Trap Hunting: Finding Personal Data Management Issues in Next Generation AAC Devices

Joseph Reddington
Royal Holloway College
Egham, UK
j.reddington@rhul.ac.uk

Lizzie Coles-Kemp
Royal Holloway College
Egham, UK
lizzie.coles-kemp@rhul.ac.uk

Abstract

Advances in natural language generation and speech processing techniques, combined with changes in the commercial landscape, have brought within reach dramatic improvements in Augmentative Alternative Communication (AAC). These improvements, though overwhelmingly positive, amplify a family of personal data use problems. This paper argues that the AAC design and implementation process needs to identify and address personal data use problems. Accordingly, this paper explores personal data management problems and proposes responses. This paper is situated in the context of AAC technology but the responses could be generalised for other communities affected by low digital literacy, low literacy levels and cognitive challenges.

1 Introduction

Electronic Augmentative and Alternative Communication (AAC) systems enable individuals with severe speech impairment to verbally communicate their needs, often using Text-to-Speech technology. Such devices are designed to give communication impaired people greater independence and improved opportunities of social integration. The devices enable users to construct utterances, many of which describe themselves or aspects of their lives, including their actions with others and, as such, can be considered 'personal data'. Recent work by Patel and Radhakrishnan (2007), Black et al. (2010), Reiter et al. (2009), and Reddington and Tintarev (2011) makes explicit use of personal data (about both the user and other parties) to improve the functionality of AAC devices. Depending on context, the use of these utterances, in an institutional setting, may be controlled under data protection legislation, or (e.g. domestically) their use may be influenced more by social norms within the context. A key factor in personal data management is the highly contextual nature of privacy related issues; privacy concerns and practices are situated in their context (Nissenbaum, 2009) and influenced by cultural issues (Milberg et al., 2000).

The diversity of technology in the AAC sector is set to increase dramatically. Apple's iPad[1] has caused a huge investment in tablet technology. Multiple, third party applications (e.g. proloque2go[2], myVoice[3], and verbally[4]) already exist that allow this new range of tablets to function as AAC devices.

The effect of this research movement maturing at a time when many new devices and producers are entering the market foreshadows probable major changes and innovations in coming years. This includes a risk of the "panoply of different privacy problems" that privacy theorist Solove (2008) foresaw as a result of diversifying and enhancing technologies.

The authors' position is that it is very timely to explore personal data management problems in this new AAC landscape and in so doing identify traps that AAC design might stumble into as this technology change gathers pace. This perspective can con-

[1] http://www.apple.com/ipad/, retrieved May 2011
[2] http://www.proloquo2go.com, retrieved May 2011
[3] http://new.myvoiceaac.com, retrieved May 2011
[4] http://verballyapp.com/index.html,retrieved May 2011

tribute to the design of technologies and governance structures that are able to both identify and respond to such traps.

As AAC devices are designed to be used in all areas of the AAC user's life, there are a broad range of personal data management problems, which are highly context sensitive and incorporate legal, social, and technical issues. This complex problem space centres on informational privacy issues that contribute to a wider family of personal data management problems that can be found in contexts of AAC use.

This paper situates the personal data management problems in the use of natural language generation and speech processing techniques in AAC. It considers all of the following as personal data: utterances constructed by the system, communication logs and re-communication of stored utterances. Following an overview of state-of-the-art AAC and discussion of how functionality development in next-generation AAC devices maps to the use of personal data, Section 2 identifies and explores personal data use problems in three AAC-specific examples. Section 3 presents possible responses to problems introduced by the examples and Section 4 considers a governance framework that enables emergent personal data management problems with future AAC devices to be identified and considers its applicability for other communities.

1.1 Personal data generated, and used, by AAC devices

Today, AAC devices may excel at needs-based communication (e.g. *"I am hungry"*, *"I'm cold"*, *"get the phone"*) but they are limited for real conversation (Soto et al., 2006). So, in the current generation of AAC devices, the implications for both personal data generation and its use are relatively small because the linguistic capabilities are small. Typical AAC devices tend towards a hierarchical structure of pages, each of which typically focuses on a context (e.g. shopping) or a category (e.g. clothes, sports), rather than observations or recent personal stories (Beukelman and Mirenda, 2005). However, Higginbotham et al. (2007) report that spontaneous conversation with typical devices is slow and difficult (new utterances are typically constructed at a rate of between 8 and 10 words per minute, slightly more if e.g. word prediction is used). Todman et al. (2008) propose utterance-based devices in which devices focus on prepared phrases to facilitate social communication rather than needs-based communication; however, in general, new utterances must be prepared in advance either by the user or a carer, with a large time and energy cost. It is this implementation of functionality designed to speed up utterance production that restricts the production of personal data rather than the underlying technology. A study by Rackensperger et al. (2005) shows that using pre-programmed phrases can reduce the ability for self-expression; as a result, the range of personal data produced is likely to be limited. As an example, there is a particular difficulty in communicating recent or single use events such as talking about one's day or talking about yesterday's television: such utterances are expensive to prepare in advance due to the potential for limited and low-probability use. Thus, AAC users tend to be passive, responding to questions with single words or short sentences, and personal stories tend to be told as a monologue or a sequence of pre-stored utterances (Soto et al., 2006).

