not enough to merely look at the smartphone itself. We should also at least include the *locks* with which the smartphone interacts and to the *managing system* connecting with the locks. As a point of reference, I will look at the lock systems of supplier Assa Abloy, which Sornes' company VingCard Elsafe is part

of, as it not only is one of the global leaders in lock systems, but also is the supplier of the Starwood Hotels and Resorts chain, one of the first hotel companies to introduce the digital key.

However, it has to be noted that the network is much broader than merely these three components. Also the cable –often forgotten when packing your bag- to connect the smart phone to the electricity grid is part of this network and the grid itself is a necessary feature for the digital key to function. Unfortunately, it falls out of the scope of this chapter to include them all in the analysis.

NFC

The locks of ASSA ABLOY that connect with the smartphone work with RFID – as they first had to interact with RFID keycards- but now they are being upgraded to become compatible with NFC (Near Field Communication) and BLE (Bluetooth Low Energy) technology.

NFC is a:

> "short range and wireless technology for data transfer without physical touch"[...] "NFC is an open standard so it can be integrated into many electronic devices. On the consumer's side the primary NFC device is a mobile phone or a tablet computer. In combination with NFC, the device will act as a smart-key to gain access to services from any other NFC device or tag" (Pesonen and Horster 2012: 11).

Madlmayr and Sharinger (2010: cited in Pesonen & Horster 2012:12) make a comparison between different wireless technologies and find that compared to Bluetooth and WiFi, NFC is superior because of its fast and automated connection. They remark that NFC could also be used to set up a Bluetooth or WiFi connection. NFC "originates in Radio Frequency Identification (RFID) Technology" but whereas with RFID "the focus is on identification, NFC is based on interaction" (Pesonen and Horster 2012: 12). Pesonen and Horster (2012:12) also list several advantages of NFC. They state that, amongst others, the technology is "compatible with existing RFID structures, tags and contactless smart cards", that the "short transmission range provides inherent security", and that "it is easy to use as users do not need to know anything about technology". The NFC chip is increasingly being integrated in mobile devices such as in the iPhone 6, iPad mini 3, and the Apple Watch. This latter gadget

even has its own integrated hotel app, which apparently is compatible with the lock system of Starwood's hotel chain (Boden 2015).

BLE

BLE –also promoted as *Bluetooth Smart*- is a low-power technology developed for short-range control and monitoring applications (Gomez et al. 2012: 11734). Bluetooth in general allows devices to communicate with each other over radio links. It is a global standard and is incorporated in almost all mobile phones, tablets, and laptops. Typically, Bluetooth enables communication over 100 metres, but by adapting the power rates, this distance can decrease to ensure the "appropriate combination of power consumption and distance" for the application that the device is intended for (Gupta 2013: 20).

The Low Energy variant is the latest enhancement and is now part of Bluetooth 4.0 specifications. One of the biggest advantages of this technology is that because of its low power feature, devices compatible with this standard are expected to function on very low power rates. Consequently, these devices will be able to

> "operate for months or even years on coin cell or smaller batteries without the need for recharging or replacing batteries. This is very useful in applications (like hotel locks! EK) where it may be difficult to recharge frequently and longer battery life is important" (Gupta 2013: 6).

Other advantages of BLE that are listed are: its small size, low cost, short range, faster connections, and security. Moreover, BLE can be built onto the existing Bluetooth infrastructure, making it easy to adopt (Gupta 2013: 7-8).

One of the most promising BLE-enabled applications is the so-called Beacon. Apple, as a frontrunner in this domain, even developed its own iBeacon. These often-small devices send out a unique identifier to a compatible app or device in the vicinity, after which a certain action can be triggered or a push notification can be sent. Apple, for instance, uses iBeacons in its stores to provide users with: extra product information tailored to the products the customers are looking at in a specific part of the store, special offers, and the opportunity to pay for the products through their phone, skipping the line in front of the checkout. In a similar fashion, iBeacons can also be used on hotel premises.

Security of BLE and NFC

Both BLE and NFC are deemed to be secure, with NFC being the most secure because of its proximity requirements. More than NFC, BLE runs the risk to interfere with other transmissions and it is also more vulnerable to DDOS (Distributed Denial Of Service) attacks. Also, researchers at Context Information Security have shown that it is rather easy to monitor and record data sent by BLE empowered devices; they even developed an Android app to demonstrate this (Lester 2015). The researchers found that although BLE devices have a random MAC address –which network protocols need to identify devices-, these MAC addresses seldom change, making them in fact a unique identifier. Although it is possible to "implement public key encryption and keep packet sizes down, while also supporting different authentication schemes"[64], the researchers found that:

> "Many BLE devices simply can't support authentication and many of the products we have looked at don't implement encryption, as this would significantly reduce battery life and increase the complexity of the application"[65].

For the hotel business –as we will see when discussing the involvement of the curators- the longer range of BLE is nevertheless very attractive as BLE beacons integrated in the hotel environment and may be used to enhance and personalize the stay of the hotel guest.

It has to be noted that NFC also has its vulnerabilities. It is for example possible to replace a NFC tag with malicious content, seducing the unaware user to download malware on its device. Or, an 'accidental' bump against a virus-infected NFC enabled device can threaten the integrity of your device and the information stored on it.

Managing systems

Functionality is also of uttermost importance when looking at the property management systems (PMS) of hotels. Increasingly, all necessary hotel operations –

[64] http://www.net-security.org/secworld.php?id=18422, accessed 10 June 2015.

[65] http://www.net-security.org/secworld.php?id=18422, accessed 10 June 2015.

from checking in guests, maintenance, to security - are brought together in one managing system. Also Assa Abloy offers a modular system, giving hotels the possibility to build a system tailored to their needs and wishes. Assa Abloy offers an online and an offline management version. The offline version is the basic software program with options such as: easy check-in, access management, payment with the keycard.

The online option gives the hotel owner the opportunity to extend the offline system by different modules such as a *Security Operations module* to detect wandering intruders (for example, when someone uses one card to try to open different doors, this card automatically is cancelled), immediately block access to certain areas, and track users (by seeing the user's last registered locations). Also, it is possible to add a module called *Frontdesk Operations,* which includes the automatically activation of keycards upon check-in and in advance sending to the guests an SMS or e-mail with their room number[66].

While it is true that the data mostly stays in the property management system (Mitchell 2006), the smartphone –or RFID-enabled keycard- itself can function as a source of new data as it may transmit the data on the activity of the guest in the hotel to the management system.

Security of management systems

It is rather difficult to assess the security of the management systems as the hotels as well as the system suppliers do their uttermost best to protect the technical specifics of their proprietary technologies (Manley 2015). Especially in the early phase of development and implementation, hotel businesses want to obtain and keep an advantage over their competitors. This leads to superlative, but rather trivial language, such as:

> "Our commitment and strength is to offer the highest reliable security for both your hotel and your guests. Through experience, knowledge and

[66] http://www.assaabloyhospitality.com/en/aah/com/press-room/product-documentation/, p19-26, Accessed 09 June 2015.

> modern technology, VingCard has continued to deliver just that in over 30 years"[67].

Harry Sverdlove (cited in White 2014), chief technology officer at the cyber security firm Bit9, comments that when it comes to security in the hospitality industry, the biggest challenge is that "convenience trumps security".

> "While encryption –which hotels say they are using for mobile keys- certainly helps, the more difficult a digital system is to access and make changes to, in general, the harder it is to breach. [...] A system flexible enough to accommodate requests for physical keys, multiple guests per room and other considerations is a priority for hotels concerned about the user's experience, but these concessions can make the system more vulnerable" (White 2014).

In general, hotel owners do claim that "the locks and mobile keys are designed to be equally secure as traditional room keys" and that they "prioritize guest and property safety above all else" (Manley 2015). All in all, it seems that when it comes to security, hotel guests simply have to trust the hotels and their system suppliers to have invested in appropriate technical security measures.

6.3.3 Curation

The most important curator of the digital hotel key is the hotel owner[68]. To understand hotel companies' reasons for developing and implementing digital keys, one has to take into account the more encompassing trends in the hospitality sector.

---

[67] http://www.assaabloyhospitality.com/en/aah/com/press-room/product-documentation/, p6, Accessed 09 June 2015.

[68] It has to be noted that there of course are also other important actors who participate in curating the digital key. Suppliers, designers, and employers also have a role in the way in which the digital key mediates the interaction. Unfortunately it falls out of the scope of this chapter to give an exhaustive description –if this would ever be possible- of all curators involved. In further research, however, the range of curators could be enlarged to describe in more detail the actors involved in curating the digital key and the influence this has on trust.

These trends are, not surprisingly, linked to dominant, technological developments in society. As we have seen in the previous chapter, in our current networked era innovations in ICT enable more personalized and pro-active services. Providing customers with digital keys to give them more control over their stay fits neatly with this trend.

A necessary condition to pro-actively cater to the personal needs of customers is the collection and analysis of large quantities of –personal- data. The hotel business, therefore, explores the potential of monitoring technologies, "not only to optimize existing processes but facilitate the creation of more meaningful and personalized services and experiences" (Neuhofer et al. 2015: 1).

By collecting all sorts of data on their customers, hotels want to offer them a personalized and comfortable stay. Moreover, hotels want to seamlessly fit in the way in which their guests go about their businesses, which increasingly is by making use of mobile devices. Based on their own research and the pilots that were conducted, hotels conclude that "guests were thrilled to be part of the next-gen way of hoteling and eager to use the technologies at more properties" (Manley 2015). Chris Holdren (cited in: Manley 2015), senior VP of global and digital at Starwood Preferred Guest & Digital explains:

> "Our tech-savvy guests manage most aspects of their life and travel from their smartphone, and many no longer want to keep track of or fumble with keycards each time they enter their room. Because of this, we are constantly working ahead of the curve to implement the latest technologies and all of our brands are constant working laboratories for the latest innovations".

The gathering of data is already part of the hotel processes for quite some time. Through customer relationship management services (CRS), provided by companies such as Libra OnDemand, hotels collect information about guests. This information is not limited to what happens within the hotel –whether or not guests make use of room service, the restaurant, special offers, if there are incidents, …- but may also include information retrieved online (Lindberg 2013).

With the arrival of the hotel app and -as a new part of that- the digital key, also *geo-location* information can now be added to the CRS database. Moreover, by also making use of iBeacons, guests -who have downloaded the designated app- can be

tracked and followed when they move around in the hotel. This information can not only be used to monitor hotel operations –for example if there is a queue in the restaurant- but iBeacons can also be programmed to send push notifications with special, targeted offers to the guest's smart phone, when she is, for example, nearby the pool or cafe.

From a user's perspective, digital tools are designed to give guests the desired "choice and control" over their stay[69]. From the curator's perspective, digital tools help to create personalized and pro-active services, which in the end must lead to more revenue for the hotels. Or, as Hilton Worldwide's global head of digital services Geraldine Calpin puts it:

> "Everything we do is designed to better serve our guests so they are more loyal to our brands –including digital tools- thereby driving business and generating revenue for owners. We expect a high return on investment from the digital tools driven by increased brand loyalty and incremental revenue from push notifications, upsell opportunities and pre-arrival requests"[70].

All in all, the implementation of the digital key should not only ensure that the customer's experience of staying in a hotel is up to date and in line with her expectations of how a hotel in the networked era functions. The digital key is also a new tool in the hands of hotels to better get to know their customers by collecting a wide range of information –including geo-location data- of their customers, which then can be used to sustain customers' loyalty to the brand and, in the end, make more money.

---

[69] http://www.hoteliermiddleeast.com/22579-exclusive-qa-hilton-worldwides-digital-check-in/1/print/, Accessed 07 June 2015.

[70] http://www.hoteliermiddleeast.com/22579-exclusive-qa-hilton-worldwides-digital-check-in/1/print/, Accessed 07 June 2015.

6.3.4 Codification

Finally, as the last part of analysing the several conceptual layers of the digital key, we will look into the privacy policies of some of the digital key pioneers in the hotel business to see in which ways they issue rules and regulations about the digital key and the collection of data. While it is true that these privacy policies have to comply with the applicable legal requirements, hotel companies might well look for the borders of what is legally acceptable. Moreover, it is not always clear-cut what is allowed and what is not. Thus, as long as companies don't get reprimanded, their interpretation of what is legally acceptable, which can be found in their issued policies, is determining.

Hotel Corporation Hilton[71] distinguishes between *personal information* and *other information*. The former is information that directly refers to a person, where the latter does not personally identify a customer[72]. This *other information* may also include, amongst others, data collected online through cookies; Hilton does not respond to "do not track" and other blocking technologies. The privacy policy of the Hilton Hotel states that personal information is being used to provide services requested by the customer and pro-active services initiated by the hotel (for example, promotions and prize draws). The other information "may be disclosed for any purpose"[73]. The Hilton hotel may also combine personal information and other information, which will then be treated as personal information.

Information on the use of the digital key, which is being collected by the hotel as well, is also categorized as 'other information'. When the app makes use of GPS, Hilton will make use of this information to locate a hotel nearby and/or to provide customers with "relevant location-based information"[74]. Hilton states that it will abide by the settings of the device when accessing these data. When they collect

[71] http://hhonors3.hilton.com/en/promotions/privacy-policy/english.html, Accessed 25 June 2015.

[72] Hilton Hotel also refers to sensitive information (health data, racial data, ethnic origin, political opinions,...). They claim not to collect this kind of information unless the customer volunteers it. Health data, for example, when voluntarily provided, may then be used to provide better services to the hotel guests.

[73] http://hhonors3.hilton.com/en/promotions/privacy-policy/english.html, accessed 25 June 2015.

[74] Idem.

location-based information, they may share it with third parties.

As the digital key can be part of a loyalty programme of the hotel brand, it may be necessary for guests to become a member of such a programme first, in order to receive a digital key. Consequently, extra information will most likely be shared, like an online profile, preferred airline partners, language preferences, and room type preferences.

Starwood, another frontrunner in the digital key domain-, lists similar activities in its privacy statement. In addition, Starwood explicitly states that they also collect information from social media platforms such as Foursquare and Facebook, as guests can connect to these platforms with their hotel-issued app. Because a company such as Facebook makes it possible for third parties like Starwood to also collect the list of Facebook friends of the hotel guests, even information about people who are not frequenting the hotel can be kept in the hotel's database.

Marriot[75], another hotel branch issuing digital keys, has included a separate paragraph in its policy about the use of beacon technology as well as a paragraph on liability concerning the use of the digital key. Concerning the former, Marriot states that when customers *opt-in* through their app –by giving consent to the sharing of information-, the hotel will collect information about them and send special offers through Bluetooth. Marriot will continue to do so, until the customer has logged out. If the app is running in the background, it will still gather information.

Concerning the latter, the statement declares that if the use or misuse of a digital key leads to the hotel guest experiencing any kind of loss, the hotel is not liable. All members of the Marriot group and connected third partners:

> "expressly exclude any liability for any direct, indirect or consequential loss or damage incurred by any user in connection with his/her use, or inability to use, a digital key, including, without limitation any liability for loss of income or revenue; loss of business; loss of profits or contracts; loss of anticipated savings; loss of data; loss of goodwill; and for any other loss or damage of any kind, however arising and whether caused by tort

[75] http://www.marriott.com/about/digital-entry-terms-of-use.mi, accessed 25 June 2015.

(including negligence), breach of contract or otherwise, even if foreseeable"[76].

In other words, whereas in 1976 the court in *Garzilli vs. Howard Johnson's Mother Lodges Inc* explicitly pointed at the hotel's responsibility of reasonable care for the safety and integrity of the room and instigated the introduction of the keycard as a safer medium than the traditional key, with the introduction of the digital key, hotel owners seemingly aim at putting this responsibility partly back into the hands of the hotel guest. Hopefully we do not need a new Connie Francis to see if this kind of exoneration holds.

As we already established in the previous paragraph, hotels always have been collecting information about their guests. The introduction of the digital key is, rather than the base line, the icing on the cake when it comes to the collecting of data. Therefore, in order to get an idea of the true range of personal information gathering, one also has to take into account the other information strategies already put in place by the hotel.

Taking into account the privacy policy of the Hilton hotel, the conclusion can be short: "at every touch point or guest interaction" personal information may be gathered[77]. This information includes –amongst others- contact information, personal characteristics, nationality, income, passport number and data and place of issue, travel history, etc. It may also include the collection and keeping of information and records "related to conversations, including recording or monitoring customer service calls"[78].

Personal information may also be obtained from third parties such as from airline and credit card partners, as well as information derived from social media sites. It may also be shared with affiliates, franchisees or business partners of the hotel. And the information will be retained "for the period necessary to fulfil the purposes outlined in this Statement (the privacy policy, EK), unless a longer retention

[76] http://www.marriott.com/about/digital-entry-terms-of-use.mi, accessed 25 June 2015.

[77] http://hhonors3.hilton.com/en/promotions/privacy-policy/english.html, accessed 25 June 2015.

[78] Idem.

period is required or permitted by applicable law"[79].

All in all, the introduction of the digital key can be seen as a new strategy in a longer tradition of hotels to gather information on their customers. With the arrival of the digital key not only is a new tool being added to the hotel's arsenal of monitoring tools, also a new sort of information is collected: geo-location data. Moreover, the digital character of the collected information enables the hotel not only to collect, but also combine, analyse and share the data, bringing forth questions concerning privacy and accountability.

6.3.5 Revised and repeated: Trust between the hotel owner and the hotel guest

With the arrival of the RFID card, and especially with the introduction of the smartphone as a digital key, the trust interaction between the hotel owner and the hotel guests has changed again. Whereas with the introduction of the bulky *keychain*, trust became distributed between the guest + key + chain, the use of the *keycard* decreased the vulnerability of the hotel owner, making trust less needed as a way to deal with the uncertain behaviour of hotel guests. With the arrival of the digital key, trust becomes important once again, as the complexity within the interaction rises, through the collection of data and the pro-active services based on these data.

However, where in the previous situations it first and foremost was the *hotel owner* who had to bear the uncertainty of not knowing for sure if customers would return the hotel key, now the vulnerability increasingly lies with the *hotel guest*. Where the hotel through the collection of information comes to know its guests a lot better (making the future more predictable and therefore less complex), what exactly happens to and with these data generally lies beyond the knowledge and influence of the hotel guest. She has to trust the hotel owner to make use of this information in a trustworthy and secure manner.

A shift in the trust relation has therefore occurred. Whereas in the previous settings the hotel owner was the principal *trustor* and the hotel guest was the *trustee*, now the *hotel guest* is the trustor and the *hotel owner* the trustee. The introduction of the digital key is not a neutral switch of instruments. It not only pre-sorts the

[79] Idem.

interaction of the hotel owner and the hotel guest. It also, because of the networked ontology of the digital key, which I analysed by making use of the 4 Cs, poses new challenges for trust.

On the context level, we see that hotel guests are provided with an application that enables them to tailor their stay to their own preferences. They do not have to physically check in at the hotel desk and they can choose themselves which specific room they want to occupy. Moreover, the interfaces of the hotel apps are generally designed in such a way that they are easy to use and self-explaining. Ostensibly, this leaves hotel guests with more control and therefore less vulnerability or complexity to resolve. By circumventing the hotel owner –or more likely, her staff- at the front desk, it seems as if the interaction between hotel owner and hotel guest –and the vulnerability attached to this interaction- no longer takes place; it all is dissolved into a digital piece of transferable code. It now merely revolves around a hotel guest opening the door of her temporary private domain with her phone. The hotel owner has, so to speak, left the building, and the digital key has seemingly enabled a more direct interaction for the hotel guest. As a result, trust shifts from the interpersonal level to the artefact, to the system itself. As the interaction between hotel owner and hotel guest has vanished, trust is not to be found at the front desk of the hotel but in the interaction of the hotel guest with the digital key.

Can the digital key be trusted? As we have seen in the previous chapter, this question is often translated in a rather superficial dual presentation: does it or does it not function properly, this digital key? Because of this dual perception of the digital key, the mediating workings of it are barely taking into account by the hotel guests.

When we look at the *construction* of the digital key, which is dominantly based on NFC and BLE, it becomes clear that through these techniques the functionality of the digital key is not limited to the opening and closing of the hotel door. The digital key downloaded on the smart phone of the hotel guest does not only enable a new association with the guest, but also enables a new association with the *curators*, the hotel owners. They make use of the digital key and the app in which this digital key resides, to monitor guests and collect data on their behaviour. Taking into account some of the interviews conducted with these curators, their main reason for introducing the digital key is to make more money by gathering data in order to provide customers with pro-active and personalized services. This may partly be in

line with the interests of the hotel guests. Based on these collected data, hotel guests can receive the special offers and personalized services they desire. It may however also conflict with their interest when the collection of data is used for alternating purposes that reach far beyond their stay at the hotel.

The quick review of the privacy policies (*codification*) of some of the hotel chains that are introducing the digital key, shows that not only can almost all data be collected and stored, it often remains rather vague as to with which parties' information may be shared and how data is being combined and mined. A similar uncertainty exists when it comes to the security of the app and the digital key itself (*construction*). Due to 'proprietary reasons' suppliers and hotels remain silent about the security measures they installed to render the digital key as safe as possible.

6.4 Conclusion

Keys are not merely instruments to open a door or lock a room. With every innovation of the hotel key the central actors, the hotel owner and hotel guest, change. With every innovation, the hotel key mediates the interaction of hotel owner and hotel guest in a different manner. Because the interaction changes, the way in which trust is being shaped changes as well. Where the bulky key chain to a certain extent reduces the uncertainty hotel owners have to deal with, they still have to trust their guests to return their key. Vulnerability remains. With the introduction of the keycard, this vulnerability by and large dissolves. Because it becomes easy to replace the keycard, trust in the hotel guest to return it becomes less urgent.

The introduction of the digital key again transforms the interaction of hotel owner and hotel guest on a fundamental level. It relocates vulnerability, making the hotel guest the trustor and the hotel owner the trustee. Just as it was the case with the keycard, the hotel owner is not depending on the hotel guest to return the digital key, which is located and remains on the smart phone of the hotel guest. However, although the digital key resides on the smart phone of the hotel guest, it is not solely her key. For the hotel owner as well, the digital key is opening doors, not to the hotel room per se, but it gives access to the personal information of the hotel guest, to her whereabouts and preferences, including her geo-location data. As a result, the hotel guest becomes more vulnerable, because she is being exposed to the monitoring workings of the digital key.

By taking into account the 4 Cs and looking beyond the mere experience of the hotel guest (context), this extra functionality of the digital key becomes visible and the new vulnerabilities within the relation of hotel owner and hotel guest can be mapped. As a result, it is also possible to foresee the challenges the digital key imposes on the trust relation of hotel owners and hotel guests.

First and foremost, to speak of trust, actors have *to be aware* of the fact that they are in a situation where something is at stake. However, how can hotel guest move beyond merely assessing the functioning of the digital key to also appraising its *mediating qualities*? From a phenomenological perspective, there seems to be a gap between the contextual level, and the other Cs that deem to be crucial to understand the mediating workings of a networked artefact such as the digital key.

Although trust often is placed in a non-reflected way, some basic assumption on the kind of situation one enters has to be present for trust or confidence to be placed. Hotel guests would have to become aware of the mediating workings of the digital key, which are now conveniently hidden away behind the interface of the app (construction) and in the hotels' privacy policies (codification). If hotel guests are not aware –not even implicitly- of their vulnerability, of the complexity introduced by the digital key, they may feel misled or tricked into a situation when something goes wrong. This would be an unwanted side effect of the introduction of the digital key for both hotel owner and hotel guest.

