SmartData:

The Need, the Goal, the Challenge

George J. Tomko, Ph.D.
Hon Kwan, Ph.D.
Don Borrett, M.D.

March 2012

SmartData: The Need, the Goal, the Challenge

New scientific discoveries and innovations in technology are the lifeblood of society's well-being and prosperity. This lifeblood depends on a continued political and cultural context of freedom, which like oxygen, supplies the energy for innovation. It is not coincidental that the western world, which enjoys the most political freedoms, has also become the most innovative and prosperous. However, the pillars of freedom, which include respect for privacy and protection of individual and property rights, are being jeopardized in the pursuit of public safety against potential terrorist threats. In this pursuit, society is experiencing a greater expansion of electronic surveillance, increased misuse of personal information, and the concomitant erosion of civil liberties. In response to terrorism, mankind has excelled at developing technologies of surveillance against an ever-expanding list of security threats identified by governments around the world. However, as these technologies become more sophisticated and incorporate recent advances in artificial intelligence, they will become a threat not only to our nation's enemies, but also to its citizens, in the potential loss of our freedoms.

In the private sector, corporations are also collecting more personal data about their customers in their quest for additional revenues and marketing advantage. This fact alone poses privacy challenges, but now governments are "encouraging" corporations to share their customers' personal data, that they have obtained through business-related transactions, in order to build personal profiles and identify potential terrorists or criminals. Governments are promulgating a widespread view that in order to protect citizens against the new threats of the 21st century, some individual freedoms must be relinquished – freedoms such as civil liberties and the right to privacy. They espouse a zero-sum paradigm wherein public safety may only be protected at the expense of our freedoms, especially privacy.

Not only do we believe that this view is flawed, but it is dangerous. It is especially flawed in that it demonstrates a fundamental ignorance of technology. Whereas adopting a zero-sum paradigm invariably leads to ongoing reductions in privacy when pitted against the need for security, the opposite may also occur - technology can indeed be designed to provide a positive-sum outcome – the basis of *Privacy by Design*. By that we mean building into technology the capability to achieve multiple functionalities – public safety and privacy, or using personal data within the constraints of privacy such that both businesses and users may benefit. Adopting a zero-sum paradigm is dangerous because curtailment of our privacy and freedoms will ultimately stifle innovation, lead to mistrust and fear of our governments and corporations, and diminish the prosperity of our society. But we believe there is a better way. In the spirit of *Privacy by Design*, that way is to use artificial intelligence to protect privacy and civil liberties – to build "SmartData," data that protects itself in a manner that is sensitive to the needs of the data subject (to whom the data relate), while enabling multiple functionalities

such as judicially authorized requests for data. But first, some background behind the need for the pursuit of a SmartData strategy.

The original Internet was one-dimensional in that it only processed text. With the introduction of the World Wide Web, a second dimension was added, allowing both text and images to be shared. Recently, Philip Rosedale, inventor of Second Life, made a compelling case that the current flat 2-D internet will be transformed into a 3-D virtual world wherein text and images will only form a subset of the total cyber-environment. He argues that since human beings evolved in a 3-D world and are by nature social animals, a corresponding virtual world would allow more familiar, efficient and social ways of exchanging information. However, there is another aspect. Up to now, users have always been "external' to the Web -- on the "outside," looking in. We interface with the Web directly through a keyboard or via a computer programmed to carry out our instructions. A similar situation exists in current 3-D virtual worlds such as Second Life where avatars are, for the most part, directed by the user or a computer-surrogate on the outside, in the "real world." But this is changing - getting "inside" the Web has already started with the introduction of agents such as viruses, worms, cookies, and Trojan horses, although mainly for malicious purposes. However, these agents are not autonomous. They are essentially "dumb" in that they can only take actions based on previously programmed code. Although they have no agency, *per se*, what is important is that the direction of "agents" moving "inside" the virtual world has already begun.

