Foundations and Trends® in Information Retrieval
Vol. 8, No. 1 (2014) 1–107

DOI: 10.1561/1500000033

LifeLogging: Personal Big Data

Cathal Gurrin
Insight Centre for Data Analytics
Dublin City University
cgurrin@computing.dcu.ie

Alan F. Smeaton
Insight Centre for Data Analytics
Dublin City University
alan.smeaton@dcu.ie

Aiden R. Doherty
Nuffield Department of Population Health
University of Oxford
aiden.doherty@dph.ox.ac.uk

Contents

Abstract

We have recently observed a convergence of technologies to foster the emergence of lifelogging as a mainstream activity. Computer storage has become significantly cheaper, and advancements in sensing technology allows for the efficient sensing of personal activities, locations and the environment. This is best seen in the growing popularity of the quantified self movement, in which life activities are tracked using wearable sensors in the hope of better understanding human performance in a variety of tasks. This review aims to provide a comprehensive summary of lifelogging, to cover its research history, current technologies, and applications. Thus far, most of the lifelogging research has focused predominantly on visual lifelogging in order to capture life details of life activities, hence we maintain this focus in this review. However, we also reflect on the challenges lifelogging poses to an information retrieval scientist. This review is a suitable reference for those seeking a information retrieval scientist's perspective on lifelogging and the quantified self.

C. Gurrin, A. F. Smeaton, and A. R. Doherty. *LifeLogging: Personal Big Data*. Foundations and Trends® in Information Retrieval, vol. 8, no. 1, pp. 1–107, 2014.
DOI: 10.1561/1500000033.

1

Introduction

Lifelogging represents a phenomenon whereby people can digitally record their own daily lives in varying amounts of detail, for a variety of purposes. In a sense it represents a comprehensive “black box” of a human’s life activities and may offer the potential to mine or infer knowledge about how we live our lives. As with all new technologies there are early adopters, the extreme lifeloggers, who attempt to record as much of life into their “black box” as they can. While many may not want to have such a fine-grained and detailed black box of their lives, these early adopters, and the technologies that they develop, will have more universal appeal in some form, either as a scaled-down version for certain applications or as a full lifelogging activity in the years to come.

Lifelogging may offer benefits to content-based information retrieval, contextual retrieval, browsing, search, linking, summarisation and user interaction. However, there are challenges in managing, analysing, indexing and providing content-based access to streams of multimodal information derived from lifelog sensors which can be noisy, error-prone and with gaps in continuity due to sensor calibration or failure. The opportunities that lifelogging offers are based on the fact that a lifelog, as a black box of our lives, offers rich contextual information, which has been an Achilles heel of information discovery. If we know a detailed *context* of the user (for example, who the user is, where she is and has been recently, what she is doing now and has done, who she is with, etc. . .) then we could leverage this context to develop more useful tools for information access; see the recent FNTIR review of Contextual Information Retrieval, Melucci (2012). This valuable contextual information provided by lifelogging to the field of information retrieval has received little research attention to date.

Before we outline the content of this review we will introduce and define what we mean by lifelogging, discuss who lifelogs and why they do so, and then introduce some of the applications and core topics in the area.

1.1 Terminology, definitions and memory

There is no universal or agreed definition of lifelogging and there are many activities which are referred to as lifelogging, each producing some form of a lifelog data archive. Some of the more popular of these activities include quantified-self analytics[1], lifeblogs, lifeglogs, personal (or human) digital memories, lifetime stores, the human black box, and so on.

In choosing an appropriate definition, we refer to the description of lifelogging by Dodge and Kitchin (2007), where lifelogging is referred to as"*a form of pervasive computing, consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive*". The unified digital record uses multi-modally captured data which has been gathered, stored, and processed into semantically meaningful and retrievable information and has been made accessible through an interface, which can potentially support a wide variety of use-cases, as we will describe later.

A key aspect of this definition is that the lifelog should strive to record a totality of an individual's experiences. Currently, it is not

[1] http://quantifiedself.com

possible to actually record the totality of an individual's experiences, due to limitations in sensor hardware. However, we take on-board the spirit of this definition and for the remainder of this review, we assume that lifelogging attempts to capture a detailed trace of an individuals actions. Therefore, much of the lifelogging discussion in this review is concerned with multimodal sensing, including wearable cameras which have driven many first generation lifelogging efforts.

Because lifelogging is an emergent area[2], it is full of terminology that is not well considered and defined. Therefore, for the purposes of this discussion, we regard the lifelogging process as having the following three core elements:

- *Lifelogging* is the process of passively gathering, processing, and reflecting on life experience data collected by a variety of sensors, and is carried out by an individual, the lifelogger. The life experience data is mostly based on wearable sensors which directly sense activities of the person, though sometimes data from environmental sensors or other informational sensors can be incorporated into the process;

- A *Lifelog* is the actual data gathered. It could reside on a personal hard drive, in the cloud or in some portable storage device. The lifelog could be as simple as a collection of photos, or could become as large and complex as a lifetime of wearable sensory output (for example, GPS location logs or accelerometer activity traces);

- A *Surrogate Memory* is akin to a digital library, it is the data from the lifelog and the associated software to organise and manage lifelog data. This is the key challenge for information retrieval, to develop a new generation of retrieval technologies that operates over such enormous new data archives. Given the term surrogate memory, we must point out that this does not imply any form of cognitive processes taking place, rather it is simply the digital li-

[2] Although lifelogging has been around for several decades in various forms, it has only recently become popular.

brary for lifelog data, which heretofore has been typically focused on maintaining a list of events or episodes from life;

It is important to consider that lifelogging is typically carried out *ambiently* or passively without the lifelogger having to initiate anything. There have been a number of dedicated individuals who are willing to actively try to log the totality of their lives, but these are still in the very significant minority. For example, Richard Buckminster Fuller manually logged every 15 minutes of activity from 1920 until 1983, into a scrapbook called the Dymaxion Chronofile, as described in Fuller et al. (2008). More recently Gordon Bell's MyLifeBits project, Bell and Gemmell (2007) combined active and passive logging by using wearable cameras and capturing real-world information accesses. Another example of active logging is Nick Feltron's Reporter app, which allows an individual to manually log whatever life activity they wish in as much detail as they desire. Reporter will periodically remind the user to 'report' on the current activities.

While such dedicated lifelogging is currently atypical, most of us often explicitly record aspects of our lives such as taking photos at a social event. In such cases there is a conscious decision to take the picture and we pose and smile for it. Lifelogging is different, in that by default it is always-on unless it is explicitly switched off and it operates in a passive manner. Therefore the process of lifelogging generates large volumes of data, much of it repetitive.Thus the contents of the lifelog are not just the deliberately posed photographs at the birthday party, but the lifelog also includes records of everything the individual has done, all day (and sometimes all night), including the mundane and habitual.

Compare this to the recently popular field of quantified self analytics. Quantified self is considered to be a movement to incorporate technology into data acquisition on aspects of a person's daily life in terms of inputs (e.g. food consumed, quality of surrounding air), states (e.g. mood, arousal, blood oxygen levels), and performance (mental and physical). While there is a level of ambiguity in terms of the cross-over between quantified self and lifelogging, this review assumes that the key difference between lifelogging and quantified self analytics is that quantified self is a domain-focused effort at logging experiences (e.g. exercise levels, healthcare indicators) with a understanding of the key goals of the effort, whereas lifelogging is a more indiscriminate logging of the totality of life experience where the end use-cases and insights will not all be understood or known at the outset of lifelogging.

Considering how to organise these vast lifelog data archives, we believe that lifelog data should be structured in a manner somewhat similar to how the brain stores memories. While a debate on human memory models is beyond the scope of this review, we select the Cohen and Conway (2008) model of human memory due to the fact that many other memory scientists who have ventured into the application of lifelogging; for example Doherty et al. (2012); Pauly-Takacs et al. (2011); Silva et al. (2013), all refer to this model. Cohen and Conway's model suggests that the memory of specific events and experiences should be called our episodic memory. It is autobiographical and personal, and can be used to recall dates, times, places, people, emotions and other contextual facts. Our semantic memory is different and is our record of knowledge, facts about the real world, meanings and concepts that we have acquired over time. While our episodic memory is personal, our semantic memory is shared with others and is independent of our own personal experiences or emotions since its contents can stand alone and are abstract. It is suggested that our semantic memory is generally derived from our episodic memory in the process that is learning new facts or knowledge from our own personal experiences, as described in Cohen and Conway (2008) For lifelogging, much of the focus thus far has been on supporting and generating surrogates of episodic memory.

Based on such a model, one would consider a typical day being segmented into a series of events of various durations. Figure 1.1 shows a timeline of a day with events represented by an image and various metadata sources. Dressing and self-grooming, preparing food, eating, travel on a bus, watching TV, listening to music, working on a computer, taking part in a meeting, listening to a presentation, doing gardening, going to a gym, and so on, are all examples of everyday events. Some of these events are regular and repetitive. For example, many of us eat the same or similar breakfasts each day at approximately the same time and in the same place. Going to a movie or attending a party is probably a rarer occurrence, perhaps weekly or monthly. While debate exists on the formation of human memories, the view presented in this review is that lifelogging creates a lifelog which is similar to the Cohen and Conway (2008) model of episodic memory. A lifelog captures the “facts” around the episodes in our lives but not their emotional interpretation.

A lifelog does not typically capture or store semantic memory, so when we want to know the capital city of Azerbaijan (Baku) or the winners of the 2000 FA Cup (Chelsea), we don’t ask a lifelog, we go to Wikipedia or we search the web. As of now, we do not refer to a lifelog for such semantic facts. Therein lies one of the real challenges in lifelogging: how to search a lifelog for relevant information given that the IR techniques we have developed over the last several decades are developed to search semantic rather than episodic memory. We shall return to this point later.

Other use-cases of lifelogging are broad and varied, such as the ability to detect and mine insights from our daily lives, in a Quantified Self type of analysis. We will return to a detailed discussion of the use-cases later. Whichever use-cases we employ, in order to maximise the potential of lifelogging (as with any technology), we should map this new technology into our lives and develop the technology in support of, rather than to try to change, our lives around the technology. Thus at the outset we should ask ourselves what are the characteristics and structures which form the organisation of our lives where we can use lifelogging to build upon.

1.2 Motivation

Lifelogging is becoming more accessible to everyone due to data capture becoming more feasible and the availability of inexpensive data storage technologies. Gordon Bell from Microsoft was one of the first to fully embrace digitising his life as part of the MyLifeBits project (Gemmell et al. (2002, 2006)) at Microsoft Research and this helped raise the profile of lifelogging. Lifelogging alone can generate large volumes of

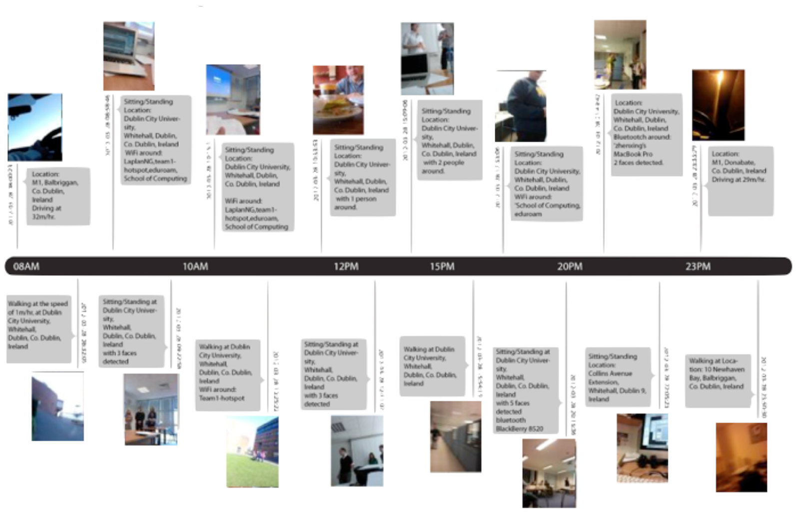

Figure 1.1: An event timeline showing key images with associated metadata from a lifelog.

data on a per person basis and a sense of this can be found when we examine the amount of information in the world in general, and also in our own personal lives, as recently discussed in The Economist (2010). When we factor in the possibilities of linking our personal lifelogs with "external data" in order to semantically enrich our lifelogs, then as an information management task it becomes a challenge to maximise its potential. Lifelogging is not a new idea, and it is not new in practice either, but apart from the media coverage generated by projects like MyLifeBits it has recently become popular for several reasons, including the following:

1. Computer storage has become incredibly cheap, both on the cloud or as personal storage. In fact we have seen exponential growth in disk storage capacity over the lifetime of digital storage;

2. We are seeing advances in sensors for sensing the person as well as sensing the person's environment which are making such sensors cheap, robust and unobtrusive;

3. There is growing social interest in the phenomenon of sens-

ing and recording oneself, the so-called quantified-self movement. Sometimes this is driven by applications like sports and health/wellness, other times it is sensing just because we can;

4. We can observe an increased openness to storing and sharing information about ourselves as can be seen in social networks.

5. New technologies such as Google Glass has brought lifelogging to the fore as a topic for public discussion.

These contributing factors evolved independently and some came together with the CARPE (Continuous ARchiving of Personal Experiences) workshop, Gemmell et al. (2004), in 2004 which brought together for the first time those whom Steve Jobs would have called the rebels, the square pegs in round holes, people like Steve Mann, Kiyoharu Aizawa, Gordon Bell, Jim Gemmell and others. This workshop in 2004 was the first real gathering of those who previously had been working independently or in isolation and suddenly as a result there was a lot of sharing of tools and experiences and lifelogging emerged as a research area.

While most of the interest in lifelogging is in either the technologies we can use, or the applications that lifelogging can be usefully used for, these do represent sizeable challenges in their own right. From an information science perspective, lifelogging presents us with huge archives of personal data, data with no manual annotations, no semantic descriptions, often raw sensor data (sometimes error-some), and the challenge is to build tools for semantic understanding of this data, in order to make it usable.

This has similarities to the early days of content-based image retrieval, but it is different in that the multimodal sensory information which forms part of the lifelog can be used to make this an opportunity for *big data* analytics. "Big data" is an often mis-used term and is unfairly associated with huge volumes of information, hence the use of the term "big". In fact "big data" isn't just about volume, it is equally about veracity (the accuracy and correctness of data which may have been eroded due to things like calibration drift in sensors), velocity (the shifting patterns and changes in data over time) and variety (the heterogeneous sources from which data is gathered). Big data is a contemporary problem and is about mining and cross-referencing information from diverse sources in order to discover new knowledge. The opportunity with lifelogging is to do this on a personal rather than on an enterprise level. Personal lifelogging can also be regarded as a new search challenge, with new use-cases defining new search and access methodologies, and providing a new opportunity to re-examine contextual IR, as described recently in Melucci (2012), with new data sources from lifelogging.

1.3 Who lifelogs and why ?

As with any new technology, there are pioneers of lifelogging like those mentioned in §1.2, and there are early adopters who take lifelogging into new applications. These applications exhibit the main advantage of concentrating on better understanding of an individual's *life* interactions, not just their activities on social media or their past search behaviour on electronic commerce sites or search engines.

However, in order to move beyond this and into a more mainstream and sustainable contribution to society, lifelogging needs to show successful application in different domains. We return to the point regarding lifelogging and quantified self analytics. The question of whether lifelogging when focused in a narrow domain is actually lifelogging is a topic for discussion, but as we will describe, the first set of lifelogging applications that are getting market traction are focused quantified self applications, perhaps because of the immediate value that can be mined from the focused data.

At the present time, there are already a large number of such applications which show successful inclusion of lifelogging technologies and concepts. Many of these are based around some form of personalised healthcare or wellness. There are already several relatively cheap products on the market which log caloric energy expenditure and types of human physical activity being performed such as the FitBit One$^{\mathrm{TM}}$worn as a clip-on device on the belt or trouser pocket, the Nike FuelBand$^{\mathrm{TM}}$worn as a bracelet, or the Lark$^{\mathrm{TM}}$, also worn as a bracelet[3] These have built-in accelerometers and gyroscopes and with a fairly simple algorithm employed, can be used to count the number of steps the wearer takes in a day. They are quite accurate at measuring some activities like walking but not so good for other activities like cycling, contact sports or swimming. They are popular because they provide real-time feedback to the user on their physical performance or they have been embedded into a gamification model and integrated with social network thereby allowing for league tables and comparisons against the self and against peers is used to incentivise exercise or even change behaviour, Barua et al. (2013).

Monitoring sleep patterns and quality has also become a consumer-level product in recent times. These sense even the most minute movements we make when we pass through the various stages of the circadian rhythm as we sleep and from that they can compute an indicator of sleep quality. Given our recent realisation of the importance of sleep as a health indicator as well as its all-round restorative properties, its no surprise that a market has quickly grown up around this. Sleep sensing devices are typically made up of a combination of accelerometers and gyroscopes, fabricated onto a small, self-contained device worn on the wrist which detects, logs and stores timestamped movement information. Alternatively, there are apps on smartphones which do the same thing but not as accurately and there is a technology which emits low-power radio waves and measures its refraction as we breathe or move, its advantage being that it is contactless and it is built into a device marketed as the Renew SleepClock from Gear4.

Healthcare self-monitoring has other, more significant applications besides a desire for personal analytics. Smoking cessation, diet monitoring for weight loss or tracking sugar intake for monitoring diabetes all have apps to record our activity, some of them to record manually and some semi-automatically. Moving to fully automatic applications, these become very challenging, because even with wearable cameras, it is difficult to automatically sense and detect what you eat and it is easy to cheat such a system. Hence most of these apps use manual lifelog-

[3] `http://www.fitbit.com`, `http://www.nike.com/fuelband` and `http://www.lark.com` respectively.

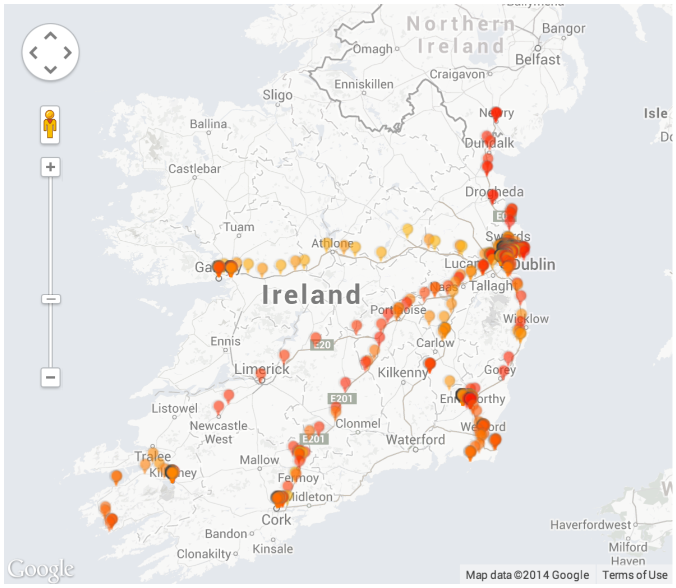

Figure 1.2: A Location Lifelog of one of the authors over a period of approximately one month in early 2014. This log was gathered automatically by the Openpaths app and uploaded to a web service, where the data can be shared with research projects.

ging techniques to record activities which are subsequently presented to the wearer as a memory from the recent past to remind him/her to manually log any missed activities.

There is also recent progress in the area of location logging, whereby apps on a smartphone make use of the inbuilt sensors to log the movements of an individual. This may be for social purposes (e.g. Foursquare checkins), for fitness purposes (any exercise mapping app), or just for lifelogging purposes (e.g. the OpenPaths app). An example of the location log for one of the authors over a one month period is shown in Figure 1.2.

However, we do note a recent movement of technology away from the focused quantified self analytics towards the idea of the totality of life experience. The Moves and SAGA smartphone apps capture in a non-visual manner all life activities (locations, activities) of the individual and present them in a basic version of a lifelog.

Recently introduced hardware devices such as the OMG Autographer and the Narrative Clip (to be discussed later) bring the idea of the capture of the totality of life experiences one step closer. Such cameras capture thousands of images per day from the wearer's viewpoint and will enable a new suite of true lifelogging technologies. One such example is triggering recall of recent memories; an application of lifelogging where the detailed lifelog acts as a memory prosthesis, thereby providing support for people with Alzheimer's or other forms of dementia. It is well-known in memory science that experiences from the past can be spontaneously re-lived based on a trigger such as an image, smell, sound or a physical object, as presented in Hamilakis and Labanyi (2008). Examples might be the smell of a pine tree which can remind a person of Christmas or a even specific Christmas from their childhood. Similarly, re-living recent experiences from a lifelog, such as the ordinary things that happened during a given day, can induce spontaneous recall, known as Proustian recall, which is discussed in Stix (2011). There have been several studies reported using visual lifelogging devices, which log and then re-play a given day for a person with memory impairment, triggering short-term recall of everyday happenings and in this way opening up cognitive pathways. Berry et al. (2009) describe studies at Addenbrooks hospital in Cambridge, UK that show measurable effects of replaying a day's activities for memory rehabilitation.

Yet while we can record a given day in very fine detail, using lifelogs for the detection of longer-term cognitive decline or gradual behaviour change, for example, is far more difficult because of the variations in our daily activities; put simply, there is no such thing as a normal day in our lives, as described in Doherty et al. (2011a).

There is also potential for lifelogging technologies to be used by organisations as a means of recording/logging the activities of employees, for various reasons, such as logging employee activities for legal/historical reasons, replacing manual record taking, logging information access activities as in Kumpulainen et al. (2009), or potentially as a new technology to support aspects of what Stein (1995) refer to as organisational memory. The idea here is to automatically capture procedures and processes for everyday activities in the workplace. While this tends to have more success for office environments where we log digital activities (web usage, emails, document accesses) rather than physical ones, there are examples of recent work with healthcare workers in clinical practice who have to log their work and record their clinical notes at the end of a shift, Kumpulainen et al. (2009), as well as lifelogging for other job-specific tasks, Fleck and Fitzpatrick (2006). Lifelogging has also been used in market research, targeting novel qualitative analysis based on analysis of subjects' lifelogs and the amount of exposure they have to advertisements, Hughes et al. (2012).

Therefore, we can see that there are a huge number of application areas for lifelogging, though many of them have been driven by throwing technology at problems rather than having the technology developed specifically to address the problem. In Chapter 5 we discuss applications of lifelogging in more detail.

For the remainder of this review, we will focus on the actual implementations of lifelogging that have heretofore been employed by researchers; therefore the focus of the review will be on visual lifelogging using wearable sensors, that aim to record the totality of an individuals experiences. We will leave aside descriptions of quantified self analytics tools and other limited forms of lifelogging.

1.4 Topics in lifelogging

The end-to-end processes of lifelogging and the applications which then use the lifelog, are complex and involve many challenges and multiple disciplines. Starting at the beginning and at the *hardware* level, are the sensors themselves which, in the case of wearable sensors, need to be robust and unobtrusive because the human body is a harsh environment for any kind of sophisticated technology. Robustness is needed because sensors can be impacted when we bump into things, they can be exposed to high levels of moisture and humidity when we get caught in the rain or even in bathrooms. They must be tolerant to drift in calibration and not require re-calibration too often if at all. Wearable sensors should also be small enough that they do not interfere with our everyday activities, and they need to have enough battery life to last at least a complete day without needing replacement batteries or re-charging. Energy scavenging is an important topic for wearable sensors and good progress is being made in this field, as shown in Kansal et al. (2007). If the wearable sensors log and record data on-board (i.e. no real-time upload) then they need enough storage capacity that data uploads are not required for several days ideally, and if they upload data wirelessly then they need to be able to take advantage of networks that come into range or to partake in ad-hoc networking. If the wearable sensors themselves support real-time upload of data, then this has a negative impact on battery life.

In terms of *software middleware*, the raw data captured from heterogeneous sensor sources has to be aligned temporally and possibly spatially as well. This requires more than just transfer from one format to another and usually needs data cleaning as well as alignment. Data quality is an important topic in areas as diverse as business informatics, Watson and Wixom (2007), and environmental sensing, Ganeriwal et al. (2008); O'Connor et al. (2009). In addition, topics such as how to dynamically compute and utilise the trust and provenance or the reliability associated with data streams which have all the issues mentioned above, come into play. In lifelogging there has been little work done in this area to date and there is much that can be learned about data quality, trust and reputation from other fields.

Once sensor data for lifelogs has been gathered, cleaned and aligned, *signal processing* is then required to analyse and structure this data. Heretofore, this as typically been structured into a data unit called an event, as shown in Figure 1.1. This automatic segmentation into events is similar to segmentation of video into shots and scenes and requires structuring personal data into discrete units. A subsequent phase of mining patterns and the deviations that those patterns can follow, would allow for the determination of their uniqueness or regularity within the lifelogger's lifestyle. It is worth noting at this point that an event is not necessarily the optimal data unit, but it is the one that has received most attention in research to date. The event segmentation models described later in this review are to be considered as early stage models. There is a lot of potential for more flexible retrieval units than events to be considered, but as of yet, this has not yet received much research attention.

This segmentation is then followed by *semantic processing* whereby we perform semantic analysis and annotation of data, including (since we focus on visual lifelogging) an analysis of visual data from wearable cameras. Ultimately this leads to a semantic enrichment of the lifelog data at the event level, or at the sub-event level, thereby helping to construct a rich lifelog.

Once a lifelog is created, we then turn our attention to how to use it and how to access it. The challenge here is learning what are the appropriate *retrieval models* for lifelogs and whether conventional information retrieval techniques, developed for accessing our equivalents of semantic memory, can find uses in information retrieval for episodic memory. Naturally such retrieval models would be based on identified use-cases, but many of the use-cases for lifelogs are as yet unknown. We do however have an early indication of use-case categorisations from the 5R's of memory access proposed by Sellen and Whittaker (2010), which are recollecting, reminiscing, retrieving information, reflecting, and remembering intentions. Each of these five R's address different access requirement for lifelogs. Once the use-cases have been defined, it then becomes important to consider the access methodologies and the HCI factors. Lifelogging is a topic which, like current and future web search, needs to support various access mechanisms to address not only the initial the 5Rs of memory access, but also to develop useful lifelogging tools for the first-generation of lifeloggers. A desktop interface to a lifelog may be useful to support detailed reflection, quantified-self style, whereas a mobile or wearable (e.g. Google Glass) interface would be needed to support real-time recollection or retrieval of information. Since the use-cases for lifelogging are not yet well defined, the access mechanisms are yet to be clearly identified, so at this early stage in lifelogging research, we do need to consider a range of commonly used access mechanisms.