To develop the potential for interaction, and therefore increase the degree to which AAC devices can support increased independence, recent research has examined the potential for location-aware devices to offer different content to the user under different conditions (Dominowska et al., 2002; Patel and Radhakrishnan, 2007), and for devices that generate new phrases automatically. In the later case: Black et al. (2010) use external data to populate a communication device, and Reiter et al. (2009) use a NLG engine to generate text from a database of personal facts. These innovations could allow users to increase social interaction and reduce the device maintenance, complementing the growing range of AAC systems with internet connectivity.

1.1.1 Impact on personal data

As the capability for interaction increases, the potential for increased personal data also increases. For example:

- utterances generated from geo-location enabled devices can potentially include information about people (data subjects) other than the

AAC user, as well as increased information about the device users themselves;

- utterances generated from input by teachers, care staff and parents can again potentially contain information about other data subjects, as well as increase the range of information about device users themselves;

- internet access as a medium brings a range of issues for personal data use in terms of the methods used to broadcast and replay utterances and it greatly increases the possibilities for data input (potentially including information about third parties) into the utterances;

- the general browsing facility of internet access increases the ability of users to communicate with the wider world, carrying with it a set of personal data management and privacy issues, much of which is the subject of on-going research (Kani-Zabihi and Coles-Kemp, 2010; Kumaraguru and Cranor, 2005; Spiekermann and Cranor, 2009).

Increasing the potential for interaction and giving more control to the AAC user will increase the range of personal data generated and hence the range of potential personal data use problems. Moreover, the increased creation of novel utterances and wider opportunity to relay such utterances potentially increase intellectual property issues.

AAC devices are designed to increase social interaction in all settings and therefore the devices, and their supporting approaches, must be equally effective in all situations. This is a challenge for any type of personal data management that includes aspects of privacy. Also, AAC users themselves develop their uses and desires for communication (Rackensperger et al., 2005). Therefore, any approach to personal data management has to be highly context sensitive and capable of responding to changing requirements.

1.2 Related work in AAC literature

Although ethics in the context of complex disabilities is well studied, there is little direct research into privacy and personal data management issues in AAC: much of the work is in small accompanying sections to other research contributions and focuses directly on personal data dissemination. For example, Smith (2005) notes that externally displayed lexicons (such as a communications board) violate some aspects of privacy and proposes finding ways to ensure that vocabulary can be delivered discreetly without affecting access. Similarly, Black et al. (2010) address privacy as part of a discussion of security. Additionally, there is some meta-work that looks at the ethics of research into AAC rather than AAC itself: Pennington et al. (2007) notes that the data collected by AAC devices makes identification of the individual trivial, especially when considering the relatively small pool of users, a theme that is also examined in work by Lesher et al. (2000) on logging output of AAC devices.

Privacy has also been raised explicitly in the AAC community by researchers considering design frameworks for next generation devices, e.g., Rackensperger et al. (2005) and DeRuyter et al. (2007). There is also a growing body of AAC research that, in discussing next generation AAC technology, raises a wide range of implicit issues related to privacy and ICT mediated communication. These issues include: anonymity; personalisation of services; identity management; autonomy; and the changing of relationship boundaries through mediation. These are topics that feature in traditional privacy research, but with added complexity.

Therefore, work on the future of AAC and internet connectivity (in particular key features highlighted in DeRuyter et al. (2007)) have great bearing on personal data management, although privacy and personal data management are not directly discussed. DeRuyter et al. (2007) discuss simplified usability, including 'embeddedness' functionality: making AAC devices unobtrusive in their environment. When simplifying usability, there is a tension between requiring user intervention and decision making automation. For example, where should consent mechanisms related to personal information disclosure be placed?

Discussions on future AAC functionality also emphasise adaptive technology that personalises AAC use so that AAC devices are able to recognise a user and adjust functionality accordingly (DeRuyter et al., 2007). However, adaptation algorithms designed

to anticipate or assess user capabilities will make adjustments to functionality based on logs of personal data and usage patterns and thus implicitly process personal data. The ability to adjust such algorithms would give users and their carers increased control over the use of this personal data. In addition, adjusting the capabilities of Internet-enabled AAC devices is likely to also result in changes to the disclosure of a user's personal data. This disclosure would be determined using a logic internal to the adaptation algorithms. Making the logic explicit to users and their carers would make both the personal data disclosure implications of adjustment visible and give greater control over disclosure.

