Ontological Foundations for Trust Dynamics: The Case of Central Bank Digital Currency Ecosystems

Chapter · May 2022

DOI: 10.1007/978-3-031-05760-1_21

3 authors:

Glenda Amaral
Banco Central do Brasil

Tiago Prince Sales
University of Twente

Giancarlo Guizzardi
University of Twente

Ontological Foundations for Trust Dynamics: The Case of Central Bank Digital Currency Ecosystems

Glenda Amaral[1], Tiago Prince Sales[1], and Giancarlo Guizzardi[1,2]

[1] Conceptual and Cognitive Modeling Research Group (CORE), Free University of Bozen-Bolzano, Bolzano, Italy
{gmouraamaral, tiago.princesales, giancarlo.guizzardi}@unibz.it
[2] Services & Cybersecurity, University of Twente, The Netherlands

Abstract. In recent years, disruptive technologies have advanced at a rapid pace. These new developments have the power to accelerate the production and delivery, improve the quality, and reduce the costs of goods and services, as well as to contribute to individual and collective well-being. However, their adoption relies largely on user trust. And trust, due to its dynamic nature, is fragile. Therefore, just as important as to build trust is to maintain it. To build sustainable trust it is fundamental to understand the composition of trust relations and what factors can influence them. To address this issue, in this paper, we provide ontological foundations for trust dynamics. We extend our previous work, the Reference Ontology of Trust (ROT), to clarify and provide a deeper account of some building blocks of trust as well as the many factors that can influence trust relations. We illustrate the working of ROT by applying it to a real case study concerning citizens' trust in central bank digital currency ecosystems, which has been conducted in close collaboration with a national central bank.

Keywords: trust dynamics · UFO · OntoUML · central bank digital currency.

1 Introduction

New and disruptive technologies have been developed at a rapid pace, affecting almost every area of our lives. Industrial robots, artificial intelligence algorithms, machine learning, big data, decentralized technologies, just to cite a few, have the power to accelerate the production and delivery, improve the quality, and reduce the costs of goods and services, as well as to contribute to individual and collective well-being. However, the adoption of these innovative technologies relies largely on user trust. And trust is highly dynamic. Trust is generally said to be one of the easiest things to lose and one of the most difficult things to win back. It may break in an instant or erode gradually. Therefore, it is important to build sustainable trust that is not easily lost. In the case of information technology systems and ecosystems, building sustainable trust involves addressing stakeholders' trust concerns from the system (or ecosystem) inception to their constant monitoring, as trust changes with time. But how to identify these trust concerns? What makes one trustworthy? And what factors can influence trust? In this paper we address these questions via an ontological analysis and representation of trust dynamics.

In a previous work [2], we proposed the Reference Ontology of Trust (ROT) - an ontologically well-founded reference model, grounded in the Unified Foundational Ontology (UFO) [15,16], which formally characterizes the concept of trust, as well as clarifies the relation between trust and risk, and represents how risk emerges from trust relations. This paper sheds new light on trust-related concepts and relations under the perspective of trust dynamics. We extend our previous work to clarify and provide a deeper account of (i) the different factors that can influence trust; (ii) the signals that trustees can emit to indicate their trustworthy behavior; and (iii) pieces of evidence that suggest that a trustee should be trusted. We validate and illustrate the use of ROT with a real case study on citizens' trust in central bank digital currency (CBDC) ecosystems, which was conducted in close collaboration with a national central bank.

This paper is organized as follows. In Section 2 we introduce the reader to the Reference Ontology of Trust. Then, in Section 3 we present the extensions to the ontology to provide ontological foundations for trust dynamics. We apply ROT in a real case study and discuss its results in Section 4. We assess related work in section 5 and conclude the paper in Section 6 with some final considerations.

2 The Reference Ontology of Trust (ROT)

The Reference Ontology of Trust[1] (ROT) is a well-founded ontology, based on the Unified Foundational Ontology (UFO) and specified in OntoUML [15]. It formally characterizes the concept of trust, clarifies the relation between trust and risk, and represents how risk emerges from trust relations [2]. ROT makes the following ontological commitments about the nature of trust:

- **Trust is relative to an intention.** An agent, the trustor, trusts an individual, the trustee, only relative to a certain intention, on which achievement she counts on the trustee. The trustor may trust in the trustee regarding a certain intention, but not bestow such a trust regarding a different one. For example, I trust my dentist to fix a cavity in my tooth, but not to fix my computer. Furthermore, such an intention is not always atomic. For example, in the trust relation "Bob trusts a certain airline to take him on his holiday trip comfortably and safely", trust is about a complex intention, composed of (i) Bob's intention of traveling; (ii) his intention of being safe; and (iii) his intention of being comfortable.
- **A trustor is an "intentional entity".** Trustors are cognitive agents, being endowed with intentions and beliefs [9].
- **A trustee does not need agency.** A trustee is an individual capable of impacting one's intentions by the outcome of its behavior [9]. A trustee may be either an agent (e.g. a person, an organization) or an object (e.g. a car, a vaccine).
- **Trust is a complex mental state.** Trust is a mental state of a trustor regarding a trustee and its behavior. It is composed of: (i) an intention; (ii) beliefs that the trustee can perform a desired action or exhibit a desired behavior; (iii) beliefs that the trustee's vulnerabilities will not prevent it from performing the desired action

[1] The complete version of ROT in OntoUML and its implementation in OWL are available at http://purl.org/krdb-core/trust-ontology.

or exhibiting the desired behavior; and (iv) if the trustee is an agent, beliefs that the trustee intends to exhibit that behavior. For example, a mother who trusts a babysitter to take care of her kids believes that: (i) the babysitter has experience in caring for children and is First Aid trained (a belief about the babysitter's capabilities); (ii) the babysitter is well and probably is not going to have health issues (a belief about the babysitter's vulnerabilities); and (iii) the babysitter is willing to take good care of her children (a belief about the babysitter's intention).