Given this lightweight summary of just some of the major topics associated with lifelogging, we can see that it represents a complex set of challenges, not just the individual challenge areas taken in isolation, but the sum of the components into a whole. In the next section we present a summary of how we have structured the remainder of this overview of lifelogging.

1.5 Review outline

This review sets out to provide a comprehensive review of lifelogging, to cover the history of the field, the technologies that are currently available and the applications for which lifelogging can be used. In the next chapter we present a history of lifelogging, covering the major contributors and their impacts, as well as the advances in capture, storage and access to lifelog data. In Chapter 3, and in particular in §3.1 we give an overview of various lifelogging devices and technologies that have been employed in the field, for both capture and storage of lifelog data. Following that, Chapter 4 looks at the challenges in organising lifelog data with a focus on identifying and annotating or indexing events. Even though lifelogging generates an autobiographical record of our episodic memories and information retrieval is traditionally applied to some form of semantic memory, we believe it is important to look at how information retrieval techniques have a role in the implementation of access mechanisms to lifelogs. In Chapter 5 we present a wide range of applications of lifelogging and in the final chapter we reach some conclusions, we generate some pointers to future work and we discuss some of the most significant challenges facing this discipline.

2

Background

2.1 History

In 1945, Vannevar Bush (1945) introduced the world to the concept of the Memex, a proposal for a hypermedia system which would allow organising all the knowledge (books, records and communications) a person might accumulate in a lifetime. This would operate as a desk-based mechanical device with levers and knobs would allow people, specifically research scientists, to create explicit links between fragments of related information in different documents that they had come across as part of their work pursuits. In doing so they were creating "trails" of information which could be left for themselves or shared with others, who could follow them at some point in the future. This proposal came about at the close of the Second World War, at a time when the technology needed to support such a visionary system had not yet been developed, hence the suggestion of levers and knobs for regulating interaction.

Memex remains a seminal contribution because it introduced new concepts such as information links or trails which are created by an individual for his/her own use or for use by others and which form part of the permanent archive or record of the documents which are linked. Memex provided an inspiration for the first generation of hypertext systems, such as the Hypertext Editing System by Nelson and van Dam in 1969 and the Xanadu system described in Nelson (1988). Xanadu introduced the terms hypertext and hypermedia and is seen as a precursor to the development of the world wide web, though it had different aims and objectives.

When proposing the Memex as a mechanised device to organise a life-time of knowledge, Bush described it as an *"enlarged intimate supplement to one's memory"* which hints at the first application of what we now call lifelogging. The vision Bush proposed was for a *device* that could store information, documents and links among those documents. One could argue that this function is indeed the purpose of a library, any library, right back to the days of ancient Greece, but what makes the Memex into a close precursor to the lifelog is the user-authored contributions to the Memex in the forms of links or of comments based on our own personal experiences. In a way, this also makes the Memex a precursor to any form of personal, user-generated content, which pervades our information landscape today.

Memex also introduced the concept of a 'camera hound', an individual who wears a head-mounted camera which can be triggered to capture a photo of anything of interest which is subsequently inserted into and indexed by the Memex. In effect Bush had described the current lifelogging solutions being brought to market in 2014, many of which utilise wearable cameras to capture life experience from the wearer's viewpoint. We include one example wearable camera photo from the Posters and Demonstrations session of ACM SIGIR 2013 in Figure 2.1.

With the advent of digital technologies, the capabilities to deliver on the Memex vision became possible. Early research into lifelogs started in the 1980s, when pioneers such as Steve Mann (2004) began developing increasingly smaller wearable sensing and lifelog capture devices with continually improving power consumption. Much of this early research was focused on developing new types of sensing and display hardware. With regard to sensing, the focus was on trying to visually capture the world that we see in our everyday lives. Interestingly,

Figure 2.1: A typical wearable camera photograph, captured at the demo/posters session of the ACM SIGIR2013 Conference in Dublin, Ireland, in July 2013.

Mann not only saw the importance of wearable sensing (especially cameras) but he instinctively understood the importance of PoV (Point of View) when gathering lifelogs; a concept that even today's first generation lifelogging companies have not fully grasped. In addition to visual information, later research, such as that by Kiyoharu Aizawa and described in Aizawa et al. (2004b) also tried to capture a diverse set of "context" types such as location (GPS), though in these early days this was primarily used as metadata for indexing and retrieval of the captured visual data.

Data capture has been led by a number of key visionaries, including Steve Mann, Gordon Bell, and to some extent, very early adapters such as Jennifer Ringley and Josh Harris. Both Ringley and Harris (neither of whom were academics) explored the social aspects of life recording and life streaming, as we will explain below.

Much of the early research into lifelogging has focused on developing technology to automatically populate a computer-based storage of life experiences (the lifelog) in as much detail as possible. Detail, and exactly how much detail is required for a lifelog, is an important point that we will return to later in this review.

Steve Steve Mann (1997) has been developing wearable computer technologies since the early 80s and has been described as the father of wearable computing. He has dedicated decades of research into developing wearable life-capture technologies and he coined the phrase *"sousveillance"* to refer to the idea that digitally capturing life experience is a form of self-surveillance, which is inherently different from the conventional state- or enterprise-sponsored surveillance for security and governance that we have become so used to. In fact, Mann has gone so far as to identify many different types of 'veillances' all grouped under the single concept of veillance and has recently hosted a conference (ISTAS 2013) to consider the issues of veillance[1].

Mann has developed many generations of wearable camera technologies and addressed many of the fundamental challenges in wearable lifelogging technologies from early personal imaging in Mann (1997), to the eye-tap system described in Mann et al. (2005), which can be viewed as a forerunner of Google Glass. Mann has, in recent years, developed many solutions to the everyday, real-world challenges of gathering lifelogs from wearable computers, such as supporting different lighting levels to avoid capture blackout or whiteout, Mann et al. (2012), and real-world use-cases, Mann et al. (2011). Within the field of wearable computing and lifelogging, Mann is considered a visionary. Aside from the many forms of veillance that he proposes, he has coined the term lifeglogging to refer to long-term (lifetime) lifelogging and lifeblogging to refer to the act of publicising and streaming out lifelog data.

However, the most influential actor in the area is Gordon Bell, co author of the book Total Recall, see Bell and Gemmell (2009), who, while VP of Engineering at Digital Equipment Corporation, oversaw the development of the VAX computer and coined Bell's law that describes how types of computing systems (referred to as computer classes) form, evolve and may eventually die out. In the early years of this century, Bell turned his attention to a Bush-esque proposal for a lifetime store

[1] `http://sites.ieee.org/istas-2013/`

Figure 2.2: Gordon Bell, lifelogging pioneer. Shown here wearing a Vicon Autographer, a Memoto, a audio recorder, a veillance conference badge (ISTAS 2013) and a pen.

of personal information. This software was called MyLifeBits and is described in Bell and Gemmell (2007); Gemmell et al. (2002, 2006). MyLifeBits acted as a lifetime store of everything; in effect, it was an attempt to fulfil Bush's 1945 Memex vision. The MyLifeBits software supported full-text search, text & audio annotations, hyperlinking, reporting, visualising and clustering between content. Bell actively lives his research and with the help of administrative assistants, Bell has captured a lifetime's worth of articles, books, CDs, letters, memos, papers, photographs (including periodic phases of SenseCam capture), pictures, presentations, home movies, videotaped lectures, and voice recordings into the MyLifeBits software. Figure 2.2 shows Gordon Bell at ISTAS 2013 wearing a myriad of lifelogging devices, including lifelogging cameras and audio recording devices.

As technology progressed, the requirement for supporting administrative assistants diminished and lifelogging to extreme levels become possible for nearly anyone who had such a desire. Cathal Gurrin, at Dublin City University, inspired by Gordon Bell's SIGIR 2004 keynote talk, took the opportunity to wear a SenseCam in mid-2006 to gather a detailed and extensive visual archive of life experience. Gathering data for 16-18 hours per day, the archive (still growing) consists of almost 14 million automatically-captured images of life-experience, along with time-aligned sensor data (locations, movements, environmental noise, temperature, and so on). This longitudinal archive, the largest the authors are aware of, has supported many different research efforts. For example Gurrin et al. (2008b) has explored the the nature of large personal archives, Gurrin et al. (2013) has explored various applications of lifelogging and the potential of the smartphone as a lifelogging tool, Doherty et al. (2012) has used the archive to explore approaches for human experience sampling, Lee et al. (2008) has viewed the lifelogging effort as a canonical data process model and Doherty et al. (2009) has explored the nature of event segmentation & decay over years of lifelog data. Such a long-term effort at lifelogging has enabled the authors of this review to gain an understanding of the use-cases for, and organisation methodologies required, to support ubiquitous lifelogging, and as such, influences this review.

Aside from such extreme early adapters, who took a grand view of the potential of lifelogging and gathered archives of all life activities in an indiscriminate manner, there were also some real-world applications of lifelogging developed that targeted specific use-cases. Kiyoharu Aizawa, from the University of Tokyo has spent almost two decades researching multimedia content processing and more recently, lifelogging. Initially Aizawa developed technologies to support summarising wearable video, Aizawa et al. (2001); Hori and Aizawa (2003) and multimodal contextual capture and retrieval of lifelogs, Aizawa et al. (2004a). Moving from generic data capture to targeted use-case inspired capture, Kitamura et al. (2008) described Foodlog, which is a use-case specific lifelog that focused on monitoring diet.

One unifying principle of these early lifeloggers was that the lifelog would be for personal use. Following the Bush vision of the personalised supplement to one's own memory, Bell, Mann, Aizawa and Gurrin, in the first decade of this century, all understood the personal nature of lifelogs and that sharing of data would need to be a carefully-controlled activity. Hence, we still do not have a freely available lifelog dataset or test collection that can be employed for comparative research. However, other actors (such as Ringley and Harris) considered lifelogging as a more public activity; lifeblogging as Steve Mann would refer to it as. The first, and most well known, of these was Jennifer Ringley who famously set up a webcam in 1996 (called Jennicam[2]) that streamed periodic photos for open-access. Jennifer was the first to present a cam-eye view of her life and addressed early legal issues concerning the right to publish content that was not deemed harmful to other adults. Ringley maintained her lifestreaming activity for seven years and eight months. Josh Harris's project in conceptual art in 2000 live-streamed 24/7 content to anyone who was interested to watch, from his home using microphones and 32 cameras[3]. Ringley and Harris could be considered to be the earliest form of lifebloggers because they made their lifelogging a public activity. While we refer to lifeglogging and lifeblogging in this review, the term that we will use is lifelogging as it encompasses all aspects of capturing life activity, whether short-term or long-term.

In recent years, there has been a significant increase in the societal awareness of lifelogging, not simply due to the Jennicam or other similar concepts such as the Big Brother TV show syndicated around the world. Recently aome well-known actors have become involved, such as AJ Jacobs, the Editor-at-Large for Esquire magazine. AJ is known for immersing himself in a project or lifestyle, for better or worse for a period of time (usually a year), then writes about what he learned. He has undertaken a year of living biblically, living healthily and most recently, nine weeks of lifelogging [4], in which he explored the impact of lifelogging on people he met and explored the potential of visual (video) lifelogging as a tool to enhance daily life.

[2] `http://www.wired.com/2010/04/0414jennicam-launches/`
[3] `http://www.imdb.com/title/tt0498329/`
[4] `http://www.esquire.com/features/overly-documented-life-0113`

These early adopters were driven by desire for understanding the technical capabilities and the potential use-cases, however, there have also been early adapters in the field of medical science. In some instances of behavioural science, manual Ecological Momentary Assessment (EMA), sometimes called Experience Sampling Method (ESM), is employed whereby an individual has to manually enter what they are doing when prompted by a cellphone, Dunton et al. (2011). It is a form of retrospective self reporting. While the benefits of this is clear for studies in behavioural science, the concept of manual entry of lifelog data is usually linked to a particular use case, such as gathering data for healthcare logging or similar quantified-self activities. One such manual lifelogging app is called Reporter, which is available for the iPhone as of May 2014. Reporter, created by Nick Feltron, the well-known designer of the Facebook Timeline is designed to help a user to collect, categorise and communicate their everyday data via a user created personal ontology of annotated concepts.

Much of the lifelogging activities we have just described were non-academic activities. Within the academic community, the first dedicated lifelogging academic event, the CARPE workshop (Gemmell et al. (2004)) took place at ACM MULTIMEDIA in 2004. At this workshop Mann (2004); Takahashi et al. (2004); Gemmell et al. (2004) all explored issues surrounding the capture of lifelog data, Jaimes et al. (2004) explored retrieval of lifelog data, Cheng et al. (2004); Aizawa et al. (2004b) explored the organisation, privacy and legal issues around lifelogging . Although the CARPE workshops no longer take place, there is a dedicated SenseCam conference that occurs every 18 months. The recent SenseCam conferences, in 2012[5] and 2013[6] have specifically explored the potential of visual lifelogging to provide insights into human behaviour. Other conference series, dealing with Pervasive and Ubiquitous Computing, Human Computer Interaction and mobile devices have also explored various aspects of lifelogging over the past decade. Many of the related publications at these conferences are referred to in this review.

[5] http://sensecam2012.dph.ox.ac.uk/

[6] http://http://sensecam.ucsd.edu

2.2 Capture, storage and retrieval advances

The Memex vision proposed by Vannevar Bush described the concept of a surrogate memory, though many decades before it would become feasible to actually achieve the vision. Only in recent years has it become feasible to actually capture, store and retrieve a lifetime of digitally sensed data. We can identify three streams of progress have brought us to this point, (1) sensing and inexpensive capture devices, (2) storage resources and cloud processing, which will enable (3) real time semantic extraction and search tools. Of the three streams, the search and semantic extraction, including the whole area of data management and organisation, are the least advanced. We will examine each of the three streams in turn.

Assuming automatic and passive capture, lifelogging is a data intensive activity that relies on sensors to gather information about us or our environment and to relay this information to a processor for analysis or storage. From a lifelogging perspective, the definition of a sensor is somewhat more broad than the traditional definition. Traditionally, one views a sensor to be some form of converter that measures a physical quantity and converts it into a signal which can be read by an observer which may be a human or an electronic instrument of some sort. For example, a conventional mercury-in-glass thermometer is an example of a sensor in that it converts the measured temperature into expansion and contraction of a liquid which, when read against a calibrated glass tube, displays a temperature reading to a human observer. There are many such examples of a conventional physical sensor and there are innumerable applications for sensors of which most people are never aware including cars, machines, aerospace, medicine, manufacturing and robotics.

In terms of lifelogging however, a sensor may be very different in that it may be a physical quantity sensor (such as the thermometer), it could also be a positional sensor (e.g. GPS), a chemical sensor (pH of perspiration on a so-called smart-watch), an environmental sensor, an acoustic sensor (microphone), an optical sensor (camera), some form of biomedical sensor or an informational (reading information) sensor. All lifelogging sensors need to deliver information in digital format to a receiver for processing and storage into a lifelog. This broad range view of sensors is more in keeping with the Memex vision.

Recent advances in digital sensing technology means that it now becomes possible to utilise ubiquitous and inexpensive digital sensors to sense the person, the environment that they are in and their access to information. In terms of lifelogging, the sensors that are used depend on the variation of lifelogging that is performed. A focussed stream of lifelogging can come from those who wish to record their activities such as their exercise regime using a device like a Polar heart rate monitor[7], or their sleep patterns using a wearable accelerometers such as the Lark[8]. Such quantified-self enthusiasts are engaging in a focused and discriminatory form of lifelogging and typically display a passion for recording data to understand themselves. The more extreme lifeloggers, such as Bell and Gurrin, typically want to capture everything about their activities from all feasible sources. This naturally imposes a sensor overhead on these extreme lifeloggers, but at the benefit of non-discriminatory capture, whereby no manual logging is typically required.

For those who dabble, or even the quantified-self enthusiasts, one or two sensors are usually sufficient to capture a log of some aspect of life experience, whereas for the extreme lifeloggers the challenge (or desire) is to capture as complete a picture of life experience as possible. Considering what we mean by ‘as complete a picture of life experience as possible’, this is entirely dependent on the state of current sensing technologies. At present (and for the foreseeable future), a suite of sensors are currently required to generate a rich lifelog, such as wearable cameras, smartphones as sensing platforms, physical body sensing, and so on. These sensors mostly sample data that is the equivalent of two of the human senses (sight and sound) and are usually accompanied by a host of additional human context and performance sensors (location, activity, etc.). Of course these sensors must be low cost and low overhead to employ, so in many cases lifeloggers currently use one or two multi-modal physical sensing devices along with informational sensors

[7] `http://www.polar.com`
[8] `http://www.lark.com`

Figure 2.3: The array of wearable sensors that lifelogging can call upon including (a) Nike ‘Fuelband’, (b) Contour wearable video camera with GPS, (c) Vicon Revue wearable camera with sensors, (d) Heart Rate Monitor display watch, (e) Autographer from Vicon, (f) GPS tracker, (g) Jawboneï£¡s ‘Up’ Fitness Tracking Bracelet, (h) Android smartphone with built-in sensors, (i) Audio recorder, (j) fitbit ‘One’ Wireless Activity Tracker and (k) fitbit ‘Zip’ Wireless Activity Tracker. Much of this functionality can be replicated by a modern smartphone.

to capture a rich record of life experience.

Examples of the physical sensors employed in lifelogging include positional sensors using GPS and Wi-Fi, Bluetooth sensors to identify other device (people) in the environment, acoustic sensors to measure noise, environmental acoustic signatures, wearable cameras to capture what we see and do, and even spoken words, along with optical sensors to sense what’s happening in the environment. Add to this informational sensors that sense our information interactions such as what webpages we visit, what documents we analyse, the emails processed, our SMS messages, etc. This gives an initial idea of the types of sensors that have become ubiquitous and inexpensive enough, so as to enable our current concept of lifelogging and go towards fulfilling the Memex vision. Ideally all of these sensors would be combined into one multimodal holistic device and an early example of one such sensing device is the SenseCam, which was seen as providing a major step forward in lifelogging research.

The SenseCam, as described in Hodges et al. (2006, 2011), was a wearable camera developed by Steve Hodges, Lyndsay Williams and others at Microsoft Research, in Cambridge, UK. This was a device incorporating an optical sensor in the form of a VGA camera with fisheye lens, an accelerometer, a light intensity meter, a thermometer and a passive infra-red (PIR) sensor to detect the presence of people. The SenseCam was worn around the neck using a lanyard (see Figure 2.4) and the readings from the various sensors could be used to determine when a picture is to be taken by them; for example no picture is taken while the wearer is turning around because the image would be blurred, and an image is taken when the PIR sensor detects the arrival of a person in front of the wearer, or there is a change in ambient light intensity indicating the wearer has moved indoors or outdoors. By default, SenseCam captured a new image about every 40 seconds unless triggered by its sensors to capture an image sooner. Captured data was stored using on-board memory which could hold up to ten 10 days worth of lifelog data, though it did require recharging nightly. A full day of wearing of a SenseCam would generate a lifelog comprising between 3,500 and 4,500 images stored at VGA resolution. Since the SenseCam was initially developed as a memory recall tool, it was considered that VGA images were sufficient in size and quality. A sample wearable camera image is shown in Figure 2.1.

What makes the development of the SenseCam such an important step is not the sensing technology that it used, but the way it was made available to the research community, initially through competition and later through direct application and finally through direct sale as the Vicon Revue. As a result, SenseCam is a notable landmark in lifelogging because of the wide range of lifelog applications in which it

Figure 2.4: The SenseCam, initially created by Microsoft is a wearable camera, worn about the neck that can capture thousands of photos daily. Shown here is the Vicon Revue, a revised version of the SenseCam created and sold by Vicon OMG.

has been used including memory rehabilitation Berry et al. (2007); Silva et al. (2013), dietary monitoring O'Loughlin et al. (2013), lifestyle analysis Doherty et al. (2011a), workplace activity recording Byrne et al. (2008a), qualitative market research Hughes et al. (2012).

The technology behind the SenseCam was licensed to a UK company Vicon, who initially developed and marketed a SenseCam look-alike device called the Vicon Review and subsequently they enhanced the functionality of this (including GPS) with a newer device called the Autographer (which went on sale in August 2013). Another similar device is the Narrative Clip (formerly the Memoto wearable camera). The Narrative Clip has gone on sale in late 2013 and incorporates a small wearable camera that clips onto the clothes of the wearer and captures over a thousand images per day using the in-built optical sensor and uploads these to the Narrative cloud-based server for on-line event segmentation and analysis. Both the Narrative Clip and the OMG Autographer, worn on a lanyard or clipped onto the clothing, capture images at multiple megapixel resolution and a year of lifelog data from either of these devices would require about 500GB of disk space to store.

Although not designed as a lifelogging device, the much anticipated (and sometimes feared) Google Glass, when used standalone, or in conjunction with a smartphone could be employed to gather a rich lifelog archive, but with the added advantage of the head-mounted camera tracking the head movements of the wearer (referred to as Point of View - PoV), thereby reducing the need for a fisheye lens as used in the SenseCam and OMG Autographer. The research of Steve Mann (predating Google Glass by as much as 25 years) has been already discussed. Mantiuk et al. (2012) have gone a step further in developing a low-cost, head-mounted video recording tool that tracks eye movement to identify the exact object receiving the wearer's attention, which could make a very effective lifelogging sensor.

From a technological perspective, many or perhaps all of the sensors we would need for wearable lifelogging are already available in a modern smartphone, as described in Reddy et al. (2007a); Burke et al. (2006); Lazer et al. (2009), although the camera on a smartphone would not typically be capturing PoV. A review of the efficacy of a smartphone as a lifelogging device can be found in Gurrin et al. (2013), which assumes that the smartphone is carried on a lanyard worn around the neck, as would be a SenseCam.

With regard to informational sensors, the best example of this sensing is the MyLifeBits project at Microsoft that has been previously discussed Gemmell et al. (2006). Another sensing modality that we have not mentioned is that of acoustic sensing. It is technically feasible now to record all-day audio on a smartphone without affecting the phone or battery performance to a notable extent. We will discuss acoustic/audio sensing in the following chapter. Finally, we have already discussed the Quantified Self sensing devices, so we will not return to that discussion here.

In summary, for sensing technologies, if we assume that wearable optical, acoustic, positional, and informational sensing are sufficient, then the capture challenge, although not solved, has been comprehensively addressed. However, we as humans, typically build our cognitive model of the real-world using five senses: sight, sound, touch, smell and taste. Current lifelog sensing technologies are focused primarily on sight, with some consideration of sound. In this respect, the current sensing technologies address only part of the problem of accurately sensing our life experience, and it is up to the semantic processing tools (within the surrogate memory) and our human understanding of a lifelog event (described in later chapters) to fill in the gaps in terms of the representation of sensed experiences, at least until touch, taste and smell sensing become cheap and feasible to use for lifelogging.

With regard to the second major challenge in lifelogging, storage, this is also considered to be a challenge that is adequately addressed at present. Hard drive capacities have been increasing since their first creation at a pace much faster than the doubling in semiconductor chip performance occurring every 18 months in Moore's Law. Walter (2005) describe Kryder's law, named after Mark Kryder who was Seagate Corp.'s senior vice president of research and chief technology officer; Kryder's Law suggests that disk capacities will continue to keep pace with the increasing demands for lifelogging data. From Bell and Gemmell (2009) we know that in 1970, a 20 MB hard drive cost twenty thousand US dollars and was the size of a washing machine. Fast forward 45 years and the typical hard drive is 3TB and costs less than a hundred US dollars. In fact, digital storage technologies are increasing in pace with our ability to gather ever-increasing quantities of data to fill the drives, but faster than our ability to actually find the content to extract the information back out.

Scientists and futurists such as Ray Kurzweil predict that the exponential growth rate of disk and processing capability will continue into the future. In his essay "The Law of Accelerating Returns"[9], Kurzweil discards the intuitive linear view of technological progress in favour of the exponentially increasing view and describes this as the 'law of accelerating returns'. Taking this viewpoint, if one was to record 16 hours per day of HD (Blu-ray quality) video data, this would require 90GB per day of storage. Extrapolating out to a full year, this suggests

[9] `http://www.kurzweilai.net/the-law-of-accelerating-returns`

a requirement of 32 TB per year. Taking a Kurzweillian viewpoint, it would be possible to store such a quantity of data on a 100 US dollar hard drive by 2017. This assumes that hard disks are the most appropriate technology for storing video, which they may not be, because there are extra running costs such as power and air conditioning, and magnetic hard disks are designed to provide access to any of the information stored in a matter of milliseconds rather than the kind of streamed access that playback/recall of a life experience may require. In reality, the different use-cases of lifelogging may require a selection of appropriate storage technologies to be employed.

Within the field of Information Retrieval, there is a long and valuable history of employing sizeable datasets and test collections, with the TREC evaluation forum the most well known. The datasets at TREC have continually been increasing in size over the years with the largest of the currently (2013) used TRECVid[10] video collections standing at 300GB (a full 5 years of a daily BBC TV series). The challenges of lifelogging for Information Retrieval becomes clear when a single lifelogger wearing an OMG Autographer or Narrative Clip would generate more than this data in any given year. These are the challenges of scale and semantics, but more of that in later chapters.