Second, as we have seen, while technology may reduce complexity, it simultaneously also brings forth new complexity. This also applies to the digital key. Not only is there the secrecy about the security measures set in place for which hotel guests have to confide in the suppliers and hotel owners. It also remains to be seen how the actions of hotel guests themselves are pre-sorted by the digital key. Will they be more careful with their phone, now it has a new function? Will they, for example, alter their password and not leave it unattended when going for an extra round at the cold buffet?

All in all, because of the networked ontology of the digital key, more associations of actants are created –to return once more to the vocabulary of Latour- and more doors than just the hotel door are being opened. Consequently, this has an impact on the trust relations that are created, even beyond the awareness of some of the actors involved.

7

Interpersonal system trust in Airbnb.

Together with the introduction of the Internet came high expectations about the possibility of instigating new forms of governance, communities, and economic development[80]. Especially in the 1990s and the beginning of the new millennium, the Internet was characterized as a technology that could break down physical boundaries, make time differences irrelevant, and facilitate direct interaction without intermediaries.

Supporters of this Open Internet perspective - also referred to as the Open or Digital Commons - include engineers, hacker groups, p2p communities, online entrepreneurs, and all kinds of political activists[81]. Obviously, this is not a homogeneous group of users, but what they nonetheless have in common is their belief in *self-regulation* (and as a consequence their dislike of governmental regulation), *in bottom-up participation,* and *problem solving*. The open internet adherers are convinced that the Internet can open up a space for people to experiment with their identity, giving them the opportunity to become who they want

[80] This chapter is partly based on, and includes sentences from Keymolen 2013.

[81] These Open Internet adherers are not necessarily academics or people who are interested in publicizing or engaging in academic debate. Nonetheless their activities on and visions of the Internet are of great importance because they co-shape the evolution of the online world. Therefore, to attend to their ideas and activities, one has to take into account non-academic sources such as blogs, online discussions, videos, etc.

to be, not constrained by the physical limits the offline world imposes on them. "All they see are your words" Turkle (1995: 184) writes. The disruptive power of the open Internet will enable people to organize themselves, cutting out the centralized, traditional powers of governments and large companies.

This utopian belief in the power of the open Internet has been widely contested by scholars from different disciplinary backgrounds (also see chapter 4). For example, Deibert et al. (2012a) state that already in 2000 this early optimistic phase of the Internet had to make way for a time where access to the Internet increasingly is being monitored, denied, and controlled by governmental actors. There also is a flow of reports, books, and articles on techno-regulation (Zittrain 2008; Lessig 2006; Goldsmith and Wu 2008; Wu 2011), and on the impact of censuring measures on human rights in cyberspace (Morozov 2011; Deibert 2008; Deibert et al. 2010; Deibert et al. 2012a), all, from different angles, contesting trust in a free and open Internet.

Notwithstanding these critical voices, the open Internet community remains strong. New concepts –although still with a reference to the 'old world'- such as *P2P community, online commons, collaborative consumption,* and *sharing economy*[82] signpost that the new, decentralized Internet-powered society finally has arrived.

One of the recent advocates of this movement is Rachel Botsman who, together with Roo Rogers, has written what has been referred to as 'the bible' of the shared economy: "*What's mine is yours. How collaborative consumption is changing the way we live*" (Botsman and Rogers 2010). Grounding their idea of *collaborative consumption* on the concept of an *Open Internet*, Botsman and Rogers claim that a new economy will arise built on the key values of "critical mass, idling capacity, belief in the commons and trust between strangers**"** (Botsman and Rogers 2010: xvi).

Amongst others, they focus on *Airbnb*, a platform for people who want to rent out their spare room or house and travellers who want to find accommodation, to show how people can *access* certain goods instead of *owning* them, not by depending

[82] Although all these names refer to slightly different parts of the open Internet movement, I have put them here together because they do all believe in the power of the Internet to leave traditional societal structures behind.

on large, centralized actors such as hotels, but by building on personal relations. Botsman and Rogers find in the Internet the possibility to overcome distances and bring people together to collaborate in an 'old-fashioned' way. Based on *interpersonal trust*, people will be able to collaborate on online platforms in a way that resembles their familiar, face-to-face interaction in small communities. They state that:

> "Online exchanges mimic the close ties once formed through face-to-face exchanges in villages, but on a much larger and unconfined scale. In other words, technology is reinventing old forms of trust" (Botsman and Rogers 2010: xiii).

Restoring interpersonal trust through the connective power of the Internet is key to the fundamental changes Botsman and Rogers foresee.

Without any doubt, trust between individuals is a necessary condition for a successful shared economy in general and even more specifically for a platform like Airbnb that brings together people from all over the world. While I therefore agree with Botsman and Rogers that trust online is essential for interpersonal interaction and that much of the complexity inherent in human interaction can be dealt with through the act of trust, the online context, however, is not a neutral environment merely facilitating interpersonal trust.

I will argue, by analysing Airbnb through the lens of the four Cs that Botsman and Rogers too narrowly focus on the context level. By not sufficiently taking into account the other Cs (curation, construction, and codification), they mistakenly believe that "technology is reinventing old forms of trust" where, in fact, a new form of trust –which I will call *interpersonal system trust*- is being established.

This is not merely an issue of deviating definitions. I will argue that by their narrow focus on the users of the platform, Botsman and Rogers broadcast a misleadingly utopian message, which eventually may backfire as it makes them blind to the challenges -and certainly also chances- the other Cs may evoke. Trust *between* the users of Airbnb but also of the users *in* Airbnb form the primary asset of its business model. Remaining blind for the way in which curation, construction, and codification play a role in the establishment of trust, may actually result in the *erosion of trust*. Therefore, a different, less utopian perspective on the influence of the online platform on trustworthy interactions has to be developed.

7.1 What is collaborative consumption?

The advocates of collaborative consumption[83] or the shared economy can be characterized as belonging to this mixed group of *netizens* that endorse the Open Internet perspective. They strongly believe that, through the Internet, interpersonal relations can be built which will support a new economic model based on *sharing*. Instead of *owning* a car, you *share* one, you no longer *buy* clothes but *swap* them, and instead of going to the bank to beg for a loan, you turn to *peer-to-peer* lending sites to look for individual investors. Where the 20th century was defined by *hyper-consumerism* based on *owning*; *collaborative consumption* or a *shared economy* based on *access* will characterize the 21-century.

As we have seen, collaborative consumption stands in the – in Internet terms - 'long tradition' of approaching the Internet as a technology to empower people. That it is a trend unlikely to fade away soon is supported by the fact that, besides Botsman and Rogers, a lot of other key-authors, such as Tapscott (2006, 2010), Chesbrough (2006), Benkler (2006), Bauwens (2012), and Rifkin (2014) write about similar developments. Moreover, there is growing attention for this phenomenon in international media. The Economist, for example, predicted one of the important trends in 2013 as the "ownerless economy expands" (Malnight and Keys 2012). And already in 2011, TIME magazine viewed collaborative consumption as one of the "10 ideas that will change the world" (Walsh 2011).

[83] With their 2010 book, Botsman and Rogers have allegedly coined the term 'collaborative consumption'. However, over the last few years, the term 'sharing economy' has increasingly become popular, more or less ousting 'collaborative consumption'. In an interview (http://magazine.ouishare.net/2014/03/communities-the-institutions-of-the-21st-century-an-interview-with-rachel-botsman-2/ accessed 15 December 2015) Botsman states that to her the 'sharing economy' is more specific than 'collaborative consumption'. As Airbnb fits the collaborative consumption paradigm and simultaneously is also part of the 'sharing economy', I will use both terms interchangeably when I analyze Airbnb.

7.1.1 Four principles of collaborative consumption

Botsman and Rogers (2010) identify four basic principles that lie at the heart of this new movement: critical mass, idling capacity, belief in the commons, and trust between strangers.

Critical mass stands for the required *momentum* to make a collaborative consumption initiative successful. For example, if I want to rent an electric saw, but I have to drive an hour to get one, this tempers my will to participate. An initiative needs enough participants – how many exactly depends on the kind of initiative - to make it attractive.

Idling capacity refers to the core assumption that there is a large offer of things and services, which by redistribution can be made useful elsewhere with the Internet as a distributor *par excellence*.

With *the belief in the commons,* Botsman and Rogers refer back to the well-known article of Garrett Hardin (1968) "The Tragedy of the Commons" which describes how people who self-govern a piece of land that no one owns, will eventually take too much, damaging all participants. However, the advocates of Collaborative Consumption assert the opposite. They claim that, especially on the Internet, it is possible to provide value to the community and at the same time enable social value to expand for oneself. A digital common can become a reality.

With *trust between strangers* we touch the central principle of collaborative consumption. On online peer-to-peer platforms, the traditional role of the middleman who enables third-party trust ceases to exist. Based on rating-systems and other reputation schemes, known from websites such as eBay, trust between strangers can be enabled.

7.1.2 The concept of trust in Collaborative Consumption

Although these four principles are all very important and lie at the heart of the movement, I will chiefly focus on trust, which probably is the most challenging one to accomplish (Brodwin 2012)[84].

[84] The biggest barrier to participate is the "… concern that a lent item would be lost/stolen (30 percent), followed by worries about trusting the network (23 percent) and privacy concerns (14

Trust as described by Botsman and Rogers (2010) seemingly has a *direct nature*. It is first and foremost something that happens between persons and therefore is in line with what in this book is called *interpersonal trust*. Botsman and Rogers refer back to times where interactions were based on strong ties of friendship, family relations, reciprocity, and reputation. Apparently, the interactions on collaborative consumption platforms resemble these by-gone interactions. For Botsman and Rogers, curators and technology only have a facilitating role. In the end it is up to the users of collaborative consumption platforms to build trust, consequently, making action possible.

With this conceptualization of interpersonal trust in a context of collaborative consumption, Botsman and Rogers deviate from the system trust described in chapter three. Chapter three discussed the arrival of large systems in late modernity, consequently making interpersonal trust relations no longer sufficient to deal with the complexity inherent in everyday life. People increasingly had to vest their trust in the systems and the experts who controlled the systems, in order to go about their business in society. The hotel business is a prime example of such a system. Where personal relations were no longer sufficient to find a place to stay while traveling abroad, hotels became the trusted parties to fill this void. With the arrival of Airbnb, this intermediary becomes redundant. Through the connective power of the Internet, people can again rely on interpersonal trust. In this sense, the trust as described by Botsman and Rogers supposedly reverts to the pre-modern, small community based idea of trust. The idea of cutting out the middleman goes hand in hand with abandoning trust in the system.

The interpersonal trust in the collaborative consumption context also differs from the concept of "face work" Giddens (1990) introduced to explain how at the entries of the system people employed by these systems are functioning as their 'human face'. These intermediaries of the system are regarded as crucial for the effectiveness of system trust. For example, flight attendants are the face of the air traffic system; our interpersonal interaction with the flight attendants enables us to trust the air traffic

---

percent)" (Bauwens et al. 2012: 135).

system. The trust we –hopefully- have in the flight crew is, however, not of a mere interpersonal character. It is not based on a shared history, friendship or family relation. Rather, there is a kind of *trust-loop* developed in the interaction between traveller and flight attendant. On the on hand, the flight attendant mediates as it were the interaction of the traveller with the air traffic system, instigating overall system trust. On the other hand, the trust a traveller has in the flight attendant is partly based on the fact that he or she works for a company that has selected and trained the attendant to become an expert in his or her work.

The interpersonal trust as described by Botsman and Rogers, therefore, is fundamentally different from this kind of *face work*. In the case of collaborative consumption, the 'flight attendants' are cut out of the interaction. Contrary to the air traffic case, on collaborative consumption platforms, there is no *face work* being done by system representatives. On the contrary, the curators of these platforms move as much as possible to the background in order to give room to the interpersonal interaction and trust building between users. From the perspective of Botsman and Rogers, the only *face work* being done in the collaborative consumption systems is by the *users* themselves. The curators provide tools to the users of the platform in order to enable "self-managed exchanges and contributions". These tools, such as a secure payment system, online personal profiles, and rating systems, which will be discussed extensively in the paragraphs on *context* and *construction*- should restore reputation mechanisms, which means that:

> "[w]e have returned to a time when if you do something wrong or embarrassing, the whole community will know. Free riders, vandals and abusers are easily weeded out, just as openness, trust and reciprocity are encouraged and rewarded" (Botsman and Rogers 2010: 92-93).

7.1.3 The concept of technology in collaborative consumption

The direct, interpersonal nature of trust in the collaborative consumption paradigm is underpinned by an *instrumental perspective on technology*. The Internet in general and the platforms of shared economy communities specifically are being approached as simple service-hatches, connecting users and facilitating their interactions.

While it is true that all technologies can be seen as 'tele-technologies' (Weibel

1992) -bridging the ontological gap human beings experience in themselves, amongst themselves, and in their relation towards the world- this bridge is always of a *temporary nature*. In the collaborative consumption paradigm, however, not much attention is being paid to the technology's own weight. Or, to rephrase it once more in Plessner's terms: they only take into account the *immediateness* of the interaction and not the *mediated* aspects of it.

This tacit presupposition translates itself in their analysis into a sole focus on the interaction of the users, the intentions of the users and how they put technology to work (the context level). The ways in which a specific, online environment is shaping the building of trust itself is not an object of analysis (construction). Because they under-conceptualize technology, they are also less aware of the *networked ontology* of the platform that enables an active role for the curators of the platform (curation) and the power struggle between different stakeholders over the control of the platform (codification).

7.1.4 Collaborative consumption: a utopian standpoint

All in all, this instrumental perspective on technology together with a rather incomplete perspective on interpersonal trust can be partly traced back to the *utopian* and rather misleading belief in technology as a means to *not only bridge but* also *overcome* the hiatus that defines human beings. Although human interactions are always simultaneously direct and indirect, human beings have the tendency to dismiss the aspect of indirectness and act as if their interactions are simply of a *direct* and stable nature. They try to set aside the triviality (*Nichtigkeit*) of their existence and flee to a *utopian world* –Plessner speaks of a *utopian standpoint*- in which they can find a final ground, a *definitivum* that provides them with a predictable environment. While Plessner describes how this desire to find a final ground leads human beings to the *domain of religion*, nowadays this domain has to move over in favour of the *domain of technology*. As de Mul (2001: 20) notes:

> "in the secular world,… the Internet functions as the 'holy grail'. It is a resource that promises us attributes which up until now belonged to God: omniscience, omnipresence, and omnipotence".

The utopian perspective of the collaborative consumption movement, trying to make

whole and direct what will always be partly broken and indirect, makes them blind to technology's own weight, the eccentric positionality of human beings and how these two shape trust. In the second part of this chapter, I will therefore look at Airbnb as a case of collaborative consumption to show that how the perspective on how trust 'works' in the shared economy may shift when one opts for a more layered perspective, taking into account the networking effects of the Internet technology on the building of trust.

7.2 Airbnb

A prime example of collaborative consumption and more specifically of the shared economy is the platform Airbnb. Airbnb started in 2008 with a couple of airbeds on the ground in the home of Airbnb founders Brian Chesky and Joe Gebbia. Running out of money, they thought that providing a place to sleep and breakfast to people, who weren't able to book a hotel because of a saturated hospitality market, would be a way to pay for their own house. They knew people were having problems finding a place to stay during the Industrial Design Conference, held in San Francisco. Therefore they listed a message on the conference website, advertising their spare room, breakfast, and good company. Different people responded and they were surprised it did not feel as if they had strangers visiting their home. An idea was born.

Although they initially encountered difficulties gathering funds for their idea, they were able to raise the necessary startup money, develop a business plan, and attract investors. In 2009 they changed airbedandbreakfast.com in Airbnb.com. It were no longer just airbeds listed on the site, but also whole houses, castles, boats, islands, etc.[85] In 2010, 210.000 users were registered on Airbnb.com. You could then find 28.000 properties in more than 157 countries, across 8.122 cities (Botsman and Rogers 2010: xi). Over the years, Airbnb grew explosively, now accounting for more than 35.000.000 registered users, 1.200.000 accommodations (of which 600 castles) across 34.000 cities in 190 countries[86]. For every booking made on the platform,

[85] https://en.wikipedia.org/wiki/Airbnb, accessed: 17 July 2015.

[86] https://www.airbnb.nl/about/about-us, accessed: 17 July 2015.

Airbnb charges a service fee, ranging from 6% to 12% of the subtotal of the booking[87].

Making use of the four core principles of collaborative consumption as defined by Botsman and Rogers (2010), it becomes possible to understand why Airbnb is a prime example of collaborative consumption.

Critical mass. Reading the short history above, it becomes clear that Airbnb has enough participants to fulfil the goal for what it was set up to do. This critical mass is important because Airbnb needs "a core group of lay and frequent users" to give body to their community (Botsman and Rogers 2010: 81). As reputation is key to the functioning of Airbnb, frequent encounters are needed to build up a profile that can function as a token of trustworthiness.

Botsman and Rogers (2010:75) chiefly focus on critical mass to describe "the existence of enough momentum in a system to make it become self-sustaining". However, they do not address the question of whether there is also *a maximum of participants* for a platform such as Airbnb to perform well. Taking into account the rapid uptake of Internet connectivity and the fact that there are no real barriers for people to sign up, no limits –besides for having a residence to rent or money to pay for it- are set. As we will see later on, growing numbers of participants also brings along risks for Airbnb: from attracting professional landlords and agencies to criminals and frauds.

Idling capacities. It is not enough to have enthusiastic participants. There also have to be enough houses and spare rooms to be rented out. If you wanted to make use of Airbnb in the starting days, you first had to check where there were hosts active, now you can pick your destiny first and then see where you want to be staying. Around the globe, you can now find a large range of properties. The founders of Airbnb saw the discrepancy between on the one hand a rather saturated and stagnated market of hotels and on the other a reservoir of dwellings around the world, which could now easily be distributed through the Internet.

Belief in the commons. Airbnb is built on a strong belief in the commons. The idea that by sharing you are not only adding value to the community but can also gain personally is at the core of Airbnb. The more people take part in Airbnb, "…the better

[87] https://www.airbnb.com/help/article/125?cref=127375e6d, accessed: 24 July 2015.

the system works for everyone- there is a 'network effect'" (Botsman and Rogers 2010: 91). So, even when you make use of Airbnb for selfish reasons (you want a cheap place to stay or some extra money in the bank), participating in Airbnb nevertheless creates value for the people involved.

Trust in strangers. Airbnb would not be able to exist if trust between strangers was unfeasible. By providing users with some tools on their platform –reviews, connection to social media account such as Google or Facebook, a safe payment method-, the most fundamental complexity of an Airbnb interaction is reduced. "Airbnb does not routinely perform background checks on its users"[88]. At the end of the day, it is up to the users to build trust.

Now that we know what the vision and premises of collaborative consumption are and how these relate to the sharing economy spurred by Airbnb, it is time to look at it from a different angle and analyse Airbnb through the conceptual lens of the four Cs. It may not come as a surprise that the *context* level of Airbnb extensively overlaps with the vision of Botsman and Rogers. As the interaction of users, the way they build up trust and belief in the commons is key to collaborative consumption, the context level fits neatly with Botsman and Rogers' perspective. The utopian belief that users can interact through Airbnb solely based on interpersonal trust is in fact the core of Airbnb's business model. The company goes to great lengths to promote this vision by presenting its users as a strong, connected, and trust-building community. However, things come to look different –and increasingly interesting- when one also takes into account the other Cs (curation, construction, and codification).

7.3 Context

The way in which a platform is designed pre-sorts the options users have in order to shape their online interactions. As a result, the platform strongly influences the way in which users are able to establish trust. Botsman and Rogers (2010) see it as the role of the curator –in this case, the company Airbnb- to create an environment in

[88] Although Airbnb does reserve the right to perform a background check nonetheless. https://www.airbnb.com/help/article/4?topic=357 Accessed on: 21 July 2015.

which trust online can thrive. It is then up to the users to take on these tools to develop an online reputation and 'materialize' their interactions, making them visible to the whole community.

Airbnb provides users with a myriad of options to gain trust. For example, they offer offline ID validation[89], users can log in to Airbnb making use of their social media accounts, Airbnb can validate the photos of the locations that are put online, etc. (also see Abramova et al. 2015: 2).

One of the most dominant tools provided to establish trust on the platform is the *online reviewing system*. Users can judge each others' reputation or trustworthiness by giving a review. Because these reviews are public to the Airbnb community, they should enable self-regulation, making it possible for users to make a better decision about who they want to rent out their place to and –from the other side- with who they want to stay. Research being done in the domain of e-commerce shows that reviews have a positive effect on the willingness of customers to interact with online vendors (McKnight et al. 2002b; also see McKnight and Chervany 2002; McKnight et al. 2002a).

The basic rationale behind this reviewing system is that people on the one hand want to safeguard their own reputation and therefore have an incentive to act trustworthily. On the other hand a good reputation brought forth by good behaviour in the past says something meaningful about the way in which a person will act in the future. Reviews are therefore often determining in accepting or denying a guest (see for example: Thomas 2014). Axelrod (1984) refers to this process as "the shadow of the future". If someone wants to establish a durable relation or wants to participate in a community for a longer period of time, it is necessary to act in a reliable way to convince people of his or her good intentions, consequently making interaction possible.

Botsman and Rogers (2010: 218-219) speak of "reputation capital". It is a currency that claims "you can trust me", and in Botsman and Rogers' view it is one of the pillars of the new, shared economy[90]. Completely in line with their belief in

[89] To verify the identity of users, Airbnb –amongst others- makes use of government-issued documents, users are asked to upload to the company.

[90] Airbnb is not the only platform and definitely not the first to have such a review system inserted

transparency within the community, Airbnb goes even a step further and offers users also the possibility to *comment on reviews*. Research indicates that in some instances a confession or apology of the host –for example, because the room was not clean- may have a positive impact on the trusting beliefs of potential guests confronted with negative reviews (Abramova et al. 2015). Studies suggest that because high ratings are judged to be central to the success of Airbnb, hosts may go to great lengths to receive excellent reviews. From rejecting guests whom they believe to be unsuitable, to starting all over again with a new property page freed from negative publicity (Zervas et al. 2014: 12).

Although it is generally accepted in e-commerce as well as in the domain of the shared economy that reputation systems are a valuable tool for users to assess the trustworthiness of their peers, recent research indicates that the impact of reviews on Airbnb has little to no effect on the behaviour of users looking for a place to stay (Ert et al. 2015). The researchers (Ert et al. 2015: 25)–to their surprise-

> "did not find evidence for the effect of online review scores on market prices. Further exploration of this result revealed that review scores had no effect on Airbnb prices because these scores were exceptionally high and thus lacked sufficient variance."

Based on an analysis of 600,000 properties listed on Airbnb, researchers found that nearly 95% received a 4.5 to 5 star rating (with 5 being the maximum). Virtually none of the analysed properties have a rating lower than 3.5 stars (Zervas et al. 2015). Rightfully so, Thomas (2014) doubts if the world could be as perfect as the Airbnb rating system wants users to believe. Although it is difficult to determine exactly why the ratings are so high, Thomas (2014: 23) argues that next to the hesitance people may feel to criticize people in public, in the shared economy -as promoted by Airbnb- there is a common feeling of belonging and positivity which people do not likely want to interfere with by giving a bad review. Ert et al. (2015) assume that the reciprocity

---

onto its platform. Mother of all online rating systems is that of eBay. This online second-hand marketplace already introduced its peer-to-peer monitoring scheme in 1996. Because of its rating system, traders can build up a reputation of a trustworthy buyer or seller, enabling new interactions.

of the rating system causes the high ratings. Mutual feedback between guest and host may result in retaliation (idem: 26), making it less likely for users to give a negative report. Moreover, because of the personal interaction that may occur within the Airbnb experience, participants are also less willing to provide negative feedback (idem: 26). All in all, the lack of diversity in the rating of the listings on Airbnb, makes the rating less valuable for users to assert the trustworthiness of the host and make a decision on where to stay.