– **Trust is context-dependent.** A trustor may trust a trustee for a given goal in a given context, but not do so for the same goal in a different context. For example, Mary may trust her brother to drive her to the train station in a sunny day, but she does not trust him to do so when it is snowing. We assume trust relations to be highly dynamic [9].
– **Trust implies risk.** By trusting, the trustor becomes vulnerable to the trustee in terms of potential failure of the expected behavior or outcome [17, p 21]. In trust relations, risk can emerge as consequence of either the manifestation of a trustee's vulnerability or the unsatisfactory manifestation of a trustee's capability. In the above-mentioned babysitter example, the babysitter getting sick during the term of the employment contract corresponds to the manifestation of a babysitter's vulnerability that prevents her from going to work and ultimately can hurt the mother's intention of having an adult to take care of her kids.
– **Trust can be quantified.** Our trust in a certain individual can increase or decrease in time, and we can trust certain individuals more than others. To account for these scenarios, ROT assumes that trust can be quantified, even if it does not commit to any particular scale or measurement strategy. In ROT, the quantitative perspective of trust is captured by means of (i) the trust degree — the extent to which the trustor trusts in the trustee; (ii) the belief intensity — the strength of a trustor's belief ; (iii) the performance level — how well the trustor believes the trustee can perform the action; and (iv) the manifestation likelihood — how strongly the trustor believes a disposition of the trustee may be manifested through the occurrence of certain events.

The aforementioned ontological commitments are captured in the ROT diagram presented in Figure 1. In the OntoUML diagrams depicting the ontology, we adopt the following color coding: substantials are represented in pink, modes and qualities in blue, relators in green, and classes whose instances might be of different ontological nature in gray. The reader interested in an in-depth description of the complete version of ROT is referred to [2].

3 Modeling Trust Dynamics

We previously mentioned that trust is composed of a trustor's intention and a set of her beliefs about the trustee and its behavior. However, several other factors that influence trust have been discussed in the literature [18]. For instance, Castelfranchi and Falcone [9] argue that "trust changes with experience, with the modification of the different

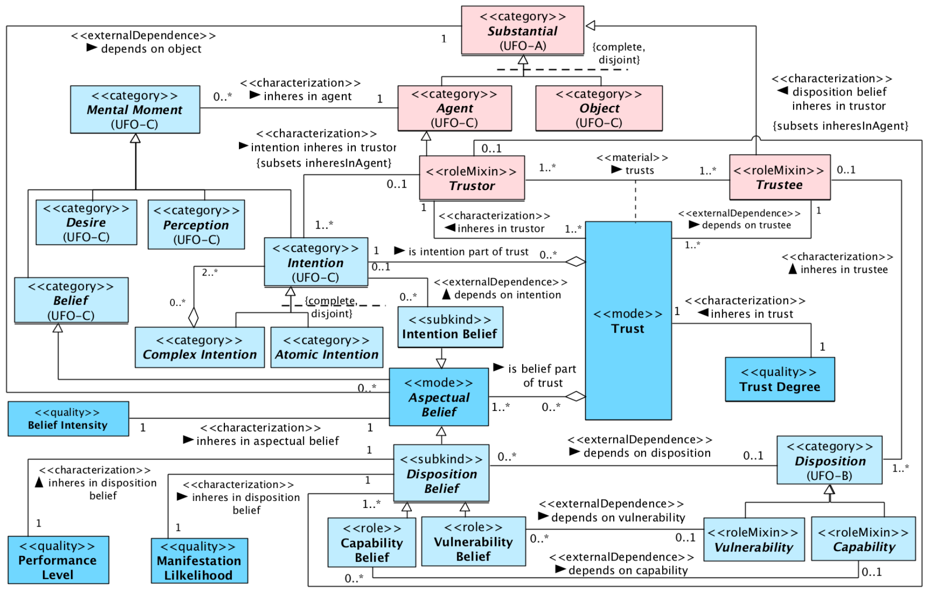

Fig. 1: Reference Ontology of Trust.

sources it is based on, with the emotional or rational state of the trustor, with the modification of the environment in which the trustee is supposed to perform, and so on". They claim that trust is a dynamic entity because it depends on dynamic phenomena.

In this section, we extend ROT to allow it to account for trust dynamics more adequately by categorizing influence relations according to the ontological nature of the factors that explain them. These factors are: (F1) trust influencing trust (Section 3.1); (F2) mental biases (Section 3.2); (F3) trust calibration signals (Section 3.3); and (F4) trustworthiness evidence (Section 3.4). In the diagrams depicted in figures 2 and 3, the colored concepts represent the extensions proposed to ROT. The concepts in white belong to the original version of the ontology. In the remaining of this section, we present a detailed description of these extensions.

3.1 Trust Influencing Trust

This category represents the situation in which a trust relationship is influenced by another trust relationship. According to Castelfranchi and Falconi [9] "in the same situation trust is influenced by trust in several rather complex ways". In the same work they also discuss the phenomenon of trust creating reciprocal trust, as well as how trust relations can influence each other. In fact, countless examples can be found in real life about trust influencing trust, either positively or negatively. For instance, citizens' trust in the central bank positively influence their trust in the national currency. People's trust in the healthcare system, in the experts defining vaccination strategies, and more generally in government bodies influence their trust in vaccines.