Of course, one can question the need to store all-day HD video content. If we drop the video to reasonable SD quality, then we can already store all of this data on a 100 US dollar hard drive, if an appropriate encoding scheme is employed. Thus it is already feasible to store long-term, all-day video from a wearable camera for an individual person. However, even if the storage technologies allow it, this is still not really feasible, due for the most part to the limitations of battery capacities on wearable devices. Hence, a technology such as the SenseCam was so important, in that it supported all-day photo capture. This would result in about 1.6 million photos being captured per year with a storage requirement of only about 32.7GB. This storage requirement is so low because the SenseCam captures highly compressed VGA resolution photos that average about 19.2KB in size. Recent refinements of the technology and advances such as the Vicon Revue, the OMG Autogra-

[10] `http://trecvid.nist.gov/`

Content Type	Volume/day	In one year	In a lifetime
HD Video	5,840 hours	32.8TB	2.65PB
Autographer Camera	1.1 million images	479.6GB	40.8TB
Audio (mono - 22KHz)	5,840 hours audio	227.8GB	19.4TB
Microsoft SenseCam	1.65 million images	30.2GB	2.6TB
Accelerometer (1 Hz)	21 million readings at 1 Hz	0.05GB	4.25GB
Locations (0.2 Hz)	3.9 million GPS points	0.01GB	1TB
Bluetooth Interactions	150,000 (estimated) encounters	2GB+	150GB
Reading Log	User dependent	1GB+	80GB

Table 2.1: An illustration of the data quantities and data sizes for a selection of lifelog data over a day, year and a lifetime (typical 85 year Japanese lifespan)

pher or the Narrative Clip wearable cameras capture data in the range of 3 to 5 megapixels per picture.

To illustrate the variety of data sizes and quantities, a summary table of a selection of lifelog data is shown in Table 2.1. In this table we include the annual storage requirements as well as a lifetime (85 years) storage requirement. Of course, extrapolating across a lifetime, when there is an assumption that data bit rates remain static is merely for illustrative purposes only. Data sources, qualities, resolutions and bit rates are constantly changing.

Hence, it is our opinion that the storage challenges of hosting extreme lifelogging data are not currently a major impediment. The disk densities and costs are still generally following a Kurzweillian model of exponential growth and with the promise of new storage technologies coming online in the future, such as solid state, phase change and even DNA-based memory, should mean that the storage of lifelogs will not be a roadblock.

We have seen capture challenges and storage challenges, processing challenges, how to index huge personal archives of data, these are all (just) technical challenges which will continually be addressed by the provision of new and novel sensing technologies as well as via advances in computer science and general information technology. However, the really big challenges in lifelogging are in using Big Data and Information Retrieval approaches to do things like combining, correlating, cross-referencing, leveraging, data mining from heterogeneous sources, learning, and from all that gleaning useful knowledge, making it discoverable and presenting it in an appropriate manner. That is the focus of the remainder of this review.

2.3 Lifelogging disciplines

As mentioned earlier, there are many application areas to which lifelogging could be usefully applied and it is this wide applicability and the wide spectrum of technologies that it encompasses that make it unique, and most challenging. From a technology perspective, lifelogging requires and can leverage developments in materials science and chemistry, leading to new sensor technologies. It has long been an area that has pushed the development of new technologies for device miniaturisation, energy management on devices and energy harvesting by devices. Enabling real-time lifelog data capture and analysis will require broad coverage, always-on, low latency networking and so developments in ad-hoc wireless sensor networks have also found application in lifelogging. Captured lifelog data requires pattern analysis and data mining on data which can be noisy and error-some in order to perform deep semantic analysis so data analytics is yet another area which contributes here.

From an applications perspective we have already mentioned several applications in medicine, Doherty et al. (2013a); Brindley et al. (2011); Kumpulainen et al. (2009); Conway and Loveday (2011); Pauly-Takacs et al. (2011); Berry et al. (2009); Kelly et al. (2012), in work applications Byrne et al. (2008a); Fleck (2005); Fleck and Fitzpatrick (2006) and in behaviour analysis Doherty et al. (2011a, 2013c), and there are really no limits to the kinds of applications – work-related as well as leisure-based — which could benefit from detailed, ambient, non-intrusive logging of the activities and behaviour of an individual.

In their book Total Recall, Bell and Gemmell (2009), identify four main areas for lifelogging applications, specifically work, health, learning and everyday life (social) uses. So when we combine all this together we find that lifelogging is an activity that requires inputs from computer science, cognitive science, artificial intelligence, bio-sensing, hardware design, HCI, and all working together in order to develop useful solutions across a huge variety of potential use-cases. There are few, if any, other developments in modern technology which require such breadth of inputs, and have potential across such a breadth of areas. With regard specifically to computer science and information retrieval, the challenges and opportunities are enormous. Just a selection of these include:

- data gathering when the data is, by its nature, private and very time consuming to generate (it takes a year to gather a year of lifelogging data);
- data analysis & semantic extraction / semantic organisation from heterogeneous sources, multimedia, text and sensors;
- search & retrieval based on, as of yet, not well understood retrieval requirements and use-cases, when the user is only likely to query the system when they can't remember the information, leading to a high likelihood of incorrect queries;
- evaluation when there are no readily available (or likely to become available) datasets and where the only person who can truly evaluate a collection is the data gatherer;
- summarisation and data mining, to support quantified-self style analysis and narrative/story-telling presentation;
- user interaction and presentation when the likely usage scenarios are unknown, potentially omnipresent and even how to support query formulation for many of the use-cases is poorly understood.

These challenges cross the entire spectrum of information retrieval and we will look at many of them in this review. In the next Chapter we present some more of the work done in lifelogging from a “big data” perspective, highlighting that lifelogging is in fact a source of personal big data.

3

Sourcing and Storing Lifelog Data

3.1 Sources of lifelog data

As we have mentioned, there has been an astonishing collection of technological advances in human sensing in recent years. From wearable cameras which sense our environment and activities, to sensing of our online digital footprints, these sensors can capture vast personal *"big data"* archives and the storage and processing of these is becoming cheap and affordable. This myriad of available sensing technologies can be used both to sense the person and to sense the environment in which they are situated. In this chapter we present a discussion of the different types of data that can be employed for lifelogging, using readily available technologies, some of which can carry semantic annotations and some of which exists as raw data.

The arrival of the smartphone (and more recently technologies such as Google Glass) really created the opportunity for mass participation in lifelogging as prior to that, all of the hardware required was specialist or proprietary. One of the popular early tools for the smartphone was the Nokia Lifeblog which was a digital photo album tool for the Nokia smartphone, designed for mobile phone photographers and bloggers. This collected photos, SMS messages and blog posts into an album which was synced with a PC or laptop and created a form of personal digital memory, albeit one where the logging and recording was based on user interactions with the device, rather than being an implicit activity of autonomously gathering data.

Since then, the range of lifelogging tools has exploded in size and Machajdik et al. (2011) have presented an overview of the different categories of lifelogging tools which we used as a starting point and extended based on our own experiences with lifelogging. These categories are described below.

- Passive Visual Capture. This refers to wearable, always-on cameras which record images or video. In general, this capture is from the point of view of the wearer, either head-mounted or worn on a lanyard around the neck. Examples of this include the SenseCam from Microsoft Research (and derivatives such as the Vicon Revue, OMG Autographer) that we have already mentioned, the EyeTap discussed in Mann et al. (2005), the DejaVue system for real-time capture and upload developed at the University of Southampton, by De Jager et al. (2011), the real-time lifelogging solution using smartphones developed by Qiu et al. (2012), the Dietsense project at UCLA which uses visual lifelogging for monitoring food intake by Reddy et al. (2007b), the WayMarkr project by Bukhin and DelGaudio (2006), eyeBlog project by Dickie et al. (2004), the Insense project by Blum et al. (2006), as well as various consumer-ready wearable cameras such as the Looxcie™[1] wearable video camera. Many of these visual capture devices include some additional sensors which are discussed below. Given our focus on lifelogging for capturing a totality of life experience, in the remaining chapters of this review, we place a heavy emphasis on visual capture technologies as a source for lifelogging.

- Personal Biometrics. There are many personal sensing devices for monitoring everyday activities and aimed at the consumer market and these are used by interested parties, such as the quantified-self community. Such devices include the Lark™wristband, the

[1] `http://www.looxcie.com/`

FitBit One™pod, the Nike Pulse™wristband, the Zeo sleep monitor (discontinued), and others, all used for monitoring activity levels (number of steps taken, distance travelled, caloric output) and sleep duration and sleep quality. These are based on a combination of basic sensors, such as accelerometer, magnetometer and gyroscopes which log activity levels on-board for subsequent docking or wireless transmission to a cloud service via a laptop or PC. Most of these are gamified and give daily performance scores, incentives to maintain levels or to improve;

- Mobile Device Context. This refers to using the smartphone to continuously and passively capture the user's context as the smartphone can now be used to record location, acceleration and movement, WiFi signal strength for indoor localisation, Woodman and Harle (2008), Bluetooth for identification of (the smartphones) of nearby users, Eagle and Pentland (2005), and so on. Smartphones with built-in sensors have allowed the development of personal sensing apps, especially using the built-in GPS, accelerometers, compass, camera, and sometimes even the microphone. Recent smartphones can also record aspects of the environment such as pressure, temperature and humidity[2]. With power-aware sensing, it is possible to gather a contextual lifelog for an entire day without impacting too negatively on battery life. Already we have seen the emergence of a first generation of lifelog apps for mobile devices which passively capture users' context data and curate it into a simple lifelog. Such apps include Moves[3] and Saga[4] which generate automatic lifelogs of a user's activity, though without relying on passive capture of visual content;

- Communication Activities. Passively logging our (electronic) communications with others such as our SMS messages, instant messages, phone calls, video calls, social network activities and

[2] Available on the Samsung Galaxy S4, announced March 2013 `http://www.samsung.com/global/microsite/galaxys4/`

[3] `http://www.moves-app.com/`

[4] `http://www.getsaga.com`

email exchanges, can also be recorded and form part of a lifelog. There are many tools to support this process, such as SMS backup tools[5], phonecall recorder apps[6], social networking histories, and so on. The advantage of logging our communication activities is that much of the data is already in text format and there is less of a semantic gap between the data and its meaning. For a discussion of the semantic gap in multimedia information retrieval, see Smeulders et al. (2000);

- Data Creation/Access Activities. Apart from the interactions we have with others while on our desktops, laptops or tablets, other activities we take part in can be monitored and can form part of our lifelogs. The Stuff-I've-Seen system from Microsoft Research, Dumais et al. (2003), is an example of such a logging tool, as is the work by, d'Aquin *et al.*, on monitoring and logging web information exchanges, d'Aquin et al. (2010). This area is sometimes known as Personal Information Management, Elsweiler and Ruthven (2007); Jones (2007), and generally only focuses on an individual's web/PC interactions. There are a number of tools that can log web pages accessed and general computer activities [7], either via apps run, keystrokes input or via screenshots of computer interactions. Steven Wolfram has logged every email he wrote since 1989 and has logged over 100 million keystrokes; he uses this to personally analyse his own 'intellectual activities[8]. Additional sources of lifelog data could include home energy consumption, distance travelled in one's car, financial history, and so on;

- Active Capture of Life's Activities. Machajdik et al. (2011) refer to either indirect or direct logging of activity but it is initiated

[5]SMS backup tools such as SMS Backup +

[6]Phone call recording apps are readily available, such as Call Recorder for Android. There are, of course, issues with non-consentual recording of voice conversations, but in the majority of US states, it is acceptable to record phone conversations if at least one member of the conversation has given consent.

[7]A popular tool to record general computer activities and web pages accessed is called RescueTime and is available for MAC OS

[8]`http://blog.stephenwolfram.com/2012/03/the-personal-analytics-of-my-life/`

by the user and covers things such as blogging or status updates on social networks or posting of videos or photos. Because these are not passive and "always on" activities it is debatable whether they are really lifelogging or not.

To this list of lifelogging activity types we also add the following:

- Environmental Context and Media. Lifelogging is mostly, but not exclusively about recording using wearable technology and an example of where it is is not is in a smart home. Sensors in our homes can record presence using passive infra-red sensors, pressure sensors in chairs or beds, activities can be inferred from monitoring of electricity, gas or water consumption in the home, and so on. While these are primarily developed to support remote monitoring of independent living by older people, e.g. work in Ahmad Jalal (2012), this can also be a form of ambient lifelogging. In addition, de Silva et al. (2007) developed a system for retrieval and summarisation of multimedia data from a home-like environment that continuously captured video and audio sequences of home activities and presented them on an interactive user interface;

- Passive Capture Audio. This involves the identification of real-world activities, audio events and activity types by audio sensing alone, e.g. Al Masum Shaikh et al. (2008). Recording actual audio is something most people are actually quite uncomfortable with and we have already mentioned phonecall recording in the previous list. Consequently, the focus here has been on using techniques like HMM classifiers to determine the *type* of activity the user is situated in. Work by Ellis and Lee (2006); Shah et al. (2012) has shown loggers are interested in using audio to identify things like location, activity type, to recognise people in the vicinity based on voice matching or perhaps keywords or phrases spoken in a dialogue. In addition, Hayes et al. (2004) developed the personal audio loop, a ubiquitous audio memory aid, while Heittola et al. (2010) developed an audio context classification system using audio event histograms;

In many cases, data gathering in lifelogging does not neatly fit into one of these categories, so we need to explore them as cross-category sensing tools. In recent years, this has meant using the mobile smartphone and its range of sensors, as the platform for one integrated unit and as a consequence, many of the above are easily available as consumer devices. For the more serious quantified-self lifeloggers or for the extreme ones, there are other examples of sources of lifelog data which can be captured. Funf Human Dynamics, as described in Dong et al. (2011), is a smartphone gathering infrastructure allowing anyone to create a sensing app and host the gathered data in a service of their choice. Qiu et al. (2012) have also used the smartphone for mobile sensing, developing a holistic lifelogging tool that is concerned with multiple sensor data gathering, upload to a storage and processing server and search / interaction through a web interface. Finally Gurrin et al. (2013) have shown that the smartphone image logging is comparative in real-world use-cases to a custom device such as a Vicon Revue.

A final category of lifelogging which gives yet other sources of lifelog data, are applications for performance monitoring, typically where we want to monitor the performance of the human body in some testing application such as in sport, or in extreme work activities. An example of the technology available is the Equivital Technology Platform™which records physiological parameters including ECG, heart rate, respiratory rate, GPS location, movement from tri-axial accelerometers, body position, motion intensity, falls detection, skin temperature and core body temperature (from a capsule swallowed which wirelessly transmits data), Galvanic Skin Response, and blood Oxygen Saturation levels (SpO_2). This is definitely not a consumer-level device but is used in sports science applications and because the data is captured passively and without intervention, it counts as a form of lifelogging. Other related devices that are aimed at the consumer market would include the Basis watch that actively records heart rate patterns, body motion, calorie expenditure by activity, multiple sleep stages, perspiration and skin temperature.

As lifelogging moves from being an extreme activity, engaged in by few early adapters, to being a more widely-practiced activity, the variety and range of sensors available will inevitably increase. Coupled with this, it is likely that a platform such as a smartphone (or whatever it morphs into over the next few years - e.g. Google Glass) will be the platform that other wearable sensors will interact with. This platform will send data, in some power-efficient manner to a server (local or cloud based, see Chapter 6 for this discussion) where it will be both stored and processed as a form of personal big data. Given the current variety of devices, the need to temporally align and utilise the data to create higher-level semantics is a key challenge for big data analytics and information retrieval.

3.2 Lifelogging: personal big data — little big data

"Big Data" applications are generally believed to have four elements which popularly characterise a big data scenario, and these are volume, variety, velocity and veracity.[9] In this section we will examine how lifelogging does, or does not conform to those four characteristics because there are certain advantages which "big data" technologies could bring to the lifelogging application.

Lifelogging is essentially about generating and capturing data, whether it comes from sensors from sensors, our information accesses, our communications, and so on. One characteristic which makes lifelogging a big data application and poses both challenges and opportunities for information retrieval, is because of the *variety* in the data sources.

Primary data includes sources such as physiological data from wearable sensors (heart rate, respiration rate, galvanic skin response, etc.), movement data from wearable accelerometers, location data, nearby bluetooth devices, WiFi networks and signal strengths, temperature sensors, communication activities, data activities, environmental context, images or video from wearable cameras, and that doesn't take into account the secondary data that can be derived from this primary lifelog data through semantic analysis All these data sources

[9] see `http://www-01.ibm.com/software/data/bigdata/` for an example of popular characterisation of what makes a big data application

are tremendously varied and different. In lifelogging, all these varied sources merge and combine together to form a holistic personal lifelog where the variety across data sources is normalised and eliminated.

This naturally poses a real challenge to automated analysis. For example, the capture frequency from sensors can range from +30Hz for devices such as accelerometers, Zhang et al. (2012), to less than one reading per day for course-grained location changes or personal encounters, Byrne et al. (2007).

The *velocity* of data refers to the subtle shifting changes in patterns within a data source or stream, and this is not much of an issue for lifelog data, yet, because most lifelogging analysis and processing work is not done in applications which require identifying a pattern or a change in real time. This is one of the trends for future work; real-time pattern analysis could potentially be employed for contextual information retrieval, healthcare monitoring and real-time interventions.

Lifelogging generates continuous streams of data on a per-person basis, however despite the potential for real-time interactions, most of the applications for lifelogging we have seen to date do not yet operate in a real-time mode. So while lifelogging does not yet have huge volume, this volume of data is constantly increasing as more and more people lifelog. For a single individual, the data volumes can be large when considered as a Personal Information Management challenge, but in terms of big-data analysis, the data volumes for a single individual are small. Considering a lifelog of many people, thousands, perhaps millions, all centrally stored by a service provider, then the data analytics over such huge archives becomes a real big-data challenge in terms of *volume* of data.

Finally, *veracity* refers to the accuracy of data and to it sometimes being imprecise and uncertain. In the case of lifelogging, because much of our lifelog data can be derived from sensors which may be troublesome, or have issues of calibration and sensor drift, as described in Byrne and Diamond (2006). Hence, we can see that lifelogging does have issues of data veracity which must be addressed. Semantically, such data may not be valuable without additional processing. In applications of wireless sensor networks in environmental monitoring, for example, trust and reputation frameworks to handle issues of data accuracy have been developed, for example RFSN (Reputation-based Framework for High Integrity Sensor Networks) developed by Ganeriwal et al. (2008). Similarly, data quality is a major issue in enterprise information processing.

As lifelog applications will become more widespread, we can see that lifelogging does indeed have a big data application with a requirement to provide facilities to extract meaning, etc. in order to create surrogate memories based on useful and meaningful lifelogs, which is both the end goal and the big challenge for information retrieval over lifelogs.

In Baeza-Yates and Maarek (2012), the three major developments in web search are highlighted; they are firstly, addressing scalability to enable huge numbers of web pages to be crawled and searched, secondly to usefully exploiting web link structure and more recently, mining the signals provided by users when interacting with a search engine. An example of the latter is the introduction of the Panda ranking factor into the Google page ranking algorithm in order to account for positive user experiences when visiting a website as exhibited by their click-through behaviour, see Kent (2012). This is effectively an application of big data, that is efficiently mining data (sensor) streams for patterns and linking and cross-correlating information sources in order to deduce new knowledge; in this case using machine learning to learn about user behaviour and experience so as to improve web search. What this means for lifelogging is that it has now been demonstrated that deep data mining and linking of information can be achieved at an enterprise scale, and so should also be achievable to an even greater degree at a personal level. This makes lifelogs a form of “little" or personal big data.

3.3 Storage models for lifelog data

There is prior work on storage models for lifelogging data, which is about data quality, data cleaning and data integration. Because of the comparative ease of working in sports and fitness as opposed to health and wellness, most of this work is focused on lifelog data gathered for monitoring sports performance, sometimes at elite or upper levels because the subjects tend to be fit and well as opposed to having medical issues. Typical of this kind of sensing is where a single integrated sensor platform is not available for monitoring participants, say combining heart rate or respiration rate with on-field location or movement data (speed, distance covered, etc.) but instead the challenges of time alignment, normalised sampling rates and handling missing or error-some data are addressed directly, see Roantree et al. (2012).

The management of sensor data quality can be achieved at the device level, as part of the data integration, or the cleaning, alignment and normalisation can be treated as a back-end process, using an integrated storage warehouse such as that based on XML and described in O'Connor et al. (2009). At the device level, FitBit One is a popular example of a device that processes its own captured data (from onboard accelerometers and gyros, with high precision timestamps) and processes this raw data into more usable information including number of steps taken by the wearer per day, number of flights of stairs climbed per day, and estimates of caloric expenditure per day while factoring in activity levels, weight, gender and age of the wearer. This is achievable because all of the sensors are on the one device and thus there are no timing issues because all sensors are synchronised off the same timing chip and data is captured, processed locally and stored on-board until it is uploaded to cloud storage wirelessly. Most wearable sensing follows this model of store on-board and upload later, though recent developments in real-time lifelogging, De Jager et al. (2011); Qiu et al. (2012) allow analysis and processing of sensor data, and mining of behaviour patterns over time, on the cloud which provides much greater value for the lifelogger.

If the lifelog integrates data from multiple independent sensor sources or hardware devices, or if it integrates external data from third parties (such as pictures of you taken by somebody else at a social event or weather conditions at the location in which the wearer was situated), then data cleaning, alignment and perhaps temporal normalisation may be required. These are all now fairly well-understood processes which can overcome problems caused by gaps or missing data as well as the issues caused by differing sampling rates. As we are seeing more and more data repositories expose their data through open APIs, including social network sites (FaceBook, LinkedIn), news sites (bbc.co.uk, cnn.com), personal media sites (Flickr, YouTube), the issues of data cleaning, alignment (temporal and spatial) and provenance will increase in importance.

Whichever approaches to gathering and storing lifelogging data are used, whether on-board storage or real-time upload, or using single integrated sensors or multiple independent platforms, there are pros and cons to each and the whole area of data quality in the case of heterogeneous sensing remains a challenge for lifelogging. In fact, mixing and aggregating multiple independent sensor data streams is a challenge not just for lifelogging but also for other real-world sensing applications such as environmental monitoring, traffic or city monitoring, or monitoring the weather.

The Table 2.1 previously presented illustrates the volumes of lifelog data that could be stored. The richness of the data gathered impacts hugely. It is trivial to store small data volumes such as location and accelerometer based data (as in the Moves app), but all-day video, on the other hand, poses major challenges for any data storage technology, not to mention the organisational challenges of adding value and making this data actually useful for the individual who is gathering it.

With regard to the question of where to store the data, there are two principal alternatives and there are pros and cons of each also. Firstly one could ask/allow the user to store the data locally on their own computers. This naturally gives the user control of their own private lifelog data, but at the expense of less security and redundancy of the data, as well as the potential for higher latency access, or no access at all, to the lifelog when away from home. However, we would not assume that most potential lifeloggers would be able to adequately secure and maintain their own lifelog data on home computers, so this may not be a reasonable option.

The second option is to store the data in a cloud hosting service, as many people now do using services such as Dropbox or Google Drive. However, then there are many privacy and security concerns about storing such vast personal archives of non-curated personal data in a cloud-based service, as will be discussed in Chapter 6. There are also major costs associated with cloud storage. Consider a company such as Narrative who offer the Narrative Clip with free online storage for a year. In a given year of typical all-day use, the Narrative Clip would generate about 500GB of visual data. This is a huge storage requirement per user and financial/business models would need to take this into account, along with the (likely and reasonable) expectation that lifelog data should be stored indefinitely. There is some potential to remove redundancy in the data and typically lifelogging creates a lot of redundancy. For example, the author in writing this section has probably captured over 100 images, all containing similar views of the laptop in a coffee shop. When considering removing redundancy, this has previously been explored in various domains such as audio or video surveillance applications; prior work in lifelogging shows how progressive redundancy elimination can optimise storage utilisation and minimise semantic loss of information, as described in Gurrin et al. (2009). However, given our definition of lifelogging, this is not ideal because it impacts negatively on the totality of capture concept. In any case, one would imagine that centralised cloud-based storage is the best option, especially when considering the types of semantic analytics that could be performed to segment, annotate and make use of the lifelog data.

4

Organising Lifelog Data

There are a number of challenges that immediately arise when organising lifelog data. These range from capture and storage to processing and presentation. We have already discussed capture and storage, i.e. making the lifelog, now we need to turn our attention to examining the prior work that has gone into making the lifelog searchable and useful; the tools that turn a lifelog into a surrogate memory. Therefore, it is important to review the state-of-the-art in terms of lifelog data processing and presentation. A key challenge for processing of lifelog data is in extracting meaningful semantics from the content; bridging the semantic gap, which was introduced by Smeulders et al. (2000). However even before we begin to extract semantics, there are other challenges to overcome. Different from other IR tasks, such as web or blog search, in lifelogging there is no concept of a document or even a generally accepted atomic unit of retrieval. As we will see later in this chapter, the unit of retrieval is heavily dependent on the use-case.

Another challenge with lifelog data is that since the majority of data originates from sensors, most of the data exists in a form which is not searchable using well-known information retrieval techniques, where the concept of relevance or degree of similarity, is integral. This naturally poses challenges for any retrieval system and there is even no single way that a user can form a query to actually extract value from a lifelog. In this chapter we explore the ways that raw lifelog data is (and could be) semantically enhanced and organised, so as to bring value to the individual who engages in lifelogging.