However, Ert et al. (2015) discern another trust tool that does seem to have an effect on the behaviour of guests on Airbnb: *the picture of the host*. Their results suggest that guests by looking at the picture of the host assess his or her trustworthiness. This *visual-based trust* is established unconsciously as only a minority (8%) explicitly mention the picture of the host as a factor of influence (Ert et al. 2015: 28). Hosts who are aware of this trust factor can based on this knowledge choose a trustworthier picture by for example uploading a picture presenting them smiling and looking straight into the camera. The findings also suggest that women are found to be more trustworthy, consequently, it may be beneficial for a heterosexual couple to put the picture of the woman online.

While the posting of the host's picture may boost trustworthiness, it may also have unwanted side effects. A recent study of Edelman and Luca (2014: 2) shows that after controlling for other factors, "non-black hosts charge approximately 12% more than black hosts for the equivalent rental". Posting a profile picture, ostensibly a neutral tool to enhance trust, in fact brings about discrimination[91]. This conflicts with the core principles of the open Internet movement, which builds strongly on the liberating and emancipating force of the Internet.

Finally, another tool to enhance trust Airbnb has implemented on its platform is the possibility for prospective guests and hosts to communicate *before* a reservation is made. Although Airbnb cannot oblige people to contact each other, they do nudge this behaviour, for example by keeping track of the host's response rate and speed. The latter, amongst others, influences the place of the host in the search

[91] It would be interesting to investigate if Airflow, Airbnb's own pricing algorithm which also advises hosts about the pricing of their listing, takes into account the ethnicity displayed on the profile picture. If this would be the case, the discrimination would in fact be partly instigated by the algorithm.

ranking and may therefore impact the likelihood of receiving bookings.

If all these trust tools do not lead to a trustworthy interaction, users have the possibility to *flag* other users. On any moment in the interaction on Airbnb, users have the possibility to click on a flag when they believe something is suspicious or inappropriate[92]. Airbnb investigates each flag on a case-by-case basis. By delegating some of the policing of the platform to its users, Airbnb makes them into 'deputy sheriffs' (Torpey 2000; Lahav 2000). Although it is not their primary responsibility, users are incorporated into the system nonetheless. This strategy of 'responsabilisation' (Garland 2001) fits the collaborative consumption approach where users should take the lead and curators should facilitate and follow.

7.4 Construction

When it comes to the 'back office' or infrastructure of Airbnb, about the way in which the company builds, maintains, and develops their platform, not much information is officially been published. Or as Mike Curtis, Airbnb's vice president of engineering puts it: "Anything that is completely core and unique to our business… we'll keep that"[93]. Airbnb considers the search algorithm, which is central to the platform, to be their intellectual property. Airbnb releases bits and bytes of information about the search algorithms only in general terms, about the way in which the company mines data, and for which purposes these data are being used. Looking at general media coverage, their own policies, reports, and blog posts, there are, however, a few broad lines that can be discerned which shed some light on these fundamental, technical workings of the platform. More specifically, I focus on those technical aspects that are closely connected to the building of trust in the community.

The Airbnb platform can be accessed on the World Wide Web by making use of a computer and it can also be downloaded as an application on mobile devices such as tablets and smart phones. On all devices, their search algorithm is one of the central

[92] https://www.airbnb.com/help/article/4. Accessed: 28 July 2015.

[93] http://blogs.wsj.com/cio/2015/06/17/airbnb-open-sources-software-to-lure-talent-amid-insane-competition/. Accessed: 05 December 2015.

operations taking place. As this algorithm is the primary tool for prospective guests to find their host, it is key to the functioning of the platform. Airbnb describes its search algorithms as working:

> "with uncertain and incomplete information to routinely understand the specifics of local markets and geography in order to estimate the quality of the platform's inventory and answer users' queries with relevant results, while keeping in mind the hosts' preferences"[94].

On their website,[95] Airbnb explains that the underlying principle guiding their approach to design the search algorithm is that they: "want to reward hosts that deliver a great experience to guests". The three main categories that affect the search are: *the quality of the listing, the ease of booking,* and *guest preferences.*

The first refers to aspects such as the way in which the booking is presented (accurate description, professional photographs), if the pricing is competitive, the quality of the reviews, and to what extent the account is verified.

The second has to do with things such as the speed and consistency of the response of the host, if the host has an updated calendar in order for guests to know when the property is available, and if the host has ever cancelled a booking (which Airbnb judges to be a very negative factor).

The third refers to, amongst others, the relevance of the location, social connections (for example if the host and guest have mutual friends, Airbnb can detect that if users link their Facebook profile to their Airbnb account), and the personalization of results (the search rank may vary from query to query).

In addition, April 2015, Airbnb started to also include the preferences of the host to decide on the ranking of the search results. As Airbnb concludes: "personalization can be effective on the buyer as well as the seller side"[96]. Consequently, the top results do no longer include listings completely tailored to the guests' preferences, but room has also been made for the wishes of the hosts, for

---

[94] http://nerds.airbnb.com/search-airbnb/?_ga=1.247282116.1978383405.1437126877. Accessed: 22 July 2015.

[95] https://www.airbnb.com/help/article/39. Accessed: 22 July 2015.

[96] http://nerds.airbnb.com/host-preferences/, Accessed: 22 July 2015.

example concerning the duration of the stay and the amount of visitors. This led to a 3,75% increase of matches on the platform (also see: DeAmicis 2015).

Next to the ranking of their search results, Airbnb, by mining the data collected on their platform, also aims at personalizing[97] the overall experience users have on the platform. For example, based on the location of the user, earlier travels and searches, Airbnb alters the welcome page on their website, highlighting locations Airbnb believes might be of interest to this individual user.

All in all, even without all the specifics on the functioning of the search algorithm and their personalization practices, it becomes clear that these techniques are in fact the backbone of the platform and have a leading role in shaping *the familiar world of Airbnb*, necessary for trust to be established. By aiming at predicting where travellers want to go, by tailoring the search results to the preferences and wishes of users, these techniques reduce the complexity inherent in a global network, which Airbnb in fact is. Without some guidance, brought forth by the search engine and the personalization of the platform, the complexity would likely be too high, making it almost impossible for users to trust and bridge the uncertainty gap. Random search results carry the risk of alienating users, where personalized search results may have already neutralized some fundamental basic uncertainties, which trust on its own would not be able to cope with. Where Airbnb in its communication always emphasizes it is a "community" of like-minded people, these techniques bring this mantra into practice by creating an online world that fits the beliefs and expectations of its inhabitants.

Although Airbnb is rather secretive about their key algorithms, they nevertheless are also open sourcing some of their code[98]. *Airflow*, for example, is such an open-source project. It is a platform that can be used to structure and analyse data. In the context of Airbnb, it "segments the raw data from customers' mouse clicks, including listings they've viewed or reservations they've made, and structures them into summaries for Airbnb staff to analyze" (Boulton 2015).

[97] For a more detailed description of personalization, I like to refer the chapter 8.

[98] http://airbnb.io/. Accessed: 23 July 2015.

Another open source project is *Airpal*[99]. This is a web-based query tool that can be used to structure and analyse data. It has a very user-friendly interface that even enables employees who are not familiar with SQL to write queries and look for data in the database of Airbnb. One out of three Airbnb employees has run a query through this system.[100]

A third project is *Aerosolve*[101], a machine-learning library. This software Airbnb uses to help it better understand "the relationship between the price of an Airbnb listing in a given market and factor in demand for that listing"(Boulton 2015).

Indirectly, these open source projects give the outside world an idea of the techniques developed and used within Airbnb. These technical tools support Airbnb in their effort to create a familiar world for their users. But also on another level, these open source projects contribute to trust relations. One of the reasons for Airbnb to make these projects publicly available is to establish *a trustful relation with the tech community* on which it leans heavily. By giving back to this community, Airbnb not only wants to build a trustful relation, it also hopes to attract engineers to come and work for the company.

7.5 Curation

For Botsman and Rogers the actors who develop, run, and maintain platforms for collaborative consumption are mainly "curators" and "ambassadors", earning money by creating "the right tools and environment for familiarity and trust to be built" (idem: 92). In the end, it is *up to the users* to establish this trust.

However, when looking into the working and functioning of Airbnb, it becomes conspicuously clear that Airbnb does much more than merely facilitate connections between travellers and hosts. Over the years, Airbnb has increasingly taken on a more active and steering role when it comes to ensuring trust in the community.

A turning point was the widely publicised case of an Airbnb-host whose house

[99] http://airbnb.io/projects/airpal/. Accessed: 23 July 2015.

[100] http://airbnb.io/projects/airpal/. Accessed: 23 July 2015.

[101] http://nerds.airbnb.com/aerosolve/. Accessed: 10 December 2015.

was completely ransacked by travellers in 2011. EJ, the pseudonym of the host, blogged about her experience and accused Airbnb of letting her stand out in the cold[102]. After this incident, Airbnb immediately invested in a dedicated *Trust and Safety team*, by adding, amongst others, former intelligence officers and government investigators to their work force (Chesky 2011; also see Gannes 2013). Since then, this team monitors the interactions online, investigates suspicious interactions, and also functions as a mediator when users have a disagreement they cannot solve themselves.

Because all interaction takes place on the platform of Airbnb, the company has a pile of data at their disposal of which they can make use to police their community. For example,

> "[i]f a host uses the words Western Union in a conversation with a guest - a sign that they may be trying to route around Airbnb's system- the company will block the message. If a host and guest are repeatedly booking rooms with one another, it could be a scam to build fake positive reviews." (Tanz 2014: 19-20).

These and other analytic data analyses provide each user with a *trust score*. If this score turns out to be too low, Airbnb will further investigate the user. In addition, every user that has been "flagged" by another user will be looked into.

So, while it is true that on the context level, users have at their disposal a range of tools to interact and build trust, *these tools are not solely their tools*. By monitoring the way hosts and guest make use of these tools, Airbnb can steer and redirect these interactions. On the one hand this sort of monitoring, is a measure that adds to the trust building in the Airbnb community. For example, by detecting and subsequently kicking frauds, criminals, and other people with malicious intends off the platform, the risks for users are mitigated, which makes it easier to act on trust. On the other hand, more subtly, Airbnb *nudges* its travellers in behaving in a trustworthy manner. For example, in 2013 Airbnb employees personally called every group of eight people and more that booked a stay in New Orleans during the Super

[102] http://ejroundtheworld.blogspot.nl/2011/06/violated-travelers-lost-faith-difficult.html. Accessed: 24 July 2015.

Bowl to wish them a nice stay and remind them to treat the properties with care (Gannes 2013).

Airbnb's active measures to ensure the security of their platform –in order to boost trust- may, however, also induce the opposite reaction. When Airbnb introduced its *verified identity programme* and users were asked to upload government documents in order to clear their identity, this led to public outcry in the community[103] [104] (Banerjee 2014; Roudman 2013). People protective of their privacy or wary of identity theft did not want to comply. This incident illustrates the challenge Airbnb faces to strike a fair and acceptable *balance between security on the on hand and privacy on the other*. Both are important to develop a familiar world where trust can be established. However, when the 'costs' of security rise to a certain level, users may start to doubt if the prospective savings of booking through Airbnb are worth the risk of identity fraud or other privacy intrusions.

Airbnb does not just monitor the platform in order to be able to vouch for a trustworthy online environment; the company also uses this information to secure their own business model and to make sure they receive their fee on the booking. This, however, may sometimes conflict with the user's interests.

For example, by blocking the possibility of exchanging phone numbers or personal information until the moment of the actual booking, they do not only want to discourage scammers, but also make sure users stay on Airbnb and pay the fee. Although Airbnb through its monitoring activities on a general, system level does cover some of the possible security issues, they do not on an individual basis screen every user or perform a background check[105]. Blocking personal information deprives hosts and guests of the possibility of using this information to check for themselves if

[103] An overview of some of the online reaction can be found here: http://blogs.law.harvard.edu/doc/2013/05/28/lets-help-airbnb-rebuild-the-bridge-it-just-burned/. Accessed: 24 July 2015.

[104] Also see comments on the Airbnb blog itself: http://nerds.airbnb.com/verified-id/. Accessed: 24 July 2015.

[105] https://www.airbnb.com/help/article/4. Accessed: 24 July 2015.

they deem the person on the other side to be trustworthy[106]. At this point, there is a conflict of interest between Airbnb –ensuring their business model by blocking the possibility to share personal details- and the interests of the user –the missed possibility of checking for themselves if someone is trustworthy enough to take the leap-.

Next to pro-actively intervening to mitigate risks for users, Airbnb also monitors its platform to identify possible negative experiences. An online blog post of an Airbnb traveller for example uncovers that Airbnb customer service may contact guests after a host has completed a refund to inquire about their stay[107]. Airbnb also has a 24/7 hot line (telephone, chat, and e-mail) to assist their users if they encounter problems they cannot solve on their own. And finally, Airbnb has also set in place different kinds of insurances for both host and guest in order to refund their users when necessary. All these extra support measures must reassure users that Airbnb has a safety net when needed.

This kind of extensive after-care can be of importance to the trust users have vested in the platform. When Airbnb as a system can convince users that their bad experience is merely an unfortunate incident and not a system failure, the loss of trust might be restricted to the interpersonal level and not become a loss of trust on the system level (Keymolen et al. 2010: 58-61).

All in all, it becomes clear that Airbnb has a very active role in building and maintaining a familiar world for its users - a role which goes far beyond merely equipping their customers with trust-tools to sort things out on their own. Taking into account the global reach of the community and the risks involved, it is difficult to imagine Airbnb taking on a different position than it does now. Although in general, interpersonal trust can reduce much complexity and enable a wide range of

[106] Airbnb does offer a possibility to work around the prohibition of exchanging phone numbers before the actual booking has been made. If a guest submits a request to make a reservation, the host can –if the guest agrees- ask Airbnb to set up a phone connection between the prospective guest and host (5 of the Privacy Policy of Airbnb. https://www.airbnb.com/terms/privacy_policy. Accessed: 24 July 2015)

[107] http://matadornetwork.com/trips/drugged-and-terrified-an-airbnb-booking-gone-wrong/. Accessed: 24 July 2015.

interaction, some sort of top-down, structuring power is needed to help turning a complex global network of hosts and guests into a familiar world.

The proposition that the shared economy is about reinventing old forms of trust is therefore false. Rather, a new form of trust is being shaped where interpersonal interaction on the context level is strongly influenced by a pro-active environment (construction) steered by the curator Airbnb. I will refer to this kind of trust as: *interpersonal system trust*, emphasizing how both the interpersonal and the system level are intertwined.

However, policing the platform and steering the interaction of their users does not come without risk for Airbnb. Not only may users perceive some measures as being too intrusive, sometimes the interest of *Airbnb as a business* may conflict with the ability of *Airbnb as a community* to build trust.

7.6 Codification

The disruptive power of Airbnb, not only deranging the hospitality sector (Zervas et al. 2014; Guttentag 2013; Ikkala and Lampinen 2015), but the local communities in which it operates as well (Gottlieb 2013; Morozov 2014), has evoked legislative and societal upheaval, which incontrovertibly has had an impact on trust amongst its users. In this section, I will look into these recent, legal and governmental developments and more specifically, I will focus on the impact these changes have on trust in Airbnb as a platform. As we will see, the codification developed by Airbnb is not necessarily in line with the codification imposed by the government.

At the basis of Airbnb lies the idea that through the Internet, platforms can be created on which individuals connect and based on mutual trust interact, –if necessary- sorting things out on their own. In line with the beliefs of the Open Internet movement, Airbnb therefore not only wants to cut out large corporate intermediaries such as hotels, it also wants as little governmental intervention and external regulation as possible. This seemingly incompatibility of traditional governmental actors and pioneers in the shared economy can be explained in two ways.

Generally it is stated by actors in the shared economy such as Airbnb and Uber –another major player and disrupting power in the taxi-business- that current legislation is out-dated and not fitted to cope with the innovation brought forth in the

domain of collaborative consumption (also see: Guttentag 2013: 8-9). Legislation should therefore be reformed to cater the innovation in this sector. Another explanation, brought forth by Benjamin Edelman, associate professor at Harvard Business School, is that these companies deliberately "tend to skirt laws" (Edelman 2015) as it gives them an advantage over their competitors and in the end a significant larger market share.

All in all, with the large expansion of Airbnb over the recent years, not only the role of the company behind the platform became more visible and imperative – as I argued in the previous sections -, also different governmental actors stepped in to regulate the sharing economy of Airbnb. Where it has clearly been the strategy of Airbnb to "to root itself as deeply as possible before confronting its legal issues" (Guttentag 2013: 10), the company can no longer avoid the question if they comply – and if not how to ensure they do- with the legal principles of the different states in which they operate.

For one thing, the way in which Airbnb is set up, does not particularly fit the heavily regulated hospitality sector (McNamara 2015). The company

> "disclaims any liability for use of its services. Instead Airbnb encourages users to be aware of their particular locality's rules, zoning restrictions and tax regulations, before placing a home or apartment up for rent on Airbnb's site" (McNamara 2015: 152).

Because local laws may vary when it comes to for example hotel and tourist taxes and regulating short-term rentals, one has to look state-by-state, or city-by-city, to see how the rules may apply to Airbnb hosts. Focussing for example on the state of New York, one of the largest Airbnb markets, citizens are allowed under state law[108] to rent out their residence for less than 30 days only when they are present in the house themselves. Professional landlords on the other hand, who are not present, are not allowed to rent out their properties for less than a month. The aim of this law is to prevent residences from becoming hotels[109].

[108] For the specific law see: http://codes.lp.findlaw.com/nycode/MDW, accessed 27 July 2015.

[109] Also see: http://time.com/money/3513420/airbnb-new-york-attorney-general-says-airbnb-is-making-millions-on-illegal-listings/. Accessed: 27 July 2015.

There are also other legitimate reasons for cities to maintain such laws. Guttentag (2013: 9) listed four: first, cities want accommodations to abide to certain safety and health standards. Second, a continuous float of tourists may be a burden for a local community. Third, an abundance of short-term rentals may negatively impact the local housing market. Fourth, the promise of profits from short-term rentals may instigate immoral behaviour (for example, landlords evicting tenants to earn more money through short-rental activities).

As some rulings show, these local laws can have severe consequences for Airbnb hosts in the state of New York. In 2013, the Environmental Control Board decided that Nigel Warren, an Airbnb host, had to pay $ 2400 for violating these state's laws[110]. After appealing, the fine was thrown out, because Warren could prove that his roommate was present when the guest was in the house and they therefore maintained a "common household", an exception allowed by the state's rules on short-term rental. The decision of the Environmental Control Board was, however, very narrow. The board made it conspicuously clear that they ruled in favour of Warren only because of the presence of his roommate. This judgement therefore did not really provide legal certainty to other Airbnb hosts who are in general not present when renting out their house. Or as the State Senator Liz Krueger, a Democrat representing Manhattan and supporter of the 2010 law on short-term rentals in reaction to this ruling states:

> "The vast majority of Airbnb's business in New York City — short-term rentals of apartments in residential buildings without any permanent residents present — remains unambiguously illegal" (cited in: Carrns 2013).

In February 2015, another landmark case *(42nd & 10th Assoc. LLC v Ikezi)* took place in New York where a tenant was evicted from his house after the court found that he was engaged in profiteering by renting out his rent-stabilized house[111]. This ruling will

[110] For the official decision and order, see: http://www.scribd.com/doc/142650911/Decision-and-Order-for-NOV-35006622J. Accessed: 27 July 2015.

[111] "Rent stabilized tenants are protected from sharp increases in rent and have the right to renew their leases." http://www.nycrgb.org/html/resources/faq/rentstab.html#exactly. Accessed: 27 July 2015.

obviously have an impact on the way in which Airbnb hosts renting out their rent-stabilized house look at their activities. Moreover, this court success also affected landlords as more of them are now preparing to sue their tenants for participating in Airbnb (Marsh 2015).

Next to individual hosts that are being prosecuted, also Airbnb itself has been –and still is being- confronted with demands from legal authorities. Although the company, up until now, has not been held liable for illegal activity on its platform, it does have access to all the data about the activities taken place on the platform. Access to this data would make it much easier to find possible wrongdoers.

In the fall of 2013, the Attorney General of the state New York, Mr. Schneiderman, issued a subpoena to receive information concerning the hosts of Airbnb. He wanted this information in order to check if Airbnb hosts residing in his state were paying taxes under the state's law. In addition, he wanted to investigate the possibility that Airbnb was being used for the exploitation of illegal hotels (also see: Macmillan and Karmin 2014).

After negotiations that took more than 6 months, Airbnb and the Attorney General reached an agreement[112]. Airbnb assented to hand over an anonymized data set of Airbnb users, stripped from personally identifiable information. In the following year, the Attorney General would start hunting down Airbnb hosts that were in violation with the state's laws. Airbnb was obliged to disclose information concerning these hosts that are under investigation, when requested by the Attorney General.

In August 2014, as a consequence of this agreement, Airbnb was asked to hand over "the unredacted, personal information on 124 individuals"[113], all having multiple listings on the platform. After informing the hosts involved, Airbnb handed over their information to the Attorney General. All in all, the investigation of the Attorney General's office resulted in a report, issued October 2014, which stated that 72% of

[112] For the official agreement regarding compliance with subpoena, see: http://www.ag.ny.gov/pdfs/OAG_Airbnb_Letter_of_Agreement.pdf. Accessed 27 July 2015.

[113] http://publicpolicy.airbnb.com/new-york-community-update/. Accessed 27 July 2015.

the listings that appear on Airbnb are in violation of the state's laws[114]. Although 94% of the Airbnb hosts have at most two listings online, the other 6% of the hosts dominated the platform with up to hundred dwellings online counting for 37% of all host revenue. These 6% hosts were regarded as commercial hosts[115].

Airbnb contests these findings by claiming that the report does not take into account the 2000 allegedly illegal listings Airbnb had already taken down.

If Airbnb wants to maintain trust in its platform, it is crucial it finds a way to come to terms with the limitations set by the law. The uncertainty deriving from the legal position of hosts can be a threat to the sharing community on which Airbnb is based. Clear, shared, and predictable rules add to the familiar world where trust can thrive. Just as trust on the interpersonal level is a strategy to deal with the complexity inherent in human life, the legal system is a strategy to deal with this same complexity on the societal level. In a society without some sort of legal system, all the complexity has to be dealt with on the interpersonal level, making it much harder to cope with risks that might affect people personally but cannot be influenced by them individually. A legal framework set in place can neutralize some of the most basic complexities, bringing forth a familiar world in which there is room for interpersonal interactions based on trust.