3.2 Mental Biases

This category represents situations in which trust is influenced by *mental moments* (a concept from UFO). *Mental moments* refer to the capacity of some properties of certain individuals to refer to possible situations of reality [16]. A *mental moment* is existentially dependent on a particular agent, being an inseparable part of its mental state (Figure 1). Examples include beliefs, desires and intentions. Beliefs have a propositional content that agents consider to be true. They can be justified by situations in reality. Examples include my belief that Rome is the Capital of Italy, and the belief that the Moon orbits the Earth. Beliefs can be formed by perceptions expressing how agents sense their environment and the things that happen around them. Desires and intentions can be fulfilled or frustrated. A desire expresses the will of an agent towards a possible situation (e.g., a desire that Italy wins the next World Cup), while an intention expresses desired states of affairs for which the agent commits to pursuing (e.g., Mary's intention of going to Paris). For an extensive discussion of mental moments, please refer to [16].

Mental moments can significantly influence trust relationships. Let us consider the example of a person who really wants to travel but cannot. One day she receives an email containing an unbelievable offer for an exotic destination that is just about to expire. Although many will immediately think it is a scam, the person's desire to travel may influence her to trust the email offer [13]. Another example would be people who are strongly committed to environmental preservation and tend to trust companies that support environmental sustainability. There is also the case of beliefs not related to specific trustees. An example discussed by [19] suggests that some religious beliefs, which prescribe honesty and mutual love, lead people to generally assume that others are usually honest, benevolent, competent, and predictable.

Another important aspect is the occurrence of events that can affect one's perception regarding a trustee. McKnight et al. [20] discuss how trust changes in response to external events and propose a model that addresses the mental mechanisms people use as they are confronted by trust-related events, which "indicates that trust may be sticky or resistant to change, but that change can and will occur" [20]. Castelfranchi and Falconi [12] claim that the success of an action performed by the trustee in order to reach a goal of the trustor depends not only on the trustee's capabilities but also on external conditions that allow or inhibit the realization of the task. To illustrate this point, the authors use the case of a violinist that will give a concert in an open environment. In general, people trust the violinist to play well. However, if it is particularly cold during the concert, their trust will decrease if they infer that the cold can hinder her ability to play. Similarly, in financial systems, the emergence of detrimental information about a financial agent can negatively affect public trust in this agent, which can lead to considerable adverse effects on one or several other financial institutions that can ultimately propagate to the entire financial system.

3.3 Trust Calibration Signals

The emission of *trust-warranting signals*, that is, signals that indicate trustworthy dispositions of a trustee, is one of the ingredients for building sustainable trust [21]. In trust relations, once the trustee's capabilities and vulnerabilities related to the beliefs of the trustor are known, it is possible to reason about the signals that the trustee should emit to indicate that it can successfully realize its capabilities and prevent its vulnerabilities from being manifested. These signals are specifically created to indicate a trustworthy behavior on the part of the trustee and therefore can influence trust. For example, information about how privacy and security measures are implemented could be provided as signals of the trustworthiness of a system. Another example is the establishment of a universal brand to create visual identity, so that users can identify the system interface elements in a clear and unambiguous way, thus facilitating the understanding and adoption of its functionalities.

Equally important are *uncertainty signals*, i.e. signals that communicate uncertainties regarding the realization of capabilities and the prevention of vulnerabilities. Some examples are the publication of uncertainties about the accuracy of scientific findings, patient communication of uncertainties on the precision of medical diagnosis, investor communication of uncertainties in forecasting financial investments returns, communication to the public about uncertainties regarding the efficacy of vaccines, among others. While trust-warranting signals contribute to trust building, uncertainty signals allow trustors to adjust their trust level appropriately to the trustee's trustworthiness, thus avoiding misplaced levels of trust. Research show that communicating uncertainty can be beneficial for maintaining trust and commitment over time [5,23]. This is because building trust that is higher than the actual trustworthiness of the trustee might set trustors' expectations too high, which may result in disappointment sooner or later.

We extend ROT to represent trust signals emitted by trustees (F3) in the following way. As illustrated in Figure 2, the `Trustee` may emit `Trust Calibration Signals` regarding its `Dispositions` (either a `Capability` or a `Vulnerability`). `Trust Calibration Signal` is specialized into `Trust-warranting Signal` and `Uncertainty Signal`. The former represents trust-warranting signals that should be emitted by the trustee in order to ensure trustworthy behavior, while the latter represents uncertainty signals emitted by the trustee, which allow trustors to adjust their trust levels[2].

3.4 Trustworthiness Evidence

Another trust influencing factor corresponds to *trustworthiness evidence*, pieces of evidence that can make a trustor believe that the trustee should be trusted. Similarly to trust-warranting signals, they suggest that a trustee can realize its capabilities and shield its vulnerabilities. However, differently from signals, which are purposefully emitted to suggest trustworthiness, evidence result from the observation of a trustees' trustworthy actions. Examples include:

- third-party certifications and credentials (e.g. John's TOEFL certification makes me believe that he can speak English, because I trust the certificate issued by a certain authority);
- performance history (e.g. accuracy of a medical diagnosis system);

[2] `Emits` is grounded on a communicative act of the trustee [16] and, hence, a *historical* relation in the sense of [14]. The propositional content of this act refers to a disposition, thus, grounding the (derived) `refers to` relation between the latter and `Trust Calibration Signal`.

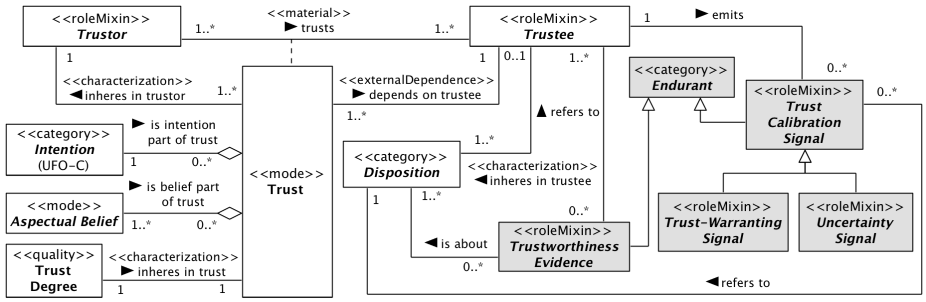

Fig. 2: ROT - Trust Calibration Signals and Trustworthiness Evidence Extensions.