Since lifelog data is passively and indiscriminately captured human data over extended periods of time, it naturally has a historical context and can be seen to reflect an individuals' memory of the past. Hence, for data organisation and search tools in lifelogging, the initial starting point has always been to frame the discussion in terms of the human memory system. Hence, in collaboration with cognitive neuropsychologists, a core set of baseline principles has been identified for a useful lifelog, based on the Cohen and Conway (2008) model of episodic memory. These principles, as described in Doherty et al. (2012), in combination with our practical lifelogging experiences, Gurrin et al. (2008a); Doherty et al. (2013b), and can be summarised thus:

- *Segmenting* raw unprocessed lifelog data into meaningful units provides us with the basic atomic unit of retrieval in lifelogging. In prior work, this unit has typically been the event, where the event is a (temporally related) sequence of lifelog data over a period of time (with a defined beginning and end). Events are not new and prior work exists on event detection and segmentation in many fields. We consider events are the starting point for developing lifelog retrieval systems in that they can provide a document-like retrieval unit, however,as we will see in this chapter, the event is not always the appropriate unit of retrieval for many use cases. For example, in quantified-self analysis, or when reminiscing about the past, or in retrieval of specific facts, a summarised or aggregate of the data is more appropriate than a listing of events. Consider, from an information retrieval perspective, that document summarisation, novelty detection or question answering are examples of techniques that remove the need for retrieval of the entire document in response to an information need, so even in information retrieval there is an acceptance that the document is not always the answer for the user. However, events have received a lot of prior attention in lifelogging research, so we will frame the discussion around the event as the retrieval unit;

- *Annotating* events (or other atomic units of retrieval) with meaningful semantics supports retrieval from a lifelog. In an effort to bridge the semantic gap between sensor data and human understanding as expressed in a user's information need, it is necessary to employ semantic extraction techniques to generate meaningful annotations for lifelog data;

- *Access and retrieval* makes a lifelog useful because since a lifelog is likely to be very large in size, it is certainly too large to be effectively accessed via a browsing methodology. Supporting appropriate search and retrieval technologies over a lifelog will help to address many of the use-cases that we will define for lifelogging, as well as those that we have yet to encounter. Note that we don't just assume search as the underlying access methodology. Just as general information retrieval covers a wide range of techniques from ad-hoc retrieval to summarisation to sentiment analysis to question answering, lifelog use cases will require different methods of accessing and querying the lifelog to extract appropriate knowledge for these use cases. For lifelogging, these methods range from straightforward ad-hoc style querying to summarisation, narrative generation and question answering; in fact the whole spectrum of information retrieval access mechanisms could be applied to help with different use-cases for lifelogging;

- *Multimodal interaction* should be supported when users want to access lifelogs. Individuals who want to access their lifelogs will not necessarily be doing so from a desktop PC or a mobile handheld devices. New technologies such as Google Glass will support omnipresent access to information and to our data archives. Lifelog access methods should cover a wide range of interaction paradigms and access devices.

The key technical advances to date associated with these four areas are now outlined.

4.1 Identifying events

With so many streams of raw data potentially being gathered in lifelogging, it is necessary to organise these streams by structuring them into something meaningful, based on their content, either at indexing time or at query time. In other applications of managing content, the "unit" of information which is to be managed is obvious. The early days of managing digital content were predominantly focused on text document retrieval in which the unit of retrieval was fixed as the document, and information retrieval researchers developed techniques for ad-hoc document retrieval. In some cases when dealing with very large text documents, sub-document units like paragraphs or structural units like sections were considered, but the important point is that the document as the unit was fixed and unambiguous and that unit did not need to be defined or identified.

Moving away from text information, we see where this simple view of the document as the unit breaks down. Consider video as an example; content-based video retrieval has received significant research attention through evaluation campaigns such as TRECVid, Smeaton et al. (2006), and PASCAL, Everingham et al. (2010). Video is composed of a hierarchy of structures from the lowest level of pixels that combine to make visual frames, that are sequentially presented at a given frame rate to constitute video shots, which are then logically combined to make scenes that then combine to become whole semantic videos that are part of video streams. In video search the unit of retrieval can be the shot, which is the default in the TRECVid ad hoc search task for many years, or it can be the scene or perhaps the whole video and there are search tasks for which the scene or the whole video is the retrievable unit of information. For other applications in video search the unit of retrieval can be a specific event in the video, such as the example of a goal being scored in a football match, or an event like a cake being baked or a shelter being made as in TRECVid multimedia event detection task.

In general, the notion of retrieving specific segments of content from a continuous stream such as in video, is addressed by structuring the stream into "segments", based on an analysis of the content.

These segments can be shots, scenes, events or any other logical unit. In some cases, the segments may even be query dependent, such as in the TRECVid Multimedia Event Detection task or in the MMM Video Browser Showdown competition as described in Schoeffmann et al. (2013).

When we consider lifelogging, the unit of retrieval or the unit of content that we wish to manipulate is quite uncertain; we have not had sufficient use-cases to define accurately what the all the potential units should be. The varied sources of data can range from, at the low-level, triaxial accelerometer readings (numeric triplets), through GPS locations (longitude, latitude, elevation and speed readings) or a temperature reading (single value) through to simple images, segments of wearable video, documents accessed, and so on. Because of this variety, it is the use-case that really does define the unit of retrieval. While this is still an undefined area, the dominant unit of retrieval that has been employed in lifelogging has been an atomic unit called the event until now.

Event detection and segmentation is not a new concept. Automatic recognition of events, where the event has some semantic significance, has been identified as important for many years in applications like managing personal photos, as described in Lim et al. (2003), indexing surveillance videos by Atrey et al. (2006) and summarising sports videos in work by Sadlier and O'Connor (2005). In fact a whole series of workshops have been held with a specific focus on the automatic identification of events from data streams, Doulamis et al. (2008); Scherp et al. (2010). However, in each application of event segmentation, one needs to know what an event is. According to Zacks and Tversky (2001) "an event is a segment of time at a given location that is conceived by an observer to have a beginning and an end". However in lifelogging, events don't necessarily even have to be in a given location; consider the event of driving to work, which will have a starting point, and end point and potentially many locations in between. Consequently, the concept of an event in lifelogging is more flexible than the above definition. Therefore, for lifelogging, the focus has, until now, been on the event as the umbrella unit, with the event merging various sources of sensed data together into one logical unit, for which we typically have of the order of a few dozen per day, which we know from Doherty and Smeaton (2008b).

For better or for worse, organising data into events has now become an accepted practice thus far for lifelogging research, as the first phase of analysis and processing of lifelog data, Doherty et al. (2012, 2011b). This mirrors efforts in other domains which organise continuous streams of data into some unit of meaningful information. For example in the memory science domain, Zacks et al. (2006) notes that "segmenting ongoing activity into distinct events is important for later memory of these events". Meanwhile in the video retrieval community Smeaton et al. (2010) suggests that "automatic shot boundary detection is an enabling function for almost all automatic structuring of video". In addition, ' Lin and Hauptmann (2006) state that for lifelogging "continuous recordings need to be segmented into manageable units so that they can be efficiently browsed and indexed'. Furthermore while analysing epidemiological accelerometer datasets, Troiano et al. (2008) states that "for comparison to physical activity recommendations, 10-min activity bouts were defined as 10 or more consecutive minutes above the relevant threshold", which ties in with the concept of an event.

A consistent event model and an event-centric notion in lifelogging is therefore a useful starting point for this discussion on organising lifelog data. Previously the dominant presentation paradigm for reviewing lifelog images was a conventional sequential "replay" or fast-forwarding of all captured images, more formally known as RSVP, Spence (2002). This clearly would not be scalable beyond a few days of lifelog data however. Following that, early applications of image clustering to the event segmentation task either defined events as being of a fixed duration of time or were adaptations of approaches used to identify scenes in video, as described in Wang et al. (2006); Yeung and Yeo (1996), or stories from conventional manually captured images. Event segmentation models over lifelog data could make use of the entire range of sensory data being gathered, which suggests that there would be great diversity in the types of approaches that could be employed. Since we are more concerned with the capture of a totality of life experience, many of the prior works we mention are reliant heavily on the processing of visual or audio data. Some of the main event segmentation techniques proposed in the field of lifelogging include:

- Ellis and Lee (2006, 2004) used audio data to segment lifelogs into distinct activities (e.g. in the house, on the subway, at the park, at work, in a meeting, etc.), with boundaries to the nearest second;
- The Princeton Approach segmented lifelog videos into clips of fixed duration as described in Wang et al. (2006), with boundaries to the nearest 5 minutes;
- The time-constrained clustering technique from Lin and Hauptmann (2006), with boundaries to the nearest second;
- Yeung and Yeo (1996) used a time constrained clustering technique to segment the data, with boundaries to the nearest second;
- Doherty's work on an early prototype and adaptation of Hearst's text-tiling technique, Doherty et al. (2007); Hearst and Plaunt (1993), with second level boundaries;
- Another approach by Doherty and Smeaton (2008b) which used sensors such as movement, light intensity and passive infra red readings from a wearable device (a SenseCam) to define event boundaries through a process of detecting activity level changes, with second level boundaries.

The Doherty and Smeaton (2008b) approach using motion sensors represents the most widely deployed event segmentation model for lifelogging to date and the source code for this has subsequently been released as open source Doherty et al. (2012) and represents a good baseline for future efforts into improving event segmentation. The output of this event segmentation model is shown in Figure 4.1, in which a day of lifelog activities are segmented into approximately twenty events, each automatically annotated with a novelty value, which are grouped

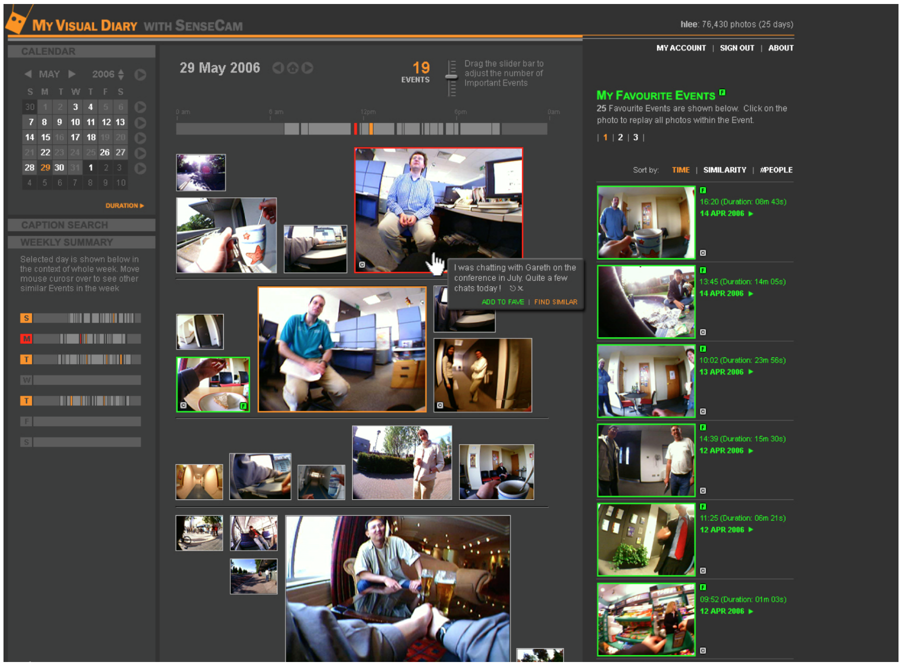

Figure 4.1: A Visual Lifelogging tool Lee et al. (2008). The screenshot shows, in the centre panel, a day, segmented into events, with each event represented by a keyframe, the size of which is dependent on the novelty of that event. On the leftmost panel is the date browsing calendar and on the right is a ranked list of similar events to the currently selected event.

into morning/afternoon/evening and night clusters and a representative keyframe chosen for each event, with links to similar events for any event selected.

Given these event segmentation models, it is still important to note that an event is a very subjective concept and it is unlikely that one event segmentation model that segments a lifelog sequence into a set of discrete events will suit all (or even many) use-cases. Given the opportunity to manually segment a lifelog representation of a day into events, two people will very likely come up with different segmentations. This is hardly surprising, especially if one of the two is the data owner. In fact it is suggested that owners of lifelog data are better judges (than non-owners) on their data as they have the best knowledge of the semantic meaning, and worldly context, of that data. However, even the same individual at different points in time is likely to come up with different segmentations, as we know from Doherty et al. (2009). This has been called "event decay" and models the fact that, with the passing of time, the human mind forgets about an event that has happened. It was found that between 20-36% of original event boundaries were not considered important a number of years later. Comparing to prior IR work on video segmentation, or event segmentation from conventional digital photo streams, Graham et al. (2002), where photo capture is naturally bursty in nature, the challenges of working with continuous lifelog streams are clearly identifiable. The idea of event segmentation and event decay needs to be considered when developing real-world lifelogging retrieval systems and it the concept of a segmented event as the atomic unit of retrieval, while the de-facto standard at the moment is likely to be replaced by other more dynamic techniques in the coming years.

4.2 Annotating events and other atomic units of retrieval

In much the same way as textual annotations are used to bridge the Semantic Gap in multimedia information retrieval, as described in Smeulders et al. (2000), lifelogging requires a similar process to bridge the semantic gap to allow a user to locate an event (or other unit) of information without engaging in an exhaustive browsing effort or relying on ineffective date/time search methods, Naaman et al. (2004); Doherty et al. (2012). Perhaps the closest analogous task is in digital photo retrieval, where past work for example that of Anguera et al. (2008); O'Hare et al. (2007) and many others, have shown the potential of using metadata or sensor data to automatically generate annotations.

So, for lifelogging, once events (or some suitable atomic unit) have been identified, there is then a need to generate meaningful semantic information for users in a phase of semantic enrichment. We now discuss some techniques used in the literature to annotate lifelog events. This can refer to internal representation which is used within a system and is not intended to be visible to users, similar to the way the internal

Sixteen most Enjoyable Activities			
Intimatcy	Socialising	Relaxing	Spirituality
Eating	Exercising	Watching TV	Shopping
Preparing Food	On the Phone	Napping	Taking care of my Children
Computer Use	Housework	Working	Commuting

Table 4.1: The sixteen most enjoyable life activities as identified by Kahneman et al. (2004) using a convenience sample of 1018 employed women in 2004. The table should be read from left to right and top down (from Intimacy to Commuting).

representation of a web page by a web search engine is not intended to be visible to the user. Alternatively, or even in parallel, it can refer to a form of manual annotation of events, similar to the way that videos or still images are manually annotated or tagged in YouTube or Flickr.

Like other forms of multimedia retrieval in which there are multiple levels of semantic enrichment, lifelogging is no different. Single sensor sources could define basic semantics such as the user activity (from accelerometers), audio events in the environment, low-level visual features (such as colour or texture) from video or image streams, or even the objects/shapes in these video or image streams. Alternatively, one can offer higher-level semantics that can characterise the lifelog data. For example, a basic, canonical set of higher-level concepts that can categorise daily life activities and that could apply to most/many people have been used to define event categories in some lifelogging work, Wang and Smeaton (2011, 2012). These categories have been determined in advance by Kahneman et al. (2004) and colleagues from work in the area of behavioural economics and are shown in Table 4.1. On the other hand, prior work also includes best-guess efforts by computer scientists, Doherty et al. (2011a), in which the events chosen (e.g. at a desk, in front of a door, driving, eating, etc.) were selected due to their potential for describing the common life activities of office-bound researchers.

Since we all lead different lives and do different things, it follows that the sets of semantic concepts which define our unique behaviours should either be unique to the user, or be so broad as to cover most of the possible concepts that exist in daily life (e.g. the behavioural economics life activities in Kahneman et al. (2004)). As implementations of concept detection become faster and more accurate, we are likely to witness a progression from small numbers of carefully chosen concepts to large-scale generic concept detection. Recent progress in video search allows a user who is searching for a known item video which s/he has seen previously and recall as something obscure, like *"Yellow single-decker bus driving towards a camera with vegetation and cityscape in the background, in the daytime"* to build a concept detector for that topic on-the-fly using positive and negative examples, Oneata et al. (2012); Chatfield and Zisserman (2012). Such dynamically constructed concept detection can use external images (for example from Google Image Search) as candidate positive (and negative) images and a classifier can be learned in real time, which is then used to process the video archive. The potential for this in lifelogging is that it could be used in a slightly refined way, to learn the most common behavioural patterns from a lifelog, for annotation.

Since lifelogging is concerned with daily life experience, we can turn to prior work concerning meaningful reflections on daily life. Lee and Dey (2007) found four general categories of cues (from fourteen experiences) were most effective: people (7 of the 14), action (4), object (2), and place (1). This is similar to the uncategorised listing of contextual data sources in early photo retrieval, see Naaman et al. (2004). It would therefore be advantageous to detect and interpret implicit semantics of lifelogging data from heterogeneous sources to explain the *Who, What, Where and When* questions which are common in everyday events, Bristow et al. (2004); Doherty and Smeaton (2010). These four primary types of context information were proposed by Dey and Abowd (1999) as the fundamental information to generate an historical diary, incorporating aspects of location, identity, time and activity. We will base our discussion of event semantics around these four facets of event description, though this is only one such categorisation and several other models of categorisation could be equally valid, if indeed a model of categorisation is required at all.

4.2.1 Annotating lifelogs - who

Attempts to identify people *who* where co-present in an event have mainly used Bluetooth sensors to scan and log co-present devices, Eagle and Pentland (2006); Lavelle et al. (2007). In addition to annotating people directly, the number and distribution of co-present people also helps determine an event's distinctiveness or uniqueness. For example, Aizawa et al. (2004b) found faces helpful to detect event importance scores. Others extended this event importance scoring approach by merging visual uniqueness, based on MPEG-7 visual features for example, with the number of encountered people in a given event, Doherty and Smeaton (2008c). Distinctiveness is also a critical issue in autobiographical memory, as found by Brewer (1988), but we don't yet understand what makes some things more distinct than others, though this is a topic in current research, Hebbalaguppe et al. (2013). The concept of distinctiveness is also taken into account in the information retrieval domain, for example web page importance or in various forms of novelty detection, Allan et al. (2003b).

Detecting people and faces is also useful in helping to select representative "keyframes" for events, should image data be present. In addition, image saliency has also been shown to be important when selecting event keyframes, as discussed in Doherty et al. (2008). These ideas were inspired from description of keyframe selection in the video domain by Girgensohn and Boreczky (2000), and represented an alternative to simpler techniques such as selecting the middle image in each event as proposed by Smeaton and Browne. (2006); Blighe et al. (2008). Keyframes can subsequently become the source data for visual analysis or for use in a browsing interface.

4.2.2 Annotating lifelogs - what

To annotate *what* type of activity is occurring in events, efforts in lifelogging have focused mainly on employing computer vision and audio processing technologies. Audio processing assumes a continual audio log of daily activities and using computer vision assumes wearable cameras, such as the SenseCam. Wearable video, would of course capture both modalities together and recent experiences suggest that the inclusion of additional sources of evidence (such as GPS location, user activity, time of the day) can all help to inform the semantic annotation process. There have been some efforts using audio, for example Kern et al. (2007) used a combination of audio activity levels and movement/accelerometer sensors in an attempt to identify what was happening at a given point in time. There have also been audio-based approaches that attempt to convert spoken words into textual transcripts, but none in the lifelogging domain, to the best of our knowledge.

Given our focus in this review on capturing a totality of life experience, computer vision based approaches will receive most attention. Such approaches have involved using general multimedia processing techniques for internal representation of events in lifelogging such as using MPEG-7 visual features, Salembier and Sikora (2002), SIFT (Scalable Invariant Feature Transformations), Lowe (2004), SURF (Speeded up Robust Features), Bay et al. (2006), search using a bag of visual words approach, Nistér and Stewnius (2006), and others. This work involves exploring image feature vector similarity options, Kokare et al. (2003); Rubner et al. (2000), and also merging different data sources together, Montague and Aslam (2001); Fox and Shaw (1993). All similar approaches generate signatures for a given image from an event. From these signatures it is possible to either support 'find visually similar' browsing interaction or they can be used as input into a higher-level event classification.

The goal of higher-level image based approaches has been to apply an automatic form of semantic labelling to map lifelog images to given concepts or activities, Doherty et al. (2011a), generally based on Support Vector Machine (SVM) learning, Joachims (2002). Since low-level features can be extracted automatically from media objects including lifelog image content, these are assumed to correspond to the semantics of the query in multimedia information retrieval, and to the semantics of the lifelog event in our case. The FnTIR review of Concept Based Video Retrieval by Snoek and Worring (2009) provides the background information on identification of semantic concepts (such as indoors, outdoors, eating, cars, explosions, etc.) on visual media including video, based on the extraction of low-level features (such as MPEG-7, SIFT/SURF, etc.), and such techniques have also been employed on visual-based lifelogs to semantically enrich the content. In multimedia information retrieval, state-of-the-art techniques use statistical approaches to map low-level features to concepts which are then fused to relate to high-level query topics, Snoek et al. (2006). The whole task is generally broken down into two steps: the detection of a set of concepts and the association of concepts with queries. Semantic concepts are usually automatically detected in a mathematical way by mapping low-level features to high-level features. One state-of-the-art approach is to apply discriminative machine learning algorithms such as SVMs to determine the most likely concepts given the extracted features Snoek et al. (2006), though we note a move towards deep learning recently. Compared to a discriminative model which is more task-oriented, generative statistical models such as Markov models try to analyse the joint probability of variables, which are also proposed in concept annotations in work by Li and Wang (2003). This modern methodology facilitates an understanding of topic queries and low-level features by analysing the mapping in a semantic way. To build a large-scale ontology and lexicon for semantic gap filling, large efforts have been made in activities like LSCOM (Large-Scale Concept Ontology for Multimedia), Naphade et al. (2006); Kennedy and Hauptmann (2006), TRECVid, Smeaton et al. (2009) and MediaMill's 101 concepts, Snoek et al. (2006). Smeaton et al. (2009) state that acceptable results have been achieved already within the TRECVid video retrieval evaluation framework for many cases particularly for concepts where there exists enough annotated training data. Based on concept detection, encouraging improvement has been reported showing the efficiency and the effectiveness of concepts for higher level retrieval, Snoek et al. (2006); Neo et al. (2006). We refer the reader to other sources, such as Snoek and Worring (2009) for a thorough overview of such techniques.

Concerning the number of concepts that would be required to support effective search and retrieval result over lifelogs, Hauptmann et al. (2007) considered that 10,000 concepts is sufficient to provide Google-quality retrieval from a video collection. We also know from Dean et al.

(2013) that 10,000 concepts is considered sufficient for a human understanding of the world. In recent work, Google researchers have illustrated that fast, accurate detection of 100,000 object classes can be performed on a single workstation. The total processing time is less than 20 seconds and the approach achieves a mean average precision of 0.16. Such technologies will be a key underlying technology in how we organise and retrieve from large-scale lifelogs in future, as we move from small numbers of concepts to systems that approximate the level of human cognition.

Most of the analysis described above employ visual and audio processing with the aim of identifying the 'what' of human lifestyle. However, when lifelogging, there are likely to be other sensor streams of lifelog data that can aid the process of automatically annotating events. This is one potential benefit of using lifelog data when compared to single/dual sensor sources such as video streams. For example, even the simple sensors on a SenseCam can aid in landmark detection, as shown by Blighe et al. (2006), or location can significantly enhance the potential of landmark detection from Zheng et al. (2009) in any visual stream. Another example is the accelerometer, which can enhance the ability to identify the physical activities of the wearer. Qiu et al. (2010) have shown very high accuracy activity identification using a combination of accelerometer and WiFi signal strength. With lifelogs, the richness of the sources of sensor data have the potential to significantly increase the depth and accuracy of automatic annotation.

4.2.3 Annotation lifelogs - where

Retrieval by location has been shown by Naaman et al. (2004) in photo retrieval to be one of the most effective selection techniques. With the advent of cameraphones with in-built GPS location sensing, the quantity of GPS-tagged photos has increased enormously. With regard to lifelogging, *where* an event has occurred would be considered as a key annotation element and the predominant form of annotating location is based on using GPS and WiFi wearable sensors. For example Aizawa et al. (2004b) captured a diverse set of "context" types using GPS location data. Lazer et al. (2009) used cell phone GPS sensors to annotate where users were for given events. Finally Gurrin and colleagues have extended this approach by also incorporating WiFi sensors to identify fine-grained whereabouts of events in Gurrin et al. (2013).

Indoor localisation has recently started to appear in healthcare applications or as a commodity service in built-up areas like supermarkets and shopping malls where understanding indoor location can have many potential benefits. For example, the “trail” of a user/shopper can be used to select content to push to the user as advertising. Indoor pressure sensors were used by de Silva et al. (2007) to identify location in a home diary application. Other more recent, and potentially most promising techniques used signal strengths from a network of wireless network access points including WiFi and Bluetooth and a recent survey of these can be found in Fallah et al. (2013). While annotation of lifelog events based on indoor location has yet to appear, the fact that the indoor location-tracking systems are becoming available so easy it is inevitable that they will be used in some lifelogging applications.

4.2.4 Annotating lifelogs - when

Naaman et al. (2004) have shown that, for photo retrieval, date/time are not usually effective data organisation or annotation tools, primarily because individuals do not normally remember the time/date of past activities, unless there is some unique cue or some memory support. When applied to lifelogging, this effectively renders any browsing mechanism ineffective when indexing large volumes of user data.

Lifelogging efforts have relied on the ready availability of built-in clocks in capture devices to annotate by time, yet, when capturing across multiple devices, or when travelling across timezones, the time across different devices may slip out of alignment. Therefore, prior work in lifelogging has not focused on developing new techniques to annotate *when* events occur. Rather, the time is utilised as a source of annotation, either in absolute value or in relative value. When time is employed, it should not only focus on objective ideas of time, but also subjective user perceptions of time. For example those in the social sciences have studied in depth the concept of cardinality of time which considers how densely-grained users perceive time to be. Such considerations may improve future efforts to represent the *when* axis of lifelog events.

4.2.5 Making use of the annotations

Naturally, automatically-computed annotations can be used to support search and retrieval from lifelog archives by indexing each annotation element in a database and supporting database-based retrieval. Aizawa (2005) have utilised spoken textual annotations to describe each event and the MyLifeBits project utilised simple narrative (location and time) to represent content. However, it is also possible to employ a text retrieval approach if the annotation elements are merged to form textual surrogate descriptions of the lifelog unit (typically the event). Byrne and Jones (2008) have suggested that narrative text should be generated to represent each lifelog unit. As described by Riedl and Young (2010), "narrative is a sequence of events that describes how the story world change over time", and we know that narratives can be used to support keyword text searches using information retrieval techniques from Jaimes et al. (2004). After all, text is the natural method that users now have to locate knowledge using web search engines.