2 Examples

Given the situated nature of informational privacy, in order to explore personal data management issues meaningfully, it is vital to situate the evaluation policy and its related personal data management issues into a particular context. We have selected three AAC-specific examples through which to explore the issues in particular contexts.

This section describes three illustrative scenarios for potential personal data use problems. They are broadly based on the categories of generated content in Reddington and Tintarev (2011) and are constructed with input from legal experts, youth work practitioners, and disability officers. The examples situate the personal data management problems before analysis in Section 3.

2.1 Example 1 - Creating Novel Utterances

The simplest, and least intrusive level of automatically generated content contains inferred utterances that can be deduced from logs of previous utterances. Thus, if a device logged the phrase "Hello Mary" and later "Thanks Mary" the phrase "Today I spent time with Mary" could be added to the list of available phrases. It is trivial to imagine other possible situations where this is applicable - "My communications unit has been away for repair", "I was up very late last night", and "I like to talk about football", are all deducible from previous utterances.

We consider an AAC user Alice, who is solely reliant on her AAC device for communication and is non-literate. Alice is able to generate novel utterances using utterance segments programmed by care staff and by taking advantage of inferred utterances that her device has been designed to provide. Alice has the right to delete certain utterances but care staff and family members are able to restore the deleted utterances. The digest of Alice's activities are backed up every day and could be placed in a catalogue of utterances that the care provider uses at promotional events or on the provider's website.

This scenario raises issues related to intellectual property rights, ownership, and the management of personal data. The management issues centre on control of data, and rights to recover deleted items.

2.2 Example 2 - Implicit and Explicit Personal Data Exchange Rules

The second and third levels of automatically generated content involve receiving data from network portals (such as the internet) and local sensors. For example: "It's very warm today", and "It rained on Friday!". Also included is media data: "On YouTube I watched the 'Star Wars Kid' video ", or "New series of Doctor Who!".

A useful context here is the "How was School Today...?" (HWST) project (Black et al., 2010; Reddington and Tintarev, 2011), which generates stories for students with complex communication needs at a special needs school. The project logs interactions with people, objects and location changes. This sensor data is supplemented with timetable information (to infer classes based on time and location) and voice recordings, before new content is generated.

Consider that Alice is in a class and that her AAC device reads information from sensors to generate novel content in a similar way to the HWST system. Also in class is Charlie, a typically developing child. Charlie's actions are also recorded by the sensors and he takes part in activities with Alice. Alice is able to report Charlie's behaviour to her parents and to other class members. Unlike when her classmate Charlie verbally communicates about his school day, Alice's utterances take a permanent form and can be replayed and reused. Charlie is not really aware of this capability and what this means. Charlie's parents are aware that Alice has some kind of communication device and that sensors are used at school but are not clear on the details. Alice's Mum puts some of Alice's stories, including the one about

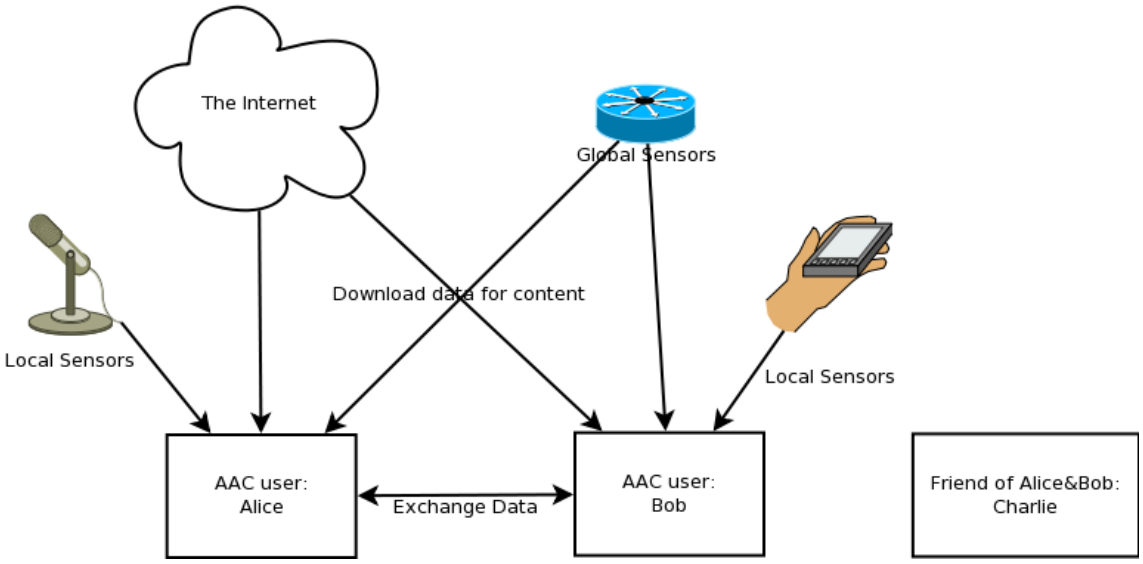

Figure 1: An example information flow

Charlie and the water fight, up on their family blog.