The upheaval caused by the investigations of the Attorney General and the different court cases form a disturbance to this familiar world. Suddenly, very fundamental and shared assertions within the Airbnb community are deprived of their self-evident character. As the reactions of Airbnb hosts on the story of Neil Warren show, although they appreciate the efforts of Airbnb to help Warren, they also worry about what the legal consequences of their own participation on the platform may be. As one Airbnb hosts writes on the Airbnb blog[116]: *"Are we going to be taken to court and fined? Should we all pull our listings till we know?"*

Airbnb cannot promise to its users that it will never hand over their data to

[114] http://www.ag.ny.gov/pdfs/Airbnb%20report.pdf. p2, Accessed 12 December 2015.

[115] Idem.

[116] http://publicpolicy.airbnb.com/huge-victory-new-york-nigel-warren-host-community/. Accessed: 27 July 2015.

officials. If Airbnb is required by law, the company will have to comply as also stated in their terms and conditions.

Just as their users, Airbnb cannot ignore the legal reality, they can, however, try to change it. Over the years, Airbnb has become a fierce defender of the possibility to collect so-called *hotel taxes* on behalf of its users[117]. The company sent a letter to all members of the New York State legislature asking to adapt the law in order to make this collecting of taxes possible. On its website, completely dedicated to the New York Airbnb community, it also mobilizes its users to take action and write to their legislator[118]. Although Airbnb never saw or presented itself as a hotel, rather its main goal was to side-step hotels altogether, "formalizing its relationship with tax collectors" could be seen as "the first step toward gaining broader legal acceptance" (Griswold 2015) and restoring trust in the community.

Where the state of New York persisted and up until now made no changes to the law, cities such as San Francisco, Portland, and –in Europe-, Amsterdam did accommodate Airbnb by requiring them to collect local taxes. While making individual arrangements with local authorities is the strategy Airbnb probably is planning to follow, it will not be the end of all legal uncertainty. For example, only recently in the Netherlands –despite the agreements with individual cities-, the federal tax authority started an investigation to see if Airbnb hosts are in compliance with national rules on income tax (van Noort 2015). It is to be expected that more of these investigations will follow in other countries as well.

As a way of concluding this section, it has to be noted that all these governmental interventions are not merely based on governmental actors who want to make sure the state –or city- does not miss out on tax money. Governmental actors also are called upon by society to intervene. Different authors (Gottlieb 2013; Guttentag 2013; Thomas 2014; Morozov 2014; Prof. J. Schor interviewed by: Bouma

[117] With some cities Airbnb has an agreement to collect tourist taxes. For example, for the city of Amsterdam (The Netherlands), Airbnb collects 5% taxes based on the price of the booking. Also in this case, Airbnb does not hand over any information about its users to the authorities. Amsterdam therefore has to trust Airbnb to collect the taxes correctly as it cannot control the collection itself.

[118] https://www.airbnbnyc.com/take-action. Accesses: 27 July 2015.

2015) argue that people are increasingly also confronted with the negative effects of Airbnb in everyday life. Only those who have something to share –or are able to pay for it- can participate, leading to a divide in local communities. Moreover, for local communities the constant arrival of new Airbnb guests can be a burden and weaken social cohesion. Especially, the exploitation of illegal hotels making use of the Airbnb platform can have a negative impact on the quality of living. Next to their agreement with Airbnb to collect taxes, the city of Amsterdam therefore also started a hotline for its citizens, which they can call if they experience nuisance or suspect illegal short-rent activities in their vicinity. Within two weeks, the hotline already led to the discovery of four illegal hotels[119].

All in all, it can be concluded that Airbnb faces some challenges to maintain trust in its platform. Not only must it find a way to come to an agreement with local laws and authorities to restore the familiar world online for its users, Airbnb is also increasingly being confronted with opponents who see trust being put to the test in local communities because of the misbalance that occurs due to the constant arrival of new Airbnb guests.

7.7 Conclusion: Why the sharing economy is not just about you and me

By analysing Airbnb, as a prime example of the collaborative consumption movement, I have showed that although trust between strangers is definitely key to the functioning of the platform, it certainly is not the whole story. By taking into account the construction, curation, and codification of the platform, a more nuanced image is painted of the way in which trust is being established.

It then becomes clear that the way in which the platform is designed and more specifically the way in which the algorithms pre-sort users' interactions and consequently have an impact on the way in which trust is being built. Moreover, the interests of Airbnb as a platform play a leading role in this trust building. As trust between users is central to the business model of Airbnb, the company goes to great length to facilitate trust. However, sometimes, the company's interest may conflict

[119] http://www.nu.nl/reizen/4096078/vier-illegale-hotels-amsterdam-opgedoken-bij-alarmlijn.html. Accessed: 28 July 2015.

with those of the users, consequently limiting the options users have to build trust. In addition, Airbnb increasingly has to deal with legal issues. Not only is this a challenge to trust vested in the platform as these legal quarrels bring along much uncertainty, but the initial self-regulation central to the collaborative consumption movement also has to make room for a more traditional, external, and top down regulation from governmental actors.

By analysing trust in the shared economy through the conceptual lens of the four Cs, I endorse the claim of the collaborative consumption movement that trust between strangers is a necessary condition for any initiative in this domain to become successful. However, I strongly disagree with the utopian belief that Internet technology enables us to restore old forms of interpersonal trust; rather a new kind of trust occurs characterized by the intertwinement of the interpersonal and the system, which I referred to as *interpersonal system trust*. If the case of Airbnb shows us anything worth to remember than it must be that *trust in the shared economy is not just about you and me*, but about you, me and the system, in all its facets, that brings us together.

8

A too familiar world.

One of the greatest challenges in the networked era is to acquire knowledge out of all the information that is piling up around us[120]. *Personalization* has been regarded as the technical solution to this problem. By filtering information based on the profiles of users, it becomes possible to create *tailored information environments*. Other authors have referred to the same phenomenon as *filter bubbles* (Pariser 2011) and *echo chambers* (Sunstein 2007).

These information environments are stretched out over the online and offline domain (if it is even possible to still clearly discern the two nowadays). On the Internet, where data are constantly multiplying, it has become impossible to efficiently search and find information without some technological assistance in the form of filtered and ranked content. In addition, rather than presenting users with

[120] This chapter is partly based on, and includes sentences from Keymolen 2014a.

random or general advertisements, items or services, online companies personalize their offers to fit the profile of the potential customer. In the offline domain, as we explored in chapter six, new technologies such as iBeacons in hotels make it possible to tailor the 'real-life' environment of hotel guests. Also the promises of ambient intelligence include the development of a *personalized* and pro-active living environment (van den Berg 2009). Because it is online that these personalized information environments are currently the most developed, they will be the main focus of my argument.

Many consider *online personalization* – the possibility to tailor online services to the individual needs and preferences of users – as one of the "Holy Grails" in the world of ICT (see: van der Hof and Prins 2008; Nabeth 2008; Chen and Stallaert 2014). The search engine Google, which provides users with search results relevant to their individual context, is a prime example of online personalization. Facebook also personalizes its services by ranking the posts on a user's timeline in order of importance, and online advertisement companies make use of *behavioural targeting* –a specific application of online personalization- to provide clients with tailored ads, sometimes following potential costumers all over the web (also referred to as *retargeting*). Obviously, online personalization has many advantages. It provides an easy retrieval of relevant information, it boosts the effect of advertisement, and it enables a more efficient and adequate way of doing business. In short, it makes online interactions run smoothly.

It is clear, nevertheless, that personalization may also cause privacy issues as it is based on the collection and analysis of large amounts of personal data (Solove 2004; Benoist 2008; Brownsword 2008; Chellappa and Sin 2005; van der Sloot and Borgesius 2012). Privacy issues may negatively impact trust. For example, when users feel online companies are not respecting their privacy, this may negatively influence the trust vested in these curators (Liu et al. 2005; Metzger 2004; Flavián and Guinalíu 2006).

Personalization, however, can also interfere with trust on another level. As personalization becomes increasingly sophisticated as well as ubiquitous, it may also fundamentally shape *the familiar world*. The way in which personalization influences the familiar world and therefore has an impact on the way trust is being established

will be the central topic of this chapter.

From a philosophical anthropology perspective, I will analyse how profiling – the current dominant technique enabling personalization- has an impact on the way in which meaning is being constituted. Where up until now, they way in which human beings perceive the world has always had an *intersubjective* character, due to personalization this perspective increasingly becomes *subjective*. And, where this world has always been characterized as one of *evolving stability*, due to personalization it becomes determined by a *fixed stability*.

While personalization on the one hand *reduces complexity* and therefore helps to create a familiar environment by predicting the needs of users, on the other hand, I will argue, a perfect –or almost perfect- personalized interface may result in a *'too familiar world'*. This *too familiar world* may strengthen on the one hand users' self-confidence by presenting them with information that affirms their initial beliefs. It may also make them less perceptive for information that challenges their behaviour. Online personalization provides users, therefore, with an *individualized familiar world* instead of a *shared familiar world* necessary for interpersonal trust to be established.

Research indicates that when people are experiencing a threat and are looking for information –for example a diagnosed patient who googles information on treatments- they have the tendency to only expose themselves to information they prefer (Liao and Fu 2013: 2366). This tendency is only being fortified by personalization. Without being confronted with alternative beliefs of others or contradicting information, it becomes more difficult to comprehend the motives of other persons or feel empathy for their considerations (Nussbaum 1998). Moreover, it might also reduce social capital (Pariser 2011) and weaken deliberative democracy (Sunstein 2007). For example, research on extreme right videos on Youtube indicates the existence of an "extreme right filter bubble" as users can "be immersed in this content following a short series of clicks" (O'Callaghan et al. 2013: 9).

All in all, personalization practices make it more strenuous to bridge the ontological distance human beings face in their interactions with others. Reverting to Plessner's distinction between the animal *Umwelt* and the human *open world*, I will argue that a personalized world of information tends to become an *Umwelt* instead of an *open world*, nudging human beings to cling to their *centric instead of their*

eccentric positionality. Personalization may hinder the bridging of the ontological distance between human beings. *In a personalized world interpersonal trust erodes.*

I will first analyse online personalization through *the lens of the four Cs*. Starting with the level of *construction*, I will look into the functioning of profiling and personalization. Then, I will look at the *curation* of the personalized interface by focussing on the personalization practices of Google. Third, I will take into account the *codification* concerning personalization. Finally, I will look at the *context* level, making use of some key concepts of mediation theory, which were already presented in chapter four. After this analysis of the personalized interface, I will examine the influence of personalization on the familiar world and interpersonal trust, elaborating on the work of philosopher Helmuth Plessner, amongst others.

8.1 Construction

Personalization can be perceived as

> "a form of user-to-system interactivity that uses a set of technological features to adapt the content, delivery, and arrangement of a communication to individual users' explicitly registered and / or implicitly determined preferences" (Thurman and Schifferes 2012: 776).

A necessary condition for online personalization is *automated profiling*. By means of algorithms, databases filled with huge sets of data are mined to create, discover, or construct knowledge (Hildebrandt 2008: 17). Profiling is used to create profiles of individual users or groups based on which personalization can take place. These profiles can be seen as "hypotheses" (idem: 18); predictions about future preferences and behaviour. Interestingly, these hypotheses are not necessarily based on a common sense expectation or on earlier-established knowledge. The hypotheses often just "emerge" in the process of gathering and analysing data (idem).

In the context of his research on behavioural targeting, Borgesius (2014) discerns 5 stages in the profiling process. First, there is the *collection of data*. In this phase, firms gather data about the behaviour of people by tracking them online. They drop for example a cookie – a small, non-intrusive text file – in the potential buyer's

browser which enables them to identify this specific device when visiting a website (Watts 2012). Etzioni (2012: 929) reports the use of "supercookies" which are not only difficult to detect but can even reinstall themselves after they are removed. Also Facebook, with its Like button implemented on many websites, is able to track the visitors of those websites even when they are not a Facebook-member themselves (Roosendaal 2010).

Second, the data is *stored*, often "tied to a unique identifier such as a cookie" (Borgesius 2014: 61). In this stage, a profile is made of a user. A profile is "a set of correlated data that identifies and represents a data subject" (Hildebrandt and Backhouse 2005: 106).

In phase 3, the collected data is *analysed*. By making use of algorithms, the data is mined, looking for correlations and patterns that may shed a light on preferences of the user. In the end it is the goal of this analysis to make a prediction; for example –when it concerns an advertisement company- about the probability a user will click on an ad.

Phase 4 is the phase of *data disclosure*. Data brokers collect and sell personal information, which other companies in their turn can use to personalize their content or services (Borgesius 2014: 70-71). It is, however, not always necessary to buy the data. Online retailers for example that want to make use of *retargeting* can also turn to companies such as Google that started testing this specific form of profiling - they refer to it as *remarketing*- in 2009 (Helft and Vega 2010) (For a legal analysis of Google's advertisement activities see: van der Sloot and Borgesius 2012).

The final phase is the phase where the targeting actually takes place. Based on the collected, stored, and analysed data, content tailored to the profile of a specific user will be displayed. Because the content is personalized, different visitors may experience different websites or different information environments, contributing to the arrival of a personalized online environment (for more emperical research in online personalization see: Mikians et al. 2012; O'Callaghan et al. 2013; Nguyen et al. 2014).

Personalization online can take on different forms. Probably one of the most well-known examples of personalization is the *recommendation tool* of *Amazon*. This algorithm allows the company to personalize its website by providing users with tailored recommendations based on their purchase history and by connecting one

item to another item (Linden et al. 2003). An example of such a recommendation could be: people who bought "Harry Potter and the Philosopher's Stone" also bought "Harry Potter and the Chamber of Secrets" (Rowling 1997, 1998).

Another example is the personalization of website content. For instance, news sites increasingly tailor their content to the individual profiles of visitors. In order to do so, they increasingly "…rely on software algorithms to predict readers' content preferences" (Thurman and Schifferes 2012: 775).

Another application of personalization is *behavioural targeting:* "the monitoring of people's online behaviour, to use the collected information to show people individually targeted advertisements" (Borgesius 2014: 30).

A sub-class of behavioural targeting, which has skyrocketed the last couple of years and is an important feature of the personalized web is *retargeting* (Beales 2010; Helft and Vega 2010; Lambrecht and Tucker 2013). Online retailers do not merely want to display a website or an ad tailored to the specific interests of their visitors. Better still, since visitors often leave the website without purchase, corporations want to *follow* visitors all over the web with personalized ads in the hope to persuade them to buy the item –or a related one- that they have shown interest for in the past.

E-advertising companies make this real-time targeting possible by monitoring online behaviour. If a potential buyer is for example looking at a pair of shoes, a cookie is placed into her browser connecting it to that pair of shoes (Steel 2007; Helft and Vega 2010). When she leaves the online shoe retailer, surfing to another website, the company is alarmed and automatically starts bidding on advertisement space on that other website, ensuring a personalized shoe-advertisement shows up when that web page has been loaded. All this happens fully-automated in a mere 6 milliseconds[121].

Borgesius (2014: 77) also reports that there are even firms whose core business it is to personalize websites based on demographic, behavioural and historical information. It is also possible to morph the design of websites. "Morphing involves automatically matching the basic 'look and feel' of a website, not just the content, to

[121] http://www.criteo.com/us/solutions. Accessed 08 January 2013.

cognitive styles" (Hauser et al 2009: 202 cited in Borgesius 2014: 77).

8.2 Curation

As we have already established in previous chapters, the vision of an open and free Internet as it was proclaimed in the early 90s can be judged as utopian and perhaps a little naïve. It is now generally recognized that online curators such as search engines, online businesses, and other information intermediaries have a big say in what kind of information a user has access to. Personalization fits this larger tendency to monitor, pre-sort, shape, and increasingly control the information environment. From a curation perspective personalization can be perceived as:

> "[…] an organisational strategy of companies, governments and other organisations to provide services by means of ICTs to a large number of individual customers worldwide on an individualised basis" (van der Hof and Prins 2008: 113).

Although users have the feeling they are anonymous online and nobody is interested in their online activities, the opposite is the case (Benoist 2008: 168). Almost 80% of the most often-visited websites use tracking technology to gather information of their visitors (Angwin 2010) and a majority of them use this information to tailor their interface to the personal profile of their users.

The basic rationale behind this personalization is that if users are presented with tailored information, they will be more interested, and hence buy a product. However, as Borgesius (2014: 36) notes, there is no consensus on the effectiveness of targeted advertisement. In their recent economics study on advertisements making use of behavioural targeting, Chen and Stallaert (2014) found that not in all cases does behavioural targeting pay off. Their research indicates that for small publishers it might be best to stay with traditional advertising, while when there is sufficient competition among similar advertisers, "the behavioural targeting revenue for the online publisher can approach double the income from traditional targeting" (Chen and Stallaert 2014: 447). In spite of these fluctuating findings, advertising companies, who are large contributors to the personalized information environment online, spent more than an estimated $ 1.3 billion in targeted advertising in 2011, and it is expected

that this figure has only risen the years that followed (Chen and Stallaert 2014: 430).

Many large curators online make use of personalization. To better understand their incentives in doing so, I will focus on Google as its business model leans in *two ways* on personalization. The company uses personalization *to tailor their search results* on the one hand and *to sell targeted advertisement* on the other. Mager (2012) speaks of the "*service-for-profile model*". A user can use the search engine free of cost because the profile Google creates is sold to profit-making corporations, or at least advertisements based on this profile are sold.

Although the average user will know Google mainly for its activities as a search engine, the company's revenue is mostly based on their advertisement activities. In 2014, the company generated 89% of its revenue from advertisers[122]. In order to both rank search results and sell advertisements, Google has to have access to a large body of *behavioural data* to create user-profiles. Taking into account their status as an "obligatory passing point" (Mager 2012: 776) for almost everyone who wants to find information online, a lack of data does not seem to be very likely.

However, little is publicly known about the way in which Google personalizes its search results (Hannak et al. 2013: 528). Google might relate a query to the user's search history and has the ability to cross-reference this information with data coming from their other services such as Gmail and Google Docs (Tene 2008: 1448). In addition, Google always makes use of *contextualization* (Enge 2011). The search engine takes into account context elements such as geography, language, and seasonality to make the interaction between its interface and the user run smoothly. In addition, even when a user is not logged in to Google, the search engine personalizes its results by making use of cookies. For a period of 180 days, a cookie linked to the user's browser keeps track of the search history.

Also, when it comes to Google's behavioural advertising program, it is not completely clear which data Google uses to build profiles of its users (van der Sloot and Borgesius 2012: 78-79). Research indicates (Gomez et al. 2009), nevertheless,

[122] http://www.sec.gov/Archives/edgar/data/1288776/000128877615000008/goog2014123110-k.htm - s9D1B941756DC95DE1882A6359272D250. Page 7, Accessed 05 August 2015.

that Google is dominant when it comes to tracking websites, as 92 out of the top 100 websites Google is able to monitor. Moreover, taking into account the privacy policies of other Google services such as YouTube, it seems that many data are gathered when making use of these other Google products (in this example, by looking at YouTube video clips) (van der Sloot and Borgesius 2012: 77-78).

All in all, it becomes clear that Google has access to a lot of data about the online behaviour of users and that all these sorting techniques enable Google to create profiles it can use to tailor its list of search results to the specific needs of the user as well as base its targeted advertisement on. Both the pre-sorting and ranking of information in order for users to find relevant information, as well as the targeted advertisements contribute to an online individualized information environment. As not just Google, but more and more online parties are using these personalization technologies, Pariser (2011: 111) states that in the end:

> "we'll increasingly be forced to trust the companies at the center of this process to properly express and synthesize who we really are".

8.3 Codification

Similar to the arrival of Airbnb, personalization, being a new technique leaning heavily on the collection and analysis of data, has also started several legal discussions. For instance, while it is clear that behavioural targeting or personalization involves the processing of data, it can be questioned if these data generally are also *personal data*. Again looking at Google, the company defines personal information as:

> "information you provide to us which personally identifies you, such as your name, email address, or billing information, or other information which can be reasonably linked to such information by Google"[123].

However, as Van Der Sloot and Borgesius (2012: 83) remark, this definition by

[123] https://privacy.google.com/about-ads.html. Accessed 5 August 2015.

Google is narrower than the definition of personal data provided by the Data Protection Directive (Directive), which is the regulatory instrument that focuses on the processing of personal data in the EU. In the Directive, personal information is defined as:

> "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity"[124].

The Article 29 Working Party, an independent European advisory body on data protection and privacy consisting of the data protection authorities of all EU member states, has elaborated on this definition, explaining that information not only relates to a person because the content is about the person, but also when information is used to evaluate or influence the behaviour of a person[125]. In addition, it is also stated that to decide if a person is identifiable all the means that reasonably can be used by either the controller –in our example Google- or anyone else to identify the person should be taken into account. Taking the previous into consideration, Van Der Sloot and Borgesius (2012) conclude that in general behavioural targeting involves the processing of personal data. They write:

> "[t]he collection and analysis of personal data of Internet users is a process that falls within the definition of processing of personal data in the Directive. Google is the controller as it determines the goal of the processing, targeted advertising, and the means by which the data are processed, such as determining the data mining techniques. In short, the Directive is applicable" (van der Sloot and Borgesius 2012: 85).

However, this broad perspective on what personal data is, as being put forward

[124] Article 2(a) of the Data Protection Directive.

[125] Article 29 Working Party, Opinion 4/2007, "On the concept of personal data", 20 June 2007.

by the Working Party, has also met resistance. In line with this perspective, for example also IP addresses are considered to be personal data. Not only has Google objected to this perspective, but also legal scholars such as Zwenne (2013: 8) argue against such a broad interpretation, as it will lead to the situation that "data protection law will apply in many situations where it is not needed".

This discussion on personal data is important for users because if personalized data are being processed and the Directive is applicable, the privacy of the users involved is protected by the standards and requirements set in the Directive[126]. Consequently, if personalization includes the processing of personal data, there has to be one of the *legal bases* listed in the Directive on which this processing takes place. The *unambiguous consent* of users –the Directive speaks of *data subjects*- as a legal basis is almost always required when curators –the Directive speaks of *data controllers*- process personal data for behavioural targeting (Borgesius 2015).

However consent is in 'crisis', Schermer et al. (2014) convincingly argue (also see: van Eijk et al. 2012). Users click 'agree' without giving it a second thought or reading the 'terms and conditions'. As a consequence, they agree to data processing without fully understanding the impact. Not only may this uninformed consent weaken the trust of users in data processing, it can also impose problems on the curators as they are processing personal data based on a shaky consent, as it may not truly reflect the wishes of the user.

All in all, it becomes clear that the uncertainty about the legal grounds for personalization makes it difficult for the average user to understand on which legal protection she can count. Van der Hof and Prins (2008: 116-117) suggest that the attention of data protection should shift from

> "individual sets of personal data towards the statistical models, profiles and algorithms with which individuals are categorized in a certain group or 'identity'. After all, these models and algorithms are privately owned and thus unavailable for public scrutiny. The interests of personal data

[126] As the directive has no direct impact but has to be adopted by EU member states in their national laws, users will have to turn to these EU member states in case of privacy violations.

> protection however, seem to require that they are made known to the public and thus are part of the public domain".

'Knowing that' personalisation takes place is of course a necessary condition for any reflection on the matter. However, 'knowing how' personalisation shapes the informational environment becomes even more important because of the increasingly ubiquitous presence of personalization and its potential impact on a variety of domains in everyday life. Returning to Google as our prime example of a curator who makes use of personalization, its privacy policy clearly states that:

> "our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection"[127].

However, what is not included in the privacy policy is the *rationale behind the personalization*. What are the mechanisms behind assigning a user a specific profile? And how does this profile lead to presenting this user with certain content and depriving her of others? To put it differently, users are notified that personalization is taking place, but the way in which this happens remains secret. What is needed, van der Hof and Prins (2008: 117) argue, are

> "instruments to enhance the visibility of and knowledge about how personal data are used and combined, on the basis of what data individuals are typified, by whom and for what purposes."