From narratology (the study of narrative form), we can identify three processes needed to generate narratives from segmented events, namely fabula, sjuzet, and discourse generation. From a lifelogging viewpoint, fabula is the raw material of the lifelog, most probably as a sequential series of meaningful sentences generated from lifelog annotations representing real-world activities, sjuzet is the combination / re-representation of the fabula to generate the lifelog narrative (most probably at the event level), which is communicated to the user by means of the discourse. We are not aware of any published work that generates narrative text from total-capture style lifelog data, however, with the popularisation of lifelogging, the generation of automatically-generated narrative diaries is likely to receive more attention.

All of this presented thus far assumes that the lifelog is a flat collection of non-related events, each represented by a textual annotation and various forms of semantically meaningful metadata. This is similar to the pre-web field of information retrieval where retrieval models were based on assuming a flat collection of text documents. While this is likely to be effective enough for many retrieval scenarios, there is significant potential in considering that a hierarchical data structure does not adequately represent the lifelog of annotated events. A lifelog is more like a densely linked hypermedia archive that provides opportunities for modeling the archive as a graph; this more closely maps onto the proposed structure of human episodic memory. We know from prior research, Richardson and Domingos (2002); Haveliwala (2002); Taneja and Gupta (2010); Gurrin and Smeaton (2004), that exploiting the linkage structure of the web allowed the PageRank algorithm, as described in Page et al. (1999), to enhance the effectiveness of large-scale information retrieval on the web. There are issues to be solved in this research regarding the subtle but important difference between organically created links by humans on the web that carry the latent implication of related content and links between lifelog events that would need to be created by automatic means, at least in the near-term.

4.3 Search and retrieval within lifelogs

Depending on the use-case for lifelogging, the search and retrieval mechanisms employed would be very different. For example, in a simple form of quantified self analytics, the access mechanism may be based on data summarisation and aggregation and there would be little need to actually locate any single sensor reading or semantically meaningful unit. However, in this review, we are taking a view that the lifelog is a media rich repository and the access mechanisms that are supported by the surrogate memory will support a large number of use-cases. This in itself is one of the major challenges facing the lifelogging community; that is effectively retrieving information that is useful to a user for any given information need, or that can be used to underpin some application like memory recall, as shown in Gemmell et al. (2004, 2006); Bell and Gemmell (2007); Aizawa et al. (2001); Tancharoen and Aizawa (2004); Tancharoen et al. (2005, 2006); Hori and Aizawa (2003); Aizawa et al. (2004a). An initial assumption would be to employ state-of-the-art techniques from database search and information retrieval to scalably index the life-experience events in lifelog and provide omnipresent access via keyword/database search, ranking though a desktop interface. This is the approach taken by the MyLifeBits project at Microsoft Research. Perhaps in the world of big-data this should move us towards NoSQL databases, with lower models of data consistency than for regular relational databases.

However, as lifelog archives grow larger, the set-based retrieval model (as used in relational databases) very quickly becomes unworkable. In a similar manner to the early progress of IR technologies, the move towards ranked output for ad-hoc queries becomes necessary. Extending such ad-hoc retrieval mechanisms by adding access via recommendations using historical and immediate context sources as drivers, coupled with multimodal and omnipresent interaction, is the most likely way for progress in user access to lifelogs.

However, there has never been any significant effort made to develop formal models for lifelog retrieval. This is most likely due to the lack of distributable test collections (more about this later) and a lack of readily available users. In most use-cases (principally in the medical domain), the focus has been on short-term data gathering with manual analysis/playback of SenseCam-type data or quantification/summarisation into charts, diagrams or other widgets aimed at increasing our understanding in the lifestyle analysis domain. Progressing from a small, dedicated group of early adapters to a scenario in which large numbers of users are gathering detailed lifelogs means that the research community must face the challenge of providing access tools, so that a lifelogger can locate any piece of knowledge from multi-year archives with the same ease and confidence as one executes a web search today.

Without the significant involvement of many users contributing information needs based on real-world use cases, an understanding of the actual information needs of a lifelogger has never been thoroughly investigated. The small number of early adapters such as Bell, Mann, Gurrin and Aziawa (see Chapter 2) have contributed to the discussion, however such a small number of (atypical) users can not sufficiently motivate the development of a broad set of effective retrieval models.

Figure 4.2: An animated playback window of lifelog data captured from a SenseCam. Shown here is a event segmentation and gesture-based playback tool for a livingroom TV, as described in Gurrin et al. (2010a)

The animation style playback of a day's events (see Figure 4.2), or a quantified-self style chart-based analysis can hardly represent a suitable retrieval model, except in very limited use cases. The first retrieval system developed for lifelog data over large archives was developed by Lee et al. (2008); Doherty et al. (2012). This lifelog search engine employed a photo search metaphor of event segmentation, event annotation and multi-axes search, similar to that of O'Hare et al. (2005). Not surprisingly, it was found that the search metaphor was significantly more useful than a time/date browsing metaphor. The search interface from Doherty's experiment is shown in Figure 4.3 and the browsing interface was previously shown as Figure 4.1.

Without having a large user base, and not wanting to rely solely on the experiences of the few early adapters, we need to look elsewhere for search use-cases and interaction metaphors. Fortunately, there has been consideration of the reasons why people would access their past memories and we can use these reasons as motivations for the types of retrieval models that should be developed. In presenting a constructive critique of lifelogging, Sellen and Whittaker (2010) propose incorpo-

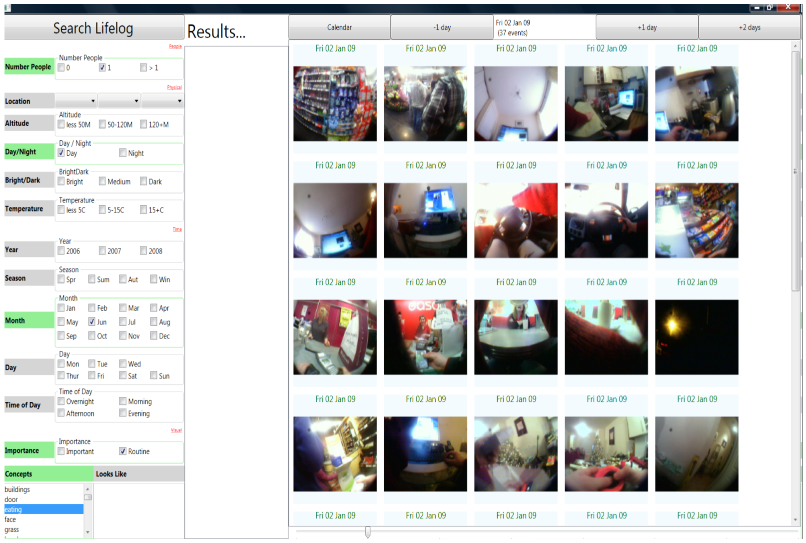

Figure 4.3: Visual lifelogging "multi-axes" browser developed in 2010. The primary design goal was to provide multi-faceted retrieval of events to support person recollection. This browser aims to support search on the "who", "what", "when" and "where" axes of retrieval, Doherty et al. (2012)

rating the psychology of memory into the design of lifelogging systems. In this work, five aspects of human memory access (called the "five R's") are proposed and these provide a first framework within which to focus lifelogging search and retrieval efforts. The five R's are recollecting, reminiscing, retrieving, reflecting and remembering intentions. Each of these can define a different reason why people want to access their memories, and by inference, their lifelogs. Each could potentially become a track in a lifelog evaluation forum, along the lines of the tracks in TREC, but at present, they are valuable to inform our understanding of the potential use-cases.

- *Recollecting* is concerned with reliving past experiences, which can be necessitated for various reasons; for example, we may want to recall who was at a party we attended, or what we said in that New York museum back in 2000. Recollecting is concerned with accessing episodic memories, those memories of our past life experience. To support a user recollecting information from vast multi-year archives will require highly accurate search engines that semantically rank content and extract just the sequence of the lifelog that is most pertinent to the user and present this sequence in as much detail as required to aid recollection. The unit of retrieval may be the event, or it may be a sub-event, depending on the user information need. In any case, this is the closest retrieval model to ad-hoc retrieval and as a starting point, event-level retrieval is certainly a target to aim for. The best example of such a retrieval system is the retrieval interface described in Doherty et al. (2012) that we have previously discussed.

In order to support recollection as a retrieval model, the event that is retrieved in response to an information need is a retrieval unit that is not a memory in itself, rather it helps to trigger the lifelogger's own memory process to recall the actual event from the user's own memory. There have been studies with memory impaired individuals, the aim of which was to use SenseCam to aid recall of past activities, Pauly-Takacs et al. (2011); Berry et al. (2009). In many of these studies the individual's organic memory was triggered to recollect past experiences by some small nugget of information from the lifelog, such as the brand of tea on the table, or the position of the car in the car park. To support user recollection will require conventional information retrieval, coupled with query-specific experience segmentation and the integration of an appropriate presentation representation.

- *Reminiscing*, which is a form of recollecting, is about reliving past experiences for emotional or sentimental reasons. It is often concerned with story-telling or sharing of life experiences with others. Byrne and Jones (2009); Byrne et al. (2011) have explored the potential of lifelogging to enhance storytelling and found visual modalities to be most fluent in communicating experience with other modalities serving to support them and that the users employed the salient themes of the story to organise, arrange and facilitate filtering of the content.

  Reminiscing will need to integrate retrieval to find the correct region of the lifelog to retrieve, which is more likely to be at the event level or at a higher level, such as the morning, the day or even longer periods, such as the holiday. This suggests that it will likely require new techniques for combining multiple relevant temporally arranged events into one coherent segment of the lifelog. Once the appropriate segment is found, narrative generation, topic tracking and detection, novelty detection and summarisation, all operating in conjunction with conventional multimedia document ranking techniques, could will be useful to provide a suitable result representation.

- *Retrieving (information)*, is a more specific form of recollecting in which we seek to retrieve specific information from the lifelog, such as an address, a document, a location, or any atomic piece of information. Such retrieval will require high precision search engines that semantically rank content and extract just the nugget of information that is most pertinent to the user's information need, i.e. not the event, but some other unit of retrieval. This is challenging because the retrieval unit may be unknown until query time, and even so, some forms of semantic enhancement may be necessary to infer an answer from the underlying data, which is composed of non-heterogeneous documents in different media types. The query will define the type of knowledge that is required; it is unlikely that a whole document would be expected in response to a query, as is the norm for web search. Rather the requirement is for a retrieval mechanism like MyLifeBits or some form of question-answering system. In any case, the retrieval is more akin to the question-answering as in Wolfram Alpha[1] than to whole document retrieval as in Google.

- *Reflecting*, is a form of quantified-self analysis over the lifelog data in order to discover knowledge and insights that may not be immediately obvious. We already see many examples of this from the quantified-self community. It is concerned with analysing patterns or viewing past experiences with a different perspective.

[1]http://www.wolframalpha.com/

This is a true big-data challenge in that it will require extracting semantics from large archives over extended periods of time. Given that it is difficult to pre-identify all types of reflection that could form user queries, the retrieval system should be flexible enough to support users constructing queries at query time and the appropriate mechanism should be provided to support this.

To get some idea of the potential types of information needs, we can examine the quantified-self movement who extract knowledge from manually, or automatically, constructed logs of activity. Supporting reflection would require retrieval based on user information needs, various forms of data analysis for semantic extraction and information summarisation to infer new semantic knowledge. An example of supporting reflection is Kelly et al. (2010) who used colour as the abstract representation of life activities by simulating a 1 pixel camera based on SenseCam archives, the Annual Reports by Nick Feltron [2], or finally, the work of O'Loughlin et al. (2013) who used wearable camera images to help participants recall their nutrition intake over the previous 24 hours.

- *Remembering Intentions*, which is more about prospective (remembering future plans) memory than episodic memory (past experiences). This is a form of planning future activities which is a life activity that everyone engages in. This assists people to remind or prompt them about tasks they would like to do (*e.g. post that letter*) given the current context, or real-time prompts on who they are talking to (*e.g. this is Paul*), or giving prompts on conversation cues (*e.g. last time here together, you had just came out of the new Batman film*). Past lifelogging efforts were exclusively focused on episodic memory as it was always a post-hoc analysis (i.e. constrained by technology); however with real-time lifelogging technology available (see §3.3) we can now consider context/situational awareness (and past history of the user) to provide prospective memory prompts as envisioned in Starner

[2] `http://feltron.com`

et al. (1997).

These five Rs of memory access provide an initial guideline to the types of retrieval that are needed to support a wide variety of information access types on lifelogs. One could envisage a TREC track to cover each of these R's at some point in the future.

Even though the five R's provide us with clues as to the types of information accesses that could be performed over lifelogs, we have yet to consider the query construction methodology, which is potentially far more complex than we are used to in current web or media retrieval. One issue that is inherent in all types of lifelog access is how to support a user making a query. Supporting a user to quickly and easily make a sufficiently rich query will not be a simple task. The only prior retrieval experimentation into accessing memories from multi-year lifelogs is from Doherty et al. (2012) and this shows that it took on average 127 seconds to find a known item from a three year archive and that search was successful 75% of the time. Exploring this 127 seconds in more detail, it is clear that the actual process of query construction takes valuable time as the user considers the information need combined with recalled aspects of past experiences to construct the complex query and then selects query factors from a query panel. See Figure 4.3 for the interface used in this experiment. Of course, the concept of a query panel for selecting dates, locations, concepts, people, etc. is not an optimised query interface. In a world of ubiquitous computing where the computer can literally be next to your face, we would simply be able to speak a query into a search engine as if we are asking a knowledgeable friend. Assuming that, the knowledgeable friend still needs to be able to disambiguate a query that will be incomplete yet detailed; not so much“videos of a cat on a piano” as “the name of the bottle of wine that I drank when at a dinner party with Jane and Kate back in Summer, two years ago (I think), at the weekend, sometime after my birthday, and it was a wet night”. One can see the challenge that the field of information retrieval needs to address.

Another factor that is easily overlooked, and one that differentiates some types of lifelog retrieval from conventional information retrieval is that the user query can be inadequate to sufficiently disambiguate one event from possibly hundreds of others like it, or worse still, the user query can be incorrect in one or more aspects. Human memory is inherently flawed and we as individuals are prone to forgetting our past experiences or merging similar experiences together. Another contributory factor in making this retrieval process more challenging is that since the lifelog is effectively a surrogate memory for the user, when an individual wishes to recollect or retrieve information, then she is likely to be querying because she can not recall the experience (or some aspect of) in the first place, so it is even more likely that the query will be incorrect in some aspects. All of this makes the retrieval process significantly more challenging than we are used to in ad-hoc serach. Without larger use-cases, though, it is difficult to draw any detailed conclusions as of yet.

Retrieval efforts on significant lifelog datasets have been sparse thus far, with the few noteworthy efforts. One such example is from Ó'Conaire et al. (2007) showing poor performance in automatically retrieving results for specific queries of interest (P@5 of 0.3). The challenge of effective and efficient retrieval remains unsolved at present.

4.4 User experience and user interfaces

It is important to understand how people will use and access their life archives, rather than simply viewing lifelogging as a new multimedia retrieval challenge and developing top N ranked-list type solutions. Given that lifelogging is a technology in its infancy, as discussed previously, it is not so easy to define all the ways that a user will interact with their lifelog content. Traditionally, information retrieval focused on ad-hoc search which generates, for a given information need a set of, or a ranked list of documents or objects, which are presented to the user as a list. However lifelogging appears to need different types of search depending on the information need such as question answering, summarisation, prompting based on contextual cues, etc. There has been some initial work on design considerations for lifelog content by van den Hoven et al. (2012) which focused on using lifelogs to help people remember in everyday situations, while Whittaker et al. (2012) proposed a set of design principles for developing lifelogging systems that support human memory. Hopfgartner et al. (2013) presented a set of user interaction templates for the design of lifelogging systems and Byrne et al. (2008b) present a set of guidelines for the presentation and visualisation of lifelog content, including the need for a simple and intuitive interface, segmenting the content into comprehensible units, aiding human memory and support for exploration ad comparison. All of these offer meaningful insights into how to design the user interface and user experience for lifelog applications.

Thus far, most lifelog interaction scenarios have focused on supporting user browsing through lifelog archives. Since there are only a handful of long-term lifelog archives available, this is understandable, and in many ways mirror the early days of video retrieval in which video *browsing* systems were initially developed before being replaced with video *search* systems. Looking initially at browsing interfaces, there has been the following prior reported work:

- A SenseCam image browser which facilitates annotations of images of interest and also rapid playback of image sequences from Hodges et al. (2006);
- An event-based SenseCam image browser which automatically segments lifelog data into events or episodes and then allows users to manually annotate those events, Doherty and Smeaton (2008a);
- An update to Doherty's event-based SenseCam image browser, where wearable camera images are combined with accelerometer data which determines event boundaries, Doherty et al. (2013c)
- A gesture-based event browser for use on the livingroom TV which utilised the Nintendo Wii platform to browse through lifelog images via a gesture interface, Gurrin et al. (2010a);
- Representing entire sequences of lifelog images as part of a '"colour-of-life" wheel, made available on touchscreens or desktops, Kelly et al. (2010);

- Asking users to manually tag eating episodes during data collection, automatically analysing the content of those foods, and then allowing users to browse through those image-based food episodes on a web browser, Kitamura et al. (2008);

- An ethnography browser segmenting lifelog data into events, automatically annotating those events for lifestyle traits, and thereafter allowing market researchers to target demographies of interest, Hughes et al. (2012);

- A touchscreen browser integrating event segmentation to allow computer illiterate older adults browse their lifelog content, Caprani et al. (2010);

- An episode based browser to review lifelog video data, Tancharoen et al. (2006);

- The ShareDay system which allows both touchscreen browsing through lifelog events, and event sharing with family and friends and multi-user augmented events, Zhou et al. (2013);

- A browser for segmenting raw data into episodes of equal duration, and thereafter facilitating query-by-example image search, Wang et al. (2006).

- Sellen's chronological browsing as part of memory experiments, Sellen et al. (2007)

- Kalnikaite's browser which combines both GPS and SenseCam images, pairing together images and GPS points if a SenseCam timestamp falls within 50 seconds of a GPS timestamp, Kalnikaite et al. (2010).

These represent different user interfaces that have been developed to browse through (mostly visual) lifelog data, however no clear experiments have been proposed to evaluate the effectiveness of these approaches in comparison to each other. The best solution may represent a combination of some of the aforementioned approaches, but this requires future research.

Moving from a browsing to a search scenario, and without sufficient past research to refer to, we look again towards the five Rs of memory access from Sellen and Whittaker (2010) to provide initial clues. Until such time as we have sufficient numbers of people maintaining personal life archives to get real-world usage data, they serve as valuable source of proposed interaction scenarios.

When recollecting, reliving past experiences, we know that visual media, especially captured from the first person viewpoint, provide very powerful memory cues and leads to what is referred to as Proustian moments of recall, Stix (2011). Any interaction mechanism that supports recollecting from lifelogs should incorporate a visual access methodology. Reminiscing, reliving past experiences, will rely heavily on storytelling and narrative generation. Retrieving (information), locating nuggets of information from the personal life archive, such as a document, a location, a sound, a recipe, and so on. In this case, the focus of the interaction methodology should be on supporting the user in the query generation process and presenting the precise nugget of information that is sought using a suitable interface. Reflecting will analyse patterns and discover knowledge and insights that may not be immediately obvious. The interaction mechanism for query generation is not yet understood, but the user should be able to define source data for analysis, for example, the activity levels correlated with location and time to identify where the user is most active on weekend mornings. The presentation of this data should be highly visual and employ interface metaphors such as timelines, charts and infographics, each of which could support clickthrough analysis of the underlying data to support drill-down reflection. Finally for Remembering Intentions, the user is concerned with reminders and a memory of future activities. The key driver to support this is real-time life experience sampling and query triggering. The key interaction challenges are how to capture the information need, understanding user context and how to present the reminder to the user using some aspect of pervasive interaction.

As can be seen, elements of assisted query formulation, engaging storytelling, summarisation, visualisation and potentially disruptive recommendation are the key research points that will need to be explored in greater detail. An underlying caveat in all of this work is that a human's ability to generate a query or to find a suitable browsing point is highly dependent on their own ability to remember details correctly from the event containing the information they are seeking and that may be a catch-22 situation.

4.5 Evaluation: methodologies and challenges

The field of Information Retrieval has a long and rich history of supporting well designed comparative evaluations. The development of standardised evaluation workshops for Information Retrieval such as TREC, CLEF and NTCIR drove the employment of standardised test collections for evaluation, building on the Cranfield model from the 1950s. This type of evaluation is based on a so-called *test collection*, which consists of a static database of information objects, a static set of *topics* representing information needs, and a static set of judgments of the relevance of the information objects to each topic. A common characteristic of these evaluation workshops is that test collections (assuming the appropriate legalities are in place) are made available to participants to support the comparative evaluation. This has served the Information Retrieval community well; for ad-hoc search, we have seen a doubling of retrieval performance within the first eight years of the TREC evaluation framework, Vicedo and Gómez (2007). Information Retrieval research is based on the availability of datasets and test collections that support comparative evaluation, although in recent years, there has been an increase in the number of non-repeatable experimentation appearing in top conferences such as SIGIR, ECIR, CIKM, and so on. This has caused considerable debate within the community.

An initial analysis of the challenges of lifelog evaluation was prepared by Jones et al. (2008), who found that there were many research questions to be explored. Based on these challenges, and incorporating our own experiences, we can identify that lifelogging poses some unique challenges for evaluation:

- The data is hugely personal to the data gatherer. A lifelog could contain sensor data representing intimate details of daily life, activities and interactions, extending over a prolonged period of time. As we have previously discussed, the sensor data could include images captured at short intervals, videos, locations, people interactions, communications (emails, SMS messages, phone logs, websites visited, etc.); in short, everything that the individual experiences and gathers during the collection of the dataset. Hence, due to the personal nature of lifelog data, as well as inherent legal issues around the release of visual, aural and sensor data of unknown individuals captured in the data, there is real and substantial challenge in sourcing datasets and test collections that could support comparative evaluation. Heretofore, most of the evaluation has been carried out using single (or few) person datasets;

- There are no lifelog crawlers to gather a lifelog dataset as the data is not likely to be publicly available. Crawled collections have, after all, been the foundation of many retrieval evaluation frameworks such as the web search tasks at TREC Bailey et al. (2003), but in the case of lifelogging, the overhead of gathering a dataset is immense. It would be necessary for individuals to diligently gather highly-personal datasets and make them available for community use;

- Most prior work has been performed on small-scale datasets (of the order of weeks or months of data). However, lifelogging is, by its nature, long-term. If one does wish to gather a large dataset, there is a time-delay involved, it takes one year for a person to gather one year of lifelog data; there are very few longitudinal archives of detailed lifelog data that even exist. There is no real value in simulation of the data as the lifelog is supposed to reflect real-life activities and any deviation from those real-life activities will naturally impact on the usefulness of the collection;

- The lifelog is most meaningful to only one person, the data gatherer. While anyone could examine the lifelog dataset and generate queries and relevance judgements, the actual usefulness of a lifelog retrieval system is not simply to generate a good quality ranked

> list; it is to support the individual in accessing past memories using carefully considered reasons, such as those provided by the five Rs that we have earlier discussed. Hence the only truly reliable judge of the accuracy of many potential retrieval approaches is the individual who gathered the data in the first place.

Hence, there are currently no defined holistic benchmarking evaluation tests in lifelogging. Even though there are numerous data sets and collections available across many media, aimed at many different information research challenges, there are none that take a holistic view and gather many sources of data into one dataset. The personal nature of the data and the various legal issues, both comprise the biggest challenge. There are, however, already some collections that could support some aspect of lifelogging research. There are user generated video collections from TRECVid and other evaluation forums which can be used to evaluate individual components of the lifelog collections. There are also digital photo collections, Bolettieri et al. (2009), and even email collections, Klimt and Yang (2004), or spoken conversation collections, such as the Apollo collection of six months of speech transcripts and annotations from the Apollo programme, as described in Oard and Malionek (2013). By employing such collections, aspects of lifelog retrieval can be evaluated, but without a holistic lifelog test collection or dataset, this is still only addressing part of the challenge.

Where lifelog system evaluations have taken place, they have usually been over small collections of data gathered by a few individuals, often a few weeks of data gathered by a few individuals (often less than ten). There are few extreme lifelogging collections in existence at all, the only ones we are aware of are the MyLifeBits data gathered by Gordon Bell and the visual (and other modality) lifelog of Cathal Gurrin. Such archives run into the millions of items and tens or hundreds of millions of sensor readings. We would consider these to be equivalent in size to the ad-hoc/Web TREC collections of a decade ago.

Notwithstanding the non-availability of shared test collections presently, it is worthwhile to explore the methodology employed in Doherty et al. (2012) when developing the evaluation framework for three years of lifelog data. This is the only example that we are aware of that takes a systematic approach to comparatively evaluating two interactive lifelog retrieval systems over a multi-year archive. Working closely with memory scientists, a three year archive of lifelog data (primarily focusing on SenseCam rather than location or bluetooth data) was chosen for the experimentation. The individual who gathered the data was asked to select fifty information needs (important memories) from the three years of lifelog data, without examining the data. This comprised the dataset, three years of data and fifty known-item queries. Following an interval of six months, enough time to forget which information needs had been chosen, the individual was presented with two retrieval systems, each of which executed 25 of the queries. Since this was a known-item search task, the mean-time elapsed until success and the number of successfully executed queries were the primary evaluation measures employed.

All of this is at odds with typical information retrieval research which is based on a history of theory followed by experimental evaluation going back to the Cranfield paradigm and then maturing into evaluation benchmarking campaigns. Lifelogging doesn't fit easily into the information retrieval model of comparative benchmarking as the standard mode of evaluation of information retrieval system performance. For lifelogging there is the challenge of moving from an individual query / test collection approach to search session support, to extended sessions based on user knowledge, experience and supporting long-term information access. So, while we can gather large amounts of personal sensed data and carry out closed evaluations of new lifelogging tools and techniques, this will not be in the spirit of information retrieval research, i.e. the repeatability of experiments. Until such time as a test collection is made available, we as a community are likely to keep focusing on component evaluations of lifelog systems as well as our own closed evaluations.