This scenario raises issues of consent to obtain data from sensor sources and of processing the sensor data. It also raises questions related to the dissemination of personal data - about the user and other data subjects. In this scenario, personal data is processed in two contexts: school and Alice's home. This shows the complex array of stakeholders involved in managing personal data. Moreover, there are questions of how non-AAC users are informed of the implications of AAC use in the school or in any other setting. Implicitly there is a problem of ensuring that AAC and non-AAC users are treated equally in terms of the personal data rights, which in turn raises issues of how verbal and AAC generated utterances are valued in the social context.

2.3 Example 3 - Control over Data Sharing

An additional level of complexity is the creation of narrative flow. Narratives are groups of messages that together relate an experience or tell a story. This adds the problem of creating a narrative structure and consistent style to the data-mining exercise (for NLG work on the importance of narrative information exchange see e.g. (Reiter et al., 2008)). An example might be:

> I had my breakfast quickly because I was excited to go to the arcade. I got on the bus, I went to the arcade, I played in the arcade and won a cuddly bear.

Now consider that Alice and Charlie are joined by Bob, who is also an AAC user on the same system as Alice. Alice and Bob's devices are capable of sharing data at all levels. At the device level, Alice and Bob share raw data to confirm, for example, that their system clocks are in sync and that they have the same records of people who are in the same room. It is also possible at the social episode level that Alice's system can import utterances from Bob's system so that Alice could say 'Bob liked the swimming'. It is important to note that in this scenario, if Alice deletes an utterance from her machine 'The teacher gave me a bad mark on my work', Bob could still use the deleted story because Alice is unable to delete the disseminations. However, data sharing is not only between device users. Data sharing could also take place between the agencies involved in Alice and Bob's care and support. Figure 1 shows this data sharing taking place on three levels: device, individual AAC user, and institutional.

This scenario raises the issues of personal data flow control and indicates that controls for the flow of personal data have to be set at all three levels. Importantly, when personal data flows are managed at these levels responses will always be sociotechnical in nature; therefore they include technical responses, governance responses and technology practice responses. This is a familiar combination of re-

sponses in privacy management (Paine et al., 2007)

3 Finding traps and responding to them

Section 2 demonstrated a family of personal data use problems. These problems address various aspects of using personal data in the context of AAC devices. Our family of problems partly relate to the oft used definition of privacy from Westin (1967): "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others". This is not a hierarchy with a root problem and a tree structure of related problems but a family of problems with complex relations and which are enmeshed rather than conceptually linear. Analytically, the family has four members: IPR; compliance responsibility; institutional personal data access and disclosure rights; and individual personal data access and disclosure rights. Each family member is addressed in this section. Whitley points out that Whitley (2009) "Wittgenstein (1956) tells us that language is a social activity and hence that specialised terms like privacy are arrived at socially.". The social construction of concepts related to personal data mean that personal data issues will be enmeshed in particular contexts and, as a result, the significance of these issues will vary from context to context.

Discussions with legal experts and practitioners (see Acknowledgements) revealed that responses to these problems occur at the institutional, individual, and technical levels, and an individual interpretation of personal data management issues underpins all responses. This is true of all personal data management issues; however, in the case of AAC users, the individual level will be a unique combination of AAC user, family, and care workers. At the next level is the sociocultural system in place within each institutional context, which contains the personal data management policies, procedures, practices and institutional values. This sociocultural system is supported by technological controls used to control personal data information flow.

Unlike spoken conversation, AAC devices create embodiments of conversations that can be permanently stored or logged. Then conversations become data that largely focuses on living individuals, either the users themselves or their family and friends. Under certain conditions processing this data will be regulated by data protection legislation. In other settings the processing will be governed more by social norms. Furthermore, the permanent nature of these embodiments means that they can carry copyright. Then there is a natural question of information flow control, the need for rights management and traditional information security issues such as confidentiality and access control. However, privacy is also an elastic concept (Allen, 1988) and is often considered wider than the Westin definition, including aspects of identity and relationship management. As the work of Smith (2005) and Rackensperger et al. (2005) shows, use of AAC devices is related to notions of self and relationship to others. The notions of self and relationship to others are a central aspect of privacy (Kani-Zabihi and Coles-Kemp, 2010) (Barnard-Wills and Ashenden, 2010) and the link between personal data use and privacy and identity issues should not be ignored when considering these personal data management problems.