In her inaugural lecture, Hildebrandt (2013a) also explicitly addresses this issue. If users –or as Hildebrandt puts it more eloquently "inhabitants of cyberspace"- are entitled to gain insight on the logic of the processing of their personal data, could companies 'hide' behind trade secrecies or property rights? In any case, as the excerpt of Google's privacy policy illustrates, companies are certainly not pro-actively

[127]http://static.googleusercontent.com/media/www.google.com/en//intl/en/policies/privacy/google_privacy_policy_en.pdf, Version 30 June 2015, p3. Accessed 12 August 2015.

providing users with this kind of information. Therefore, if users want to know, they would have to actively request for information first. And then, even if users receive information on the rationale behind the processing of their personal information, given the complexity of these systems and algorithms, would they be able to fully comprehend it? Hildebrandt (2013a: 19) argues that it may be more effective to rethink how "legal protection can be incorporated as a default in the architecture of cyberspace" (also see: Hildebrandt 2015). Anticipating the second part of this chapter, it will be this lack of transparency concerning the logic behind the personalization which hinders users to look beyond their filter bubble.

8.4 Context

How does personalization affects users' experience? Authors like Latour (1992), Ihde (1990), and more recently Verbeek (2011b) have convincingly argued that technologies are not just neutral instruments performing a pre-defined task, but are artefacts that also influence the actions and experiences of their users in often unforeseen ways. Personalization co-shapes the way in which Internet users perceive reality. This co-shaping of users' experiences and actions is also referred to as "technological mediation" (see also chapter five). It is important to understand that this technological mediation is two-fold. Technology and users have a permanent stake in shaping each other. More than the "*building bricks"*, they are the "*products"* of their interaction (Verbeek 2000: 183). Therefore, online personalization is not just about a *personalized interface* but also about a *personalized user*. In the interaction, the user is often unconsciously –and sometimes even unwillingly- shaping the interface based on her online behaviour. Conversely, the interface, presenting the online world in a personalized manner, is affecting the user by pre-sorting her choices and actions (Pariser 2011).

In all mediation a *translation* takes place (Ihde 1990). Some aspects of the online world are *amplified*, while others are *reduced*. Looking at the personalized interface, it even is its principal goal to *amplify* the information that fits the profile of the user and to *reduce* information that is irrelevant to it. The personalized interface pre-sorts a specific kind of interpretation and shapes what counts *as real* (also see: Verbeek 2011b).

The way in which smart artefacts *mediate* interactions or *open up* the world

for human beings can take different forms. In chapter five, I introduced the *immersive hermeneutic relation* to define the interaction of users with their personalized information environment. This *immersive hermeneutic relation* shares some of the characteristics of the *hermeneutic relation* as described by Ihde but it also differs from it, mainly because of the networked ontology key to smart artefacts and services.

In a hermeneutic relation, an artefact represents reality in such a way that its users have access to it by engaging with the concerning artefact. When a user does a Google search on her computer, the computer is the "*object of perception*", however, it is simultaneously "referring beyond itself to what is not immediately seen" (Ihde 1990: 82). The screen of the computer represents the online world to the user. It represents in a sensible way the bits and bytes which cannot be perceived with the naked eye. In this hermeneutic relation of user and personalized interface, the user can, so to speak, read herself into any possible, online situation without actually being there (see: Ihde 1990: 92). The computer is not transparent but *opaque* and the user has to master certain skills in order to interact with the device. It is not *through* but *by* (Verbeek 2000: 142) the ability of the interface to visualize the online world that it becomes meaningful to the user.

However, this kind of hermeneutic relation also falls short on certain levels to capture the distinctiveness of the interaction with the personalized information environment. The online environment in which the user reads herself into also has an *immersive character* (also see Verbeek 2015: 219). She gets deeply engaged with the online service or the interaction with others through an online platform while smart algorithms constantly monitor her actions and simultaneously adapt and personalize her interface.

Moreover, the interface of the computer does not merely represent a specific aspect of the world, but it refers to a *different reality*! The online personalized world includes algorithms (construction), references to virtual contexts (my Facebook timeline), and connections with the offline world (the pictures of my friends represent real persons, the information I find on Google shapes my experience of the 'offline' world). The interface that enables me to read myself into the online context is simultaneously an integral part of this context. As Søraker (2012) argues, in the networked era it becomes increasingly difficult to clearly distinguish between

fundamental concepts such as *technology* and *world,* since the interface that mediates the online experience simultaneously is the online world itself.

> "Virtual worlds are both worlds and technologies; the computer simulation is both the underpinning of the virtual world and the means of mediation" (Søraker 2012: 504).

When technology and world become intertwined on such a fundamental level as is the case in the personalized information environment, it becomes increasingly difficult for users to evaluate the influence of the technology on their perception of reality. It is impossible to have a 'naked perception' or a non-mediated perception of the online world based on which a user can judge if its representation is sufficient or fair. Without the interface, there is no online world. Van den Hoven (1998) speaks of "artificial authorities" to emphasize the reliance of users on their devices to function properly. The average user might be able to read the interface, but not to explore the inner workings of the underlying profiling technologies. Van der Hof and Prins (2008: 121) warn:

> "Personalization… may force individuals into restrictive two-dimensional models based on the criteria set by technology and of those who own and apply the technology".

The impossibility to see through the functioning of the underlying profiling technologies also shows that, although we speak of a *personalized* interface or a *personalized* information environment, the *person concerned* has not much control over her information environment. The mediation of perception taking place through the interface is a mediation enabled by other parties, mostly information intermediaries. These companies have their own interests, which do not necessarily align with the interests of the user (Mager 2012). Introna and Nissenbaum (2000: 175) predict that:

> "… information seekers on the Web, whose experiences are mediated through search engines, are most likely to find large, popular sites whose designers have enough technical savvy to succeed in the raking game, and especially those sites whose proprietors are able to pay for various means of improving their site's positioning. Seekers are less likely to find less

> popular, smaller, sites, including those that are not supported by knowledgeable professionals. When a search does yield these sites, they are likely to have lower prominence in rankings".

All in all, we can conclude that from a user's perspective the personalized interface mediates –just as other artefacts do- the way in which reality is presented. Personalization makes it possible to amplify some information, while reducing others. However, particular to this mediation is that it becomes increasingly difficult for users to take a step back and reflect upon the mediation, as the world and technology - in the case of the personalized information environment - are fundamentally intertwined. As personalization is mostly in the hands of online curators, users have to rely on them to present the online world in a fair manner. Moreover, where artefacts such as a thermometer or a compass, open up the world in a more or less similar way to their users, the personalized interface mediates the online world in a unique way to its users. Consequently it can no longer be taken for granted that what I see will be the same as what others see.

8.4.1 Four Cs: where are we now?

Analysing online personalization through the lens of the four Cs provides us with a picture of a wide variety of online curators that use profiling techniques to tailor their content and services to the user's preferences. There is debate on the legal underpinning of activities such as behavioural targeting as it is not always clear if the data that are processed are personal data and if the consent often required to process personal data in this matters is given in a well-informed way. There is also the appeal to shift perspective and, instead of focussing on the data, to look more to the way in which users can regain control over what happens to them based on this data processing and to make the logic behind profiling and the personalization of content more transparent. Users may benefit from personalization because it enables them to retrieve information more easily and make their interaction online more efficient. However, because the processes underpinning their information environment are opaque, they cannot critically reflect upon it nor assess the impact it has on their everyday life.

What does this all have to do with trust?

If we can see online personalization as the first step towards a personalized 'lifeworld' in the networked era, this development –when it has reached its full growth- could fundamentally change the character of the familiar world that is a necessary condition for trust to be established. The implicit, shared assumption, which partly constitutes the familiar world, namely that all human beings perceive the world more or less in a similar way, will erode. Increasingly, human beings will reside in their own personalized information environment, their own "filter bubble" (Pariser 2011), gradually lacking the shared background necessary for trust to thrive.

I will first recapture the role and function of the familiar world for trust. Next, making use of Plessner's distinction between the open world of human beings and the "Umwelt" of animals, I will analyse the impact of profiling and personalization on the familiar world.

8.5 A familiar world

As we have seen in the second and third chapter, a familiar, shared background of experience is a necessary condition for trust to be established. Trust can only reduce complexity in a world that is already to a certain extent familiar. There are two main sources for this complexity: the *other* and *time*, revealing a social and a temporal level in the complexity of the world.

On the social level, complexity comes into the world because of the possibility of unanticipated actions by fellow human beings, constituting a source of insecurity. We cannot read the minds of others. They have to a certain extent the freedom to act in ways that cannot be foreseen by others.

On a temporal level, human beings are aware of the discrepancy between possible futures and the one future that will become reality. In the present they have to cope with an over-complex and undetermined future. Therefore, trust also has to do with anticipating the future. Trust is "to behave as though the future were certain" (Luhmann 1979: 10).

If people now constantly had to consider the possibility that they perceive the world in ways radically different than others do or that natural laws were not universal but susceptible to change, they would become paralyzed because of such an uncanny environment. The complexity brought forth by *time* and the *other* would

simply be too overwhelming. There has to be some familiarity first in order for human beings to be able to trust - to act as if they know for sure what the future will bring. If there was no familiarity, the hiatus located in the *self*, between *the other and me*, and between *the world and me* would be impossible to bridge.

Trust can only take place in a familiar world in which existence is already structured in a pre-reflexive way. We take the presence of the world, our fellow human beings and the objects we encounter for granted. In everyday life, we do not doubt their existence. We expect to see and experience the world in a way similar to our fellow human beings. They are, so to speak, "presupposed and co-experienced" (Luhmann 1979: 18). Our experience of the world automatically entails the intersubjective constitution of meaning. "There is no differentiation in the operation of constituting meaning and world, which brings everybody together in a diffuse consensus" (Luhmann 1979: 18). Plessner refers to this condition as the "Mitwelt". The way human beings are in the world, even when they are alone, is always a being-together-with-others. The familiar world is always *an intersubjective world*.

Although this familiar world invokes stability, it does not entail that it is unchangeable. On the contrary, the familiar world as it is an intersubjective world is always the result of the *coming together of perspectives*. As Hildebrandt (2015: 183) puts it: "[w]e are forever guessing each other's interpretations". The familiar world always remains intertwined with the open, complex world, which it, to a certain extent, regulates. Therefore, the stability provided by the familiar world is always an *evolving stability*.

It has to be noted that this complexity is not merely a burden, a hurdle we have to take in order to live our lives. *This complexity is also productive*, as it persuades us to act, to be creative, to imagine (for an analysis on the importance of imagination see: Schinkel 2014). For example, while it is true that our fellow human beings by their – to us- unpredictable behaviour add to the complexity we have to deal with in everyday life, simultaneously, it is also through their presence, through their perspective on the world that "man's environment becomes man's own world" (Luhmann 1979: 7). By presenting us with other perspectives of the world, they make us aware of the world's horizon of infinite possibilities.

8.6 Profiling

In the familiar world of human beings complexity always shines through, even more so in late modern society where high consequence risks are part of everyday life (chapter three)[128]. As a result, human beings are constantly engaged in all sorts of 'complexity-reducing activities'. They trust (of course), they rely on their social roles, on the security brought forth by institutions, on the control they gain by using technology, and on the structuring effect of the law.

Interestingly, if we approach profiling no longer as a mere technological process, but look at it from a more functionalistic perspective, it can also be categorized as a 'complexity-reducing activity" human beings engage in. The core activity of this technology, namely to automatically categorize and generalize information, is not merely confined to machines. Perhaps counter-intuitively, the non-reflective profiling of algorithms resembles the way *living nature*, including human beings, interact with their world on a daily basis. In order to hold their ground, plants, animals, and human beings all make use of what we might call *biological profiling* to filter their overly complex environment (Hildebrandt 2008: 25-30). In a routine-like manner, they are "[...] extracting relevant information from the environment" in order to adapt themselves to this environment and survive (idem 2008: 26). In line with Hildebrandt (2008: 24) we can say that "[...] profiling is not only a part of professional and everyday life but also a constitutive competence of life itself in the biological sense of the word".

8.6.1 Animal and human profiling

To understand how profiling is in fact an important element in the everyday life of all living nature, we turn again to the work of Helmuth Plessner (1975). According to

[128] Nevertheless, in daily life people may often 'forget' that the routines inscribed in their bodies are human-made and therefore changeable (Plessner 1978). And although all interaction is mediated, human beings experience it as direct, dismissing possible side effects of the mediating artefacts at hand. Human beings tend to uphold a *utopian belief* in a stable and unchangeable world, steered by universal rules.

Plessner, animals are 'captured' in a "*Funktionkreis*". They are aware of their environment as far as their building scheme permits them. Consequently, the information they receive while profiling their environment can only be of use in a specific situation, for example, when they perceive an enemy close by and have to choose between fleeing and fighting. Although especially higher mammals have a certain awareness of their environment, they cannot reflect upon their choices. They cannot break out of the actual situation, sit down, and wonder how to bring their strategies to perfection based on the gathered information over time. All information that is acquired by profiling their environment must fit into their pre-existing knowledge frame. Not aware of a past or future, non-human animals live "here and now" in an "*Umwelt*", a closed environment limited by their building scheme (Plessner 1975).

Just like other animals, profiling by human beings often takes place in a routine-like manner. We rely on the predictability of the social roles we all play and the shared background of values and rules. Human beings tend to forget that the routines they follow and the rules that are set in society in fact are social constructs and not natural laws. To reduce the complexity inherent in human life, human beings are in an often-unconscious way generalizing and categorizing the information around them. This is what constitutes the familiar world. From this perspective, the *familiar world* of human beings is also an un-reflected world. The third anthropological law of Plessner captures this human urge to live *as if* the world is a well-ordered place, by speaking of a *utopian standpoint* people strive for. There is, however, always the possibility of questioning routines and changing them.

Notwithstanding the fact that human beings mostly act without giving it a second thought, this does not mean their reflexive attribute is unimportant or even superfluous. On the contrary, according to Plessner, human beings differ from other animals because they are "conscious of their consciousness". Human beings are aware of the fact that *they* are the ones who are profiling the world. Because of their eccentric positionality they are aware of the world's contingency, confronted with the fickleness of *time* and the *other*. They do not live in a pre-existing, fixed environment, tuned to their building scheme as other animals do. The familiar world of human beings is indissolubly connected to the complex world it orders. Complexity always shines through. The boundaries of the familiar world may be structuring but they are

never concluding. Human beings have to mould their own world through culture, language, and technology. They have the ability to break out of an actual situation and become aware of its contingency. This second-order awareness means that, so to speak, from a distance human beings can look back and reflect upon the course of action, able to consider possible alternatives. Often, this awareness is triggered by conflicting opinions of others, challenging the existent knowledge frame. Eventually, this confrontation might lead to the adjustment of an initial set of beliefs. De Mul and Van Den Berg (2011) refer to this process of evaluating internal and external motives as "the reflexive loop".

The most fundamental difference between the animal *Umwelt* and the *familiar world* of human beings is that the *Umwelt* cannot be shattered, - there is no world beyond the familiar world, so to speak - where the *human familiar world*, on the other hand, always remains a world 'under-construction'.

8.7 A too familiar world online

When we now take this distinction between an open world and a *closed Umwelt* and look at the functioning of automated personalization online, we can determine that these techniques invite users to live in a *closed Umwelt rather* than in an *open world*. This shift may affect the familiar world in two important ways: it may lead to the construction of a *subjective* instead of an *intersubjective* familiar world. And, it may result in a *fixed* instead of an *evolving* stability.

First, by feeding users a string of information that only affirms their pre-existing, individual beliefs, profiling technologies build an online world, which resembles the closed world of animals, determined by their *Funktionkreis*. In personalized information environments, the action of human beings is constantly interpreted and anticipated by algorithms. While these algorithms may have access to the [data on] users' behaviour, these users do not have access to those profiling techniques and therefore "no way of guessing how we are being 'read' by our novel smart environments" (Hildebrandt 2015: 183; also see: Hildebrandt 2013b: 19-22). As a result, meaning is *no longer intersubjectively* constituted but *subjectively*. The personalized information environment reflects the user's individual preferences and is no longer the result of the 'coming together of perspectives'. It can no longer be presupposed that the way in which I perceive the world is more or less in line with the

way others perceive it. This basic and fundamental uncertainty which familiarity normally neutralizes is under scrutiny. Online, people reside in a *personalized Umwelt* or a 'too familiar world'. Contrary to the cultural and open world, this is not a *shared familiar world [Mitwelt]* in which meaning is inter-subjectively constituted.

Second, as personalization is about predicting users' preferences and pro-actively adapting the content and services based on these predictions, the chance of being confronted with deviating opinions of others or conflicting information becomes in general far less likely. Solving the 'problem' of contingency – as we have seen in chapter five - by putting all faith in the computational powers to predict and consequently control the future, may turn out to be an impoverishment of our lives. As it happens, it is often these moments of disturbance and conspicuous complexity shining through in the familiar world that fire up the eccentric positionality of human beings, instigating creativity and new perspectives. Taking into account that "[…] conscious reflection is the incentive to create new habits […]"(Hildebrandt 2008: 27), personalization is wired to spur *stagnation* or *fixed stability* instead of *evolving stability*.

All in all, the personalization of the lifeworld, especially when this bearing is not just constricted to the online domain, will lead to a personalized information environment. It may well be that on a *personal level* trust in the self or self-confidence will be strengthened as the beliefs of human beings are constantly conformed and reaffirmed. However, on the *interpersonal level*, the ontological distance human beings have to bridge may enlarge. Consequently, interpersonal trust may erode in a personalized information environment.

8.8 Challenges for trust

Similar to the other chapters, I want to speak of *challenges for trust* – and not merely of 'problems' or 'adversities' (although these are definitely there) - as I strongly believe that we still have room to manoeuvre and design, sell, regulate, and interact with these technologies in such a way that the too familiar world I have sketched does not necessarily have to become reality. Some may argue that a too familiar world due to personalization will not arrive in any case, as technology will never reach the perfection needed to speak of such a fundamental shift. Some research indicates that

only 17,50% of the search results online currently are personalized and the obvious personalization we know from retargeting, users in general do not find difficult to identify as such.

However, one can wonder how perfect this technology has to be in order to have the effect of rendering the familiar world subjective and fixed. In addition, where one, isolated personalized domain in life may not immediately lead to a personalized information world, an accumulation of personalization practices may well have this impact. Where it becomes increasingly difficult to change technologies once they are integrated in everyday life -especially if the Internet of Things as described in chapter five takes flight - we now, with online personalization as a prime example of what may be yet to come, have the opportunity to carefully think through how we want personalization technologies to be designed. On several levels, first initiatives are already being folded out. The interplay of personal, technical, curation, and regulation strategies might make a difference.

8.8.1 Strategies for making the familiar world not too familiar

By changing their interaction with the personalized interface, users are able to adjust its workings. This "domestication" (Silverstone and Hirsch 1992; Frissen 2004, 1994) of artefacts often occurs when the artefact is embedded in daily practice. In his book, Pariser (2011) recommends several personal strategies to replenish the filter bubble with new and diverse information.

By altering her daily routines online, a user can open up the personalized interface and indirectly persuade it to build in new elements of information. Or as Pariser (2011: 223) states: "[…] varying your path online dramatically increases your likelihood of encountering new ideas and people." Another strategy is to prefer websites that are transparent about the profiling technologies they use to websites that are not. By being conscious about the kind of interfaces one uses, the influence of a personalization can be minimized.

However, a necessary condition for successfully getting around personalization practices is some basic knowledge on how profiling technologies work. *The user should become emancipated.* If users are sleepwalking into a personalized information environment, they cannot change their routines. Unfortunately,

knowledge about online personalization is often absent. Pan et al. (2007) for example show how college students are not aware of the ranking strategy of Google and blindly trust the search engine by clicking on the first search results that pop up, even when the abstract seems less relevant.

In his pamphlet "Program or be programmed", Rushkoff (2010) makes a stand against digital illiteracy and encourages the development of basic programming skills for all Internet users. Having insight into the basic workings of programming must strengthen users to use personalized interfaces in a more informed way. Developing digital literacy or e-skills is also on Europe's digital agenda. It is assumed that children can benefit more from the Internet when they are better able to recognize and deal with online risks such as a biased online environment (de Haan 2010; Sonck et al. 2011).

However, even if users become more aware of personalization practices and want to be able to evaluate, adapt, protest, or block certain practices, they also need to have the tools to act upon this knowledge. Further developing the legal framework may enable users to do so. Hildebrandt (2011a) develops the idea of 'legal protection by design'. With this concept, she does not refer to some sort of 'top-down' regulation or invisible disciplining of users. Rather, it refers to "a new articulation of legal protection" (Hildebrandt 2013a: 20); for example, by designing technology in a way that access to personal data is facilitated and the logic behind the automated decision is made comprehensible to the user. We should:

> "develop intuitive interfaces with which citizens can gain insight into the multiple manners in which they are 'being read' by their smart environments. This should give them the means to come to grips with potential consequences" (Hildebrandt 2013a: 19-20).

On a similar note, Koops (2011) argues that it is not very effective for users to control the process of collecting and managing data as such. It would be more useful to make the process of decision-making transparent. Users should be able to control how companies, but also governments make use of personal data.

Also technical measures may help to counter a too familiar world. In line with the call for more transparency for users, researchers (Nagulendra and Vassileva 2014)

developed an interactive visualization to make users more aware of the personalization and filtering online. The results of their research show that such visualization leads amongst others to increased users' awareness.

Another strategy to counter the too familiar world could be *programmed serendipity:* the intentional replenishing of the personalized interface with random information. By including a portion of information that not directly derive from the personal profile of the user in the interface, the personalized information environment may again become more evolving instead of stable and fixed.

The question remains however, based on which parameters this 'un-personalized' stream of information should be built. *Sheer randomness* – as the opposite of personalization - could easily result in information that is of no interest to the user at all. With Gadamer (1972), we could say that to get the conversation started, we should find ourselves between the limits of 'strangeness and familiarity'. If the random information is completely strange to the user, she will probably not be interested nor make an effort to evaluate it. If the information is completely familiar, no repositioning will take place either. Programmed serendipity therefore is, to a certain extent, depending on the same personalization techniques it is supposed to counterbalance. To replenish the interface with information that will catch the attention of the user, some basic interests of the user simply have to be known first. Eventually, taking Gadamer's limits a step further, it might come down to finding the right balance between random and personalized information.

Computer scientists and programmers have taken on the task of finding this balance and safeguarding serendipity in the online world (Maccatrozzo 2012; Campos and De Figueiredo 2002). For example Campos and De Figueiredo (2002) have investigated the possibility of *programming for serendipity*. They developed a software agent called Max "that browses the web in order to find information that might stimulate the user, especially information that the user is not focused upon" (idem 2002: 52). Making use of, amongst others, the user's profile and a lexical database, Max formulates suggestions based on the generation of alternatives, the selection of also less significant concepts, replacing selected concepts by other, related concepts, and random stimulation (idem 57).

Also Helberger (2011), who addresses the problem of personalization first and foremost from a policy perspective, sees concrete design principles as a manner of ensuring diverse information exposure online. She speaks of *diversity by design* and

analyses four different conceptions of exposure diversity which could inform the design of internet technologies such as Electronic Programme Guides and search engines, namely: “Discovering the Difference, Exposure to Diverse Media Outlets, Promoting Personal Autonomy, and Encouraging Serendipitous Discoveries” (Helberger 2011: 464).