- track record (e.g. reviews from service recipients and statistics on its experience);
- recommendations (e.g. my brother trusts a car mechanic and recommends his services to me);
- reputation records (e.g. positive evaluations received by an Uber driver);
- availability (e.g. a medical doctor you rarely succeed to make an appointment with is not trustworthy);
- past successful experiences (e.g. all the products I purchased at Amazon arrived on time and in perfect condition);
- transparency (e.g. offering information on what an artificial intelligence system is doing, as well as rationale for its decisions (aka explainability));
- longevity (e.g. indication that a vendor has been in the market for a long time and that it is interested in continued business relationship with the client); and
- risk mitigation measures, which indicate that one is actively trying to prevent the manifestation of one's vulnerabilities.

Ontologically speaking, a piece of `Trustworthiness Evidence` is a *social endurant*, typically a *social relator*[3] (e.g. a relator binding the certifying entity, the certified entity and referring to a capability, vulnerability, etc.), but also documents (*social objects* themselves) that represent these *social entities* (e.g., in the way a marriage certificate documents a marriage as a social relator). As illustrated in Figure 2 we extended ROT to model `Trustworthiness Evidences` (F4) as "roles" played by *endurants* (objects, relators, etc.) related to a `Disposition` of the `Trustee`[4].

To represent the role of influences, we included the `Influence` relator, which connects the sources of influence to the aspectual beliefs of the trustor under their influence (Figure 3). We distinguish `Influence` according to the source of influence into: (i) `Trust Influence`, associated to a `Trust` relationship (F1); (ii) `Mental Moment`

[3] Briefly speaking, a relator (a concept from UFO) is an entity that is existentially dependent on at least two individuals, thus, mediating or binding them [14].

[4] Playing of the "role" of `Trustworthiness Evidence` for a particular focal disposition is dependent on the belief of trustors, whose propositional content makes that connection between that player and that disposition. The `is about` relation in this model is, thus, derived from the propositional content of that belief. The `refers to` relation connected to the trustee is derived from the relation between that focal disposition and its bearer.

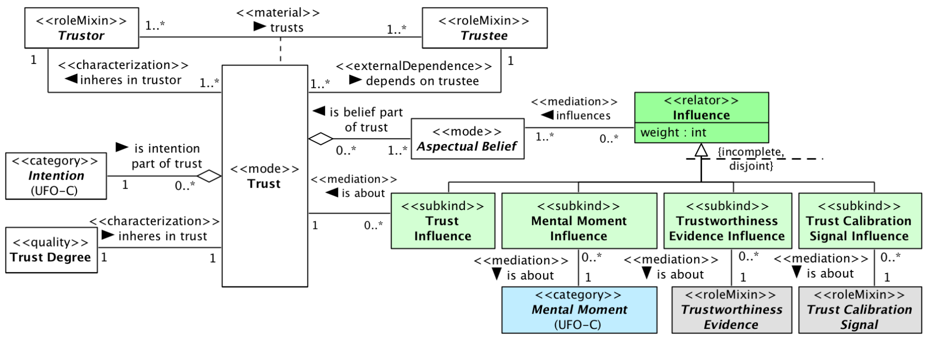

Fig. 3: ROT - Influences Extensions.

`Influence`, associated to a `Mental Moment` (F2); (iii) `Trust Calibration Signal Influence`, associated to a `Trust Calibration Signal` (F3); and (iv) `Trustworthiness Evidence Influence`, associated to a `Trustworthiness Evidence` (F4). The property `weight` corresponds to the weight of an influence over a particular belief, as certain influences may weight more heavily than others.

4 Case Study: Citizens' Trust in CBDC Ecosystems

In this section, we present a real-world case in which we use ROT to model citizens' trust in CBDC ecosystems. We worked in close collaboration with a national central bank to analyse the case and instantiate the ontology to represent the dynamics of citizens' trust in CBDC ecosystems. Due to the sensitivity of this topic (the development of CBDC ecosystems is in full swing, and their design is not finished yet) and specific request of the central bank, the only information we can disclose is that the contributing central bank's context is a country with between 50 and 300 million citizens.

4.1 Research Method

We conducted a case study, in collaboration with a national central bank, regarding citizens' trust in CBDC ecosystems. This methodological approach is particularly appropriate when the focus is investigating a contemporary phenomenon in depth and within its real-life context, and the investigator has no control over actual behavioral events [24]. The research procedure we employed was adapted from [24].

We started by planning the case study. We defined its purpose — *to verify if the Reference Ontology of Trust can properly represent real world situations, or more specifically if it can model citizens' trust in CBDC ecosystems* — identified the areas of interest, namely, economics, financial citizenship and information technology, and selected the interviewees. We also obtained the necessary authorizations from the central bank to carry out the study.

In the collect stage, we gathered information from documentation and interviews. First, documents describing and documenting information on citizens' trust in CBDC

Table 1: Questions related to key ontology concepts.

Question	ROT Concept
Citizens trust the CBDC ecosystem to...	Intention
Citizens trust the CBDC ecosystem because they believe that it can...	Belief Capability
Citizens trust the CBDC ecosystem because they believe that it has mechanisms to prevent...	Belief Vulnerability
How can the CBDC ecosystem indicate that it is trustworthy?	Trust Calibration Signal
What pieces of evidence show the CBDC ecosystem is trustworthy?	Trustworthiness Evidence
What can influence citizens' trust in the CBDC ecosystem?	Influence

ecosystems were collected from the literature [4,6,7,8,11,22] and from the central bank's website, to deepen our knowledge about the topic. Based on this documentation, we created an initial version of the ontology instantiation, to be validated and complemented at the interviews stage. Then, we conducted interviews with central bank experts in the areas of interest, namely, economics, financial citizenship and information technology. The questions that would guide the interviews with the stakeholders (Table 1) were defined based on the main concepts of the Reference Ontology of Trust (Sections 2 and 3). The ontology served as guidance for our work from the beginning of the case study, helping us focus on the domain being investigated and supporting the creation of the interview questions. As shown in Table 1, these questions are actually formulated based on the concepts from ROT (see column 2).