3.1 A Family of Personal Data Use Problems

Practically, any technical, regulatory, or social response to personal data use issues in AAC must be able to operate in a range of contexts and support a user as they blend and adjust contexts. For example, an AAC user may use their device at home, in formal education, in youth group activities, and in social settings. These different contexts may include many of the same people, but the personal data control requirements and the regulatory and personal data management frameworks are likely to differ from context to context. A further level of complexity in the case of AAC users is that capabilities and backgrounds differ widely within the AAC community (DeRuyter et al., 2007) and any personal data management approach has to adjust to these varying capabilities and different perspectives.

Technical responses would primarily be formed by meshing the AAC functionality into the underlying technical architecture of each device. Technical responses include personal information flow control over the network; encryption of sensitive utterances, e.g. health or financial information (such as credit card numbers), stored on the AAC device; access control to location databases and so on.

3.1.1 Management of IPR

Automatically generated text in AAC devices can be coupled or merged with input from other sources, increasing the ability of users to develop additional novel utterances. Given the digital nature of the utterances, there is potentially a close comparison with music copyright, which has three sets of rights: mechanical, rights related to the creation of the lyrics and performing. Using music copyright as the parallel, consider the situation where an AAC user, Alice say, imports text from a novel under copyright (mechanical rights) and adapts it by adding other text and other copyright material in order to create her own monologue (intellectual property rights). Another AAC user, Bob say, then downloads Alice's monologue and performs it through his AAC device at a concert for a school (performing rights). Clearly the majority of instances carry an implicit rights clearance, particularly in the content and performing rights elements of this example. However, if the monologue was posted on YouTube and then became sampled by a recording artist or made into a digital novel, rights clearance may not apply. Communicating the rules relating to copyright and ensuring understanding can be problematic.

Social and institutional responses to IPR problems are largely related to awareness training and the agreement of 'ground rules' or social contracts in communities such as schools and youth clubs where the legal issues and social expectations are made clear. The traditional methods for negotiating and agreeing ground rules is heavily based on the use of informational literature, one-to-one and group discussion (Barnard-Wills and Ashenden, 2010). These methods do not translate well into an environment where users may have cognitive development issues, or may be non-literate. It could be envisaged that guardians and parents would be used to negotiate and agree the ground rules and then left with the task of communicating the ground rules to their dependents. The difficulty in this is that at the same time, AAC users can become very skilled in the use of technology and may well develop practices that involve copyright material, in a way that their guardians have not been able to communicate effectively. In order to respond to this mismatch of capabilities, methods of engagement need to be sought that ensure AAC users are as integral as possible to the establishment of such rules.

3.1.2 Management of compliance responsibility

Due to the digital nature of AAC utterances, personal data output by a device is regulated by data protection legislation when being processed in the context of institutions such as schools, health, or social service. In the UK, this legislation is the Data Protection Act 1998. Under the Act there are eight principles of personal data management and the requirement that there must be a *data controller* who is responsible for compliance with the legislation. The term 'data subject' denotes individuals to whom the personal data relates. If Alice and Bob were young adults with sufficient cognitive abilities they would likely be the data controllers. However, as speech, language and communication disabilities are regularly a pan-disability, Alice and Bob may also be cognitively impaired and a parent or guardian is likely to be regarded as the data controller.

Typically, the mechanism for specifying compliance requirements is via the creation of a compliance schedule. In the case of AAC use, a compliance schedule for AAC devices is likely to be between the institution (school or health services) and the parents. The compliance schedule would establish the responsibility for data processing and agree the relationship between parents and institutions. Note that the AAC user's capabilities for technology can potentially exceed that of their guardians and parents. The relationship the AAC user has to the technology is quite possibly very different from that of the parent or guardian. If effective compliance management is to be achieved, new engagement methods need to be sought to ensure that AAC users are actively engaged in the establishment of compliance schedules. The connection between the individual, the institution (school) and privacy legislation is illustrated in Figure 2.