All in all, it becomes clear that *programmed serendipity* could help to safeguard the *open character* of the online world, but only if curators are willing to cooperate. To a certain extent, their willingness depends on –legal- regulation.

As a way of conclusion, I would like to stress that these strategies for making the familiar world not too familiar should be approached as a whole. While I may be optimistic about the possibility of consciously shaping and designing our smart artefacts and environments, I am absolutely pessimistic as it comes to the leverage of individual actors. The room to manoeuvre for the average user is small and superficial if she has no meaningful legal and technical tools at her disposal to take a stance against wrongful and/or unwanted personalization of both commercial and governmental actors. And as the interests of curators and users of technology do not always align, it may well be that the interests of the curators precede those of the users, if a clear and adequate legal framework is not set in place. In a too familiar world we gain stability and predictability, but it may also turn out to be an obstacle for trust as a productive way of dealing with the complexity inherent in human life.

9

Epilogue

I began this book by saying that every philosophy starts in wonder. Throughout this book, I have found some fruitful entries (at least that is what I hope!) to begin to understand the way in which people nowadays vest their trust in the wide variety of smart artefacts they use to build and cherish their relations and environment.

I approached trust as a strategy to reduce complexity inherent in human life. But also the activity of analysing trust, –of trying to wrap my mind around trust by capturing it in a net of concepts, neatly distributed over eight chapters- is in fact a complexity-reducing strategy in itself.

By trying to define it, to understand it, giving it a place in a pre-existing framework or familiar world of philosophical and sociological theory, my goal was unmistakably to reduce the complexity that surrounds trust.

Nevertheless, I never aspired for trust to become a docile concept, obedient to all my whims. I never aimed at undoing all of the wonder that surrounds it. Taking my own research to heart, I know better than to believe it is even possible to take away all complexity; or all wonder. I therefore like to think of this book as being successful if it has reduced the complexity surrounding trust to such a level that a productive stance can be developed to analyse, face, and even resolve trust-related issues in the networked era.

By developing the 4 Cs framework (Context, Construction, Curation, and Codification) I on the one hand want to make trust and how it functions more insightful, diminishing its 'fuzziness'. On the other hand I want to provide a tool of analysis that does justice to the complexity of trust. By looking at trust from different angles, by uncovering the different levels represented by the four Cs, turning trust as

a diamond, as it were, round and round, I want to show its different facets.

The 4 Cs framework can be used to analyse existing practices in the networked era, as I have illustrated in the chapters 6,7, and 8. Interestingly (and slightly personally disturbingly), by applying the framework, I stumbled upon my own shortcomings. Although I believe that in the previous chapters the framework has sufficiently proven its value by bringing to the fore the intertwining of the different levels of trust in the networked era, it actually deserves more than merely an overambitious philosopher in technology to operationalize it. The framework actually calls for multi-disciplinary research in which scholars from other disciplines -such as social scientists, legal scholars, and technologists- participate in order to further deepen the analysis of the 4 Cs.

Next to a tool of analysis, the framework can also be put to use when designing smart artefacts, environments, and services that foster trust. Although this book does not provide clear-cut answers as to which specific mechanisms have to be implemented to ensure such trust-enabling practices, by discerning the 4 Cs it, nevertheless, pre-sorts which aspects have to be taken into account. This makes the 4 Cs framework useful to ethicists and policymakers.

Developing a framework not just to *analyse* but also *evaluate* trust in the networked era, presumes that there is something at stake; and that we could do better. Throughout the book I spoke of "challenges for trust", emphasizing the room we have to manoeuvre, rethink and redesign our relationships mediated by smart artefacts and services in the networked era. Let me first elaborate a bit on whom I refer to with "we" before looking at what I believe have become the biggest challenges for trust in the networked era.

Before really delving into the subject of trust in the networked era, I was very optimistic about the possibilities for the individual user to reinvent –and keep reinventing- his or her interactions in the networked era. In line with the ideas lying at the heart of the Open Internet movement, I was convinced that smart artefacts and the Internet in general were carriers of freedom, enlarging the palette of actions of human beings. Acknowledging Plessner's first anthropological law of human beings, that they are "artificial by nature", I of course was aware that the own weight of artefacts also steers our actions, but I was nevertheless more focussed on -and convinced of- the power of human beings to create their environment and the way in

which this was fortified by new technological developments.

However, during the research for this book I came to see that the room for average users to actively shape their mediated interactions is actually rather limited. The interpretative flexibility or multistability on the context level is restricted and even under siege due to the strong tendency of other major actors such as governments and companies to increasingly steer and control the context level. Moreover, users often are not even aware that they are visible to –and easily manipulated by- curators, putting them in a situation of *invisible visibility*, reinforcing the power imbalance.

This observation immediately proves the necessity to analyse trust in the context of all the Cs and not merely focus on the context level, which regularly happens in trust research. When one only looks at the context level, for example when analysing the introduction of the digital key app in the hotel business or a collaborative consumption platform such as Airbnb, it mistakenly seems as if interpersonal trust is reintroduced by smart technologies and transparency, user-friendliness, and direct contact are fostered, where actually, this utopian belief in restoring direct and trustworthy interaction is mediated by technologies and monetized by curators. Though trust on social network sites and other interactive platforms may resemble interpersonal trust, it is in fact mediated through-and through. Therefore, I called it *interpersonal system trust* to emphasize the intertwinement of the interpersonal and the system.

Is this intertwinement of the interpersonal and the system the end of trust in the networked era? I would say on the contrary. It actually urges us to expand our perception of the "we" and to make sure that our analysis not only focuses on the users but includes also the actors on the other levels, like the companies, governments, and designers who are creating and curating the artefacts. Too often, the responsibility of creating trustworthy interactions is first and foremost being put on the shoulders of average users. They should vest their trust wisely and if they cannot or do not want to be vulnerable to possible risks they should simply stop making use of smart artefacts and services.

In the networked era or hyperhistory, however, retreating from technology is no meaningful option as technology in general and ICTs specifically are at the core of daily life. Cutting out these technologies would saddle people with too high a price to pay. If you would even consider holding users responsible for their use of smart

artefacts than at least they should be provided with a choice that exceeds the 'take it or leave it' option. It is only when the responsibility to create a familiar world is shared by all actors operating on the 4 C levels that the room for the average users to shape their relations with these powers can become meaningful.

All in all, as smart artefacts and smart environments, characterized by their *radical nearness* in the networked era, increasingly become central to everyday life, users should be empowered not to just *adopt* but also *adapt* them. Discussions should therefore not merely focus on the way users interact with technology but also include the role of other actors such as companies, designers, and regulators in contributing to the familiar world and the design of trust-enabling smart artefacts.

Throughout this book, I have put forward several challenges to trust; coming to the end, I will concentrate on the ones I consider to be the most critical.

First, the familiar world, necessary for trust to thrive, is under threat. The basic idea that human beings perceive the world in more or less similar ways, reducing this first and radical complexity of living in an unpredictable environment shared with fickle others, becomes pressured when this environment one-sidedly starts interpreting human beings in order to establish personalized and pro-active life-worlds. The basic starting point that human beings should be understood as always in interaction with their environment, must be comprehended as an interaction that goes both ways: environment and human beings are constantly shaping each other. Meaning and identity are not self-imposed but emerge out of the interaction.

Currently however, we see a move toward a smart environment and smart artefacts that are constantly interpreting and interacting with human beings; while human beings can only guess based on which –automated- processes this interpretation takes place. Consequently, they lose the ability to co-shape meaning and identity. This might result in a too familiar world as illustrated in chapter 8 which analysed online personalization. But it might also lead to a completely unfamiliar world where people get trapped because there is no room to relate in a meaningful way to this smart environment.

Second, the growing aversion to accept uncertainty as part of human life shoves away trust as a manner to deal with uncertainty by embracing it as part of every interaction. The unceasing, recurring utopian belief in technology as a solution

to all problems human beings face seems to have reached a new climax with the arrival of Big Data. The conviction that data-driven decision making, information-fuelled services, and smart devices will solve 'the problem of contingency' and will bring us a society built on 'omnipotence and omniscience', makes us blind for the artefacts' own weight and the impoverishment of human life if we for the sake of certainty (or safety) willingly suppress our eccentric positionality and the productive openness it brings along.

To understand the way in which human beings and artefacts shape each other it can be very illuminating to look at both their default settings. Originally, the default setting is the way in which a device or service is being programmed when it leaves the factory or the [online] shop. Research indicates that 95% of the users do not bother to change their settings; consequently, the power of the defaults should not be underestimated[129].

I believe that human beings also have a default setting and this default setting is: trust. Trust does not come by foot and leave by horse, as Thorbecke presumably claimed. To trust is one of the most fundamental actions human beings undertake to carve out their lives. Moreover, trust is robust; it can take a hit. It enables us to act, create, and take a chance without becoming paralyzed by all the possible futures lying ahead of us. I repeatedly stated that trust is a fiction as it is a kind of pretending to know what will happen while we actually don't have a clue.

Paradoxically, this makes trust the most real fiction I can imagine.

[129] http://www.theguardian.com/technology/2013/dec/01/default-settings-change-phones-computers, accessed 10 December 2015.

10

Bibliography

Abbate, J. (1999). *Inventing the internet.* Cambridge, Massachusetts: The MIT Press.

Abramova, O., Shavanova, T., Fuhrer, A., Krasnova, H., & Buxmann, P. (2015). Understanding the sharing economy: The role of response to negative reviews in the peer-to-peer accommodation sharing network. *ECIS 2015 Completed Research Papers, Paper 1*.

Achterhuis, H. (2001). *American philosophy of technology: The empirical turn.* Bloomington and Indianapolis: Indiana University Press.

Adams, M. (2004). Whatever will be, will be: Trust, fate and the reflexive self. *Culture & Psychology, 10*(4), 387-408.

Akrich, M. (1992). The de-scription of technical objects. In W. Bijker, & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change.* (pp. 205-224). Cambridge (MA): MIT Press.

Akrich, M., & Latour, B. (1992). A summary of a convenient vocabulary for the semiotics of human and nonhuman assemblies. In E. Bijker, & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 259-264). Cambridge: The MIT Press.

Ambrosino, B. (2014). Why the disappearance of hotel room keys marks the end of hospitality. http://qz.com/177505/why-the-disappearance-of-hotel-room-keys-marks-the-end-of-hospitality/. Accessed 09 June 2015.

Anderson, C., & Wollf, M. (2010). The web is dead. Long live the internet. http://www.wired.com/magazine/2010/08/ff_webrip/. Accessed 04 September 2012.

Angwin, J. (2010). The web's new gold mine: Your secrets. http://online.wsj.com/article/SB10001424052748703940904575395073512989404.html. Accessed 18 December 2012.

Anthony, S. (2014). You can now open hotel rooms with just your smartphone – and bypass check-in, too. http://www.extremetech.com/extreme/193450-you-can-now-open-hotel-rooms-with-just-your-smartphone-and-bypass-check-in-too. Accessed 10 June 2015.

Arnoldi, J. (2001). Niklas Luhmann. *Theory, culture & society, 18*(1), 1-13.

Asselt, M. B. A. (2000). *Perspectives on uncertainty and risk.* Dordrecht: Kluwer.

Aupers, S. (2002). The Revenge of the machines: On modernity, digital technology and animism. *Asian Journal of Social Science, 30*(2), 199-220.

Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Bacharach, M., & Gambetta, D. (2001). Trust in signs. In K. S. Cook (Ed.), *Trust in society.* New York: Russel Sage Foundation.

Baier, A. (1986). Trust and antitrust. *Ethics, 96*(2), 231-260.

Banerjee, P. (2014). Why I won't be using Airbnb's online verification system http://www.theglobeandmail.com/globe-investor/personal-finance/household-finances/why-i-wont-be-using-airbnbs-new-online-verification-system/article19290229/. Accessed 24 July 2015.

Barlow, J. P. (1996). Declaration of the independence of cyberspace. https://projects.eff.org/~barlow/Declaration-Final.html. Accessed 15 December 2014.

Bauwens, M., Mendoza, N., & Iacomella, F. (2012). A synthetic overview of the collaborative economy. Orange Labs, P2P Foundation.

Beales, H. (2010). The value of behavioral targeting. http://corp1.ewr1.turn.com/sites/default/files/wp-content/uploads/2010/06/Beales_NAI_Study.pdf. Accessed 12 December 2014.

Beck, U. (1992a). From industrial society to the risk society: Questions of survival, social structure and ecological enlightenment. *Theory, culture & society, 9*(1), 97-123.

Beck, U. (1992b). *Risk society: Towards a new modernity.* London: SAGE Publications.

Beck, U. (1994). The reinvention of politics. Towards a theory of reflexive modernization. In G. U. Beck, A., S. Lash (Ed.), *Reflexive modernization*. Cambridge: Polity Press.

Bednarz, J. (1984). Complexity and intersubjectivity: Towards the theory of Niklas Luhmann. *Human Studies, 7*(1), 55-69.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom.* New Haven Conn.: Yale University Press.

Benkler, Y. (2011). *The penguin and the Leviathan: The triumph of cooperation over self-interest.* New York: Crown Business.

Benoist, E. (2008). Collecting data for the profiling of web users. In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen. Cross-Disciplinary perspectives.* (pp. 169-184). Amsterdam: Springer Netherlands.

Bijker, E., & Law, J. (1992). *Shaping technology/building society: Studies in sociotechnical change.* Cambridge: MIT Press.

Bijker, W. (1995). *Of bicycles, bakelites, and bulbs. Towards a theory of sociotechnical change.* Cambridge, Massachusetts: The MIT Press.

Bijker, W. (2001). Understanding technological culture through a constructivist view of science, technology, and society. In S. Cutcliffe, & C. Mitcham (Eds.), *Visions of STS. Counterpoints in science, technology, and society studies* (pp. 19-34). Albany: State University of New York Press.

Blum, A. (2012). *Tubes: A journey to the center of the Internet.* New York: Ecco.

Boden, R. (2015). Apple demonstrates smartwatch keys and payments. http://www.nfcworld.com/2015/03/09/334534/apple-demonstrates-smartwatch-keys-and-payments/. Accessed 10 June 2015.

Borgesius, F. J. (2014). *Improving privacy protection in the area of behavioural targeting*. University of Amsterdam, Amsterdam.

Borgesius, F. J. (2015). Personal data processing for behavioural targeting: Which legal basis? *International data privacy law, first published online June 23, 2015*, 1-14.

Borgmann, A. (1987). *Technology and the character of contemporary life: A philosophical inquiry.* Chicago: University of Chicago Press.

Borgmann, A. (2000). The moral complexion of consumption. *Journal of Consumer Research, 26*(4), 418-422.

Borgmann, A. (2009). Focal things and practices. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 56-75). Lanham: Rowman & Littlefield Publishers.

Borsari, A. (2009). Notes on a "Philosophical Anthropology" in Germany. An introduction. *Iris, 2036-3257*(I), 113-129.

Botsman, R., & Rogers, R. (2010). *What's mine is yours: The rise of collaborative consumption.* New York: Harper Business.

Boulton, C. (2015). Airbnb Open Sources Software to Lure Talent Amid 'Insane' Competition http://blogs.wsj.com/cio/2015/06/17/airbnb-open-sources-software-to-lure-talent-amid-insane-competition/. Accessed 23 July 2015.

Bouma, K. Deeleconomie vergroot ongelijkheid. (2015, 13 November). *de Volkskrant,* p. 27.

boyd, d. (2014). *It's complicated: The social lives of networked teens.* New Haven, London: Yale University Press.

Bradach, J. L., & Eccles, R. G. (1989). Price, authority and trust: From ideal types to plural forms. *Annual Review of Sociology, 15*, 97-118.

Brey, P. (2009). Philosophy of technology meets social constructivism: A shopper's guide. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 98-111). Lanham, Maryland: Rowman & Littlefield Publishers.

Briscoe, N. (2000). Understanding the OSI 7-layer model *PC Network Advisor*(120), 13-14.

Brodwin, D. (2012). The Rise of the Collaborative Consumption Economy. http://www.usnews.com/opinion/blogs/economic-intelligence/2012/08/09/how-collaborative-consumption-reinvigorates-our-economy. Accessed 04 December 2012.

Broeders, D. (2014). Investigating the place and role of the armed forces in Dutch cyber security governance. Rotterdam: Erasmus University Rotterdam.

Broeders, D. (2015). The public core of the internet: An international agenda for internet governance. *WRR-Policy Brief no.2*. The Hague: WRR.

Brown, C. (2011). NFC room keys find favour with hotel guests. http://www.nfcworld.com/2011/06/08/37869/nfc-room-keys-find-favour-with-hotel-guests/. Accessed 10 June 2015.

Brownsword, R. (2008). Knowing me, knowing you – Profiling, privacy and the public interest. In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen. Cross-Disciplinary perspectives.* (pp. 345-363). Amsterdam: Springer Netherlands.

Buytendijk, F. J. J. (1938). *Grondproblemen van het dierlijk leven.* (Fundamental Problems of Animal Life.)Antwerpen/Brussel/Nijmegen: Standaard Boekhandel/ Dekker & van Devegt.

Campos, J., & De Figueiredo, A. D. (2002). Programming for serendipity. In *Proc. AAAI Fall Symp. on Chance Discovery, 2002* (pp. 48-60): American Association for Artificial Intelligence.

Carrns, A. (2013). Tenant's fine for renting to tourist is overturned. http://www.nytimes.com/2013/10/01/nyregion/tenants-fine-for-renting-to-tourist-is-overturned.html?partner=rss&emc=rss&_r=1&. Accessed 27 July 2015.

Castells, M. (1996). *The rise of the network society.* Oxford: Blackwell Publishers.

Castells, M. (1999). *The information age: Economy, society and culture.* Oxford: Blackwell Publishers.

Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management, 6*(2-3), 181-202.

Chen, J., & Stallaert, J. (2014). An economic analysis of online advertising using behavioral targeting. *Mis Quarterly, 38*(2), 429-449.

Chesbrough, H. W., Vanhaverbeke, W., & West, J. (2006). *Open innovation: Researching a new paradigm.* Oxford: Oxford University Press.

Chesky, B. (2011). On Safety: A Word From Airbnb. http://techcrunch.com/2011/07/27/on-safety-a-word-from-airbnb/. Accessed 12 December 2015.

Coeckelbergh, M. (2012). Can we trust robots? *Ethics and Information Technology, 14*, 53-60.

Cohen-Almagor, R. (2010). Responsibility of and trust in ISPs. *Knowledge, Technology & Policy, 23*(3-4), 381-397.

Coleman, J. S. (1982). Systems of trust: A rough theoretical framework. *Angewandte Sozialforschung, 10*(3), 277-299.

Coleman, J. S. (1994). *Foundations of social theory.* Cambridge, MA: Harvard University Press.

Corbey, R. (1986). Plessner, Scheler en de menselijke geest. *Tijdschrift voor Filosofie*(48), 49-65.

de Haan, J. (2010). NL Kids online. Den Hague: Sociaal Cultureel Planbureau.

de Laat, P. (2005). Trusting virtual trust. *Ethics and Information Technology, 7*, 167-180.

de Laat, P. (2012). Navigating between chaos and bureaucracy: How Open-content communities are backgrounding trust. In Gordana Dodig-Crnkovic, Antonino Rotolo, Giovanni Sartor, Judith Simon, & C. Smith (Eds.), *Social Computing, Social Cognition, Social Networks and Multiagent Systems Social Turn, 2012* (pp. 67-72). Birmingham: The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

de Mul, J. (1994). *Toeval. Inaugurale rede.* Rotterdam: Rotterdamse Filosofische Studies.

de Mul, J. (2001). Afstand in filosofisch perspectief. In V. J. J. M. Bekkers, & B. Foederer (Eds.), *ICT, afstand en compliance. Internet en Openbaar bestuur* (pp. 17-23). Den Haag: Belastingsdienst.

de Mul, J. (Ed.). (2002). *Filosofie in cyberspace. Reflecties op de informatie -en communicatietechnologie.* Kampen: Klement.

de Mul, J. (2003). Digitally mediated (dis)embodiement. Plessner's concept of excentric positionality explained for cyborgs. *Information, Communication & Society, 6*(2), 247-266.

de Mul, J. (2014a). *Destiny Domesticated. The rebirth of tragedy out of the spirit of technology.* New York: SUNY Press.

de Mul, J. (2014b). Philosophical anthropology 2.0. Reading Plessner in the age of converging technologies. In J. De Mul (Ed.), *Plessner's philosophical anthropology. Perspectives and prospects.* (pp. 457-475). Amsterdam: Amsterdam University Press.

de Mul, J., & van den Berg, B. (2011). Remote control: Human autonomy in the age of computer-mediated agency. In M. Hildebrandt, A. Rouvroy (Ed.), *Law, Human Agency and Autonomic Computing* (pp. 46-63). New York: Routledge.

DeAmicis, C. (2015). Airbnb Now Factors In Host Preferences. http://recode.net/2015/04/15/airbnb-now-factors-in-host-preferences/. Accessed 22 July 2015.

Deibert, R. (2008). *Access denied: The practice and policy of global Internet filtering.* Cambridge, Mass.: MIT Press.

Deibert, R. (2013). *Black code: The battle for the future of cyberspace.* Plattsburgh, NY: Signal/ McClelland & Stewart.

Deibert, R., Palfrey, J., Rohozinski, R., & Zittrain, J. (2012a). *Access contested: Security, identity, and resistance in Asian cyberspace information revolution and global politics.* Cambridge, MA: MIT Press.

Deibert, R., Palfrey, J., Rohozinski, R., Zittrain, J., & OpenNet Initiative. (2010). *Access controlled: The shaping of power, rights, and rule in cyberspace.* Cambridge, Mass.: MIT Press.

Deibert, R., Rohozinski, R., & Crete-Nishihata, M. (2012b). Cyclones in cyberspace: Information shaping and denial in the 2008 Russia-Georgia war. *Security Dialogue, 43*(1), 3-24.

Denardis, L. (2012). Hidden levers of internet control. *Information, Communication and Society, 15*(5), 720-738.

Denardis, L. (2014). *The Global War for Internet Governance.* New Haven and London: Yale University Press.

Desteno, D. (2014). *The truth about trust.* New York: Hudson Street Press (Penguin Group).

Dewandere, N. (2015). Rethinking the human condition in a hyperconnected era: Why freedom is not about sovereignty but about beginnings. In L. Floridi (Ed.), *The Onlife Manifesto. Being human in a hyperconnected era* (pp. 195-215). Cham: Springer.

Dreyfus, H. (2009). Heidegger on gaining a free relation to technology. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (2 ed., pp. 25-33). Lanham: Rowman & Littlefield Publishers.

Edelman, B. (2015). Digital business models should have to follow the law, too. https://hbr.org/2015/01/digital-business-models-should-have-to-follow-the-law-too. Accessed 27 July 2015.

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of airbnb.com. *Harvard Business School NOM Unit Working Paper*(14-054).

Ellul, J. (1990). *The technological bluff.* Grand Rapids, Mich.: W.B. Eerdmans.

Enge, E. (2011). How Google does personalization with Jack Menzel. http://www.stonetemple.com/how-google-does-personalization-with-jack-menzel/. Accessed 06 January 2012.