We conducted three individual interviews in the form of guided conversations, one for each expert of the areas of interest, namely, economics, financial citizenship and information technology. During the interviews we presented the initial version of the ontology instantiation to be validated and gathered information based on the aforementioned questions (Table 1). The interviews were recorded (audio) and with their feedback and validation we have improved both the ontology and the ontology instantiation, and presented them again to the central bank in a validation meeting. We had in total four sessions with the central bank, in which the modeling of citizens' trust in CBDC ecosystems were discussed in detail.

In the analyze stage, we concluded the final version of the ontology instantiation. In addition, to demonstrate the contribution and applicability of our ontology to the modeling practice, we used the ontology instantiation as a domain model to create a goal model for this case using the *i** framework [10]. Finally, we shared the results with the central bank team.

4.2 Research Context: CBDC Ecosystems

Recent innovations in the financial industry, such as cryptocurrencies, blockchains and distributed ledger technologies, smart contracts, and stablecoins have fostered the creation of financial products and services on top of decentralized technologies, giving rise to the concept of Decentralized Finance (DeFi) [25] — the decentralized provision of financial products and services. This disruption, alongside the entry of big techs into payments and financial services, pushed central banks to investigate new forms of digital money and prepare the grounds for central bank digital currencies (CBDCs). A CDBC is a form of digital money, denominated in the national unit of account, which is a direct liability of the central bank, such as physical cash and central bank settlement accounts [1].

In general, a CBDC ecosystem would comprise elements and functions similar to traditional payment systems, with central banks facing many of the practical policy questions around access, services and structure they currently do. According to [3], at the center of any CBDC ecosystem would be a CBDC "core rulebook" outlining the legal basis, governance, risk management, access and other requirements of participants in the CBDC ecosystem. Participants in the CBDC system could include banks, payment service providers, mobile operators and fintech or big tech companies, which would act as intermediaries between the central bank and end users. This broader ecosystem would be complemented by a legal and supervisory framework and contractual arrangements between end users and their intermediaries [3]. Currently, all CBDC ecosystems are still under design. The initiatives around the world are either at the stage of experimentation, proof-of-concept, or pilot arrangements.

Consumer demand for CBDC is an important element that determines how widely a CBDC would be used. Therefore, the successful implementation of a CBDC crucially depends on citizens' motivation to adopt this new digital form of public money, which is directly related to their trust[5] in the CBDC ecosystem.

4.3 Modeling Citizen's Trust in CBDC Ecosystems

Ontology Instantiation We adopt the following coding to refer to instances of key ROT concepts hereafter: **INT** for intention; **BEL** for disposition belief; **CAP** for capability; **VUL** for vulnerability; **TS** for trust-warranting signal; **US** for uncertainty signal; **TE** for trustworthiness evidence; and **INF** for influence.

Both the literature research and the interviews showed that *citizens trust the CBDC ecosystem to preserve their privacy* (INT1). Privacy emerges as a key feature, which can be confirmed both indirectly — by the presence of comments on the importance of privacy — and directly — by ranking privacy first, among many other features [11]. Citizens who trust the CBDC ecosystem believe that *it safeguards their privacy* (BEL1.1). This belief is related, for example, to the *CBDC ecosystem's capability to comply with the General Data Protection Regulation (GDPR)*[6] *and other privacy laws and regulations* (CAP1.1).

Interviewees also expressed that it is important that citizens feel safe to perform digital transactions in the ecosystem. *Citizens trust the ecosystem to safely make transactions using CBDCs* (INT2). They believe that *the ecosystem is safe* (BEL2.1) and that *it will be able to prevent security breaches* (BEL2.2). The former belief (BEL2.1)

[5] Agustín Carstens, the General Manager of the Bank for International Settlements, in a recent speech at the Goethe University's ILF conference on "Data, Digitalization, the New Finance and Central Bank Digital Currencies: The Future of Banking and Money" explicitly defended that "the soul of money is trust." (https://www.bis.org/speeches/sp220118.htm)

[6] https://gdpr-info.eu/

is related to the *ecosystem's capability to have security mechanisms* (CAP2.1), while the latter (BEL2.2) is related both to *possible security breaches* (VUL2.1)— which correspond to a vulnerability — and to the *ecosystem's capability to quickly react to risk events on security* (CAP2.2). In addition, *the existence of a cybersecurity policy* (TE2.1) is an example of trustworthiness evidence related to the capability CAP2.1.

Another aspect that emerged from the interviews and the literature is the importance of providing a simple experience for end users. *Citizens trust the CBDC ecosystem to make transactions using CBDCs easily* (INT3) and they believe both that *the ecosystem is easy to access and use* (BEL3.1) and that *it is easy to onboard the CBDC ecosystem* (BEL3.2). A possible capability of the ecosystem, related to these beliefs, is *to meet minimum usability criteria* (CAP3.1). The *existence of a manual with minimum usability requirements, which must be followed by all participants of the ecosystem* (TE3.1) is an example of trustworthiness evidence related to the capability CAP3.1. *The establishment of a universal brand to create visual identity* (TS3.1), *advertising campaigns in the media and social networks using everyday examples* (TS3.2), and *documentation available* (TS3.3) are examples of trust-warranting signals related to capability CAP3.1.