3.1.3 Management of institutional personal data access and disclosure rights

A set of rights must be agreed as part of the compliance schedule when AAC devices are used in school and healthcare settings. Many AAC devices have their data backed up to a central database. An

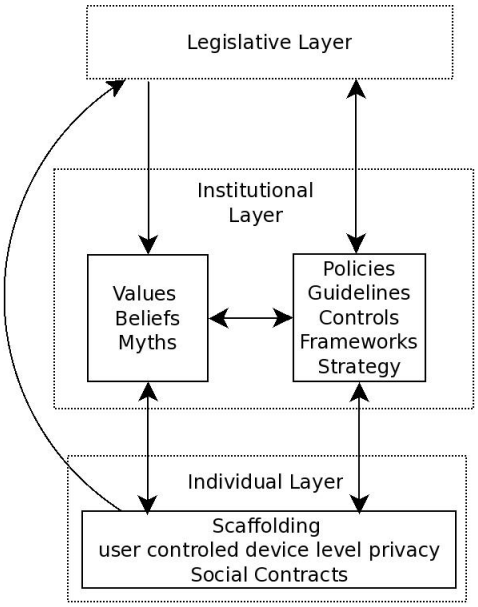

Figure 2: This diagram adapts the characterisation of institutional culture found in (Allaire and Firsirotu, 1984)

issue arises as to who has the right to back up or access the AAC data. AAC devices that can restore data that a user has deleted raise particular problems, which relate to who has the right to restore the data and the subsequent disclosure rights that this individual would have. Problems also occur as to whether other AAC users have the right to download content from another AAC device and the subsequent disclosure rights that this would afford.

AAC users will typically have considerable intervention from education and health support workers. Unlike spoken forms of conversations in other care situations, AAC utterances have a digital embodiment. This allows different teams in the care, education and support of the user to easily share utterances, and it may be deemed to make care more effective to do so. From an institutional perspective, data sharing policies should be set up to state which aspects of AAC utterances can be shared, the people that such utterances can be shared with, and a need for transparency in the logic used to interpret the utterances. In addition, the compliance schedule could specify which data transfers are permitted.

3.1.4 Management of individual personal data access and disclosure rights

Whilst many institutional issues are related to personal data use, importantly, AAC users are likely to use devices for personal data disclosure outside of the institutional context as part of family and social life. In this instance processing is controlled by social norms and practices that could be considered a social contract (Milne and Gordon, 1993).

From a social perspective, developing social contracts or norms organically responds to problems related to publishing of data about other data subjects, misuse of the AAC user's personal data by friends and family, and unintentional copyright infringements. In the scenario of AAC use, these social contracts and norms are re-enforced with education and awareness briefings (Bogdanovic et al., 2009) that are typically driven by the education and health institutions. As part of these ground rules, the degree of anonymity in any disclosures and the rights of non-AAC users to have their personal data removed from an AAC device are agreed or follow a socially accepted path. From a technical perspective, the device interface could be developed to include utterances about information disclosure and feelings of privacy. The log files could also include information disclosure and processing comments that practitioners and family members might wish to discuss or consider. Role play games could also be considered as a way of re-enforcing and encouraging ground rules.

4 AAC personal data management framework

As illustrated in Figure 2, personal data management within the AAC context is complex and any response to a personal data management problem has both technical, governance and cultural aspects. These responses have to be adaptive to differing levels of capabilities and different contextual requirements. Any technical response has to be scalable to enable users with different privacy and technical requirements to have access to their personal data controlled accordingly so that, where practical, users are able to have some control over their personal data. This scaleability can, in part, be addressed by the design and implementation of the personal data management framework.

4.1 Extending the existing framework

The personal data management problems related to AAC use have links with work in the mainstream privacy and consent research communities. Section 2 illustrates that AAC use adds additional layers of complexity to privacy and consent issues and, as a result, adds additional requirements to any personal data management framework. Due to space constraints the factors are merely highlighted to note that each is a large piece of research in its own right.

The required extensions fall into three areas:

4.1.1 Technical capability

Technical capability, in addition to education, is a factor in assessing ability to manage privacy (Coles-Kemp et al., 2010; Kumaraguru and Cranor, 2005; Buchanan et al., 2007) because a relatively sophisticated level of technical capability is required to implement the privacy controls. Technical capability and education levels are likely to be lower, on average, in AAC users.

4.1.2 Family roles

AAC users typically remain 'scaffolded' by family and the family will therefore remain involved in decisions about personal data disclosure. Whilst, family plays an important role at the start of an individual's internet journey[5], typically this intervention recedes over time and the design of privacy controls does not traditionally cater for varying levels of user independence in decision making. This needs to be addressed by the management framework.

4.1.3 Governance system design

Responses to personal data management issues use a governance system composed of policy, regulation, and practices to support the use of privacy enhancing technologies. Engagement with such a system is notoriously inconsistent because of language and conceptual complexities (Kani-Zabihi and Coles-Kemp, 2010; Bogdanovic et al., 2009; Bonnici and Coles-Kemp, 2010; McDonald and Cranor, 2008; McDonald and Cranor, 2009). It is reasonable to assume that such a governance system would require specific modifications for the AAC community to make policies more understandable, to allow for adaptations in privacy and internet safety education and to enable the role of family in decision support. However, it should also be kept in mind that similar modifications could be made for other communities with lower levels of digital literacy, literacy and cognitive challenges. Whilst the problems themselves are AAC-specific and the problems are brought about, in part, by the direction of development of AAC technology, the governance responses respond to underlying problems found in a range of communities.