5

Lifelogging Applications

As we have seen at the beginning of this volume, lifelogging represents a phenomenon whereby people can digitally record their own daily lives in varying amount of detail, for a variety of purposes, thereby generating the ultimate "black box" of life activities. Lifelogging offers potential for mining or inferring knowledge about how we live our lives. Aside from the early adopters and extreme lifeloggers that we mentioned earlier, many of whom wish to gather their own detailed digital trace of their lives, a scaled-down version of lifelogging has also been ongoing thus far for some applications and scenarios.

We will now summarise the most important of these, dividing them into personal or individual applications where the logging is done solo, and applications aimed at populations or groups of people, for the purpose of some greater societal or organisational good. Many of these lifelogging applications come from the healthcare or quantified-self domains. We discuss the future relevance of lifelogging to information retrieval at the end of this chapter.

5.1 Personal lifelogging applications

We define personal lifelogging applications as those where a single individual uses lifelogging tools to record information about him/herself for his/her own benefit and while some of the lifelog data may ultimately be shared with others, this is not the primary purpose of the lifelogging activity. While some may choose to lifelog with a view to building up an archive of life activities, as one would maintain a diary, we expect that most people will be motivated to lifelog based on an expectation of some benefit. At present, there are few actual applications of lifelogging technologies; those that are being employed currently are now described.

5.1.1 Self-Monitoring of activities

Earlier, in sections 1.3 and 3.1 on "who lifelogs" and on "sources of lifelog data" respectively, we introduced some of the many devices now available as consumer products for logging caloric energy expenditure. These include the FitBit One$^{\mathrm{TM}}$, Nike Fuelband $^{\mathrm{TM}}$, and Lark$^{\mathrm{TM}}$as examples. The whole experience is usually wrapped into a gamification model, Deterding et al. (2011). In such a model, the user can choose daily or weekly targets to be reached, there are badges to be won as rewards, there is often integration with social networks and comparisons can be made against oneself and against others. Indeed, if we go back to the time when passive (or any form of digital) lifelogging was not possible we find there is a lot of prior work which studies self-monitoring, in those days based on active rather than passive logging, and normally using a written diary. Nelson and Hayes (1981) have neatly summarised the theory behind why some people self-monitor, and in particular have studied reasons why self-monitoring can be used to alter behaviour. Although this is an article from more than three decades ago, its findings are still valid in the world of digital lifelogging. From Michie et al. (2011), we know information obtained through self-observation provides feedback and helps to illustrate when behaviour deviates from a given standard of performance thus triggering the subject to take some actions, namely to change behaviour. Lifelogging, by its inherent passive nature, has the potential to make self-monitoring more accessible to many individuals.

Monitoring of dietary intake has been manually logged by people since well before digital lifelogging but the difficulty is in remembering all of the food and drink we consume during the day because of our tendencies to eat outside regular mealtimes. Automatically sensing and then logging what we eat is difficult to automate, as shown in Amft and Troster (2009), since there is no non-invasive wearable sensor which can reliably detect this with high recall at present. As a result, most diet-based lifelogging uses wearable cameras to record the wearer's day and from this to offer memory triggers for manual logging. This depends upon the wearer wanting to generate a truthful and accurate log of food intake and so there must be some enveloping incentive for the wearer such as a target weight loss or improvement on sedentary behaviour.

The DietSense project by Reddy et al. (2007a) at UCLA, makes use of a mobile phone with a camera embedded to capture pictures of the wearer's day automatically. The images collected act as the log of the wearer's mealtimes and are used in collaboration with the subject to analyse the diet intake in order to give feedback and to improve diet choices. More recent work on the same topic is reported in O'Loughlin et al. (2013) where the SenseCam was used, again as a memory trigger for manual logging. Gemming et al. (2013) adds further methodological rigour to the issue by comparing wearable camera assisted recall of diet to traditional self-report measures. Finally, Aizawa et al. (2013) utilised foodlogs to estimate food balance for personal dietary monitoring and in later work, Aizawa (2013) presents the multimedia food log as a means of making societal donations.

Smoking cessation is another healthcare target that has been the subject of lifelogging research by Stanley and Osgood (2011) and like eating or drinking, it is difficult to automatically sense smoking without some kind of invasive wearable sensor. Thus, as in diet monitoring, the use of lifelogging to assist with smoking cessation has been as a tool to record the subject's day and from the reminiscence through the lifelog, to offer triggers or rewards to help with the smoking cessation goal.

Another form of self monitoring is sleep monitoring, both the time spent asleep and the quality of sleep. There are many devices on the market that monitor sleep quality, based on wearable accelerometers, devices placed beside the bed or even phone apps that monitor the phone's accelerometer. While sleep monitoring, like accelerometer-based activity/exercise recognition, is not a visual lifelogging activity, we include it in this discussion because it has provided one of the first use-cases for passive monitoring of an individual's activities for some perceived knowledge gain. In all of these examples, a clear benefit is provided to the individual who is logging their activities and they make a choice to do so in order to gain extra awareness or knowledge of their life activities.

5.1.2 Memory assistance

One of the early targets for personal lifelogging, especially visual lifelogging from wearable cameras, was helping to overcome the difficulties some people have with short term memory recall, especially for people with Alzheimer's and other dementias. This exploited the well-known phenomenon of *Proustian Moments* as described in Stix (2011), where a trigger of some kind – a smell, sound, image, object, etc. – causes a spontaneous recollection of something from our past. The theory is that re-living a recent experience by examining images from a lifelog can induce a Proustian recall and there has been much work using the SenseCam especially, to lifelog those people with memory impairments, especially episodic memory impairment, e.g. Lee and Dey (2008), and then to re-play their day thus triggering short-term recall and re-opening some of their cognitive pathways, albeit briefly.

Studies at Addenbrook's hospital in Cambridge, U.K. and elsewhere show measurable effects of improved memory after replay of a given day's activities, though on a small sample size, as described in Berry et al. (2009); Pauly-Takacs et al. (2011); Silva et al. (2013). In Vemuri et al. (2004); Vemuri and Bender (2004) a memory re-finding use of lifelogging which is called "iRemember" is presented, where audio clips are used as the main information key to navigate memory and once again these exploit the phenomenon of Proustian recall.

Using lifelogs in memory rehabilitation and memory assistance, is an active area of research though progress has to be slow and measured, and based on statistically significant improvements across populations rather than anecdotal evidence of whatever good it may bring. We know from Brindley et al. (2011) that memory and memory rehabilitation is a complex process, and not very well understood, and while the technology is there to digitally record our lives, how to present them back to people with memory impairments in such a way that they do bring benefit is a subject of ongoing investigation. One good example of recent research following this methodologically rigorous route is described in Silva et al. (2013), which suggests that thirty healthy adults experienced enhancements to short-term cognitive functioning by reviewing wearable camera images.

5.1.3 Longer-Term assisted living

While wearable lifelogging has benefited from the availability of cheap and reliable sensor technology, similar sensors have also been used to lifelog directly from the environments in which we live. There have been many projects worldwide based on ambient assisted living which have instrumented homes using sensors like Passive Infra Red detectors, contact sensors on doors, fridges, microwaves, ovens, sensors monitoring energy and water usage, sensors on light switches, pressure sensors in furniture and beds, and so on. The incentive for this work is to allow older adults to have independent lives in their own homes both for their own well-being as well as for economic reasons.

One example project which uses longer-term lifelogging is at the Great Northern Haven in Dundalk, Ireland, where 16 apartments which support independent living of older adults, have been instrumented with a selection of over 100 such sensors per home. This is described in O'Brien et al. (2012). Occupants are monitored in a completely passive and ambient way to detect potential alarm situations such as falls, but more interestingly their well-being and behaviours can also be studied longitudinally and over time such as months or years we can detect changes in their movement, occupancy, eating habits, sleeping, social interaction with others, etc. This is a form of lifelogging in that it is passive, not initiated by the user, and although does not come from wearable sensors, it does generate a form of lifelogs for the individual which can be queried or otherwise processed by interested parties.

5.2 Population-based lifelogging applications

Population-based lifelogging applications are those where the lifelogs are processed and combined to allow us to infer something about the population of users as a group. Sometimes the populations originate from a single organisation, like a corporation, sometimes they are a sample of research subjects, sometimes they are self-selected such as a family or group of friends. We now give three diverse examples of such lifelogging, in a healthcare clinical practice, in qualitative market analysis and in family reminiscing.

One of the earliest uses of lifelogging across a population has been to capture and processes procedures in a corporate or other workplace environment. Past work has been reported in Byrne et al. (2008a); Kumpulainen et al. (2009), where healthcare workers in a clinical practice would typically log their work at the end of their shift but in this case they used visual lifelogs to trigger their own recall of their day. In particular this was used in an analyses to better understand the information needs of clinicians in hospitals in Finland. This information was subsequently used for process improvement in information access tools for those clinicians.

A second example of population-based lifelogging is in qualitative market research where market research firms are interested in qualitative research based on participants recording lifelog data as to characterise their lifestyles across a demography of participants and to use this (visual) data to determine their exposure to particular brand names and logos, e.g. Berry et al. (2010); Hughes et al. (2012). Such analyses are based on earlier established work in the memory domain, and in computer vision using SIFT/SURF computer vision techniques to search for behavioural real-world activities and also logos/brands, as described in Hughes et al. (2012). This work takes the Gordon Bell approach of capturing everything because we can, and thereafter finding some particular use-cases for it later. In particular, when a wearer is exposed to a brand name or logo, market research is very interested in what the wearer was doing, who s/he was with, etc., and this level of qualitative market research cannot be reliably captured on large-scale by typical manual surveys or questionnaires.

A third example of population-based lifelogging is where family memories can be created as the aggregation of individual family members' contributions from their own digital memories of some family event or celebration, e.g. Petrelli and Whittaker (2010). Caprani et al. (2010) worked with older people using the SenseCam and then sharing their experiences with others, Lindley has also done similar work in Lindley et al. (2011) and de Silva et al. (2007) developed an interactive multimedia diary for the home that was evaluated with real-families in a replicated home environment. For the most part, these family stories were created just for posterity, however Kikhia et al. (2010) present work on creating digital life stories from groups of individuals for use in memory support.

Lifelogging has societal applications in terms of better fidelity in measuring the behaviour of groups of individuals in a given population which helps inform policies for tasks like transport planning, environment understanding, and relationships between lifestyle exposures and disease outcomes. For example the relationships between lifestyle behaviours and health outcomes are usually based on self-reported data which is prone to measurement error. As we saw earlier, lifelogging sensors, such as wearable cameras and their associated software tools have developed to the point that they are well-suited to measure physical activity, sedentary behaviour, active travel, and nutrition-related behaviours across populations of users and this has been discussed in Doherty et al. (2013a). We have also seen population-level studies of social dynamics by using mobile Bluetooth proximity as a sensor in lifelogging and described in Eagle and Pentland (2006).

Some specific examples of wearable camera lifelogging tools used by public health researchers and others to help inform policy decisions include:

- Kelly et al. (2011, 2012), who used wearable cameras to identify self-reporting errors in travel behaviour in both adults and adolescents;

- Doherty et al. (2013c), who were able to identify sedentary, light, moderate, and vigorous intensity physical activities through a combination of accelerometers from which they can automatically identify episodes of interest and cross-correlate with images from wearable cameras to manually classify those episode types;
- Kerr et al. (2013), who employed an annotation framework to manually categorise sedentary behaviours in fine-grained detail to better identify factors that may be driving such behaviour;
- Reddy et al. (2007a), who were able to identify self-reporting errors in a behaviour study of the nutrition intake of igroups of ndividuals using a mobile phone with a camera embedded to capture pictures automatically;
- O'Loughlin et al. (2013); Gemming et al. (2013), both of whom found that they were able to help participants to identify forgotten calories through using wearable camera images as memory prompts;
- May and Warrington (2012), who employed lifelogging tools in the sports domain, where the activity levels, rest times, etc. for search and rescue helicopter crews were recorded in order to help determine training needs.

All these are examples of group or population-level rather than individual lifelogging applications, where the individual lifelogs are processed and information is extracted and aggregated across individuals and into groups into something useful for the underlying policy question.

5.3 Potential applications of lifelogging in information retrieval

Initial applications of lifelogging are firmly routed in the domain of healthcare or quantified-self analysis and in general they rely on a single or reduced set of sensors (e.g. accelerometer only, wearable camera only, etc.). The potential for lifelogging is far greater than these initial use-cases would lead the reader to believe. Bell and Gemmell (2009) in their book "Total Recall" provide a positive and optimistic bigger picture of the potential of lifelogging. The reader of Total Recall is introduced to the belief that lifelogging will revolutionise healthcare, productivity, learning and social society. Each of these four areas of potential are examined in detail in the book. Considering this viewpoint, coupled with the optimistic outlook of Ray Kurzweil concerning the exponential growth rate of information technologies, then we can begin to imagine a world in which the detailed digital tracing of everything we do gets captured into our lifelog, processed to extract semantics and meaning and used to support us in our daily life. As of yet, nobody has managed to create such a detailed lifelog, though there are efforts underway, as outlined earlier.

While the potential applications of lifelogging are not yet well understood, it is possible that, given appropriate capture and access/interaction hardware (e.g. Google Glass is one potential technology), that armies of app developers will come up with ingenious applications and tools that fulfil some, or all of, the lifelogging visions of people such as Bell and Gemmell. Inevitably, search and information retrieval are the fundamental enablers for many kinds of lifelogging applications.

We have described earlier how the five Rs of memory access provide initial cues for how these search tools could be developed. However, possibly the biggest immediate potential for lifelogging technologies within an information retrieval framework is in providing new sources of data for contextual models of retrieval.

Models of information retrieval are an essential underpinning to progressing the area and information retrieval models for have been around for decades going back to, Kuhlthau (1991), Järvelin and Wilson (2003), Bates (2002), and Ingwersen and Järvelin (2005). As recently as 2011, Järvelin produced a summary review article on the topic in Järvelin (2011). These models of the information seeking process in information retrieval have always acknowledged the capturing and leveraging context as being one of the grand challenges in information retrieval, Allan et al. (2003a); Callan et al. (2007); Allan et al. (2012). Even in the area of multimedia information retrieval, it is now recognised that in order to make a quantum leap that supports the high demands of users, we need to make a significant step in the area of understanding users, understanding what they are doing, and why. Hanjalic (2012) refers to this as requiring a *utility-centered* research focus which we can interpret as simply a need to know more about the people who are doing the searching.

The FnTIR review of Contextual IR byMelucci (2012), introduces contextual search within a computational framework based on contextual variables, contextual factors and statistical models. It describes how statistical models can process contextual variables to infer the contextual factors underlying the current search context. Then we take it a bit further by introducing the idea of using context in addition to content in applications like managing our own personal digital photos, which has been strongly advocated in Sinha and Jain (2008), where they coined the term *"contenxt"*. In that work though, they do not make much progress on defining how to capture contenxt when it applies to users who are searching for photos and instead they focus on capturing the contenxt in which the photos have been taken/captured, such aspects as location and time.

Since the aim of lifelogging is to gather a detailed digital trace, then it follows that a real-time lifelog could be the ultimate form of context to help interpret the user, the user's environment and the information needs of the user. There is potential not only to know what the user is doing at any point in time, but also what the user is looking at, hearing, experiencing; add to that the historical context (the current state of the user is stored at every moment prior to that) and the model of the user can become the ultimate form of context. The information retrieval system will know the user in extreme detail and be able to tailor information to the user, either in response to an information need (most probably spoken) or through some form of real-time context-aware recommendation engine. It is believed that we are moving into a world where search is ubiquitous, omnipresent and intelligent, then lifelog sourced contenxt will be an enabling technology. Our information retrieval technologies will operate with ubiquitous/pervasive computing concepts to generate a new retrieval environment in which the user's everyday activities becomes an extreme form of user context. This will provide new challenges for information retrieval, ubiquitous computing and user experience modelling. In many ways, we are at the point in time when advances in sensing, storage and search all lead to a potential paradigm shift in how we access information and knowledge, with lifelogging technologies providing a platform upon which these information access advances are built.

6

Conclusions and Issues

We have observed a convergence of technologies to foster the emergence of lifelogging as a mainstream activity. In order for this potential to be realised, there are a number of issues to be addressed. In this final chapter, we will explore some of these issues, which are based on a combination of the scientific literature and personal experience of real-world lifelogging.

6.1 Issues with lifelogging

Lifelogging has the potential to both excite as well as instill fear and concern, as with any new technology. As far back as 1888, when Kodak produced the world's first portable camera, the advent of this camera technology was so sudden and pervasive that Kodaks were banned in some public places in the US and some users of cameras were referred to as "camera fiends" who were supposedly seen near beach resorts awaiting the arrival of unsuspecting female bathers. The recent announcement of Google Glass has caused a similar reaction with the device being banned from certain bars, casinos, drivers being fined for wearing Glass while driving with even a ban being considered for drivers on UK roads. Add to that the new terminology of “glasshole” to describe the type of person that constantly interacts with their Glass, ignoring the outside world, and we can see that both the early cameras and the latest wearable computing devices have received similar societal responses. Consider also the reaction of people to first generation pagers, mobile phones, SMS-enabled phones, the Sony Walkman, tablet PCs and we can see that many new technologies were initially seen as elitist, not part of mainstream society and inherently disruptive; that is until they became more affordable and mass-adoption took place as the technology moved into the early/late majority adoption. Therefore, when discussing the potential of lifelogging, it important to consider the issues that raise concerns, both technically and societally at this point in time, as lifelogging is currently at the innovators phase in the technology adoption lifecycle model, as described in Rogers (2003).

There is a data capture challenge in supporting individuals to capture detailed digital traces of their life activity into the lifelog. As more early adaptors begin to gather lifelogs, the available technologies will also increase and the human effort required to gather lifelogs will be reduced. It is likely that, over time, additional sources of lifelog data will come on-stream so our semantic understanding tools and information retrieval techniques need to be flexible to increasing volumes and types of lifelog data.

Societal acceptance of lifelogging is also a major issue. There has been a noticeable polar positive/negative response to the impending release of Google Glass, which, although not a lifelogging technology in its own right, has the potential to become a powerful lifelogging tool as app developers (e.g. Saga, Moves) port their lifelog technology to glassware. Judging by past history, it is likely that lifelogging as a new technology will become more acceptable as it becomes more mainstream and as the positive benefits of the technology become better understood. This, however, is dependent on the ready availability of sufficient novel applications that put lifelogging’s positive benefits at the core of the societal debate, as opposed to the privacy debate at present.

There are many privacy concerns to be addressed as lifelogging becomes more commonplace; Kelly et al. (2013); Price (2010); Jacquemard et al. (2013) introduce these issues. For basic forms of lifelogging, such as quantified-self analysis using small personal sensors such as the FitBit, then there is no privacy concern as the monitoring occurs inwards (i.e. towards the individual); however when using a broader set of sensors (camera, microphone, etc.) then the sensing looks both inwards and outwards with inevitable consequences on privacy beyond the lifelogger. This means that we can identify three actors in lifelogging; the lifelogger, the bystanders encountered by the lifelogger and society as a whole. Wearing a SenseCam or any visual capture device in a public place will inevitably mean capturing the images of bystanders who have not given permission to capture their images; a similar issue occurs with microphone recording or video recording. This means that the very act of passing by a lifelogger will result in the capture, storage and processing of a bystander's image. Considering society as an actor in lifelogging, society as a whole needs to understand the implications of lifelogging, both positive and negative.

As lifelogging becomes more mainstream, O'Hara et al. (2009) points out that privacy concerns of society may be offset by the empowerment of the individual as new lifelogging applications come onstream, or we may develop lifelogging systems that integrate "privacy by design" into the development process. Privacy by design is a proposed framework for ubiquitous computing, first introduced by in Langheinrich (2001), in which privacy and data protection are embedded as core considerations throughout the entire life cycle of a technology, from the early design stage to their deployment, use and eventual disposal. Privacy by design principles are based on seven foundations from Cavoukian (2010); proactive not reactive, privacy as the default configuration, privacy embedded into the design, privacy as additional (not reduced) functionality, end-to-end data security, visibility/transparency and respect for the privacy of the individual user. These seven principles provide an initial set of guidelines for developing privacy-aware lifelogging systems. While privacy by design is obviously a positive concept, it has received criticism for bring vague and lacking detail in how to actually implement such a concept while meeting the functional requirements of the system under development; for details see Gürses et al. (2011). With regard to lifelogging, there is an inevitable tradeoff between privacy and functionality, and where lifelogging settles on this tradeoff is yet to be seen. In any case, the concept of privacy by design is likely to become a core component of any large-scale lifelogging solution and it is something that information retrieval developers need to take into account when developing lifelogging organisation and retrieval tools.

One of the key issues that we have not touched yet is the ethics of lifelogging. There has been research into the ethical issues surrounding lifeloging as in Kelly et al. (2013); Jacquemard et al. (2013). For example, if my lifelog contains data of someone committing a crime (even a minor infringement such as somebody parked on a double-yellow line) then am I obliged to report it? It will be important to critically reflect on, and analyse the ongoing technological and scientific developments and the Ethical, Legal, and Social Issues (ELSI) that might arise from lifelogging. Initial thoughts in Kelly et al. (2013) propose a framework that is motivated by principles such as: respect for autonomy; beneficence; non-maleficence; and justice and is related to, though not linked to, the concept of privacy by design.

In lifelogging, there are also data ownership and access challenges, in terms of where to store the data and who owns it, who can access it and how long to keep it? The private nature of lifelog data means that for some users, the idea of hosting it in a cloud-based service would not be acceptable. For others, the convenience of hosting the vast archives on a cloud infrastructure where the service provider takes responsibility for data hosting, backup, security and providing the retrieval facilities would be far more important than any perceived loss of control of the data in the cloud. In reality, for most lifeloggers, the challenges of self-hosting the data would be too great without some secure personal semantic memory being made available.

Naturally, the storage location of the data has implications on how we develop lifelogging solutions; for example, if the data gathering device extracts basic semantics and metadata from the captured data streams and uploads only this data to a organisation/retrieval server, leaving the actual data on a local machine, then it is potentially impossible to re-analyse the lifelog data to extract new semantics once the data has been initially uploaded. In addition, if the lifelog is hosted across different locations, there could be the challenge of supporting federated search or cross-archive federated retrieval. Other issues such as supporting retrieval across multiple (e.g. shared within a family) lifelogs would also pose significant distributed retrieval challenges. There are also issues such as long term storage of the data and the suggestion that data degradation/forgetting should be supported.

Other issues such as ownership of the lifelog within a lifelogger's lifetime are more straightforward. It is our assumption that the lifelog is owned by the data gatherer. Questions remain on what to do with lifelog data when someone dies. Should the data be deleted, or passed onto a trusted (or named) relative? If passed on, the question of how long to keep the data arises; should the data be kept forever or for one (or more) generations? Digital storage advances means that the data should cost little to maintain indefinitely. Potentially, a lifelog could be retained forever, once digital library concerns of long-term preservation and format updating have been considered. Lifelogs of past individuals could provide a valuable historical context, however it must be remembered that they could contain a complete digital trace of life, the most detailed digital footprint, which raises privacy concerns for the lifelogger, even after the lifelogging has ceased. Massimi has led research into some of these 'after-lifelogging' issues, see Massimi and Baecker (2010) and this has also been explored in various movies and broadcast media.

Sharing of lifelogs also needs some consideration. If a lifelog (or part of) is shared with family members, then how does this impact on the retrieval algorithms employed and the presentation of query results? If sharing of a lifelog is allowed according to certain access policies, then these would need to be caluclated in real-time (as policies change) and this would impact on the indexes available for retrieval.

Lifelogging must also deal with the issue of allowing individuals to forget or suppress past bad experiences. Mayer-Schonberger (2011) tells us that forgetting is beneficial to humans and has a natural function.

Does lifelogging negate the ability to forget, or does lifelogging simply provide the facility to retrieve if desired? The concept of forgetting has attracted comment in lifelogging circles and by some is seen as the antithesis of forgetting, as explained by O'Hara et al. (2009) while other efforts have been aimed at modelling the human experience of forgetting in surrogate memories, as explored in Gurrin et al. (2010b).

We also identify a lack of systematic lifelogging research to date. Throughout this volume, we have alluded to the fact that lifelogging is a technology in its infancy. Factors such as the lack-of-maturity of the data gathering technologies, the resultant overhead of gathering a test collection and the challenges of getting enough users has resulted in lifelogging not receiving much research attention to date, especially from the information retrieval community. We hope, going forward, that this will change and that information retrieval technologies developed to provide retrieval on semantic memory can be refined and made applicable to the episodic memory of lifelogging. We note that there is no guarantee of how effective our retrieval technologies for semantic memory will work on episodic lifelogs, or more specifically, do the traditional retrieval techniques developed over decades for searching our equivalent of semantic memory (facts, documents, web pages, etc) have applicability in some adjusted way to allow search through a lifelog which is a collection of episodic memories? In Chapter 4, we discussed the importance of events in lifelogs, a direct application of episodic principles; there is no equivalent in conventional information retrieval, so there is a body of research to be done to modify the information retrieval models to operate effectively on episodic lifelog data. There are a diverse set of proposed solutions to interact with lifelog data. However no compelling experiments have been carried out to evaluate the effectiveness of these approaches in comparison to each other. The best solution may represent a combination of some of the aforementioned approaches, but this requires significant levels amounts of research.