Low cost was another attribute mentioned both in the literature and by the interviewees. *Citizens trust the CBDC ecosystem to make transactions using CBDCs at a low cost* (INT4) and they believe both that *it will be offered at a low cost to its users* (BEL4.1) and that *they will not need to buy a new device to make transactions in the CBDC ecosystem* (BEL4.2). The former belief (BEL4.1) is related to the *ecosystem's capability to have lower costs for consumers and merchants* (CAP4.1), while the latter (BEL4.2) is related to the *ecosystem's capability to operate using existing, accessible technology* (CAP4.2).

An additional valuable feature identified in the collect phase is the ability to make offline payments. This feature might be particularly relevant during outages and in environments where internet availability is limited or unreliable [4]. *Citizens trust the CBDC ecosystem to make transactions wherever they need* (INT5) and they believe that *they will be able to access the system from any place* (BEL5.1). This belief is related to the *ecosystem's capability to support offline transactions* (CAP5.1). Note that intention INT5 (offline access) conflicts with intention INT4 (low cost), as technology to support offline capacity may incur additional costs.

The interviews also showed that *citizens trust the CDDC ecosystem to make transactions instantly on a 24/7 basis* (INT6). In other words, users who trust the ecosystem believe that *it is able to make instantaneous transactions* (BEL6.1) and that *it will be available when they need* (BEL6.2). These beliefs are related to the ecosystem's capability to *meet high availability parameters and processing time limits* (CAP6.1). Examples of trustworthiness evidence related to this capability are the existence of *a service level agreement that establishes high availability parameters and processing time limits* (TE6.1) and *statistics on the functioning of the ecosystem showing that this service level agreement has being fulfilled* (TE6.2). Information about *instability* (US6.1) and *low response times* (US6.2) are examples of uncertainty signals.

A further important aspect identified both in the literature and in the interviews is currency acceptance. *Citizens trust the CDBDC ecosystem to make transactions using a widely accepted currency* (INT7). And they believe that *the CBDC ecosystem oper-*

ates with a digital currency widely accepted (BEL7.1). This relates to the *capability to operate using a legal tender currency* (CAP7.1).

Equally important is the stability of the currency purchasing power. *Citizens trust the CDBDC ecosystem to make transactions using a stable currency* (INT8). And they believe that *the CBDC purchasing power has stability* (BEL8.1). This belief relates to the ecosystem's *capability to have proper mechanisms to ensure stability of CBDC purchasing power* (CAP8.1).

Finally, it was also mentioned that *citizens trust the CBDC ecosystem to have access to better financial products and services offerings* (INT9). Therefore, they believe that *they will have access to more product and service offers customized to their needs* (BEL9.1). This relates to the *ecosystem's capability to provide better customized services and products offerings* (CAP9.1). Once more, it is possible to observe the existence of conflicting intentions: the intention just mentioned (INT9) conflicts with privacy preservation (INT1), as to propose better financial services offerings, financial institutions in the ecosystem need to have access to more (private) information about the citizen.

It is important to note that in the trust relation between citizens and the CBDC ecosystem, trust is about a complex intention, composed of the aforementioned intentions (INT 1 to 9).

Furthermore, it is possible to observe the existence of trust influences. For example, *citizens' trust in a country's monetary system* (INF1) positively influences their trust on the CBDC ecosystem, just as *their trust in the central bank* (INF2) does.

Figure 4 shows a graphical representation of the ontology instantiation focusing on usability (INT3). The detailed diagrams presenting the complete case study can be found at `https://purl.org/krdb-core/rot-cbdc-case-study`.

Goal Model We use the ontology instantiation as a domain model to create a goal model for this case using the *i** framework [10], presented in Figure 5. The model shows the goals that citizens delegate to the CBDC ecosystem (through the i* dependency relation). Citizens and the CBDC ecosystem are represented as actor and agent, respectively. Citizens' intentions are represented as quality dependences. Conflicting intentions are represented in yellow, circled by a red dashed line. Entities represented in green and yellow were obtained directly by mapping elements from the ontology instantiation. For each of them, more specific goals, qualities, tasks and resources were identified and are represented in blue. Besides dependencies, the goal model depicts the internal perspective of the CBDC ecosystem. Beliefs are represented as goals or qualities that contribute to (help) the achievement of higher level goals. For example, beliefs such as *the ecosystem is safe* (BEL2.1), *the ecosystem is easy to use* (BEL3.1), *the CBDC is widely accepted* (BEL7.1) were represented as qualities that contribute to the ultimate goal of *being trustworthy*. Capabilities, trust calibration signals and trustworthiness evidence are represented as goals, qualities, tasks or resources that contribute to (help) the achievement of higher level goals. For example, the goal *support offine transactions* (CAP5.1) helps the achievement of *be available*. The tasks *meet minimum usability criteria* (CAP3.1) and *keep visual identity* (TS3.1) help the achievement of *be easy to use*. The resource *cybersecurity policy* (TE2.1) contributes to *comply with cy-*