We believe that the next evolution in the internet will be the introduction of intelligent, embodied agents within 3-D virtual worlds. In turn, these agents will be connected to the digital cloud and have access to a global network of data. These agents, we predict, will become our acting-surrogates, thus producing more productive way of exchanging and processing information. The 3-D virtual internet has the potential to inspire totally new innovations, as did the flat web. However, the need for privacy and security in such a cloud-based virtual environment will be enormous. Although a full blown 3-D virtual internet is still in the future, the introduction of cloud computing has already exacerbated the difficulty of securing privacy and protecting personal data in cyber-environments, especially when there are competing forces interested in accessing personal information. Governments, public officials and businesses seek unfettered access to such data, for a variety of purposes. On the other hand, consumers only generally wish to divulge their personal information for specific purposes, after which they want the data destroyed. But the difficulty with current data protection schemes, caught in the tug-of-war of competing interests, is that once the data is in plain digital text, it can be easily copied and disclosed, against the expressed wishes of the data subject. Personal information, once released for a singular purpose, may now become lost forever in the "cloud-based virtual worlds," potentially subject to unending secondary uses.

These difficulties, although tempered by regulatory policies and legislation, can never be completely surmounted because their source arises from the way in which data has existed since the advent of digital databases - as passive in nature, merely bits and bytes on a storage device. At its core, this is the precise problem we are facing: the personal information of an individual, as represented by a binary string of data residing in the cloud, is not capable of protecting itself against unauthorized, secondary uses. In an attempt to overcome likely infringements, a tangled web of international legal provisions and commercial contracts has been established to address the various privacy and proprietary concerns of different jurisdictions. However, the prospect of what we face is not encouraging - not only a global, legal bureaucratic nightmare but also a technical morass of different systems and standards, all trying to interface with each other. Moreover, all of this is overshadowed by the nightmarish prospect of heavy-handed governments motivated by Orwellian "good" intentions, infringing on our privacy and freedoms.

While no system can solve all of these issues, we propose that the objective of transforming personal data from a passive string of bits into an “active” form capable of protecting itself will circumvent many of the legal and technological issues. A potential benefit is that the regulatory framework and legal structures between parties need no longer be the first line of defense: they will be transformed into serving as the backstop, in the same way that commercial establishments treat criminal and tort laws against theft of merchandise as a secondary line of defense – with the primary line of defense being “technological”: a secure building, the installation of anti-theft systems, the presence of security staff, guard dogs, and so forth. Unless we are able to solve the privacy, security and public safety problems in a digitally connected world in an analogous technological manner that satisfies users, businesses and governments, the innovations themselves may be curtailed since users will not trust the systems, and businesses may refrain from using them. Or far worse, society will creep toward an authoritarian hell along a road that is paved with the seductive good intentions of greater public safety.

Accordingly, the purpose of *Privacy by Design* is to proactively instantiate Fair Information Practices into the core of all data-related functions or services on the Web. To accomplish this, we want to first build a computational foundation for agents to learn how to protect personal information in conjunction with security, and bind it to “data purpose.” We then want to use this same computational foundation to expand into other data-related domains such as search, medical diagnostics, electronic health records, social networks and so on, such that all of these new data-related fields have privacy incorporated into their core processing - the essence of Privacy by Design. Since contextual processing is mandatory for effective privacy protection, as well as in other data-related applications, we foresee that it could serve as a platform technology for expansion into these other areas.

The question before us, then, is: How do we use artificial intelligence to orchestrate data to protect itself, yet provide for the multi-functionalities required for the greater good of society? The path, we suggest, is to build artificial agents that act as intelligent custodians of personal information and have these agents serve as each individual's own personal online digital surrogate. By making the data part and parcel of an intelligent agent, in a manner similar to "privacy-aware" and concerned individuals, personal information would only be divulged when it was safe, appropriate to do so or judicially warranted. Furthermore, in cases where applicable, by only divulging personal information in an "analog" format and not as a plaintext digital string, the prospects of unauthorized sharing would be further decreased. These surrogate agents would store personal information in "memory" or in a cloud-based "locked vault" whereby only they (the agent and designated individuals) have the key. The goal is to develop SmartData to release personal information based on accepted privacy practices, user preferences, and experiences of previous requests and releases.