Finally, to end this discussion, we turn to schools of lifelogging. Here we can identify two major schools of lifelogging. Gordon Bell's school is that we record everything just because we can, as described in Total Recall, Bell and Gemmell (2009), the assumption being that this data might very well be useful later. The counter-view is proposed in Sellen and Whittaker (2010), which is that one should record only what we want and when we want it; this is more like Quantified-Self analysis. Both have their pros and cons, but our understanding at this point is that the Sellen & Whittaker viewpoint doesn't capture the case of something growing in importance or significance later as you can't determine the importance at the time of capture. This has additional drawbacks in that data never captured in the first place can never be re-processed to extract semantics or meaning after the fact as new semantic analysis tools or use-cases emerge. In reality, on a practical level, you only get one shot at sampling life experience data, so we consider it best to gather the data with the future in mind, and not the present use-cases.

6.2 Future directions

Lifelogging is a technology in its infancy. There are numerous challenges mentioned above, each of which demands attention. Nothwithstanding these, and with the caveat that we are only beginning to explore the concept of lifelogging, we suggest some potential future directions for lifelogging to follow:

- Enhanced Capture. Currently our sensing technologies are still relatively rudimentary. We can sense aspects of a person's environment, their actions/interactions, a lot of what they see and hear, and to some extent their interests (via information interactions). However, this only provides a snapshot of life activities for the lifelog, and it remains for the surrogate memory to generate semantic interpretations of this data and make it useful for the user. More detailed human sensing (via next generation wearable sensors, Brain Computer Interfaces, etc.) would allow for much more detailed capture of the semantics of life, where concepts such as sentiment/mood/emotion could conceivably be captured into the lifelog. While this may seem far-fetched, continual advances in Brain Computer Interface technology slowly edge us closer to this point. The end point of this effort would be the modeling of both episodic and semantic memory within lifelogs, whereas currently lifelogging focuses on episodic memory logging;

- Integration of lifelogging context information with search engines. Lifelogging, as discussed earlier, provides a rich and novel source of data for contextual information retrieval. It is likely that new approaches to contextual understanding for omnipresent and intelligent search will emerge from both real-time lifelog streams as well as historical context from lifelog archives. Utilising this data will provide both new challenges and new opportunities to the Information Retrieval researcher;

- Use-cases. As lifelogging becomes a more popular activity, the use cases will become more clear. This will inform development of capture technologies but most importantly, from the retrieval point of view, it will inform the information needs for real-world users, as opposed to relying on the early adopters or the five Rs of memory access. The book "Total Recall" provides a good initial overview of the potential use cases;

- Anonymisation of Lifelogs. With privacy concerns in mind, the issue of anonymising unknown passers-by in lifelogs will surely receive research attention. It is our belief that this should be a access-time, dynamic process, based on user access policies, as opposed to a non-reversible capture-time process.

- Recreating the Person. A more far-fetched concept, but one that is receiving attention and funding at present is the idea that the lifelog can be used to recreate the individual in digital form (an avatar) by using the detailed trace from the lifelog as source data. Some futurists (e.g. Ray Kurzweil) have gone so far as to suggest that the human is an information processing machine, the memory of which can be replicated / enhanced indefinitely inside of an information processing machine that relies on lifelog data as the source of memory data;

- Humanising Technology. A final and perhaps more down-to-earth concept is that lifelog principles can be applied to inanimate items to provide a semblance of episodic memory or personality; an early prototype is described in Yang et al. (2012) which uses lifelogging technology to put a human-like interface to a standard coffee machine. This has most potential when combined with research into the development of humanoid robotics.

6.3 Conclusion

It is difficult to predict whether lifelogging will become mainstream or not. The first generation of use cases are only beginning to be understood and the potential is not yet clear. We may find that wearable sensing that impacts on bystanders and society as a whole (specifically cameras) is a step too far for society, or the benefits of extreme lifelogging may not live up to the potential. On the other hand, the new lifelogging technologies could be embraced as society has embraced smartphones and social networking in recent years. We have yet to see.

A lot is dependent on the first generation use cases, which require the input of information retrieval to develop effective lifelogging applications of surrogate memories. The effectiveness of the underlying information organisation tools will be key in whether lifelogging takes off as a commonly employed technology. As the field of information retrieval inevitably helped to progress web search, a similar requirement would be made on information retrieval to help progress lifelogging from a topic for early adopters to become a widely used and beneficial technology. It may appear at present that the future uptake of extreme lifelogging would be very small; however we can already see the uptake of the initial quantified-self and health-based applications and if this trend continues to increase, then lifelogging may be following a few years behind. If that is the case, lifelogging will surely keep the Information Retrieval scientist busy for the foreseeable future.

Acknowledgements

Authors AD and AS were supported by Science Foundation Ireland under Grant Numbers 07/CE/I1147 and SFI/12/RC/2289. Author AD was also a Marie Curie postdoctoral research fellow supported by the Irish Health Research Board under grant number MCPD/2010/12, and is now supported by the Li Ka Shing Foundation. Author CG was supported by Science Foundation Ireland under grant Nos. 07/SK/I1186 and 11/RFP.1/CMS/3283 as well as by the Norwegian Research Council under project number 174867. The authors wish to thank the large number of postdoctoral researchers, PhD students and research assistants who have worked with us over the last decade on various lifelogging projects.

References

Ahmad Jalal, Jeong Tai Kim, T.-S. K. (2012). Development of a life logging system via depth imaging-based human activity recognition for smart homes. In *SHB2012 - 8th International Symposium on Sustainable Healthy Buildings*.

Aizawa, K. (2005). Digitizing Personal Experiences: Capture and Retrieval of Life Log. In *Proceedings of the 11th International Multimedia Modelling Conference, 2005. MMM 2005.*, pages 10–15.

Aizawa, K. (2013). Multimedia foodlog: Diverse applications from self-monitoring to social contributions. *ITE Transactions on Media Technology and Applications*, 1(3):214–219.

Aizawa, K., Hori, T., Kawasaki, S., and Ishikawa, T. (2004a). Capture and efficient retrieval of life log. In *Proceedings of the Pervasive Workshop on Memory and Sharing of Experiences*, pages 15–20, Linz/Vienna, Austria.

Aizawa, K., Ishijima, K.-I., and Shiina, M. (2001). Summarizing Wearable Video. In *International Conference on Image Processing (ICIP)*, pages 398–401.

Aizawa, K., Maruyama, Y., Li, H., and Morikawa, C. (2013). Food balance estimation by using personal dietary tendencies in a multimedia food log. *Multimedia, IEEE Transactions on*, 15(8):2176–2185.

Aizawa, K., Tancharoen, D., Kawasaki, S., and Yamasaki, T. (2004b). Efficient retrieval of life logs based on context and content. In *CARPE'04: Proceedings of the the 1st ACM workshop on Continuous Archival and Retrieval of Personal Experiences*, pages 22–31, New York, NY, USA. ACM.

Al Masum Shaikh, M., Molla, M., and Hirose, K. (2008). Automatic life-logging: A novel approach to sense real-world activities by environmental sound cues and common sense. In *Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on*, pages 294–299.

Allan, J., Aslam, J., Belkin, N., Buckley, C., Callan, J., Croft, W., Dumais, S., Fuhr, N., Harman, D., Harper, D., et al. (2003a). Challenges in information retrieval and language modeling. *SIGIR Forum*, 37(1):31–47.

Allan, J., Croft, B., Moffat, A., and Sanderson, M. (2012). Frontiers, challenges, and opportunities for information retrieval: Report from SWIRL 2012 the second Strategic Workshop on Information Retrieval in Lorne. *SIGIR Forum*, 46(1):2–32.

Allan, J., Wade, C., and Bolivar, A. (2003b). Retrieval and novelty detection at the sentence level. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and Development in Informaion Retrieval (SIGIR)*, SIGIR '03, pages 314–321, New York, NY, USA. ACM.

Amft, O. and Troster, G. (2009). On-body sensing solutions for automatic dietary monitoring. *Pervasive Computing, IEEE*, 8(2):62–70.

Anguera, X., Xu, J., and Oliver, N. (2008). Multimodal photo annotation and retrieval on a mobile phone. In *Proceedings of the 1st ACM international conference on Multimedia information retrieval*, MIR '08, pages 188–194, New York, NY, USA. ACM.

Atrey, P. K., Maddage, N. C., and Kankanhalli, M. S. (2006). Audio based event detection for multimedia surveillance. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–V. IEEE.

Baeza-Yates, R. and Maarek, Y. (2012). Usage data in web search: benefits and limitations. In *Proceedings of the 24th international conference on Scientific and Statistical Database Management*, SSDBM'12, pages 495–506, Berlin, Heidelberg. Springer-Verlag.

Bailey, P., Craswell, N., and Hawking, D. (2003). Engineering a multi-purpose test collection for web retrieval experiments. *Inf. Process. Manage.*, 39(6):853–871.

Barua, D., Kay, J., and Paris, C. (2013). Viewing and controlling personal sensor data: What do users want ? In *The 8th International Conference on Persuasive Technology*, pages 15–26, Sydney, Australia. Springer Berlin Heidelberg.

Bates, M. (2002). Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research*, 3:1–15.

Bay, H., Tuytelaars, T., and Van Gool, L. (2006). SURF: Speeded Up Robust Features. In *Proceedings of the 9th European Conference on Computer Vision (ECCV'06)*, pages 404–407.

Bell, G. and Gemmell, J. (2007). A digital life. *Scientific American*, 296:58–65.

Bell, G. and Gemmell, J. (2009). *Total Recall: How the E-Memory Revolution Will Change Everything.* Penguin Books.

Berry, E., Byrne, D., Doherty, A. R., Gurrin, C., and Smeaton, A. F. (2010). *Proceedings of the Second Annual SenseCam Symposium (SenseCam 2010). 16-17 September 2010.* Dublin City University.

Berry, E., Hampshire, A., Rowe, J., Hodges, S., Kapur, N., Watson, P., Smyth, G. B. G., Wood, K., and Owen, A. M. (2009). The neural basis of effective memory therapy in a patient with limbic encephalitis. *Neurology, Neurosurgery, and Psychiatry with Practical Neurology*, 80(3):582–601.

Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., and Wood, K. (2007). The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis. *Neuropsychological Rehabilitation*, 17(4):582–601.

Blighe, M., Doherty, A., Smeaton, A., and O'Connor, N. (2008). Keyframe detection in visual lifelogs. In *1st International Conference on Pervasive Technologies Related to Assistive Environments (PETRA)*, Athens, Greece.

Blighe, M., le Borgne, H., O'Connor, N. E., Smeaton, A. F., and Jones., G. J. (2006). Exploiting Context Information to aid Landmark Detection in SenseCam Images. In *ECHISE 2006 - 2nd International Workshop on Exploiting Context Histories in Smart Environments - Infrastructures and Design, 8th International Conference of Ubiquitous Computing (Ubicomp 2006).*

Blum, M., Pentland, A., and Troster, G. (2006). Insense: Interest-based life logging. *IEEE Multimedia*, 13(4):40–48.

Bolettieri, P., Esuli, A., Falchi, F., Lucchese, C., Perego, R., Piccioli, T., and Rabitti, F. (2009). CoPhIR: a Test Collection for Content-Based Image Retrieval. *arXiv preprint arXiv:0905.4627*, abs/0905.4627.

Brewer, W. F. (1988). *Practical Aspects of Memory: Current Research and Issues*, chapter Qualitative analysis of the recalls of randomly samples autobiographical events, pages 263–268. Wiley.

Brindley, R., Bateman, A., and Gracey, F. (2011). Exploration of use of SenseCam to support autobiographical memory retrieval within a cognitive-behavioural therapeutic intervention following acquired brain injury. *Memory*, 19(7):745–757.

Bristow, H. W., Baber, C., Cross, J., Knight, J. F., and Woolley, S. I. (2004). Defining and evaluating context for wearable computing. *International Journal of Human Computer Studies*, 60(5-6):798–819.

Bukhin, M. and DelGaudio, M. (2006). WayMarkr: Acquiring perspective through continuous documentation. In *MUM '06: Proceedings of the 5th international conference on mobile and ubiquitous multimedia*, page 9, New York, NY, USA. ACM.

Burke, J., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., and Srivastava, M. B. (2006). Participatory Sensing. Boulder, Colorado, USA.

Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176(1):101–108.

Byrne, D., Doherty, A. R., Smeaton, A. F., Jones, G. J., Kumpulainen, S., and Järvelin, K. (2008a). The SenseCam as a Tool for Task Observation. In *HCI 2008 - 22nd BCS HCI Group Conference*.

Byrne, D. and Jones, G. J. (2009). Exploring narrative presentation for large multimodal lifelog collections through card sorting. In *Proceedings of the 2nd Joint International Conference on Interactive Digital Storytelling: Interactive Storytelling*, ICIDS '09, pages 92–97, Berlin, Heidelberg. Springer-Verlag.

Byrne, D. and Jones, G. J. F. (2008). Towards computational autobiographical narratives through human digital memories. In *Proceeding of the 2nd ACM international workshop on Story representation, mechanism and context*, pages 9–12. ACM.

Byrne, D., Kelliher, A., and Jones, G. J. (2011). Life editing: third-party perspectives on lifelog content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1501–1510, New York, NY, USA. ACM.

Byrne, D., Lavelle, B., Jones, G. J., and Smeaton, A. F. (2007). Visualising bluetooth interactions: Combining the arc diagram and docuburst techniques. In *HCI 2007 - Proceedings of the 21st BCS HCI Group Conference*.

Byrne, D., Lee, H., Jones, G. J. F., and Smeaton, A. F. (2008b). Guidelines for the presentation and visualisation of lifelog content. In *iHCI 2008 - Irish HCI 2008*.

Byrne, R. and Diamond, D. (2006). Chemo/bio-sensor networks. *Nature Materials*, 5(6):421–424.

Callan, J., Allan, J., Clarke, C., Dumais, S., Evans, D., Sanderson, M., and Zhai, C. (2007). Meeting of the MINDS: an information retrieval research agenda. *ACM SIGIR Forum*, 41(2):25–34.

Caprani, N., Doherty, A. R., Lee, H., Smeaton, A. F., O'Connor, N. E., and Gurrin, C. (2010). Designing a touch-screen SenseCam browser to support an aging population. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 4291–4296, New York, NY, USA. ACM.

Cavoukian, A. (2010). Privacy by design: The 7 foundational principles. implementation and mapping of fair information practices. *Information and Privacy Commissioner of Ontario, Canada*.

Chatfield, K. and Zisserman, A. (2012). VISOR: Towards On-the-Fly Large-Scale Object Category Retrieval. In *Asian Conference on Computer Vision*, Lecture Notes in Computer Science. Springer.

Cheng, W. C., Golubchik, L., and Kay, D. G. (2004). Total recall: are privacy changes inevitable? In *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, CARPE'04, pages 86–92, New York, NY, USA. ACM.

Cohen, G. and Conway, M. A. (2008). *Memory in the real world*. Routledge.

Conway, M. A. and Loveday, C. (2011). SenseCam: The Future of Everyday Memory Research. *Memory*, 19(7):685–807.

d'Aquin, M., Elahi, S., and Motta, E. (2010). Personal monitoring of web information exchange: Towards web lifelogging. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*.

De Jager, D., Wood, A. L., Merrett, G. V., Al-Hashimi, B. M., O'Hara, K., Shadbolt, N. R., and Hall, W. (2011). A low-power, distributed, pervasive healthcare system for supporting memory. In *Proceedings of the First ACM MobiHoc Workshop on Pervasive Wireless Healthcare*, page 5. ACM.

de Silva, G. C., Yamasaki, T., and Aizawa, K. (2007). An interactive multimedia diary for the home. *Computer*, 40(5):52–59.

Dean, T., Ruzon, M., Segal, M., Shlens, J., Vijayanarasimhan, S., and Yagnik, J. (2013). Fast, Accurate Detection of 100,000 Object Classes on a Single Machine. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)*, Washington, DC, USA.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. (2011). Gamification. Using game-design elements in non-gaming contexts. In *PART 2: Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, pages 2425–2428. ACM.

Dey, A. K. and Abowd, G. D. (1999). Towards a better understanding of context and context-awareness. In *HUC'99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 304–307, London, UK. Springer-Verlag.

Dickie, C., Vertegaal, R., Fono, D., Sohn, C., Chen, D., Cheng, D., Shell, J. S., and Aoudeh, O. (2004). Augmenting and sharing memory with eyeblog. In *CARPE'04. Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 105–109, New York, NY, USA. ACM.

Dodge, M. and Kitchin, R. (2007). "Outlines of a world coming into existence": Pervasive computing and the ethics of forgetting. *Environment and Planning B*, 34(3):431–445.

Doherty, A. and Smeaton, A. (2008a). Automatically segmenting lifelog data into events. In *Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 20–23. IEEE.

Doherty, A. R., Byrne, D., Smeaton, A. F., Jones, G. J., and Hughes, M. (2008). Investigating keyframe selection methods in the novel domain of passively captured visual lifelogs. In *CIVR '08: Proceedings of the 2008 international conference on Content-based image and video retrieval*, pages 259–268, Niagara Falls, Canada. ACM.

Doherty, A. R., Caprani, N., Ó'Conaire, C., Kalnikaite, V., Gurrin, C., Smeaton, A. F., and Connor, N. E. O. (2011a). Passively recognising human activities through lifelogging. *Computers in Human Behavior*, 27(5):1948–1958.

Doherty, A. R., Gurrin, C., and Smeaton, A. F. (2009). An investigation into event decay from large personal media archives. In *Proceedings of the 1st ACM international workshop on Events in multimedia*, EiMM '09, pages 49–56, New York, NY, USA. ACM.

Doherty, A. R., Hodges, S. E., King, A. C., Smeaton, A. F., Berry, E., Moulin, C. J., Lindley, S., Kelly, P., and Foster, C. (2013a). Wearable cameras in health: The state of the art and future possibilities. *Am J Prev Med*, 44(3):320–323.

Doherty, A. R., Kelly, P., and Foster, C. (2013b). Wearable Cameras: Identifying Healthy Transportation Choices. *IEEE Pervasive Computing*, 12(1):44–47.

Doherty, A. R., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., Hamilton, A., and Foster, C. (2013c). Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. *International Journal Behavioural Nutrition Physical Activity*, 10:22.

Doherty, A. R., Moulin, C. J., and Smeaton, A. F. (2011b). Automatically assisting human memory: A SenseCam browser. *Memory*, 7(19):785–795.

Doherty, A. R., Pauly-Takacs, K., Caprani, N., Gurrin, C., Moulin, C. J. A., O'Connor, N. E., and Smeaton, A. F. (2012). Experiences of Aiding Autobiographical Memory Using the SenseCam. *Human-Computer Interaction*, 27(1-2):151–174.

Doherty, A. R. and Smeaton, A. F. (2008b). Automatically segmenting lifelog data into events. In *WIAMIS: 9th International Workshop on Image Analysis for Multimedia Interactive Services*, pages 20–23, Washington, DC, USA. IEEE Computer Society.

Doherty, A. R. and Smeaton, A. F. (2008c). Combining face detection and novelty to identify important events in a visual lifelog. In *CIT: 8th International Conference on Computer and Information Technology, Workshop on Image- and Video-based Pattern Analysis and Applications*, Sydney, Australia. IEEE Computer Society.

Doherty, A. R. and Smeaton, A. F. (2010). Automatically augmenting lifelog events using pervasively generated content from millions of people. *Sensors*, 10(3):1423–1446.

Doherty, A. R., Smeaton, A. F., Lee, K., and Ellis, D. P. (2007). Multimodal segmentation of lifelog data. In *RIAO 2007 - Large-Scale Semantic Access to Content (Text, Image, Video and Sound)*, Pittsburg, PA, USA.

Dong, W., Lepri, B., and Pentland, A. S. (2011). Modeling the co-evolution of behaviors and social relationships using mobile phone data. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia*, MUM '11, pages 134–143, Beijing, China. ACM.

Doulamis, A., van Gool, L., Nixon, M., Varvarigou, T., and Doulamis, N. (2008). Area '08: Proceedings of the 1st acm workshop on analysis and retrieval of events/actions and workflows in video streams. In *Proceedings of the 16th ACM International Conference on Multimedia*, MM '08, pages 1147–1148, New York, NY, USA. ACM. 433087.

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., and Robbins, D. C. (2003). Stuff I've seen: a system for personal information retrieval and re-use. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 72–79, ACM. New York, NY, USA.

Dunton, G. F., Liao, Y., Intille, S. S., Spruijt-Metz, D., and Pentz, M. (2011). Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. *Obesity (Silver Spring)*, 19(6):1205–1212.

Eagle, N. and Pentland, A. (2005). Social serendipity: Mobilizing social software. *Pervasive Computing, IEEE*, 4(2):28–34.

Eagle, N. and Pentland, A. S. (2006). Reality mining: Sensing complex social systems. *Personal Ubiquitous Comput.*, 10(4):255–268.

Economist, T. (2010). The data deluge. *Special Supplement*.

Ellis, D. P. and Lee, K. (2004). Minimal-impact audio-based personal archives. In *CARPE'04: Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 39–47, New York, NY, USA. ACM.

Ellis, D. P. and Lee, K. (2006). Accessing minimal-impact personal audio archives. *IEEE Multimedia*, 13(4):30–38.

Elsweiler, D. and Ruthven, I. (2007). Towards task-based personal information management evaluations. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '07, pages 23–30, Amsterdam, The Netherlands. ACM.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The PASCAL Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2):303–338.

Fallah, N., Apostolopoulos, I., Bekris, K., and Folmer, E. (2013). Indoor human navigation systems: A survey. *Interacting with Computers*, 25(1):21–33.

Fleck, R. (2005). Exploring SenseCam to Inform the Design of Image Capture and Replay Devices for Supporting Reflection. *Cognitive Science Research Paper - University of Sussex CSRP*, 576:22.

Fleck, R. and Fitzpatrick, G. (2006). Supporting collaborative reflection with passive image capture. In *Supplementary proceedings of COOP'06*, pages 41–48.

Fox, E. and Shaw, J. (1993). Combination of multiple searches. In *TREC 2: Text REtrieval Conference*, pages 243–252, Gaithersberg, Md., USA.

Fuller, R. B., Hays, K. M., and Miller, D. A. (2008). *Buckminster Fuller: starting with the Universe*. Yale University Press.

Ganeriwal, S., Balzano, L. K., and Srivastava, M. B. (2008). Reputation-based framework for high integrity sensor networks. *ACM Transactions on Sensor Networks*, 4(3):1–37.

Gemmell, J., Bell, G., and Lueder, R. (2006). MyLifeBits: a personal database for everything. *Communications of the ACM*, 49(1):88–95.

Gemmell, J., Bell, G., Lueder, R., Drucker, S., and Wong, C. (2002). MyLifeBits: fulfilling the Memex vision. In *Proceedings of the tenth ACM international conference on Multimedia*, pages 235–238. ACM.

Gemmell, J., Williams, L., Wood, K., Lueder, R., and Bell, G. (2004). Passive capture and ensuing issues for a personal lifetime store. In *CARPE'04: Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 48–55, New York, NY, USA. ACM.

Gemming, L., Doherty, A. R., Kelly, P., Utter, J., and Ni Mhurchu, C. (2013). Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *European journal of clinical nutrition*, 67(10):1095–1099.

Girgensohn, A. and Boreczky, J. (2000). Time-constrained keyframe selection technique. *Multimedia Tools Appl.*, 11(3):347–358.

Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T. (2002). Time as essence for photo browsing through personal digital libraries. In *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 326–335, New York, NY, USA. ACM.

Gurrin, C., Aarflot, T., Sav, S., and Johansen, D. (2009). Addressing the challenge of managing large-scale digital multimedia libraries. *JDIM*, 7(5):262–270.

Gurrin, C., Byrne, D., Jones, G. J. F., and Smeaton, A. F. (2008a). Architecture and challenges of maintaining a large-scale, context-aware human digital memory. In *Visual Information Engineering, 2008. VIE 2008. 5th International Conference on*, pages 158–163.

Gurrin, C., Lee, H., Caprani, N., Zhang, Z., O'Connor, N., and Carthy, D. (2010a). Browsing large personal multimedia archives in a lean-back environment. In *Advances in Multimedia Modeling*, volume Boll, Susanne and Tian, Qi and Zhang, Lei and Zhang, Zili and Chen, Yi-PingPhoebe of *Lecture Notes in Computer Science*, pages 98–109. Springer Berlin Heidelberg.

Gurrin, C., Lee, H., and Hayes, J. (2010b). iforgot: a model of forgetting in robotic memories. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, HRI '10, pages 93–94, Piscataway, NJ, USA. IEEE Press.

Gurrin, C., Qiu, Z., Hughes, M., Caprani, N., Doherty, A. R., Hodges, S. E., and Smeaton, A. F. (2013). The SmartPhone as a Platform for Wearable Cameras in Preventative Medicine. *Am J Prev Med*, 44(3):308–313.

Gurrin, C. and Smeaton, A. (2004). Replicating web structure in small-scale test collections. *Information Retrieval*, 7(3-4):239–263.

Gurrin, C., Smeaton, A. F., Byrne, D., O'Hare, N., Jones, G. J., and O'Connor., N. E. (2008b). An examination of a large visual lifelog. In *AIRS 2008 - Asia Information Retrieval Symposium*.

Gürses, S., Troncoso, C., and Diaz, C. (2011). Engineering privacy by design. *Computers, Privacy & Data Protection*.

Hamilakis, Y. and Labanyi, J. (2008). Introduction: time, materiality, and the work of memory. *History & Memory*, 20(2):5–17.

Hanjalic, A. (2012). New grand challenge for multimedia information retrieval: bridging the utility gap. *International Journal of Multimedia Information Retrieval*, 1:139–152.

Hauptmann, A., Yan, R., and Lin, W.-H. (2007). How many high-level concepts will fill the semantic gap in news video retrieval? In *Proceedings of the 6th ACM international conference on Image and video retrieval*, CIVR '07, pages 627–634, New York, NY, USA. ACM.

Haveliwala, T. H. (2002). Topic-sensitive pagerank. In *Proceedings of the 11th International Conference on World Wide Web*, WWW '02, pages 517–526, New York, NY, USA. ACM.