[5]UK online Centres (2010) "Digital engagement understanding customers", a study (available for download at www.ukonlinecentres.com/research/research/centres-research)

5 Conclusions

Advances in text-to-speech technology and mobile computing have made a range of AAC devices available to the public. Advances in natural language generation and speech processing techniques have co-incided with changes to the commercial landscape to bring dramatic advances in AAC capabilities within reach. These advances in AAC design, though overwhelmingly positive, do result in a family of personal data use problems that were not encountered with previous generations of the devices. This paper argued that AAC devices can only significantly support users with communication difficulties to achieve greater independence and social inclusion if their design and implementation both addresses and identifies personal data problems.

Acknowledgments

The authors wish to thank the staff of the HWST project, the Natural Language Generation Group at Aberdeen and the Sociotechnical Group, within the Information Security Group at Royal Holloway. The work was developed with input from legal experts, youth work practitioners, and disability officers: in particular Robert Carolina, Amanda Gerry, and Karen Wood are gratefully acknowledged. Similarly, the insights of Nava Tintarev, Sarah Moffat, and Margaret Mitchell made this work possible. This paper was produced in collaboration with the Visualisation and Other Methods of Expression (VOME) project, which is supported by the Technology Strategy Board; the Engineering and Physical Sciences Research Council and the Economic and Social Research Council [grant number EP/G00255X/1].

References

Y. Allaire and M.E. Firsirotu. 1984. Theories of organizational culture. *Organization studies*, 5(3):193.

A.L. Allen. 1988. *Uneasy access: Privacy for women in a free society*. Rowman & Littlefield Pub Inc.

S. Balandin, N. Berg, and A. Waller. 2006. Assessing the loneliness of older people with cerebral palsy. *Disability & Rehabilitation*, 28(8):469–479.

D. Barnard-Wills and D. Ashenden. 2010. Public sector engagement with online identity management. *Identity in the Information Society*, pages 1–18.

DR Beukelman and P. Mirenda. 2005. *Augmentative and alternative communication: Supporting children and adults with complex communication needs3rd ed*. Paul H. Brookes, Baltimore, MD.

R. Black, J. Reddington, E. Reiter, N. Tintarev, and A. Waller. 2010. Using NLG and sensors to support personal narrative for children with complex communication needs. In *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies*, pages 1–9, Los Angeles, California, June. Association for Computational Linguistics.

D. Bogdanovic, C. Crawford, and L. Coles-Kemp. 2009. The need for enhanced privacy and consent dialogues. *Information Security Technical Report*, 14(3):167–172.

C.J. Bonnici and L. Coles-Kemp. 2010. Principled Electronic Consent Management: A Preliminary Research Framework. In *2010 International Conference on Emerging Security Technologies*, pages 119–123. IEEE.

T. Buchanan, C. Paine, A.N. Joinson, and U.D. Reips. 2007. Development of measures of online privacy concern and protection for use on the internet. *Journal of the American Society for Information Science and Technology*, 58(2):157–165.

L. Coles-Kemp and E. Kani-Zabihi. 2010. On-line privacy and consent: a dialogue, not a monologue. In *Proceedings of the 2010 workshop on New security paradigms*, pages 95–106. ACM.

L. Coles-Kemp, Y.L. Lai, M. Ford, and C. Hyperion. 2010. Privacy on the Internet: Attitudes and Behaviours.

J. Cornwell, I. Fette, G. Hsieh, M. Prabaker, J. Rao, K. Tang, K. Vaniea, L. Bauer, L. Cranor, J. Hong, et al. 2007. User-controllable security and privacy for pervasive computing. In *Mobile Computing Systems and Applications, 2007. HotMobile 2007. Eighth IEEE Workshop on*, pages 14–19. IEEE.

F. DeRuyter, D. McNaughton, K. Caves, D.N. Bryen, and M.B. Williams. 2007. Enhancing AAC connections with the world. *Augmentative and Alternative Communication*, 23(3):258–270.

E. Dominowska, D. Roy, and R. Patel. 2002. An adaptive context-sensitive communication aid. In *Proceedings of the 17th Annual International Conference Technology and Persons with Disabilities*.

P. Golle, F. McSherry, and I. Mironov. 2008. Data collection with self-enforcing privacy. *ACM Transactions on Information and System Security (TISSEC)*, 12(2):1–24.