The three components necessary for achieving SmartData are: (1) securing personal data; (2) embedding data access rules (based on, for example, data subjects preferences, Fair Information Practices, local regulations and potential judicial warrants) "within the agent;" and (3) responding to requests for information contingent on its access rules, background/context and ongoing experience. This is the long-term vision of SmartData. However, this symposium will not focus on developing novel security techniques since there are adequate protocols for encryption and secure storage of data using existing methods. We will copy and implement those for time being. Instead, we will focus on creating agents that can respond to requests for information within an appropriate context and set of rules, and then make decisions based on that "enriched request." We believe that it is this contextually, enriched component which comprises the novel feature of SmartData. Accordingly, our *primary objective* in this symposium will be to discuss methodologies (equipment, software, evolutionary algorithms, neural net strategies, learning rules, etc.,) to develop contextual processing properties in artificial agents that will eventually have practical applications.

In order to appreciate the challenges in designing SmartData, we will look at a human model of the process of privacy protection which can be used across a wide spectrum of applications such as online social networks, electronic healthcare records and the internet in general. The first requirement is that the data subject has secure custody of her personal information, either in her "memory" or in a secured location. Assuming that this requirement is satisfied, the typical process comprises:

- The data subject receives a request for information from a second party;
- Although it may seem obvious in the case of a human, the data subject has to "know" that it has the information;
- The data subject makes a decision whether to release the requested information;

- That decision is based on a number of factors which include: a rule set which, in most cases is intuitive, resulting from education and experience; the identity of the requestor; the background/context of the request, e.g., she is applying for a passport or trying to join a social network; her attitude toward the requestor which itself is also a function of context, and her experiences with them such as what she may have heard or read about them; and her attitude in general with respect to releasing information about herself;
- As a result of that release, the data-subject may also at some point receive feedback about the success or failure of the release. For example, she may experience a great deal of spam just after the release, or no spam whatsoever. This will be registered as experience for future decisions.

The data-subject may also voluntarily post personal information or photos on a social media site, having certain expectations about who will be allowed to share this information. Here, the data subject is dependent on the policies of the social media website. In some cases, her expectations may conflict with the policies and practices of the web-site which may have been made explicit in the "fine print," but which she has never read. The goal of SmartData, in this scenario, is to effectively transfer control of the policies underlying the sharing of her personal information from the website back to the data subject. In effect, a plain digital copy of her personal photos, for example, would never be up-loaded to the website, only the binary string of data representing her SmartData, which houses and protects her personal information.

To set the stage for the requirements of an artificial agent, consider the situation where the data-subject has 40 items of personal information. These items could be grouped into categories such as, employment history, medical records, financial records, personal photos and so forth. These 40 items represent over one trillion (2^40-1) combinations of data that could be requested. Although in practice many combinations of items within the same category, such as medical tests, would be grouped together to decrease the total combinations, when viewed across the entire spectrum of personal data, the number of combinations would still be astronomical. Similarly, the agent has to take into account the identity of the requestor and background for the requests, e.g., a passport agency issuing a new passport, physicians in various fields such as neurology and psychiatry that are treating the data subject, also insurance companies, employers, government agencies, and so forth. Furthermore, the data subject whose personal information SmartData is protecting may have specific preferences or attitudes about the requestor of the information. Similarly, she may have general attitudes about sharing data which may influence the decision to release. So the job of SmartData would be: given one of these requests, Ri, in conjunction with the identity of the requestor, Ij, the background/context of the request, Bk and the agents "attitudes," Al, make a decision, Dijkl as

to the release of the personal data. Let's assume, at this time, that the decision is binary (release or don't release) versus release of partial data.