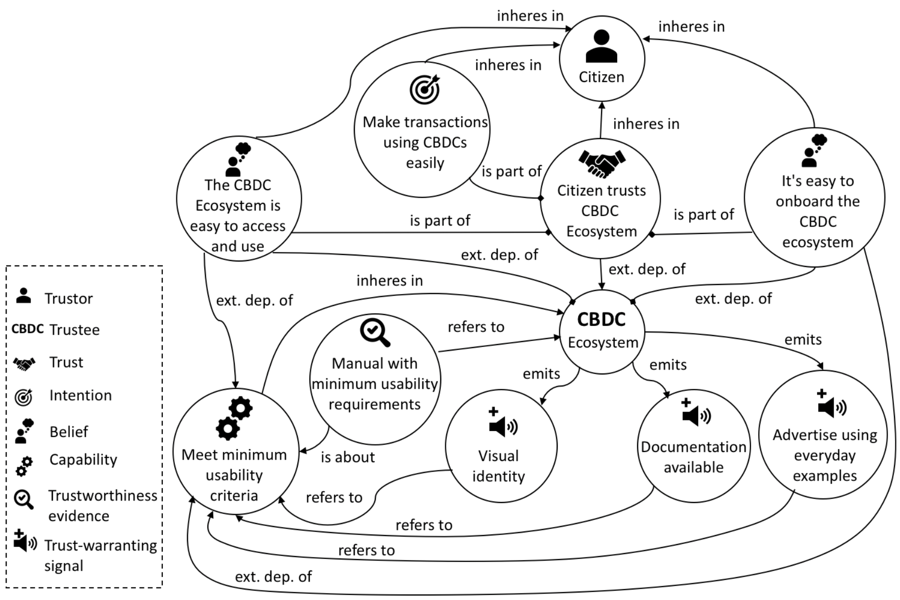

Fig. 4: Graphical representation of the ontology instantiation focusing on usability.

bersecurity policy. The resource *manual with minimum usability requirements* (TE3.1) contributes to the task *meet minimum usability criteria*, which in turn contributes to *be easy of use*. Conversely, vulnerabilities can be represented as goals, qualities, tasks or resources that negatively impact (hurts) the achievement of higher level goals. Finally, influences are represented as contribution links that help or hurt the achievement of higher level goals (help for positive influences and hurt for negative ones). The mapping between the ROT concepts and their representation in the i* Goal Model is presented in Table 2.

4.4 Discussion

In the validation session, the central bank experts of the aforementioned areas of interest were unanimously of the opinion that the ontology was capable of capturing all the important aspects of citizens' trust in CBDC ecosystems (perceived *usability* and *usefulness* of the approach). It was also mentioned by the interviewees that, when designing the CDBD ecosystem, it is useful to understand the intentions of the users that are related to their trust in the ecosystem, so that we can identify, at a very early stage, capabilities required to create a trustworthy and efficient environment, possible vulnerabilities that should be dealt with, as well as how to properly communicate about the ecosystem trustworthiness. Being able to identify citizens' goals provides a broad view of how CBDC can be successfully implemented, from a trustworthiness perspective. By eliciting goals, we can also identify the conflicting goals, and consequently we can be more proactive in resolving possible design issues.

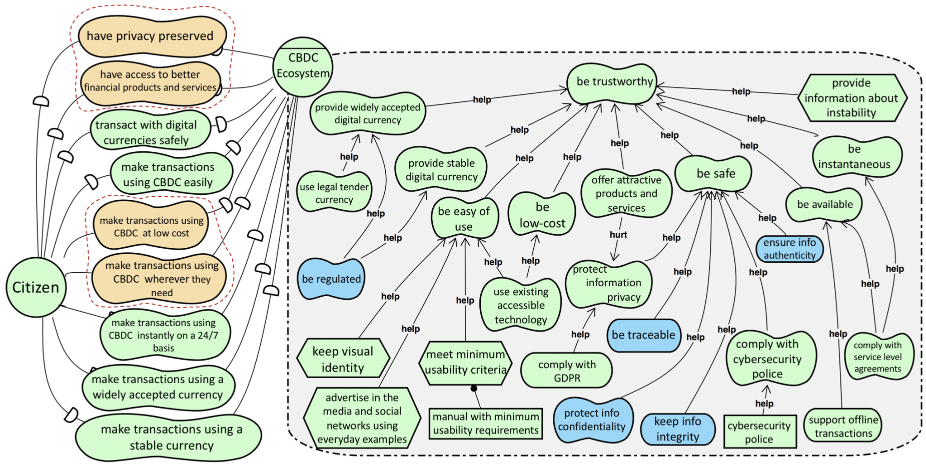

Fig. 5: A fragment of the goal model of the CBDC ecosystem.

Table 2: Representation of ROT concepts in i* Goal Model.

ROT Concept	Representation in i* Goal Model
Trustor	Actor, Agent, Role
Trustee	Actor, Agent, Role
Intention	Goal dependence, Quality dependence
Belief	Goal, Quality
Capability	Goal, Quality, Task, Resource
Vulnerability	Goal, Quality, Task, Resource
Trust Calibration Signal	Goal, Quality, Task , Resource
Trustworthiness Evidence	Goal, Quality, Task, Resource
Influence	Contribution link

An interesting finding was that once we understand what composes citizens' trust and which factors may influence it, we can take these requirements into account since the ecosystem's inception, thus enabling trustworthiness by design. Furthermore, it allows the identification of potential risks in advance and the definition of risk mitigation strategies. This is because if we know which capabilities and vulnerabilities are related to the trustor beliefs, we can reason about what can go wrong with the realization of the capabilities and the manifestation of the vulnerabilities, which will hurt the intentions of the trustor. These unwanted events correspond to risk events for which mitigations strategies may be defined in advance. Another interesting finding is that trust relations require constant monitoring as the ecosystem is very dynamic and is constantly changing. And changes in the environment can influence user trust.

5 Related Work

There are some works in the literature that address the dynamic nature of trust. Riegelsberger et al. [21] propose a framework on the mechanics of trust, in which they identify contextual (temporal, social, and institutional embeddedness) and intrinsic (ability and motivation) properties that warrant trust in another actor, which they name trust-warranting properties. They also describe how the presence of these properties can be signaled. In their model, they identify two broad categories of signals: symbols and symptoms, which are analogous to the ROT concepts of trust-warranting signals and trustworthiness evidence, respectively. Despite this similarity, their work differs from what we propose here, as they do not consider uncertainty signals and other factors that may influence trust, such as other trust relations and the mental state of the trustor. Also, they do not provide an ontological account for the concepts represented in their model.