D. J. Higginbotham, H.Shane, S. Russell, and K.Caves. 2007. Access to AAC: Present, past, and future. *Augmentative and Alternative Communication*, 23(3):243–257.

M.A. Kamp, P. Slotty, S. Sarikaya-Seiwert, H.J. Steiger, and D. Hanggi. Traumatic brain injuries in illustrated literature: experience from a series of over 700 head injuries in the asterix comic books. *Acta Neurochirurgica*, pages 1–5.

E. Kani-Zabihi and L. Coles-Kemp. 2010. Service Users Requirements for Tools to Support Effective On-line Privacy and Consent Practices. In *Procedings of the 15th Conference on Secure IT Systems, Nordic 2010*.

C.M. Karat, C. Brodie, and J. Karat. 2006. Usable privacy and security for personal information management. *Communications of the ACM*, 49(1):56–57.

P. Kumaraguru and L.F. Cranor. 2005. Privacy indexes: A survey of westins studies. *Institute for Software Research International*.

G.W. Lesher, G.J. Rinkus, B.J. Moulton, and D.J. Higginbotham. 2000. Logging and analysis of augmentative communication. In *Proceedings of the RESNA Annual Conference*. Citeseer.

A.M. McDonald and L.F. Cranor. 2008. The cost of reading privacy policies. *ACM Transactions on Computer-Human Interaction*, 4(3):1–22.

A. McDonald and L. Cranor. 2009. An empirical study of how people perceive online behavioral advertising.

S.J. Milberg, H.J. Smith, and S.J. Burke. 2000. Information privacy: Corporate management and national regulation. *Organization Science*, pages 35–57.

G.R. Milne and M.E. Gordon. 1993. Direct mail privacy-efficiency trade-offs within an implied social contract framework. *Journal of Public Policy & Marketing*, 12(2):206–215.

H. Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford Law Books.

C. Paine, U.D. Reips, S. Stieger, A. Joinson, and T. Buchanan. 2007. Internet users' perceptions of privacy concerns' and privacy actions'. *International Journal of Human-Computer Studies*, 65(6):526–536.

R. Patel and R. Radhakrishnan. 2007. Enhancing Access to Situational Vocabulary by Leveraging Geographic Context. *Assistive Technology Outcomes and Benefits*, page 99.

L. Pennington, J. Marshall, and J. Goldbart. 2007. Describing participants in AAC research and their communicative environments: Guidelines for research and practice. *Disability & Rehabilitation*, 29(7):521–535.

T. Rackensperger, C. Krezman, D. Mcnaughton, M.B. Williams, and K. D'silva. 2005. When I first got it, I wanted to throw it off a cliff: The challenges and benefits of learning AAC technologies as described by adults who use AAC. *Augmentative and Alternative Communication*, 21(3):165–186.

J. Reddington and N. Tintarev. 2011. Automatically generating stories from sensor data. In *Proceedings of the 15th international conference on Intelligent user interfaces*, pages 407–410. ACM.

S. Reilly, J. Douglas, and J. Oates. 2004. *Evidence-based practice in speech pathology*. Whurr, London.

E. Reiter, F. Portet A. Gatt, and M. van der Meulen. 2008. The importance of narrative and other lessons from an evaluation of an NLG system that summarises clinical data. In *International Natural Language Generation Conference (INLG)*, pages 147–156.

E. Reiter, R. Turner, N. Alm, R. Black, M. Dempster, and A. Waller. 2009. Using NLG to help language-impaired users tell stories and participate in social dialogues. In *European Workshop on Natural Language Generation (ENLG-09)*.

M.M. Smith. 2005. The dual challenges of aided communication and adolescence. *Augmentative and Alternative Communication*, 21(1):67–79.

D.J. Solove. 2008. *Understanding privacy*. Harvard university press.

G. Soto, E. Hartmann, and D. Wilkins. 2006. Exploring the elements of narrative that emerge in the interactions between an 8-year-old child who uses an AAC device and her teacher. *Augmentative and Alternative Communication*, 22(4):231–241.

S. Spiekermann and L.F. Cranor. 2009. Engineering privacy. *Software Engineering, IEEE Transactions on*, 35(1):67–82.

J. Todman, N. Alm, J. Higginbotham, and P. File. 2008. Whole utterance approaches in AAC. *Augmentative and Alternative Communication*, 24(3):235–254.

A.F. Westin. 1967. *Privacy and freedom*, volume 97. London.

E.A. Whitley. 2009. Informational privacy, consent and the. *Information security technical report*, 14(3):154–159.

L. Wittgenstein. 1956. Philosophical investigations.(trans. GEM Anscombe) Basil Blackwell.