Talking to GNOMEs: Exploring Privacy and Trust Around Internet of Things Devices in a Public Space

Richard Milton
Bartlett Centre for Advanced Spatial Analysis
University College London
WC1E 6BT, England
richard.milton@ucl.ac.uk

Andy Hudson-Smith
Bartlett Centre for Advanced Spatial Analysis
University College London
WC1E 6BT, England
a.hudson-smith@ucl.ac.uk

Boyana Buyuklieva
Bartlett Centre for Advanced Spatial Analysis
University College London
WC1E 6BT, England
boyana.buyuklieva.14@ucl.ac.uk

Steven Gray
Bartlett Centre for Advanced Spatial Analysis
University College London
WC1E 6BT, England
steven.gray@ucl.ac.uk

Duncan Hay
Bartlett Centre for Advanced Spatial Analysis
University College London
WC1E 6BT, England
d.hay@ucl.ac.uk

CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada
ACM ISBN 978-1-4503-5621-3/18/04.
https://doi.org/10.1145/3170427.3188481

Abstract

Privacy issues can be difficult for end-users to understand and are therefore a key concern for information-sharing systems. This paper describes a deployment of fifteen Bluetooth-beacon-enabled 'creatures' spread across London's Queen Elizabeth Olympic Park, which initiate conversations on mobile phones in their vicinity via push notifications. Playing on the common assumption that neutral public settings promote anonymity, users' willingness to converse with personified chatbots is used as a proxy for understanding their inclination to share personal and potentially disclosing information. Each creature is linked to a conversational agent that asks for users' memories and their responses are then shared with other creatures in the network. This paper presents the design of an interactive device used to test users' awareness of how their information propagates to others.

Author Keywords

Chatbots; 3D printing; Public engagement; Natural language; Bluetooth beacons; Eddystone; Privacy.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces (D.2.2, H.1.2, I.3.6), Natural Language; H.5.m. [Information Interfaces and Presentation (e.g. HCI)]: Miscellaneous

Figure 1: Models of a gnome, bat, otter and honey bee. Fifteen were 3D printed and deployed around the Queen Elizabeth Olympic Park with Eddystone Beacons for the general public to interact with. The models were painted in bright colours by local children from a community group between the ages of 5 and 15.

Introduction

This paper describes ‘Tales of the Park’, a research project conducted by University College London at the Queen Elizabeth Olympic Park in East London, designed to explore security and trust around Internet of Things technology. The project took the form of a deployment of fifteen devices called ‘GNOMES’, or ‘Geo-located Natural-language Objects for Memory Evocation’: 3d-printed creatures containing low-energy Bluetooth beacons which, through push notifications and a simple chatbot interface, solicited memories about the Olympic Park from members of the public via their smartphones. These memories were shared with other visitors to the park and amongst the GNOMES themselves, creating an intentionally ‘leaky’ (and ‘creepy’) [8] device network designed to explore the data collection and sharing issues implied by Internet of Things (IoT) technologies. In an effort to continue the ‘gaming and playfulness’ theme of the experiment, the website has been designed to look like a computer game. Users can only interact with the creatures if they are within 200 metres.

Background

Consumer electronic devices like the Amazon Alexa, Apple’s Siri, Microsoft’s Cortana, or Google’s digital assistant, have brought conversational user experiences to the general public. However, it is often not clear to the users how these new computing interfaces are using natural language understanding, giving rise to privacy issues. While ‘talking’ to an agent like Alexa, the user experience is of a one to one conversation, much like humans converse with one another. What is happening in reality is a many to one interaction between the user and a machine holding multiple simultaneous conversations. Information is able to ‘leak’ into other conversations and into the fabric of our own lives. Users may not be aware of the sophistication of the infrastructure required to provide these services; moreover, whilst

Figure 2: Loki the gnome, which is located outside the ViewTube cafe area near the cycle-hire. The plinth is 1.1 metres high.

they might suspect that data about them are being gathered, end user licence agreements are often unclear with regards to how their personal data is being used.