Castelfranchi and Falcone [9] made an important contribution with their theory of trust. ROT relies largely on their theory to formalize the general concept of trust and the concept of social trust. In their work, they present trust dynamics in different aspects: (i) how trust changes on the basis of the trustor's experiences, which is related to the ROT concepts of *trustworthiness evidence* and *influence*; (iii) how trust is influenced by trust; how diffuse trust diffuses trust (that is how A's trusting B can influence C trusting B or D, and so on); and (iv) how trust can change using generalization reasoning (the fact that it is possible to predict how/when an agent who trusts something/someone will therefore trust something/someone else, before and without a direct experience). These last three aspects are related to *trust influences* in ROT. Although it is rather comprehensive, their proposal does not mention the emission of signals to communicate uncertainties nor to indicate trustworthy behavior on the part of the trustee.

6 Conclusions

In this paper, we presented an ontological analysis of the factors that influence trust as well as other trust-related concepts, such as pieces of evidence that indicate a trustee's trustworthiness and the signals that the trustee may emit to indicate trustworthy behavior. To validate the ontology and demonstrate the applicability of our proposal, we conducted a real case study concerning citizens' trust in CBDC ecosystems. The case study experience confirmed that ROT can properly represent trust in this context, and suggests it could be used to represent other real cases. We acknowledge that our case study has some limitations as we only took the central bank's view regarding citizens' trust in CBDC ecosystems. Nevertheless, we rely on documentation and information from surveys conducted with citizens from the literature [4,6,7,8,11,22]. As future work, we plan to apply ROT to support trustworthiness by design, so that trust can be part of the design of ecosystems since their inception and be prioritized in all aspects of the ecosystem.

Acknowledgments

CAPES (PhD grant# 88881.173022/2018-01) and NeXON project (UNIBZ). The authors would like to thank the Central Bank for their time and support.

References

1. Amaral, G., Sales, T.P., Guizzardi, G.: Towards Ontological Foundations for Central Bank Digital Currencies. In: Proc. 15th VMBO (2021)
2. Amaral, G., Sales, T.P., Guizzardi, G., Porello, D.: Towards a Reference Ontology of Trust. In: International Conference on Cooperative Information Systems. pp. 3–21. Springer (2019)
3. Bank for International Settlements et al.: Central bank digital currencies: system design and interoperability. Bank for International Settlements (2021)
4. Bank for International Settlements et al.: Central bank digital currencies: users needs and adoption. Bank for International Settlements (2021)
5. Batteux, Eleonore, e.a.: The negative consequences of failing to communicate uncertainties during a pandemic: The case of COVID-19 vaccines
6. Bijlsma, M., van der Cruijsen, C., Jonker, N., Reijerink, J.: What triggers consumer adoption of CBDC? De Nederlandsche Bank Working Paper (709) (2021)
7. Boar, C., Holden, H., Wadsworth, A.: Impending arrival–a sequel to the survey on central bank digital currency. Bank for International Settlements (107) (2020)
8. Boar, C., Wehrli, A.: Ready, steady, go?-Results of the third BIS survey on central bank digital currency. Bank for International Settlements (2021)
9. Castelfranchi, C., Falcone, R.: Trust theory: A socio-cognitive and computational model, vol. 18. John Wiley & Sons (2010)
10. Dalpiaz, F., Franch, X., Horkoff, J.: iStar 2.0 Language Guide. CoRR **abs/1605.07767** (2016), `http://arxiv.org/abs/1605.07767`
11. European Central Bank: Eurosystem report on the public consultation on a digital euro. European Central Bank (2021)
12. Falcone, R., Castelfranchi, C.: Trust dynamics: How trust is influenced by direct experiences and by trust itself. In: Proc. AAMAS 2004. pp. 740–747. IEEE (2004)
13. Fischer, P., Lea, S.E., Evans, K.M.: Why do individuals respond to fraudulent scam communications and lose money? The psychological determinants of scam compliance. Journal of Applied Social Psychology **43**(10), 2060–2072 (2013)
14. Fonseca, C. et al.: Relations in ontology-driven conceptual modeling. In: 38th International Conference on Conceptual Modeling (ER). vol. 11788, pp. 28–42. Springer (2019)
15. Guizzardi, G.: Ontological foundations for structural conceptual models. Telematica Instituut, University of Twente, The Netherlands (2005)
16. Guizzardi, G. et al.: Grounding software domain ontologies in the Unified Foundational Ontology (UFO). In: Proc. 11th CIBSE. pp. 127–140 (2008)
17. Luhmann, N.: Trust and power. John Wiley & Sons (2018)
18. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. Academy of management review **20**(3), 709–734 (1995)
19. McKnight, D.H., Chervany, N.L.: Trust and distrust definitions: One bite at a time. In: Trust in Cyber-societies, pp. 27–54. Springer (2001)
20. McKnight, D.H., Liu, P., Pentland, B.T.: How events affect trust: A baseline information processing model with three extensions. In: IFIP International Conference on Trust Management. pp. 217–224. Springer (2012)
21. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: The mechanics of trust: A framework for research and design. International Journal of Human-Computer Studies **62**(3), 381–422 (2005)
22. Söilen, K.S., Benhayoun, L.: Household acceptance of central bank digital currency: the role of institutional trust. International Journal of Bank Marketing (2021)
23. Tomsett, Richard et al.: Rapid trust calibration through interpretable and uncertainty-aware AI. Patterns **1**(4), 100049 (2020)
24. Yin, R.K.: Case Study Research: Design and Methods (Applied Social Research Methods). Sage Publications (2008)
25. Zetzsche, D.A. et al: Decentralized finance (DeFi). IIEL Issue Brief **2** (2020)