Erikson, E. H. (1950). Growth and crises of the "healthy personality." In M. J. E. Senn (Ed.), *Symposium on the healthy personality* (pp. 91-146). Oxford: Josiah Macy, Jr. Foundation.

Ert, E., Fleischer, A., & Magen, N. (2015). Trust and reputation in the sharing economy: The role of personal photos on Airbnb. *Available at SSRN 2624181*.

Ess, C., & Thorseth, M. (2011). *Trust and virtual worlds.* New York: Peter Lang.

Ess, C. M. (2010). Trust and new communication technologies: Vicious circles, virtuous circles, possible futures. *Knowledge, Technology & Policy, 23*(3-4), 287-305.

Etzioni, A. (2012). The privacy merchants: What is to be done? *University of Pennsylvania Journal of Constitutional Law, 14*(4), 929-951.

Fischer, J. (2006). Philosophische anthropologie - ein wirkungsvoller Denkansatz in der deutschen Soziologie nach 1945 [Philosophical anthropology - an important approach in post-war german sociology]. *Zeitschrift Für Soziologie, 35*(5), 322-347.

Fischer, J. (2014). Philosophical anthropology. A third way between Darwinism and Foucaultism. In J. de Mul (Ed.), *Plessner's philosophical anthropology. Perspectives and prospects* (pp. 41-56). Amsterdam: Amsterdam University Press.

Flavián, C., & Guinalíu, M. (2006). Consumer trust, perceived security and privacy policy. *Industrial Management & Data Systems, 106*(5), 601-620.

Floridi, L. (2015a). Commentary on the online manifesto. In L. Floridi (Ed.), *The online manifesto*. Cham: Springer.

Floridi, L. (2015b). Hyperhistory and the philosophy of information policies. In L. Floridi (Ed.), *The Onlife Manifesto. Being human in a hyperconnected era* (pp. 51-64). Cham: Springer.

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century.* New York: Farrar, Straus and Giroux.

Frissen, V. (1994). The domestication of the telephone: Domestic technology and everyday life- mutual shaping processes. *COST A*, 28-30.

Frissen, V. (1997). Gender, ICTs and everyday life: Mutual shaping processes. *European Commision*.

Frissen, V. (2004). *De domesticatie van de digitale wereld*. Erasmus University Rotterdam, Rotterdam.

Fuglslang, L. (2001). Three perspectives in STS in the policy context. In S. Cutcliffe, & C. Mitcham (Eds.), *Visions of STS. Counterpoints in science, technology, and society studies* (pp. 35-50). Albany: State University of New York Press.

Fukuyama, F. (1995). *Trust.* London: Hamish Hamilton.

Gadamer, H.-G. (1972). *Wahrheit und Methode.* Tübingen: Mohr.

Gannes, L. (2013). After home-in trashing incident, Airbnb builds an in-house enforcer team. http://allthingsd.com/20130716/after-home-trashing-incident-airbnb-builds-an-in-house-enforcer-team/. Accessed 24 July 2015.

Garfinkel, H. (1963). A conception of, and experiments with, 'Trust' as a condition of stable concerted actions. In O. J. Harvey (Ed.), *Motivation and Social Interaction: Cognitive Determinants* (pp. 187-238). New York.

Garland, D. (2001). *The Culture of Control. Crime and Social Order in Contemporary Society.* Chicago: The University of Chicago Press.

Gasser, U., Faris, R., & Heacock, R. (2013). Internet Monitor 2013. Cambridge: The Berkman Center for Internet and Society at Harvard University.

Gergen, K. (1991). *The saturated self: Dilemmas of identity in contemporary life.* New York: Basic books.

Giddens, A. (1990). *The consequences of modernity.* Cambridge, UK: Polity Press in association with Basil Blackwell, Oxford, UK.

Giddens, A. (1991). *Modernity and self-identity, self and society in the late modern age.* Stanford: Stanford University Press.

Giddens, A., & Pierson, C. (1998). *Conversations with Anthony Giddens: Making sense of modernity.* Stanford, Calif.: Stanford University Press.

Giordano, S. (1997). Hotel room keys an endangered species. http://www.bizjournals.com/seattle/stories/1997/09/01/focus5.html?page=all. Accessed 09 June 2015.

Goffman, E. (1959). *The presentation of self in everyday life.* New York: Doubleday.

Goffman, E. (1967). *Interaction ritual; essays on face-to-face behavior.* Garden City, N.Y.: Anchor Books.

Goffman, E. (1990). *Stigma. Notes on the management of spoiled identity.* London: Penguin Books.

Goldsmith, J. L., & Wu, T. (2008). *Who controls the Internet? Illusions of a borderless world.* New York: Oxford University Press.

Gomez, C., Oller, J., & Paradells, J. (2012). Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors, 12*(9), 11734-11753.

Gomez, J., Pinnick, T., & Soltani, A. (2009). KnowPrivacy. Berkeley: UC Berkeley School of Information.

Good, D. (1988). Individuals, interpersonal relations, and trust. In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 31-48). Oxford: Basil Blackwell.

Gottlieb, C. (2013). Residential Short-Term Rentals: Should Local Governments Regulate the 'Industry'? *Planning & Environmental Law, 65*(2), 4-9.

Grabner-Kräuter, S. (2009). Web 2.0 social networks: The role of trust. *Journal of business ethics, 90*(4), 505-522.

Gralla, P. (2007). *How The Internet Works.* (eighth ed.)Indiapolis: Que Publishing.

Grene, M. (1966). Positionality in the philosophy of Helmuth Plessner. *The Review of Metaphysics, 20*(2), 250-277.

Grene, M. (1995). *A philosophical testament.* Chicago: Open Court.

Griswold, A. (2015). Why Airbnb desperataley wants to pay hotel taxes. http://www.slate.com/articles/business/moneybox/2015/02/airbnb_hotel_taxes_why_does_the_sharing_economy_startup_want_to_pay_them.html. Accessed 27 July 2015.

Grodzinsky, F. S., Miller, K. W., & Wolf, M. J. (2011). Developing artificial agents worthy of trust: "Would you buy a used car from this artificial agent?". *Ethics and Information Technology, 13*(1), 17-27.

Gupta, N. (2013). *Inside Bluetooth Low Energy.* Boston/London: Artech House.

Guttentag, D. (2013). Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, 1-26.

Habermas, J., & Luhmann, N. (1972). *Theorie der Gesellschaft oder Sozialtechnologie.* Frankfurt a.M.: Suhrkamp.

Hahn, A. (2004). Der Mensch in der deutschen Systemtheorie. In U. Bröckling, A. T. Paul, S. Kaufmann, & W. Eßbach (Eds.), *Vernunft-Entwicklung-Leben. Schlüsselbegriffe der Moderne. Festschrift für Wolfgang Eßbach* (pp. 279-291). München: Fink.

Hannak, A., Sapiezynski, P., Kakhki, A. M., Krishnamurthy, B., Lazer, D., Mislove, A., et al. (2013). Measuring personalization of web search. In *22nd international conference on World Wide Web, Rio de Janeiro, 2013* (pp. 527-538). Brazil: International World Wide Web Conferences Steering Committee.

Hardin, G. (1968). The tragedy of the commons. *science, 162*(3859), 1243-1248.

Hardin, R. (2001). Conceptions and explanations of trust. In K. S. Cook (Ed.), *Trust in society* (pp. 3-39). New York: Russel Sage Foundation.

Hardin, R. (2002). *Trust and trustworthiness.* New York: Russell Sage Foundation.

Hardin, R. (2004). Internet capital. *Analyse & Kritik, 26*, 122-138.

Hardin, R. (2006). *Trust.* Cambridge: Polity Press.

Harris, P. L. (2012). *Trusting what you're told. How children learn from others.* Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Heidegger, M. (2010). *Being and time.* Albany: State University of New York Press.

Helberger, N. (2011). Diversity by Design. *Journal of Information Policy, 1*, 441-469.

Helft, M., & Vega, T. (2010). Retargeting ads follow surfers to other sites. http://www.nytimes.com/2010/08/30/technology/30adstalk.html?_r=0. Accessed 01 August 2013.

Henslin, J. M. (1967). Trust and the cab driver. In M. Truzzi (Ed.), *Sociology and Everyday Life* (pp. 40-88). New York: Russel Sage Foundation.

Hildebrandt, M. (2008). Defining Profiling: A New Type of Knowledge? In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen. Cross-Disciplinary perspectives* (pp. 17-45): Springer Netherlands.

Hildebrandt, M. (2011a). Legal protection by design. Objections and refutations. *Legisprudence, 5.2*, 223-248.

Hildebrandt, M. (2011b). Privacy na de 'computationele wending'? In V. Frissen, L. Kool en M. van Lieshout (Ed.), *Jaarboek ICT en Samenleving. De Transparante Samenleving.* (Vol. 8, pp. 29-48). Gorredijk: Media Update Vakpublicaties.

Hildebrandt, M. (2013a). The rule of law in cyberspace. Nijmegen.

Hildebrandt, M. (2013b). Slaves to big data. Or are we? (Esclavos de los macrodatos. ¿O no?). *IDP. Revista de internet, Derecho Y Política, 17*, 7-44.

Hildebrandt, M. (2015). The public(s) onlife. A call for legal protection by design. In L. Floridi (Ed.), *The onlife manifesto. Being human in a hyperconnected era* (pp. 181-194). Charm: Springer.

Hildebrandt, M., & Backhouse, J. (2005). D7.2: Descriptive analysis and inventory of profiling practices. FIDIS.

Hirschmann, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states.* Cambridge, Mass: Harvard University Press.

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth.* Bloomington: Indiana University Press.

Ihde, D. (1993). *Philosophy of technology: An introduction.* New York: Paragon House.

Ikkala, T., & Lampinen, A. (2015). Monetizing network hospitality: Hospitality and sociability in the context of Airbnb. In *Computer Supported Cooperative Work & Social Computing, Vancouver, 2015* (pp. 1033-1044). New York: ACM.

Introna, L. D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. *The Information Society, 16*(3), 169-185.

Jalava, J. (2003). From norms to trust: The Luhmannian connections between trust and system. *European Journal of Social Theory, 6*(2), 173-190.

Joerges, B. (1999a). Do politics have artefacts? *Social studies of science, 29*(3), 411-431.

Joerges, B. (1999b). Scams cannot be busted: Reply to Woolgar & Cooper. *Social studies of science*, 450-457.

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Straus and Giroux.

Kaplan, D. M. (2009). *Readings in the philosophy of technology.* Lanham: Rowman & Littlefield Publishers.

Kelly, K. (2010). *What technology wants.* New York: Viking.

Keymolen, E. (2008). *Vol Vertrouwen. Over online (on)zekerheid en de brug van het vertrouwend handelen*. Erasmus University Rotterdam, Rotterdam.

Keymolen, E. (2013). Trust and technology in collaborative consumption. Why it is not just about you and me. In R. Leenes, & E. Kosta (Eds.), *Bridging Distances in Technology and Regulation* (pp. 135-150). Tilburg: Wolf Legal Publishers.

Keymolen, E. (2014a). A moral bubble. The influence of online personalization on moral repositioning. In J. de Mul (Ed.), *Plessner's philosophical anthropology. Perspectives and prospects* (pp. 387-406). Amsterdam: Amsterdam University Press.

Keymolen, E. (2014b). Vertrouwen en big data: "opening the black box". In P. J. Dijkman, V. Frissen, & J. Prij (Eds.), *Biopolitiek: De macht van big data* (Herfst 2014 ed., pp. 90-96, Christen Democratische Verkenningen). Amsterdam: Uitgeverij Boom.

Keymolen, E., & Broeders, D. (2013). Innocence Lost: Care and control in dutch digital youth care. *British Journal of Social Work, 43*(1), 41-63.

Keymolen, E., & Prins, C. (2011). Jeugdzorg via systemen. De verwijsindex risicojongeren als spin in een digitaal vangnet. In D. Broeders, C. Cuijpers, C. Prins (Ed.), *De staat van informatie* (pp. 293-248, WRR Verkenningen). Amsterdam: AUP.

Keymolen, E., van den Berg, B., Prins, C., & Frissen, V. (2010). Vertrouwen in hybride ketens. Een onderzoek in het kader van de Alliantie Vitaal Bestuur. Den Haag: Alliantie Vitaal Bestuur.

Kiran, A. H., & Verbeek, P.-P. (2010). Trusting our selves to technology. *Knowledge, Technology & Policy, 23*(3-4), 409-427.

Kleinrock, L. (2010). An early history of the internet [History of Communications]. *Communications Magazine, IEEE, 48*(8), 26-36.

Knight, F. H. (1921). *Risk, uncertainty and profit.* Boston, New York: Houghton Mifflin Company.

Kockelkoren, P. (2014). The quest for the sources of the self, seen from the vantage point of Plessner's material a priori. In J. De Mul (Ed.), *Plessner's Philosophical Anthropology. Perspectives and prospects.* (pp. 317-334). Amsterdam: Amsterdam University Press.

Kohn, M. (2008). *Trust: Self-interest and the common good.* Oxford: Oxford University Press.

Kopetz, H. (2011). *Real-time systems. Design principles for distributed embedded applications.* New York: Springer.

Lahav, G., V. Guiraudon (2000). Comparative perspectives on border control: Away from the border and outside the state. In *The wall around the West. State borders and immigration controls in North America and Europe* (pp. 55-77). Lanham: Rowman and Littlefield publishers.

Lahno, B., & Matzat, U. (2004). From the editors. Trust and community on the Internet. Opportunities and restrictions for online cooperation. *Analyse & Kritik. Zeitschrift für Sozialtheorie, 26*(1), 1-6.

Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research, 50*(5), 561-576.

Latour, B. (1990). Technology is society made durable. *The Sociological Review, 38*(S1), 103-131.

Latour, B. (1992). Where are the missing masses. In E. Bijker, & J. Law (Eds.), *Shaping Technology/building Society: Studies in Sociotechnical Change* (pp. 225-258). Cambridge: MIT Press.

Latour, B. (1993). *We have never been modern.* Harvard: Harvard University Press.

Latour, B. (1994). On technical mediation. Philosophy, sociology, genealogy. *Common Knowledge, 3*(2), 29-64.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., et al. (2009). A brief history of the internet. *SIGCOMM Comput. Commun. Rev., 39*(5), 22-31.

Lessig, L. (2001). *The future of ideas.* New York: Random House.

Lessig, L. (2006). *Code: Version 2.0.* New York: Basic Books.

Lester, S. (2015). The emergence of Bluetooth Low Energy. http://www.contextis.com/resources/blog/emergence-bluetooth-low-energy/. Accessed 10 June 2015.

Lewis, J. D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces, 63*(4), 967-985.

Lewis, T. G. (2006). *Critical infrastructure protection in homeland security: Defending a networked nation.* Hoboken, New Jersey: John Wiley & Sons.

Liao, Q. V., & Fu, W.-T. Beyond the filter bubble: Interactive effects of perceived threat and topic involvement on selective exposure to information. In ACM (Ed.), *SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 2013* (pp. 2359-2368). Paris: ACM.

Lijmbach, S. (2002). The natural-scientific and phenomenological approaches to animals. In *Life Energies, Forces and the Shaping of Life: Vital, Existential* (pp. 101-115): Springer.

Lindberg, P. J. (2013). What your hotel knows about you. http://edition.cnn.com/2013/02/26/travel/what-your-hotel-knows/. Accessed 18 June 2015.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE, 7*(1), 76-80.

Lindwer, M., Marculescu, D., Basten, T., Zimmerman, R., Marculescu, R., Jung, S., et al. (2003). Ambient intelligence visions and achievements: Linking abstract ideas to real-world concepts. In *Design, Automation and Test in Europe Conference and Exhibition, Munchen, 2003* (pp. 10-15): IEEE.

Liu, C., Marchewka, J. T., Lu, J., & Yu, C.-S. (2005). Beyond concern—a privacy-trust-behavioral intention model of electronic commerce. *Information & Management, 42*(2), 289-304.

Luhmann, N. (1964). *Funktionen und Folgen formaler Organisation.* Berlin: Duncker & Humblot.

Luhmann, N. (1970). Soziologie als Theorie sozialer Systeme. In *Soziologische Aufklärung. Aufsätze zur Theorie sozialer Systeme* (Vol. 1, pp. 113-136): Opladen.

Luhmann, N. (1979). *Trust and power. Two works by Niklas Luhmann.* (H. Davis, Trans)New York: John Wiley & sons Ltd.

Luhmann, N. (1988). Familiarity, confidence, trust: Problems and alternatives. In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations.* (pp. 94-107): Blackwell Publishers.

Luhmann, N. (1990). Technology, environment and social risk: A systems perspective. *Industrial Crisis Quarterly, 4*, 223-231.

Luhmann, N. (1995). *Social Systems.* Stanford: Stanford University Press.

Luiijf, E., Burger, H., & Klaver, M. (2003). Critical infrastructure protection in the Netherlands: A quick scan. In U. E. Gattiker (Ed.), *EICAR Copenhagen, 2003.*

Lyon, F., Möllering, G., & Saunders, M. (2012). *Handbook of research methods on trust.* Cheltenham UK, Northampton MA USA: Edward Elgar Publishing.

Maccatrozzo, V. (2012). Burst the filter bubble: Using semantic web to enable serendipity. In *The Semantic Web–ISWC 2012* (pp. 391-398): Springer.

Macmillan, D., & Karmin, C. (2014). New York's supreme court hosts Airbnb case. http://www.wsj.com/articles/SB10001424052702304049904579516074190780960. Accessed 27 July 2015.

Madlmayr, G., & Scharinger, J. (2010). Neue Dimension von mobilen Tourismusanwendungen durch Near Field Communication-Technologie. In *mTourism* (pp. 75-88): Springer.

Mager, A. (2012). Algorithmic ideology. *Information, Communication & Society, 15*(5), 769-787.

Manley, B. (2015). Issues loom for keyless entry in hotels. http://www.hotelnewsnow.com/Article/15618/Issues-loom-for-keyless-entry-in-hotels. Accessed 11 June 2015.

Marcuse, H. (2009 [1964]). The new forms of control (One Dimensional Man). In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 34-42). Lanham: Rowman & Littlefield Publishers.

Marsh, J. (2015). Landlords planning more evictions after Airbnb ruling. http://nypost.com/2015/02/21/landlords-planning-more-evictions-after-airbnb-ruling/. Accessed 27 July 2015.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think.* Boston: Houghton Mifflin Harcourt.

McGeer, V. (2004). Developing trust on the internet. *Analyse & Kritik, 26*, 91-107.

McKnight, D. H., & Chervany, N. L. (2002). What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce, 6*, 35-60.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002a). Developing and validating trust measures for e-commerce: An integrative typology. *Information systems research, 13*(3), 334-359.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002b). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *The Journal of Strategic Information Systems, 11*(3), 297-323.

McLeod, C. (2014). Trust. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*.

McNamara, B. (2015). Airbnb: A Not-So-Safe Resting Place. *J. on Telecomm. & High Tech. L., 13*, 149-170.

Mead, G. H., & Morris, C. W. (1934). *Mind, self & society from the standpoint of a social behaviorist.* Chicago: The University of Chicago Press.

Mead, G. H., Morris, C. W., Brewster, J. M., Dunham, A. M., & Miller, D. L. (1938). *The philosophy of the act.* Chicago: The University of Chicago Press.

Meijboom, F. L. B. (2008). *Problems of trust. A question of trustworthiness.* Utrecht: Universiteit Utrecht.

Metzger, M. J. (2004). Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer ‐ Mediated Communication, 9*(4), 00-00.

Meyrowitz, J. (1985). *No Sense of Place. The impact of electronic media on social behavior.* Oxford: Oxford University Press.

Meyrowitz, J. (2005). The rise of glocality. New senses of place and identity in the global village. In K. Nyiri (Ed.), *A sense of place: The global and the local in mobile communication.* (pp. 21-30). Vienna: Passagen.

Mikians, J., Gyarmati, L., Erramilli, V., & Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks, 2012* (pp. 79-84): ACM.

Misztal, B. A. (1996). *Trust in modern societies: The search for the base of social order.* Cambridge: Polity Press.

Misztal, B. A. (2001). Normality and trust in Goffman's theory of interaction order. *Sociological Theory, 19*(3), 312-324.

Mitchell, R. L. (2006). It's just the key to your room. Computerworld surveys 100 hotel card keys to explode an urban myth. http://www.computerworld.com/article/2561071/security0/it-s-just-the-key-to-your-room.html. Accessed 07 June 2015.

Möllering, G. (2001). The nature of trust: From Georg Simmel to a theory of expectation, interpretation and suspension. *Sociology, 35*(2), 403-420.

Möllering, G. (2005). The Trust/Control duality. An integrative perspective on positive expectations of others. *International sociology, 20*(3), 283-305.

Möllering, G. (2006). *Trust: Reason, routine, reflexivity.* Amsterdam: Elsevier.

Morozov, E. (2011). *The net delusion: The dark side of internet freedom.* New York: Public Affairs.

Morozov, E. De deeleconomie is gedwongen. (2014, 17 Oktober 2014). *NRC Handelsblad,* p. 10.

Mueller, M. (2010). *Networks and States: The global politics of internet governance.* Massachusetts: MIT Press.

Nabeth, T. (2008). Reply: Online personalisation. For the bad or for the good? In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen. Cross-Disciplinary perspectives.* (pp. 124-127). Amsterdam: Springer Netherlands.

Nagulendra, S., & Vassileva, J. (2014). Understanding and controlling the filter bubble through interactive visualization: A user study. In *Proceedings of the 25th ACM conference on Hypertext and social media, 2014* (pp. 107-115): ACM.

Naughton, J. (2012). *From Gutenberg to Zuckerberg.* New York, NY: Quercus.

Neuhofer, B., Buhalis, D., & Ladkin, A. (2015). Smart technologies for personalized experiences: A case study in the hospitality domain. *Electronic Markets*, 1-12.

Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). Exploring the filter bubble: The effect of using recommender systems on content diversity. In

Proceedings of the 23rd international conference on World wide web, 2014 (pp. 677-686): ACM.

Nissenbaum, H. (2001). Securing trust online: Wisdom or oxymoron? *Boston University Law Review, 81*, 101-131.

Nissenbaum, H. (2004). Will security enhance trust online, or supplant it? In R. M. Kramer, & K. S. Cook (Eds.), *Trust and distrust in organizations: Dilemmas and approaches* (Vol. 7). New York: Russell Sage Foundation.

Nooteboom, B. (1997). Grondslagen en grenzen van vertrouwen. *Filosofie in bedrijf, 25*, 7-14.

Nooteboom, B. (2002). *Trust: Forms, foundations, functions, failures and figures.* Cheltenham: Edward Elgar Publishing.

Nooteboom, B., & Six, F. (2003). *The trust process in organizations: Empirical studies of the determinants and the process of trust development.* Cheltenham: Edward Elgar Publishing.

Nussbaum, M. C. (1998). *Cultivating humanity: A classical defense of reform in liberal education.* Cambridge, Mass.: Harvard University Press.

O'Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2013). The extreme right filter bubble. *arXiv preprint arXiv:1308.6149*, 1-10.

O'Hara, K. (2012). Trust in social machines: The challenges. In A. R. Gordana Dodig-Crnkovic, Giovanni Sartor, Judith Simon, and Clara Smith (Ed.), *Social Computing, Social Cognition, Social Networks and Multiagent Systems Social Turn -SNAMAS 2012, Birmingham, UK, 2012*: AISB/IACAP World Congress 2012.