The “Tales of the Park” project sought to address these issues by using a network of devices which use natural language interfaces to solicit memories from park visitors. These “memories of place” were then shared between the devices and the visitors using the natural language interface, demonstrating to users the way in which data can flow around a device network. Autotopography, or objects which constitute a physical map of memory, was investigated in ‘Tales of Things and electronic Memory’ (TOTeM) [3], where a web-based system of tagging physical objects with memories was developed. Autotopography refers to the, “spatial, local and situational “writing” of the self’s life

One to One and One to Many Interactions

The one to one and one to many relationship is the topic of Balbi and Kittler's history of human communications [2]. Starting from Plato's ancient Greece and ending in the modern digital era, they conclude with the observation that, "historical sources indicate that the two patterns of communication are inherently interwoven". If Alexa is listening to what we are saying, then it is able to affect our shopping habits on Amazon. What we search for on Amazon and Google affects what other people see in their recommendations, in fact the speech recognition ability of these devices is only possible because of the large databases of audio samples from multiple conversations which provide the training data for deep learning algorithms, for example, the 'Deep Speech' architecture presented in [5]. In order for these systems to be able to converse with humans, they first need to learn from human conversation, both in speech recognition and in conversational analysis.

in visual art" [1]. Hoven and Eggen give an overview of the field of augmented memory systems [10], arguing that autobiographical memories are more than a list of memories, being triggered by cues that connect memories together, but which change as we add to them.

Our experiment explores users' willingness to share memories of place, by crowd-sourcing 'personal' and 'social' memories. A personal memory [9] is described by Thorne as "specific events from one's past that are part of the autobiographical memory", while a social memory [11] is one that is shared within a group. Thorne goes on to say that "the landscape of personal memory has got a bump", a reference to how unevenly the memories we choose to recall are distributed throughout our lives. Of interest here is the observation that only "3% of the events were rated as highly memorable after 300 days" and "any personal memories lasting more than a year are rare events". By adding location, it can be argued that both personal and social are linked, so a personal memory like, "I went swimming in the pool here" becomes part of the social history of the place where it is shared with other members of the public.

Our justification for using language and memory can be expressed as follows:

> "The main claim for the use of narrative in educational research is that humans are storytelling organisms who, individually and socially lead storied lives."

(Clandinin and Connelly [4, pp2])

It is this method of conversation by taking turns at question and answer which poses some of the biggest challenges to our agents. In 'Conversational Analysis as Social Theory' [6], the problem of "turn-taking" and how we decide who should talk next is covered. In the human to human conversations analysed, conversations break down and then are repaired. We aim to observe this effect in our agents' conversations in order to improve their ability to persuade people to offer their memories.

To conclude, in "Minds, Brains and Programs" [7], John Searle puts forward the 'Chinese room' argument. Put simply, a computer following a program that manipulates Chinese symbols could conceivably produce the appearance of understanding the language, but it does not genuinely understand Chinese. Our conversational agents use Google's Dialogflow service, which is trained using 'intents' and 'actions', which we define ourselves, together with a larger corpus of conversational training data which is a black box to us. The 'program' here is a neural network that has been trained to respond to the 'intent' with the correct 'action'. How well the agent is able to cope in the real world and sustain a conversation on a topic is where AI meets HCI.

Design

The creature models are 3D printed using white PLA and painted by schoolchildren, aged 5-15, at a local community group. Given the amount of PLA used, the approximate cost of a model can be estimated at 4 UK pounds[1]. Figure 2 shows the design of the plinth with a painted gnome inside. As the fifteen creatures are spread around the park, locations vary between exposed hilltop, entrances to venues, football stadium, lakeside and residential. The distance between the furthest two creatures is 2 kilometres, requiring approximately 2 hours to walk around all fifteen deployment sites. This limited the choice of technology to

[1] 4 UK pounds is equivalent to about 4 US Dollars, or a cappuccino in a coffee shop.

Beacon Technology

Bluetooth Beacons: Low power devices which are behind Google's 'Physical Web', they transmit messages to mobile devices in the local area.

Physical Web: When enabled on Android or iphone, the 'physical web' enables the phone to receive message packets from nearby beacons, alerting the user to physical presence of interactive objects via alerts.

BLE: Bluetooth **L**ow **E**nergy, a wireless protocol that allows messages to be transmitted using significantly less power than the standard Bluetooth protocol. This enables novel applications that allow mobile devices to communicate frequently using small message packets.

Eddystone: Open source platform from Google which includes a URL in the bluetooth packet. Apple's proprietary protocol is called 'iBeacon'.

weather-proof, vandal-proof and able to run on battery for the 3 month deployment.