Thus far, we have only outlined a static situation based on "historical data" where new personal data is not added, and backgrounds and individual attitudes remain constant. However, privacy laws and regulations, personal information, backgrounds and those wanting access to data are constantly changing. In dynamic environments, either more items are added or new requestors or backgrounds come into existence. When things change because of a change in background, say a new law, somehow the agent has to update the decisions associated with the trillions of potential requests as they arrive. Similarly, if for example, new medical diagnostic tests are added, these new tests have to be integrated within the existing set of medical data. The new data may generate new combinations of requests requiring decisions as to the release. The attitudes of the individual about whom the data relate may also change over time as she has good or bad experiences (feedback) with previous releases of her personal information. Moreover, in real life, these decisions are not just algorithmic; they are also normative, and as such, dynamic. These are the properties of adaptive online processing that must be incorporated within the SmartData agent.

Therefore, there are five main challenges in our SmartData project:

(1) The computational burden of large and expanding domains which, in our model, equates to responding to the large number of different requests for information;

(2) The incorporation of backgrounds/contexts into the large number of potential request/decisions;

(3) Learning adaptive on-line responses to requests as a function of previous experiences. These responses should eventually incorporate privacy practices and the preferences and attitudes of the data subject;

(4) In a dynamic environment, the agent, based on its current knowledge and previous experiences, ultimately has to be able to learn, or adapt to, novel unstructured situations or environments on its own. As a result, different rules may have to be learned or domains of application would have to be modified;

(5) The binary string representing SmartData must be capable of being stored in the cloud and downloaded into relevant reconfigurable hardware so as to "activate" the agent to receive potential requests for data.

There is yet another factor that we have to consider during our symposium and the research that follows, which will bear on the eventual success of SmartData. We know that the practices

and procedures involved in safeguarding privacy and security on the Web are derived from individuals' concerns and solutions in the "real" world. These concerns and solutions are themselves derived from the social and cultural environments in which we live. Therefore, if an agent in cyberspace is to function autonomously and effectively it must first "understand" the specific social and cultural environment of humans within the domain in which it will operate. An "electronic healthcare agent" may not need to understand the social and cultural environment of professional basketball; it must, however, understand the environment within the domain of healthcare. Hence, as discussed in later sections, we posit that the agent must be evolved within a simulated virtual world that presents the relevant attributes of the domain in which it will operate so that the proper cognitive characteristics will be selected for the job. As we will explain later, one cannot just "program in" relevant contexts as has been the practice in standard AI when applied to narrow and static domains. Furthermore, SmartData agents must at some point in their evolutionary cycle inhabit a world with other agents in order to allow for inter-subjective cooperation and competition which, as has been demonstrated in our evolution, gives rise to particular social practices and cultures. These are formidable challenges, requiring considerable innovation - a great deal of scientific ground-breaking will have to occur. However, these very attributes position it among the most exciting research one could think of undertaking! And it comes with enormous payoffs – privacy and civil liberties for one, but also the myriad of innovative spin-offs of processing information contextually in a manner that is natural for biological agents.

This brings me to a definition of SmartData, the subject matter of this symposium: SmartData consist of Internet-based autonomous agents who act as a data subject's online surrogate, securely storing one's personal information, and intelligently disclosing it based upon the context of the data request, and in accordance with the user's instructions. Our approach (in the next section) is based on the premise that natural evolution remains the best roadmap available for building artificial agents that possess the property of contextual processing (what we mean by "context" will be more fully explained shortly). In effect, we are attempting to shrink the security perimeter from a mass of collective personal data stored in a database, down to a single individual's sphere of personal data. One's personal data will then be wrapped in a "cloak of intelligence" such that this entity, SmartData, becomes the individual's virtual proxy in cyberspace, controlling the release of their data. SmartData proactively builds privacy and security in, right from the outset, so that nothing is treated as an afterthought. It embodies a foundation of control and trust within the technology itself, incorporating the principles of purpose specification, personal consent, security, and use limitation. We believe that by incorporating the advances made in the technology of simulating virtual worlds, together with the ideas emerging from the field of evolutionary robotics and embodied cognition within a framework of dynamical systems, we can begin to make progress toward this ultimate goal. We believe it is well worth the effort!