O'Neill, O. (2002a). *Autonomy and trust in bioethics.* Cambridge: Cambridge University Press.

O'Neill, O. (2002b). *A question of trust.* Cambridge: Cambridge University Press.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*(3), 801-823.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you.* New York: Penguin Press.

Paul, A. T. (2001). Organizing Husserl: On the phenomenological foundations of Luhmann's system theory. *Journal of Classical Sociology, 1*(3), 371-394.

Pesonen, J., & Horster, E. (2012). Near field communication technology in tourism. *Tourism Management Perspectives, 4*(0), 11-18.

Pettit, P. (1995). The cunning of trust. *Philosophy and Public Affairs, 24*(3), 202-225.

Pettit, P. (2004). Trust, reliance and the internet. *Analyse & Kritik, 26*, 108-121.

Pieters, W. (2011). Explanation and trust: What to tell the user in security and AI? *Ethics and Information Technology, 13*(1), 53-64.

Plessner, H. (1975). *Die Stufen des Organischen und der Mensch; Einleitung in die philosophische Anthropologie.* Berlin: De Gruyter.

Plessner, H. (1978). *Hoe de mens bestaan kan.* Alphen aan de Rijn: Samson Uitgeverij.

Plessner, H. (2003). *Die Frage nach der Conditio humana.* (Gesammelte Schriften VIII).1961 Frankfurt am Mein: Suhrkamp.

Poggi, G. (1979). Introduction. In T. G. P. Burns (Ed.), *Trust and Power (Niklas Luhmann)*. Chichester: John Wiley & Sons Ltd.

Prins, C., & Keymolen, E. (2011). Jeugdzorg en privacy: meten met verschillende maten In V. Frissen, L. Kool en M. van Lieshout (Ed.), *De transparante samenleving* (Jaarboek 2011 ICT en Samenleving). Amsterdam: Media Update Vakpublicaties.

Putnam, R. (2000). *Bowling alone: The collapse and revival of American community.* New York: Simon and Schuster.

Rawlinson, K. (2014). Turkey blocks use of Twitter after prime minister attacks social media site. http://www.theguardian.com/world/2014/mar/21/turkey-blocks-twitter-prime-minister. Accessed 12 December 2015.

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier.* Reading, Mass.: Addison-Wesley Pub. Co.

Richards, N. M., & King, J. H. (2013). Three paradoxes of big data. *Stanford Law Review Online, 66*, 41.

Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest Law Review, 49*, 393-432.

Rifkin, J. (2014). *The zero marginal cost society: The internet of things, the collaborative commons, and the eclipse of capitalism.* New York: Palgrave Macmillan.

Roosendaal, A. (2010). Facebook tracks and traces everyone: Like this! *Tilburg Law School Legal Studies Research Paper Series, 3*(2011), 1-9.

Roudman, S. (2013). Airbnb is disruptive, but Is it getting "creepy" now, too? http://techpresident.com/news/23949/airbnb-causes-user-uproar-new-id-policy. Accessed 24 July 2015.

Rowling, J. K. (1997). *Harry Potter and the philosopher's stone.* London: Bloomsbury.

Rowling, J. K. (1998). *Harry Potter and the chamber of secrets.* London: Bloomsbury.

Rushkoff, D. (2010). *Program or be programmed: Ten commands for a digital age.* Berkeley, CA: Counterpoint.

Schermer, B. W., Custers, B., & van der Hof, S. (2014). The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology, 16*(2), 171-182.

Schermer, B. W., & Lodder, A. R. (2014). Internet governance. In S. van der Hof, A. R. Lodder, & G. J. Zwenne (Eds.), *Recht en Computer* (6 ed., pp. 1-24, Recht en Praktijk, Vol. 4). Deventer: Kluwer.

Schinkel, W. (2014). *Over het nut en nadeel van de sociologie voor het leven.* Amsterdam: Boom.

Schütz, A. (1967 [1932]). *The phenomenology of the social world.* Evanston: North-Western University Press.

Seligman, A. B. (1997). *The problem of trust.* Princeton: Princeton University Press.

Seppanen, R., K. Blomqvist, & Sundqvist, S. (2007). Measuring inter-organisational trust: A critical review of the emperical research in 1990-2003. *Industrial Marketing Management, 36*(2), 453-486.

Sherry, J. E. (1993). *The Laws of Innkeepers: For Hotels, Motels, Restaurants, and Clubs.* Ithaca, New York: Cornell University Press.

Silverstone, R., & Hirsch, E. (1992). *Consuming technologies: Media and information in domestic spaces.* London ; New York: Routledge.

Simon, J. (2013). Trust. In D. Pritchard (Ed.), *Oxford Bibliographies in Philosophy*. New York: Oxford University Press.

Simonelis, A. (2005). A concise guide to the major internet bodies. *Ubiquity, 6*(5), 16-22.

Singer, P. W., & Friedman, A. (2014). *Cybersecurity and cyberwar: What everyone needs to know.* Oxford: Oxford University Press.

Solove, D. J. (2004). *The digital person: Technology and privacy in the information age.* New York: New York University Press.

Sonck, N., Livingstone, S., Kuiper, E., & de Haan, J. (2011). Digital literacy and safety skills. *EU Kids Online*(EU Kids Online Network).

Søraker, J. H. (2012). Virtual worlds and their challenge to philosophy: Understanding the "intravirtual" and the "extravirtual". *Metaphilosophy, 43*(4), 499-512.

Sornes, T. (1979). Lock arrangement employing mechanically acting code card and key card. Google Patents.

Specker, J., Focquaert, F., Raus, K., Sterckx, S., & Schermer, M. (2014). The ethical desirability of moral bioenhancement: A review of reasons. *BMC Medical Ethics, 15*(1), 67-84.

Steel, E. (2007). How Marketers Hone Their Aim Online. http://www.wsj.com/articles/SB118221104155539813. Accessed 12 December 2015.

Sunstein, C. R. (2007). *Republic.com 2.0.* Princeton: Princeton University Press.

Swierstra, T., Boenink, M., Walhout, B., & van Est, R. (Eds.). (2009). *Leven als bouwpakket. Ethisch verkennen van een nieuwe technologische golf.* Den Haag: Rathenau Instituut.

Sztompka, P. (1999). *Trust: A social theory.* Cambridge: Cambridge University Press.

Taddeo, M. (2009). Defining Trust and e-trust: From old theories to new problems. *Technology and Human Interaction, 5*(2), 23-35.

Taddeo, M. (2010a). Modelling trust in artificial agents, a first step toward the analysis of e-trust. *Minds and Machines, 20*(2), 243-257.

Taddeo, M. (2010b). Trust in technology: A distinctive and a problematic relation. *Knowledge, Technology & Policy, 23*(3), 283-286.

Taddeo, M. (2011). The role of e-Trust in distributed artificial systems. In C. Ess, & M. Thorseth (Eds.), *Trust and Virtual Worlds* (pp. 75-88). New York: Peter Lang.

Taddeo, M., & Floridi, L. (2011). The case for e-trust. *Ethics and Information Technology, 13*(1), 1-3.

Tanz, J. (2014). How Airbnb and Lyft finally got Americans to trust each other. http://www.wired.com/2014/04/trust-in-the-share-economy/. Accessed 26 April 2015.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything.* New York: Portfolio.

Tapscott, D., & Williams, A. D. (2010). *Macrowikinomics: Rebooting business and the world.* New York: Portfolio/Penguin.

Tene, O. (2008). What Google knows: Privacy and internet search engines. *Utah Law Review, 4*, 1433-1492.

The-Online-Initiative (2015). The onlife manifesto. In L. Floridi (Ed.). Cham: Springer.

Thies, C. (2009). *Einführung in die Philosophische Anthropologie.* Darmstadt: WBG.

Thomas, C. (2014, 25 September). Iedereen kapitalist. De deeleconomie en de idylle van het dorpsplein. *De Groene Amsterdammer,* pp. 20-25.

Thurman, N., & Schifferes, S. (2012). The future of personalization at news websites: Lessons from a longitudinal study. *Journalism Studies, 13*(5-6), 775-790.

Tomlinson, J. (1994). A phenomenology of globalization? Giddens on global modernity. *European Journal of Communication, 9*(2), 149-172.

Torpey, J. (2000). *The Invention of the Passport: Surveillance, Citizenship, and the State.* Cambridge: Cambridge University Press.

Turilli, M., Vaccaro, A., & Taddeo, M. (2010). The case of online trust. *Knowledge, Technology & Policy, 23*(3-4), 333-345.

Turkle, S. (1984). *The second self: Computers and the human spirit.* New York: Simon and Schuster.

Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet.* New York: Simon & Schuster.

Turkle, S. (2011). *Alone together. Why we expect more from technology and less from each other.* New York: Basic Books.

van den Berg, B. (2009). *The Situated Self*. Erasmus University Rotterdam, Rotterdam.

van den Berg, B., & Keymolen, E. (2013). Techniekfilosofie: Het medium is de maat. *Wijsgerig Perspectief, 53*(1), 8-17.

van den Hoven, M. J. (1998). Moral responsibilty, public office and information technology. In I. Snellen, W. van de Donk (Ed.), *Public Administration in an Information Age* (pp. 97-112). Amsterdam: IOS Press.

van der Hof, S., & Prins, C. (2008). Personalisation and its influence on identities, behaviour and social values. In M. Hildebrandt, & S. Gutwirth (Eds.), *Profiling the European citizen. Cross-Disciplinary perspectives* (pp. 111-127). Amsterdam: Springer Netherlands.

van der Sloot, B., & Borgesius, F. Z. (2012). Google and personal data protection. In A. Lopez-Tarruella (Ed.), *Google and the law* (pp. 75-111). Den Hague: Asser Press/ Springer.

van Dijk, J. (2012). *The Network Society.* (3rd ed.)London: Sage.

van Eeten, M. J., & Mueller, M. (2013). Where is the governance in internet governance? *new media & society, 15*(5), 720-736.

van Eijk, N., Helberger, N., Kool, L., van der Plas, A., & van der Sloot, B. (2012). Online tracking: questioning the power of informed consent. *info, 14*(5), 57-73.

van Est, R., Rerimassie, V., van Keulen, I., & Dorren, G. (2014). Intieme technologie. (T. Assessment, Trans.). Den Haag: Rathenau Instituut.

van Noort, W. Huis te huur staan op Airbnb? De Belastingdienst kijkt mee. (2015, 13 July). *NRC Handelsblad,* p. 5.

Vanderstraeten, R. (2002). Parsons, Luhmann and the theorem of double contingency. *Journal of Classical Sociology, 2*(1), 77-92.

Verbeek, P.-P. (2000). *De daadkracht der dingen.* Amsterdam: Boom.

Verbeek, P.-P. (2011a). *De grens van de mens: Over techniek, ethiek en de menselijke natuur.* Rotterdam: Lemniscaat.

Verbeek, P.-P. (2011b). *Moralizing technology: Understanding and designing the morality of things.* Chicago ; London: The University of Chicago Press.

Verbeek, P.-P. (2015). Designing the public sphere: Information technologies and the politics of mediation. In L. Floridi (Ed.), *The Onlife Manifesto. Being human in a hyperconnected era* (pp. 217-227). Cham: Springer.

Vermaas, P. E., Tan, Y.-H., van den Hoven, J., Burgemeestre, B., & Hulstijn, J. (2010). Designing for trust: A case of value-sensitive design. *Knowledge, Technology & Policy, 23*(3-4), 491-505.

Watts, C. (2012). A brief introduction to retargeting. http://www.retargeter.com/retargeting/a-brief-introduction. Accessed 08 january 2013.

Weckert, J. (2011). Trusting software agents. In C. Ess, & M. Thorseth (Eds.), *Trust and Virtual Worlds* (pp. 89-119). New York: Peter Lang.

Weibel, P. (1992). New space in the electronic age. In E. Bolle (Ed.), *Book for the unstable media* (pp. 65-75). Den Bosch: V2.

Weiland, J. S. (1999). *De mens in de filosofie van de twintigste eeuw.* Rotterdam: Aula.

White, M. C. (2014). Skipping the front desk, and checking In with a click. http://www.nytimes.com/2014/11/04/business/hotels-test-turning-guests-smartphonoes-into-room-keys-.html?_r=0. Accessed 09 June 2015.

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121-136.

Winner, L. (2001). Where technological determinism went. In S. Cutcliffe, & C. Mitcham (Eds.), *Visions of STS. Counterpoints in science, technology, and society studies* (pp. 11-18). Albany: State University of New York Press.

Wittgenstein, L. (2009 [1953]). *Philosophical investigations.* Oxford: Wiley-Blackwell.

Woolgar, S., & Cooper, G. (1999). Do artefacts have ambivalence? Moses' bridges, Winner's bridges and other urban legends in S&TS. *Social studies of science*, 433-449.

Workinggroup (2005). Report on internet governance. Château de Bossey.

WRR (2008). *Onzekere veiligheid. Verantwoordelijkheden rond fysieke veiligheid.* Den Haag/Amsterdam: WRR.

Wu, T. (2011). *The master switch: The rise and fall of information empires.* New York: Alfred A. Knopf.

Zand, D. E. (1972). Trust and managerial problem solving. *Administrative Science Quarterly, 17*(2), 229-239.

Zervas, G., Proserpio, D., & Byers, J. (2014). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Boston U. School of Management Research Paper.*(No 2013-16), Available at SSRN: http://ssrn.com/abstract=2366898.

Zervas, G., Proserpio, D., & Byers, J. (2015). A first look at online reputation on Airbnb, where every stay is above average. *Available at SSRN:* http://ssrn.com/abstract=2554500.

Zittrain, J. (2008). *The future of the internet and how to stop it.* New Haven [Conn.]: Yale University Press.

Zwenne, G. J. (2013). *Diluted privacy law*. Leiden University, Leiden.

11

Samenvatting

Ondanks de berichtgeving over de NSA die op grote schaal het Internet surveilleert, de cybercriminaliteit die groeit en de online platforms die winst maken door persoonlijke data te verzamelen en verkopen, lijken mensen niet het vertrouwen te verliezen in de online wereld. In tegendeel. Mensen maken in toenemende mate gebruik van diensten en producten online. Als verwondering ten grondslag ligt aan filosofisch onderzoek, dan is de volgende paradoxale verwondering het startpunt voor dit boek:

> "hoe kan het dat mensen zo ogenschijnlijk eenvoudig vertrouwen stellen in hun slimme internet-apparaten, terwijl er duidelijke redenen zijn om dit niet te doen."

In mijn onderzoek ben ik erachter gekomen dat deze tegenstelling in feite de brandstof is waarop vertrouwen draait: niet met zekerheid weten wat de toekomst brengt en desalniettemin handelen. Vertrouwen is handelen *alsof* je zeker weet wat de toekomst brengt terwijl *je zeker weet dat je nooit zeker zal weten* hoe zaken zullen lopen. In die zin kan vertrouwen dan ook gezien worden als een sprong, het overbruggen van een hiatus, een verschil tussen de stand van zaken in het heden en die in de toekomst. De relatie tussen vertrouwen en *contingentie* is dan ook cruciaal. Immers, als we wel zeker zouden weten hoe de toekomst zich ontplooit, dan was vertrouwen overbodig. Dan zouden we geen sprong hoeven wagen, we zouden niet hoeven doen *alsof* we zeker weten hoe het verder gaat, *we zouden het gewoon weten*. Vertrouwen is een strategie die het mogelijk maakt om te gaan met onzekerheid *inherent aan het menselijk bestaan*. Vertrouwen heft die onzekerheid niet op maar

reduceert het tot een dragelijk niveau, waardoor die onzekerheid uiteindelijk haar problematisch karakter verliest en handelen mogelijk wordt (Möllering 2006:6).

Gebaseerd op het vroege werk van de socioloog Niklas Luhmann over vertrouwen – die de functie van vertrouwen benoemt als het reduceren van complexiteit-, plaats ik vertrouwen in een *familie van concepten* die in samenhang vertrouwen duiden. De leden van deze familie zijn: *interpersoonlijk vertrouwen*, *systeem vertrouwen* (in het Engels: *system trust* oftewel *confidence*), een *vertrouwde wereld* en de *reductie van complexiteit*. De aanname is dat vertrouwen –zowel vertrouwen tussen mensen alsook in systemen- zich altijd afspeelt in een al enigszins vertrouwde wereld.

Om te begrijpen hoe vertrouwen 'werkt' wanneer er slimme apparaten en online omgevingen de interactie mediëren, heb ik het *4 Cs raamwerk* ontwikkeld die het mogelijk maakt vertrouwen te analyseren op een gelaagde wijze die recht doet aan de netwerk-ontologie van deze artefacten. Het 4 Cs raamwerk bestaat uit de volgende onderdelen: *Context, Curatie, Codificatie en Constructie*.

Context refereert aan de wijze waarop gebruikers hun interacties ervaren gemedieerd door slimme artefacten.

Curatie staat voor de actoren die de slimme artefacten en internet omgevingen vormgeven en beheren. Voorbeelden hiervan zijn overheden en private partijen (zoals Google of Facebook).

Codificatie staat voor de regels die curatoren opstellen voor het gebruik van slimme artefacten en internet omgevingen. Afhankelijk van de casus waarnaar gekeken wordt kan het met name gaan om wetgeving, maar ook afspraken tussen betrokken partijen over het gebruik van het artefact of de dienst, het privacy beleid,…

Constructie, tenslotte, verwijst naar het ontwerp van het artefact zelf. Wat is mogelijk en onmogelijk voor de gebruiker? Worden er data verzameld en hoe wordt daar mee omgegaan?

Dit 4 Cs raamwerk voorziet in een analytisch kader waarmee casussen geanalyseerd kunnen worden zoals gedaan is in hoofdstuk 6,7, en -in meer algemene zin- in 8. Anderzijds kan het ingezet worden in meer evaluatieve praktijken waarbij bijvoorbeeld ethici of beleidsadviseurs het 4Cs raamwerk gebruiken om vertrouwen in bepaalde diensten te beoordelen. Of het kan gebruikt worden als instrument om

vertrouwenswaardige artefacten en diensten te ontwikkelen door vertrouwen op het niveau van de 4 Cs in het ontwerp te waarborgen.

Het proefschrift is als volgt opgebouwd:

Hoofdstuk 1 betreft de introductie.

In Hoofdstuk 2 wordt vertrouwen op het ontologisch niveau geanalyseerd. Op basis van het werk van Niklas Luhmann en Helmuth Plessner, wordt het overbruggen van de hiatus als belangrijkste verklarende kracht verder uitgediept.

In Hoofdstuk 3 wordt vertrouwen in een sociologisch/historisch kader geplaatst waarbij wordt geargumenteerd dat interpersoonlijk vertrouwen in de laat-moderne tijd onder invloed van grote, mondiale systemen zoals het bankensysteem en het luchtverkeer steeds vaker plaats maakt voor systeem vertrouwen.

In Hoofdstuk 4 staat het Internet als infrastructuur centraal. De vraag die ten grondslag ligt aan dit hoofdstuk is of het Internet kan functioneren als een vertrouwde wereld. Cruciaal hiervoor is de rol van de curatoren van het Internet en in welke mate hun eigen belangen in lijn zijn met het creëren en onderhouden van een stabiel Internet.

In Hoofdstuk 5 wordt gekeken naar het micro-niveau, naar de ervaring van de gebruikers van het Internet en slimme artefacten. Gebaseerd op onder andere het werk van Helmuth Plessner en mediatie theorie analyseert dit hoofdstuk hoe vertrouwen tot stand komt in het netwerktijdperk.

Hoofdstuk 6 behandelt de eerste casus en laat zien hoe het 4 Cs raamwerk kan gebruikt worden om vertrouwen in een specifieke gebruikerscontext te duiden. De casus betreft de introductie van digitale hotelsleutels op smartphones en toont hoe door het in gebruik nemen van nieuwe technologieën in de hotel business, de vertrouwensrelatie tussen hotel en klant verandert.

Hoofdstuk 7 focust op het platform Airbnb dat wordt gebruikt om mensen die een

kamer of huis verhuren voor tijdelijk verblijf in contact te brengen met mensen die op zoek zijn naar zo een plek. De idee van de “nieuwe digitale economie” die aan dit initiatief ten grondslag ligt is dat door middel van technologie, oude vormen van interpersoonlijk vertrouwen hersteld kunnen worden. Door in de analyse het 4 C raamwerk toe te passen, wordt het echter duidelijk dat niet een oude vorm van vertrouwen wordt hersteld maar een nieuwe vorm tot stand komt: interpersoonlijk systeem vertrouwen.

Hoofdstuk 8 gaat dieper in op de ontwikkeling om in toenemende mate informatie-omgevingen (online, maar in de toekomst zeker ook vaker offline) te personaliseren. In dit hoofdstuk staat de vraag centraal of een proactieve, gepersonaliseerde informatie-omgeving ook leidt tot een vertrouwde wereld waarin vertrouwen kan floreren. De conclusie is dat personalisatie kan leiden tot een té vertrouwde wereld die wel zeer herkenbaar is voor het individu maar waar het aan gedeelde waarden en perspectieven ontbreekt. Bovendien worden mensen door de omgeving ‘gelezen’ zonder dat ze zij zelf in staat zijn de omgeving ‘te lezen’ waardoor de interactie tussen mens en omgeving in toenemend mate eenzijdige van aard wordt.

Tenslotte wordt in de epiloog teruggeblikt op de belangrijkste ontwikkelingen en uitdagen voor vertrouwen in het netwerktijdperk.

12

About the author

Esther Keymolen (1982) studied Music (Bachelor) at Codarts Rotterdam and Philosophy (Bachelor with honours, Master with distinction) at the Erasmus University in Rotterdam. During her study she worked as a student-assistent to Prof. Jos de Mul and Prof. Mireille Hildebrandt.

From 2008-2011, she has worked as a scientific staff member at the Scientific Council for Government Policy (WRR). She co-authored the book iGovernment and conducted research in the domain of digital youth care.

In 2011 she started her PhD research, funded by TNO, at the Faculty of Philosophy, in the research group Man and Culture. During her time as a PhD candidate, she taught several courses and presented her work at various national and international conferences. She has been the representative for the Faculty of Philosophy in the PhD Council of the Erasmus Graduate School of Social Sciences and the Humanities. She also was the chairperson of the PhD council of the Dutch Research School of Philosophy (OZSW).

As of July 2014, she is a lecturer and academic coordinator of the Advanced Master Programme Law and Digital Technologies at the University of Leiden (eLaw).

Key publications

Keymolen, E. (2014). A moral bubble. The influence of online personalization on moral repositioning. In J. de Mul (Ed.), *Plessner's philosophical*

anthropology. Perspectives and prospects (pp. 387-406). Amsterdam: Amsterdam University Press.

Keymolen, E., & Broeders, D. (2013). Innocence Lost: Care and Control in Dutch Digital Youth Care. *British Journal of Social Work, 43*(1), 41-63.

Keymolen, E. L. O., Prins, J.E.J, Raab, C. (2012). Trust and ICT: New Challenges for Public Administration In v. d. Donk, W., Thaens, M. (Ed.), *The Coming of Age of ICT in Public Administration* (pp. 21-35). Amsterdam: IOS Press.

van den Berg, B., & Keymolen, E. (2013). Techniekfilosofie: Het medium is de maat. *Wijsgerig Perspectief, 53*(1), 8-17.

van der Hof, S., & Keymolen, E. (2010). Shaping minors with major shifts: Electronic child records in the Netherlands. *Information Polity, 15*(4), 309-322.