The Eddystone URL beacons used are self-contained devices, utilising the Bluetooth Low Energy (BLE) protocol to send messages. A platform called "OpenBeacons" (http://openbeacons.com) is used, which allows both Eddystone beacons and iBeacons, to be registered and configured remotely once they have been deployed in the field. The site acts as a transparent proxy which redirects to the application's URL, which the beacons are programmed to broadcast to the mobile clients. It also provides some basic analytics on the use of a particular beacon. The beacon is configured with a specific identifier key which is used by the OpenBeacons platform to redirect to a specific URL when visited by a browser. To the client, the beacon's URL looks like the target URL as the OpenBeacons platform makes a request on behalf of the client and returns the header information for the target web application.

Users converse with the creatures on their mobile phones using agents built using Google's Dialogflow service. The state diagram in figure 3 shows the flow of conversation from the initial, *"Hello, what's your name?"*. The "UI" blocks in the diagram denote user interactions handled by Dialogflow, which might consist of a number of backward and forward exchanges of text before moving to the next state. The natural language processing is handled by Dialogflow's 'intents' and 'actions', although storage of memories and requests for park information can be handled by the website externally. A requirement of the design was not to send potentially disclosive user memories to a 3rd party website. All memories are manually approved by a human before being shared with others. A profanity filter is used to prevent people swearing at the creatures, eliciting a response of the form, *"This is a family-friendly experience, so keep*

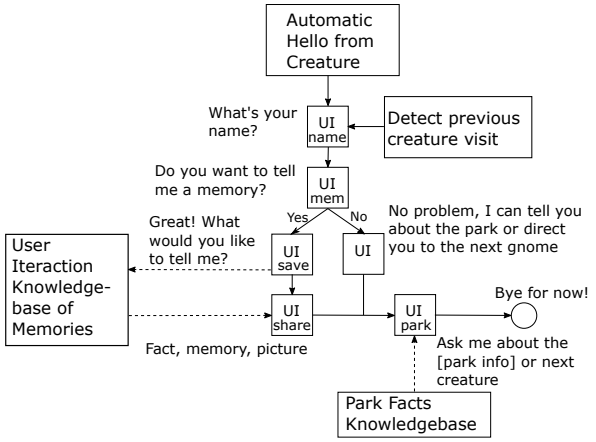

Figure 3: State transition diagram for creature conversation. The 'UI' states denote a User Interaction in natural language using Google's Dialogflow web service.

your language clean!". The talesofthepark.com website uses cookies to determine when a user has already spoken to a creature, triggering the line, *"Hi there Robert! I'm Khadija, I see you've just spoken to Loki"*.

The use of Dialogflow's web service greatly simplified the development of the conversational agents. Every creature uses the same agent program, with specialisation achieved via a webhook to fetch location specific chat from a web service on talesofthepark.com. For example, Zack the otter is outside the Aquatic Centre, so all his park related facts and chat relate to swimming and the pool.

The next context expects a memory, which is stored in the "User Interaction Knowledgebase", but an input from the user of a form similar to, "my name is..." would be detected as a name entry, taking the context back to the name input

Figure 4: Screenshots from the mobile phone interface.

stage. It is an interesting feature of Dialogflow that it does not require a perfect word match for the "user says" field, for example, "hi", "helo" and "hello" would all match for a "hello" intent.

Results

During the 3 month deployment in the park, a total of 4,007 lines of conversation between the creatures and people in the park were recorded, along with 106 memories left. This can be broken down into 3,224 lines of conversation from the creatures and 783 lines of conversation from the users talking to them in 174 distinct chains of conversation.

Interaction rates for the creatures are low compared to the expected footfall for the locations. Rosie the bee, located at the foot of the ArcelorMittal orbit attraction, is the most popular creature, with 29 conversations. An average attendance figure at the QEOP Stadium is 56,000 every other week and, from the park's attendance figures for 2016, there were 101,633 visitors to the ArcelorMittal Orbit that year. A rough calculation puts the interaction rate at 0.17% of visitors talking to Rosie.

Given the nature of the data, a qualitative analysis is required. For example, *"I got married in the Olympic park [date redacted], fish island and Formans after the Orbit"*, is an example of personal disclosure, identifying a place, time and event. In the case of, *"Emma: I've come to the park to do fieldwork for geography gcse"*, we suspect that "Emma" is female and 14-15 years old, which is the age for taking GCSE exams. Going further, we could suppose that she is from a local school, which is leading towards personal disclosure. In all of these cases, social media is an obvious place to look to fill in the missing information. It might even turn out that none of the memories left are true, in which case we gain no information. On a slightly different thread, the memory, *"It's confusing if you're on a bike because of the different levels of roads and paths through the park"*, is an example of a social memory about place. There are two such memories about the difficulties of navigating the park on a bike, so verification might eventually be achieved through quantity of self-supporting data. In the case of, *"Jeff Goldblum: I love your paint job"*, we immediately assume the name to be fake, but there is no possibility of verifying the content of the memory, which is a personal statement. Finally, *"This site used to be a large scrapyard"* and *"Trains used to be built here"* are the only two examples of historical memories about place in the dataset. These are exactly the type of memories that the project was designed to elicit.