Hayes, G. R., Patel, S. N., Truong, K. N., Iachello, G., Kientz, J. A., Farmer, R., and Abowd, G. D. (2004). The personal audio loop: Designing a ubiquitous audio-based memory aid. In *In Proceedings of Mobile HCI*, pages 168–179. Springer Verlag.

Hearst, M. A. and Plaunt, C. (1993). Subtopic structuring for full-length document access. In *SIGIR '93: Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 59–68, New York, NY, USA. ACM.

Hebbalaguppe, R., McGuinness, K., Kuklyte, J., Healy, G., O'Connor, N., and Smeaton, A. (2013). How interaction methods affect image segmentation: user experience in the task. In *The 1st IEEE Workshop on User-Centered Computer Vision (UCCV), 16-18 Jan 2013, Tampa, Florida, U.S.A.*. IEEE.

Heittola, T., Mesaros, A., Eronen, A., and Virtanen, T. (2010). Audio context recognition using audio event histograms. In *18th European Signal Processing Conference*, pages 1272–1276.

Hodges, S., Berry, E., and Wood, K. (2011). SenseCam: A wearable camera that stimulates and rehabilitates autobiographical memory. *Memory*, 7(19):685–696.

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., and Wood, K. (2006). SenseCam: A retrospective memory aid. In *UbiComp: 8th International Conference on Ubiquitous Computing*, volume 4602 of *LNCS*, pages 177–193, Berlin, Heidelberg. Springer.

Hopfgartner, F., Yang, Y., Zhou, L., and Gurrin, C. (2013). *Semantic Models for Adaptive Interactive Systems*, chapter User Interaction Templates for the Design of Lifelogging Systems, pages 187–204. Springer London.

Hori, T. and Aizawa, K. (2003). Context-based video retrieval system for the life-log applications. In *MIR '03: Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval*, pages 31–38, New York, NY, USA. ACM Press.

Hughes, M., Newman, E., Smeaton, A. F., and O'Connor, N. E. (2012). A lifelogging approach to automated market research. In *Proceedings of the SenseCam Symposium 2012*.

Ingwersen, P. and Järvelin, K. (2005). *The Turn: Integration of information seeking and retrieval in context*. Springer.

Jacquemard, T., Novitzky, P., O'Brolcháin, F., Smeaton, A. F., and Gordijn, B. (2013). Challenges and opportunities of lifelog technologies: A literature review and critical analysis. *Science and engineering ethics*, pages 1–31.

Jaimes, R., Omura, K., Nagamine, T., and Hirata, K. (2004). Memory cues for meeting video retrieval. In *In CARPE'04: Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 74–85. ACM Press.

Järvelin, K. (2011). IR research: systems, interaction, evaluation and theories. In *Proceedings of the 33rd European conference on Advances in information retrieval*, ECIR'11, pages 1–3, Berlin, Heidelberg. Springer-Verlag.

Järvelin, K. and Wilson, T. (2003). On conceptual models for information seeking and retrieval research. *Information Research*, 9(1):9–1.

Joachims, T. (2002). SVM light, http://svmlight.joachims.org.

Jones, G. J., Gurrin, C., Kelly, L., Byrne, D., and Chen, Y. (2008). Information access tasks and evaluation for personal lifelogs. In *2nd International workshop on Evaluating Information Access (EVIA)*, Tokyo, Japan.

Jones, W. (2007). Personal information management. *Annual Review of Information Science and Technology*, 41(1):453–504.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., and Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306:1776–1780.

Kalnikaite, V., Sellen, A., Whittaker, S., and Kirk, D. (2010). Now let me see where i was: Understanding how lifelogs mediate memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2045–2054, New York, NY, USA. ACM.

Kansal, A., Hsu, J., Zahedi, S., and Srivastava, M. B. (2007). Power management in energy harvesting sensor networks. *ACM Transactions on Embedded Computing Systems (TECS)*, 6(4):32.

Kelly, P., Doherty, A. R., Berry, E., Hodges, S., Batterham, A. M., and Foster, C. (2011). Can we use digital life-log images to investigate active and sedentary travel behaviour ? Results from a Pilot Study. *International Journal Behavioural Nutrition Physical Activity*, 8:44.

Kelly, P., Doherty, A. R., Hamilton, A., Matthews, A., Batterham, A. M., Nelson, M., Foster, C., and Cowburn, G. (2012). Investigating the error on self-reported journey durations in school children. *Am J Prev Med*, 43(5):546–550.

Kelly, P., Doherty, A. R., Smeaton, A. F., Gurrin, C., and O'Connor, N. E. (2010). The colour of life: novel visualisations of population lifestyles. In *MM '10 Proceedings of the international conference on Multimedia*. ACM.

Kelly, P., Marshall, S., Badland, H., Kerr, J., Oliver, M., Doherty, A. R., and Foster, C. (2013). Ethics of using wearable cameras devices in health behaviour research. *Am J Prev Med*, 44(3):314–319.

Kennedy, L. and Hauptmann, A. (2006). LSCOM lexicon definitions and annotations version 1.0. Technical report, Columbia University.

Kent, P. (2012). *Search engine optimization for dummies*. Wiley.

Kern, N., Schiele, B., and Schmidt, A. (2007). Recognizing context for annotating a live life recording. *Personal and Ubiquitous Computing*, 11(4):287–298.

Kerr, J., Marshall, S. J., Godbole, S., Chen, J., Legge, A., Doherty, A. R., Kelly, P., Oliver, M., Badland, H., and Foster, C. (2013). Using the SenseCam to improve classifications of sedentary behavior in free-living settings. *American Journal of Preventive Medicine*, 44(3):290–296.

Kikhia, B., Hallberg, J., Bengtsson, J. E., Savenstedt, S., and Synnes, K. (2010). Building digital life stories for memory support. *Int. J. Comput. Healthc.*, 1(2):161–176.

Kitamura, K., Yamasaki, T., and Aizawa, K. (2008). Food log by analyzing food images. In *Proceedings of the 16th ACM international conference on Multimedia*, MM '08, pages 999–1000, New York, NY, USA. ACM.

Klimt, B. and Yang, Y. (2004). Introducing the Enron corpus. In *First Conference on Email and Anti-Spam (CEAS)*.

Kokare, M., Chatterji, B., and Biswas, P. (2003). Comparison of similarity metrics for texture image retrieval. In *TENCON 2003: Conference on Convergent Technologies for Asia-Pacific Region*, pages 571–575.

Kuhlthau, C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5):361–371.

Kumpulainen, S., Järvelin, K., Serola, S., Doherty, A. R., Smeaton, A. F., Byrne, D., and Jones, G. J. F. (2009). Data collection methods for task-based information access in molecular medicine. pages 1–10.

Langheinrich, M. (2001). Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. In Abowd, G., Brumitt, B., and Shafer, S., editors, *Ubicomp 2001: Ubiquitous Computing*, volume 2201 of *Lecture Notes in Computer Science*, pages 273–291. Springer Berlin Heidelberg.

Lavelle, B., Byrne, D., Gurrin, C., Smeaton, A. F., and Jones, G. J. (2007). Bluetooth Familiarity: Methods of Calculation, Applications and Limitations. In *MIRW 2007 - Mobile Interaction with the Real World, Workshop at the MobileHCI07: 9th International Conference on Human Computer Interaction with Mobile Devices and Services*.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Alstyne, M. V. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, 323(5915):721–723.

Lee, H., Smeaton, A. F., O'Connor, N. E., Jones, G. J., Blighe, M., Byrne, D., Doherty, A., and Gurrin., C. (2008). Constructing a SenseCam Visual Diary as a Media Process Multimedia Systems. *Multimedia Systems Journal, Special Issue on Canonical Processes of Media Production*, 14(6):341–349.

Lee, M. L. and Dey, A. K. (2007). Providing good memory cues for people with episodic memory impairment. In *Assets '07: Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 131–138, Tempe, Az., USA. ACM.

Lee, M. L. and Dey, A. K. (2008). Using lifelogging to support recollection for people with episodic memory impairment and their caregivers. In *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*, HealthNet '08, pages 14:1–14:3, New York, NY, USA. ACM.

Li, J. and Wang, J. Z. (2003). Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1075–1088.

Lim, J.-H., Tian, Q., and Mulhem, P. (2003). Home photo content modeling for personalized event-based retrieval. *Multimedia, IEEE*, 10(4):28–37.

Lin, W.-H. and Hauptmann, A. (2006). Structuring continuous video recordings of everyday life using time-constrained clustering. In *Multimedia Content Analysis, Management, and Retieval: SPIE-IST Electronic Imaging*, volume 6073, pages 111–119.

Lindley, S. E., Glancy, M., Harper, R., Randall, D., and Smyth, N. (2011). "Oh and how things just don't change, the more things stay the same": Reflections on SenseCam images 18 months after capture. *International Journal of Human-Computer Studies*, 69(5):311–323. Designing for Reflection on Personal Experience.

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110.

Machajdik, J., Hanbury, A., Garz, A., and Sablatnig, R. (2011). Affective computing for wearable diary and lifelogging systems: An overview. In Mayer, H., Uray, M., and Ganster, H., editors, *Machine Vision-Research for High Quality Processes and Products-35th Workshop of the Austrian Association for Pattern Recognition*. Austrian Computer Society.

Mann, S. (1997). Wearable computing: A first step toward personal imaging. *Computer*, 30(2):25–32.

Mann, S. (2004). Continuous lifelong capture of personal experience with EyeTap. In *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, CARPE'04, pages 1–21, New York, NY, USA. ACM.

Mann, S., Fung, J., Aimone, C., Sehgal, A., and Chen, D. (2005). Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences. In *Alt. Chi, Proc. CHI 2005*.

Mann, S., Huang, J., Janzen, R., Lo, R., Rampersad, V., Chen, A., and Doha, T. (2011). Blind navigation with a wearable range camera and vibrotactile helmet. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1325–1328, New York, NY, USA. ACM.

Mann, S., Lo, R., Ovtcharov, K., Gu, S., Dai, D., Ngan, C., and Ai, T. (2012). Realtime HDR (High Dynamic Range) video for EyeTap wearable computers, FPGA-based seeing aids, and glasseyes (EyeTaps). In *Electrical Computer Engineering (CCECE), 2012 25th IEEE Canadian Conference on*, pages 1–6.

Mantiuk, R., Kowalik, M., Nowosielski, A., and Bazyluk, B. (2012). Do-it-yourself eye tracker: low-cost pupil-based eye tracker for computer graphics applications. In *Proceedings of the 18th international conference on Advances in Multimedia Modeling*, MMM'12, pages 115–125, Berlin, Heidelberg. Springer-Verlag.

Massimi, M. and Baecker, R. M. (2010). A death in the family: Opportunities for designing technologies for the bereaved. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1821–1830, New York, NY, USA. ACM.

May, G. and Warrington, G. (2012). Assessment of physical activity in search and rescue operations using accelerometer based technologies. In *SenseCam Symposium 2012*, Oxford, UK.

Mayer-Schonberger, V. (2011). *Delete: the virtue of forgetting in the digital age*. Princeton University Press.

Melucci, M. (2012). *Contextual Search: A Computational Framework.*, volume 6. NOW Publishers.

Michie, S., van Stralen, M., and West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1):1–12.

Montague, M. and Aslam, J. A. (2001). Relevance score normalization for metasearch. In *CIKM '01: Proceedings of the tenth international conference on Information and knowledge management*, pages 427–433, New York, NY, USA. ACM Press.

Naaman, M., Harada, S., Wang, Q., Garcia-Molina, H., and Paepcke, A. (2004). Context data in geo-referenced digital photo collections. In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 196–203, New York, NY, USA. ACM.

Naphade, M., Smith, J. R., Tesic, J., Chang, S.-F., Hsu, W., Kennedy, L., Hauptmann, A., and Curtis, J. (2006). LSCOM: Large-Scale Concept Ontology for Multimedia. *IEEE Multimedia*, 13(3):86–91.

Nelson, R. O. and Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification*, 5(1):3–14.

Nelson, T. H. (1988). Managing immense storage. *BYTE*, 13(1):225–238.

Neo, S.-Y., Zhao, J., Kan, M.-Y., and Chua, T.-S. (2006). Video retrieval using high level features: Exploiting query matching and confidence-based weighting. In *CIVR'06: Proc. Conf. Image and Video Retrieval*, pages 143–152. Springer-Verlag, Tempe, AZ, USA.

Nistér, D. and Stewnius, H. (2006). Scalable recognition with a vocabulary tree. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2161–2168, New York, USA.

Oard, D. W. and Malionek, J. (2013). The apollo archive explorer. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pages 453–454. ACM, ACM.

O'Brien, A., McDaid, K., Loane, J., Doyle, J., and O'Mullane, B. (2012). Visualisation of movement of older adults within their homes based on pir sensor data. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*, pages 252–259. IEEE.

Ó'Conaire, C., O'Connor, N. E., Smeaton, A. F., and Jones., G. J. (2007). Organising a daily Visual Diary Using Multi-Feature Clustering. In *SPIE Electronic Imaging - Multimedia Content Access: Algorithms and Systems (EI121)*, San Jose, CA.

O'Connor, E., Hayes, J., Smeaton, A. F., O'Connor, N. E., and Diamond, D. (2009). Environmental monitoring of galway bay: fusing data from remote and in-situ sources. In *SPIE Europe Remote Sensing*, pages 74780X–74780X. International Society for Optics and Photonics.

O'Connor, M. F., Conroy, K., Roantree, M., Smeaton, A. F., and Moyna, N. M. (2009). Querying XML data streams from wireless sensor networks: An evaluation of query engines. In *Research Challenges in Information Science, 2009. RCIS 2009. Third International Conference on*, pages 23–30. IEEE.

O'Hara, K., Tuffield, M., and Shadbolt, N. (2009). Lifelogging: Privacy and empowerment with memories for life. In *Identity in the Information Society*, volume 1.

O'Hare, N., Gurrin, C., Jones, G. J., and Smeaton., A. F. (2005). Combination of content analysis and context features for digital photograph retrieval. In *2nd IEE European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies*, pages 323–328, Washington, DC, USA. IEEE Computer Society.

O'Hare, N., Gurrin, C., Jones, G. J. F., Lee, H., O'Connor, N. E., and Smeaton, A. F. (2007). Using text search for personal photo collections with the mediassist system. In *SAC '07: Proceedings of the 2007 ACM symposium on Applied computing*, pages 880–881, New York, NY, USA. ACM Press.

O'Loughlin, G., Cullen, S., McGoldrick, A., O'Connor, S., Blain, R., O'Malley, S., and Warrington, G. D. (2013). Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. *American Journal of Preventive Medicine*, 44(3):297–301.

Oneata, D., Douze, M., Revaud, J., Jochen, S., Potapov, D., Wang, H., Harchaoui, Z., Verbeek, J., Schmid, C., Aly, R., McGuiness, K., Chen, S., O'Connor, N. E., Chatfield, K., Parkhi, O., Arandjelovic, R., Zisserman, A., Basura, F., and Tuytelaars, T. (2012). AXES at TRECVid 2012: KIS, INS, and MED. In *TRECVID Workshop*, Gaithersburg, Md., USA.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web.

Pauly-Takacs, K., Moulin, C. J., and Estlin, E. J. (2011). SenseCam as a rehabilitation tool in a child with anterograde amnesia. *Memory*, 19(7):705–712.

Petrelli, D. and Whittaker, S. (2010). Family memories in the home: contrasting physical and digital mementos. *Personal Ubiquitous Comput.*, 14(2):153–169.

Price, B. A. (2010). Challenges in eliciting privacy and usability requirements for lifelogging. ACM.

Qiu, Z., Doherty, A. R., Gurrin, C., and Smeaton, A. F. (2010). Turning raw sensecam accelerometer data into meaningful user activities. In *SenseCam 2010: Proceedings of the Second Annual SenseCam Symposium*, Dublin, Ireland.

Qiu, Z., Gurrin, C., Doherty, A., and Smeaton, A. F. (2012). A Real-Time life experience logging tool. *Advances in Multimedia Modeling*, pages 636–638.

Reddy, S., Parker, A., Hyman, J., Burke, J., Estrin, D., and Hansen, M. (2007a). Image browsing, processing, and clustering for participatory sensing: lessons from a dietsense prototype. In *EmNets'07: Proceedings of the 4th workshop on Embedded networked sensors*, pages 13–17, Cork, Ireland. ACM Press.

Reddy, S., Parker, A., Hyman, J., Burke, J., Estrin, D., and Hansen, M. (2007b). Image browsing, processing, and clustering for participatory sensing: lessons from a dietsense prototype. In *EmNets '07: Proceedings of the 4th workshop on Embedded networked sensors*, pages 13–17, New York, NY, USA. ACM.

Richardson, M. and Domingos, P. (2002). The intelligent surfer: Probabilistic combination of link and content information in pagerank. In *Advances in Neural Information Processing Systems '14*, volume 14.

Riedl, M. O. and Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39(1):217–268.

Roantree, M., Shi, J., Cappellari, P., O'Connor, M., and Moyna, N. M. (2012). Data transformation and query management in personal health sensor networks. *Journal of Network and Computer Applications*, 35(4):1191–1202.

Rogers, E. M. (2003). *Diffusion of Innovations*. The Free Press, 5 edition.

Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vision*, 40(2):99–121.

Sadlier, D. A. and O'Connor, N. E. (2005). Event detection in field sports video using audio-visual features and a support vector machine. *Circuits and Systems for Video Technology, IEEE Transactions on*, 15(10):1225–1233.

Salembier, P. and Sikora, T. (2002). *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley & Sons, Inc., New York, NY, USA.

Scherp, A., Jain, R., Kankanhalli, M., and Mezaris, V., editors (2010). *EiMM '10: Proceedings of the 2nd ACM international workshop on Events in multimedia*, New York, NY, USA. ACM. 433107.

Schoeffmann, K., Ahlström, D., Bailer, W., Cobârzan, C., Hopfgartner, F., McGuinness, K., Gurrin, C., Frisson, C., Le, D.-D., Fabro, M., Bai, H., and Weiss, W. (2013). The video browser showdown: a live evaluation of interactive video search tools. *International Journal of Multimedia Information Retrieval*, pages 1–15.

Sellen, A. J., Fogg, A., Aitken, M., Hodges, S., Rother, C., and Wood, K. (2007). Do life-logging technologies support memory for the past?: an experimental study using SenseCam. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90, New York, NY, USA. ACM.

Sellen, A. J. and Whittaker, S. (2010). Beyond total capture: a constructive critique of lifelogging. *Comm. ACM*, 53(5):70–77.

Shah, M., Mears, B., Chakrabarti, C., and Spanias, A. (2012). Lifelogging: Archival and retrieval of continuously recorded audio using wearable devices. In *Emerging Signal Processing Applications (ESPA), 2012 IEEE International Conference on*, pages 99–102.

Silva, A. R., Pinho, S., Macedo, L., and Moulin, C. J. (2013). Does SenseCam improve general cognitive performance? *Am J Prev Med*, 44(3):302–307.

Sinha, P. and Jain, R. (2008). Semanics in digital photos: A contextual analysis. *International Journal of Semantic Computing*, 2(3):311–325.

Smeaton, A. F. and Browne., P. (2006). A usage study of retrieval modalities for video shot retrieval. *Information Processing and Management*, 42(5):1330–1344.

Smeaton, A. F., Over, P., and Doherty, A. R. (2010). Video shot boundary detection: Seven years of TRECVid activity. *Computer Vision and Image Understanding*, 114(4):411–418.

Smeaton, A. F., Over, P., and Kraaij, W. (2006). Evaluation campaigns and TRECVid. In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, MIR '06, pages 321–330, Santa Barbara, California, USA. ACM.

Smeaton, A. F., Over, P., and Kraaij, W. (2009). High-level feature detection from video in TRECVid: a 5-year retrospective of achievements. In Divakaran, A., editor, *Multimedia Content Analysis, Theory and Applications*, pages 151–174. Springer Verlag, Berlin.

Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1349–1380.

Snoek, C. G. M. and Worring, M. (2009). Concept-based video retrieval. *Foundations and Trends in Information Retrieval*, 2(4):215–322.

Snoek, C. G. M., Worring, M., van Gemert, J. C., Geusebroek, J.-M., and Smeulders, A. W. M. (2006). The challenge problem for automated detection of 101 semantic concepts in multimedia. In *Proceedings of the 14th annual ACM international conference on Multimedia*, MULTIMEDIA '06, pages 421–430, New York, NY, USA. ACM.

Spence, R. (2002). Rapid, serial and visual: A presentation technique with potential. *Information Visualization*, 1(1):13–19.

Stanley, K. G. and Osgood, N. D. (2011). The potential of sensor-based monitoring as a tool for health care, health promotion, and research. *The Annals of Family Medicine*, 9(4):296–298.

Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., Picard, R. W., and Pentland, A. (1997). Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments*, 6(4):386–398.

Stein, E. W. (1995). Organization memory: Review of concepts and recommendations for management. *International Journal of Information Management*, 15(1):17–32.

Stix, G. (2011). Photographic memory: Wearable cam could help patients stave off effects of impaired recall. *Scientific American*.

Takahashi, M., Ito, S., Sumi, Y., Tsuchikawa, M., Kogure, K., Mase, K., and Nishida, T. (2004). A layered interpretation of human interactions captured by ubiquitous sensors. In *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, CARPE'04, pages 32–38, New York, NY, USA. ACM.

Tancharoen, D. and Aizawa, K. (2004). Novel concept for video retrieval in life log application. In *PCM: Pacific Rim Conference on Multimedia*, pages 915–923.

Tancharoen, D., Yamasaki, T., and Aizawa, K. (2005). Practical experience recording and indexing of life log video. In *CARPE '05: Proceedings of the 2nd ACM workshop on Continuous archival and retrieval of personal experiences*, pages 61–66, New York, NY, USA. ACM.

Tancharoen, D., Yamasaki, T., and Aizawa, K. (2006). Practical life log video indexing based on content and context. In *Multimedia Content Analysis, Management, and Retieval: In Proceedings of SPIE-IST Electronic Imaging.*

Taneja, H. and Gupta, R. (2010). Web information retrieval using query independent page rank algorithm. In *Advances in Computer Engineering (ACE), 2010 International Conference on*, pages 178–182.

Troiano, R. P., Berrigan, D., Dodd, K. W., Masse, L. C., Tilert, T., and McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*, 40:181–188.

van den Hoven, E., Sas, C., and Whittaker, S. (2012). Introduction to this special issue on designing for personal memories: Past, present, and future. *Human-Computer Interaction*, 27(1-2):1–12.

Vemuri, S. and Bender, W. (2004). Next-generation personal memory aids. *BT Technology Journal*, 22(4):125–138.

Vemuri, S., Schmandt, C., Bender, W., Tellex, S., and Lassey, B. (2004). An audio-based personal memory aid. In *Ubicomp*, pages 400–417.

Vicedo, J. L. and Gómez, J. (2007). Trec: Experiment and evaluation in information retrieval: Book reviews. *J. Am. Soc. Inf. Sci. Technol.*, 58(6):910–911.

Walter, C. (2005). Kryder's law. *Scientific American*, 293(2):32–33.

Wang, P. and Smeaton, A. F. (2011). Aggregating semantic concepts for event representation in lifelogging. In *Proceedings of the International Workshop on Semantic Web Information Management*, SWIM '11, pages 8:1–8:6, New York, NY, USA. ACM.

Wang, P. and Smeaton, A. F. (2012). Semantics-based selection of everyday concepts in visual lifelogging. *International Journal of Multimedia Information Retrieval*, 1:87–101.

Wang, Z., Hoffman, M. D., Cook, P. R., and Li, K. (2006). Vferret: content-based similarity search tool for continuous archived video. In *CARPE '06: Proceedings of the 3rd ACM workshop on Continuous archival and retrival of personal experences*, pages 19–26, New York, NY, USA. ACM.

Watson, H. J. and Wixom, B. H. (2007). The current state of business intelligence. *Computer*, 40(9):96–99.

Whittaker, S., Kalnikaitė, V., Petrelli, D., Sellen, A., Villar, N., Bergman, O., Ilan, B., Clough, P., Brockmeier, J., and Whittaker, S. (2012). Socio-technical lifelogging: Deriving design principles for a future proof digital past. *Human-Computer Interaction (Special Issue on Designing for Personal Memories)*, 27(1-2):37.62.

Woodman, O. and Harle, R. (2008). Pedestrian localisation for indoor environments. In *Proceedings of the 10th international conference on Ubiquitous computing*, UbiComp '08, pages 114–123, New York, NY, USA. ACM.

Yang, Y., Zhou, R., and Gurrin, C. (2012). A mechanical memory: Prototype digital memories in everyday devices. In *6th Irish Human Computer Interaction Conference (iHCI2012)*, Galway, Ireland.

Yeung, M. and Yeo, B.-L. (1996). Time-constrained clustering for segmentation of video into story units. *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, 3:375–380.

Zacks, J. M., Speer, N. K., Vettel, J. M., and Jacoby, L. L. (2006). Event Understanding and Memory in Healthy Aging and Dementia of the Alzheimer Type. *Psychology and Aging*, 21(3):466–482.

Zacks, J. M. and Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1):3–21.

Zhang, S., Rowlands, A. V., Murray, P., and Hurst, T. L. (2012). Physical activity classification using the GENEA wrist-worn accelerometer. *Med Sci Sports Exerc*, 44(4):742–748.

Zheng, Y.-T., Zhao, M., Song, Y., Adam, H., Buddemeier, U., Bissacco, A., Brucher, F., Chua, T.-S., and Neven, H. (2009). Tour the world: Building a web-scale landmark recognition engine. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)*, pages 1085–1092.

Zhou, L., Caprani, N., Gurrin, C., and O'Connor, N. E. (2013). ShareDay: A novel lifelog management system for group sharing. In Li, S., Saddik, A., Wang, M., Mei, T., Sebe, N., Yan, S., Hong, R., and Gurrin, C., editors, *Advances in Multimedia Modeling*, volume 7733 of *Lecture Notes in Computer Science*, pages 490–492. Springer Berlin Heidelberg.