Conclusion

An unexpected result of the study, which comes from analysing the individual conversations in detail, is that it could loosely be described as an "in the wild Turing Test". There is a definite split between people who converse with the creatures as if they were talking to a human, compared to those who understand its limitations and use a more 'command and control' style. As we do not know the demographic of our users, this could turn out to be a children versus adults phenomena, or it could be down to user conditioning based on the success of devices like the Amazon Alexa. Finally, in order to classify the memories that were left, a framework built around the three dimensions of, "semantic, spatial and temporal" was used for viewing potentially disclosive data. Essentially, this is searching and filtering of data, so, "I got married on [date] at [place]" is enough to limit the search results to the point of personal disclosure. The 'semantic' dimension relates to a fact that narrows the search space, so 'GCSE' is a qualification that 14-15 year old children take, which inadvertently discloses age. Given the limited nature of this study, it was surprising just how much personal information was being 'leaked' to the creatures, information

which suggests that users either did not understand exactly what they were giving away, or accepted it as normal.

Acknowledgements

The authors would like to thank EPSRC for funding this research through the Privacy, Ethics, Trust, Reliability, Acceptability and Security (PETRAS) grant (EP/N02334X/1), Cyber Security of the Internet of Things and IoT in the Park projects. Also, we would like to thank the London Legacy Development Corporation (LLDC) and Queen Elizabeth Olympic Park, without whose support this could not have happened. The OpenBeacons platform was funded through the Google Internet of Things Research Technology Awards Pilot study.

REFERENCES

1. M. Bal. 2002. "Autotopography: Louise Bourgeois as Builder.". *Biography* 25, 1 (2002), 180–202. DOI: http://dx.doi.org/doi:10.1353/bio.2002.0002 Project MUSE.
2. Gabriele Balbi and Juraj Kittler. 2016. One-to-One and One-to-Many Dichotomy: Grand Theories, Periodization, and Historical Narratives in Communication Studies. *International Journal of Communication* 10, 0 (2016). http://ijoc.org/index.php/ijoc/article/view/4405
3. Ralph Barthel, Kerstin Leder Mackley, Andrew Hudson-Smith, Angelina Karpovich, Martin de Jode, and Chris Speed. 2013. An internet of old things as an augmented memory system. *Personal and Ubiquitous Computing* 17, 2 (01 Feb 2013), 321–333. DOI: http://dx.doi.org/10.1007/s00779-011-0496-8
4. F. Michael Connelly and D. Jean Clandinin. 1990. Stories of Experience and Narrative Inquiry. *Educational Researcher* 19, 5 (1990), 2–14. http://www.jstor.org/stable/1176100
5. Dario Amodei et al. 2015. Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. *CoRR* abs/1512.02595 (2015). http://arxiv.org/abs/1512.02595
6. John Heritage. 2009. *Conversation Analysis as Social Theory*. Wiley-Blackwell, 300–320. DOI: http://dx.doi.org/10.1002/9781444304992.ch15
7. John R. Searle. 1980. Minds, Brains and Programs. *Behavioral and Brain Sciences* 3, 3 (1980), 417–57.
8. Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2347–2356. DOI: http://dx.doi.org/10.1145/2556288.2557421
9. Avril Thorne. 2000. Personal Memory Telling and Personality Development. *Personality and Social Psychology Review* 4, 1 (2000), 45–56. DOI: http://dx.doi.org/10.1207/S15327957PSPR0401_5 PMID: 15710560.
10. Elise van den Hoven and Berry Eggen. 2008. Informing augmented memory system design through autobiographical memory theory. *Personal and Ubiquitous Computing* 12, 6 (01 Aug 2008), 433–443. DOI:http://dx.doi.org/10.1007/s00779-007-0177-9
11. Eviatar Zerubavel. 1996. Social memories: Steps to a sociology of the past. *Qualitative Sociology* 19, 3 (01 Sep 1996), 283–299. DOI: http://dx.doi.org/10.1007/BF02393273