Transparent Accountable Inferencing for Privacy Risk Management

Accepted for presentation at 2006 AAAI Spring Symposium "The Semantic Web meets eGovernment

D. J. Weitzner
MIT CSAIL DIG
djweitzner@csail.mit.edu

H. Abelson
MIT EECS & CSAIL DIG
hal@mit.edu

T. Berners-Lee MIT
CSAIL DIG
timbl@w3.org

C. P. Hanson
MIT CSAIL & CSAIL DIG
cph@csail.mit.edu

J. Hendler
Univ Maryland
hendler@umd.edu

L. Kagal
MIT CSAIL DIG
lkagal@csail.mit.edu

D. L. McGuinness
Stanford KSL
dlm@ksl.stanford.edu

G. J. Sussman
MIT EECS & CSAIL DIG
gjs@mit.edu

K. Krasnow Waterman MIT
Sloan Fellow; CSAIL DIG
kkw@sloan.mit.edu

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge MA 02139

Abstract

There is an urgent need for transparency and accountability for government use of large-scale data mining systems for law enforcement and national security purposes. We outline an information architecture for the Web that can provide transparent access to reasoning steps taken in the course of data mining, and accountability for use of personal information as measured by compliance with rules governing data usage. Legislative debates and judicial oversight will determine how large and how fast the expansion of data mining power available to homeland security and crime prevention efforts will be. Our approach to the privacy challenges posed by data mining is to concentrate on transparency and accountability in the *use* of personal information. As technology designers, we can help society feel more certain that data mining power is used only in legally-approved ways, and that the data that may give rise to adverse consequences for individuals is based on inferences that are derived from accurate data. We can meet these goals by making sure that the architecture of new Web technologies provides transparency into the inferencing mechanisms and creates technical means for assuring that government data mining efforts are accountable for improper use of data.

Transparency and Accountability Design Goals

Understanding how to incorporate transparency into decentralized information systems will be critical in helping society to manage the privacy risks that accrue from the explosive progress in communications, storage, and search technology. A prime example of a growing, decentralized information system is the World Wide Web, especially as it is augmented with structured data capabilities and enhanced reasoning power. As the Web gets better and better at storing and manipulating structured data it will become more like a vast global spreadsheet or database, than merely a medium for easy exchange and discovery of documents. Technologies such as XML, Web Services, grids, and the Semantic Web all contribute to this transformation of the Web. While this added structure increases inferencing power, it also leads to the need for far greater transparency and accountability of the inferencing process. Today, when an individual or an enterprise uses a single, self-contained set of data and applications, the controls necessary to assure accuracy and contextualize the results of queries or other analyses are available and generally well understood. But as we leave the well-bounded world of enterprise databases and enter the open, unbounded world of the Web, data users need a new class of tools to verify that the results they see are based on data, which is from trustworthy sources and is used according to agreed upon institutional and legal requirements. Hence, we must develop technical, legal and policy foundations for transparency and accountability of large-scale aggregation and inferencing across heterogeneous data sources.

Current investigations of the impact of data mining on privacy have generally focused on limiting access to data at the point of collection or storage. Much effort has been put into the application of cryptographic and statistical techniques to construct finely tuned access-limiting mechanisms. Yet for all this emphasis on access restriction, the reality is that the Web is making it increasingly difficult to limit access to data, while at the same time making it increasingly easy to aggregate data from multiple sources of information and to do searching and inferencing based on these aggregations. And so, in the long run, access restriction alone cannot suffice neither to protect privacy nor to ensure reliable conclusions. This must be augmented by increased transparency and accountability of the inferencing and aggregation process itself.

Addressing transparency and accountability requires a scalable framework for reasoning over applicable privacy policies in a Web-scale information space as well as public policy models that are suitably designed to regulate government and citizen behavior. On the technology front, we draw upon the Semantic Web effort which is laying the foundations for tagging, classifying and expressing logical relationships at Web scale. We combine this with technology for automated deduction and justification of

conclusions across large-scale databases and multiple reasoning systems. On the policy front, we explore analogues to current-day Fourth Amendment protections that consider not only access to information, but also the ways in which diverse information sources are aggregated and the ways in which implications are drawn. Technical design and eventual legal implementation of transparency and accountability will draw on technical expertise in reasoning systems and decentralized information architecture [WHBC05], along with perspective on the interaction of technical architecture and public policy.

The Transparency and Accountability Research Agenda

The research agenda to advance transparency and accountability has both technical and public policy components. In order to assess the utility of our technical approach, we analyze a hypothetical drawn from the public policy debate over the design of the proposed airline passenger screening systems (CAPPS, CAPPS II, and now Secure Flight). A critical privacy question raised regarding this particular data mining application is whether data collected in the course of assessing security risks can then be used for other law enforcement purposes. The hypothetical is designed to provide a variety of levels of complexity in proof-checking and to ensure that both positive and negative results can be identified.

Our hypothetical involves airline passenger John Doe who's Passenger Name Record (PNR) has been provided to the Transportation Security Administration (TSA) of the Department of Homeland Security as part of its test of the Secure Flight program. [SF01] TSA gathers additional information and tries to determine if he is a match against the same name in a list of known or suspected terrorists from the National Counterterrorism Center. In our hypothetical there is an incomplete PNR due to a computer error and the match is a "possible." The information is forwarded to the appropriate region's Joint Terrorism Task Force, which investigates further. [JT01] In the course of investigation, an Internet search reveals that he is the subject of an open warrant for unpaid child support. [NY01] John Doe is arrested under the federal "deadbeat dad" law. [CS01] While this hypothetical oversimplifies the actual process, it provides the necessary components for building and testing the proof-checking mechanisms.

Architectural Components

Our technical goal is to make the reasoning conducted over the Semantic Web transparent to the users as well as others with an interest in the legality of inferences, including institutions responsible for legal oversight (courts) an, in some cases, the subjects of those inferences. The Semantic Web will enable us to draw logical inferences over the information represented on the Web. In order to make use of these inferences, and in order to integrate these inferences into a larger set of personal, social, and legal expectations, we need to understand the basis of the inferences we see. Should we trust a given assertion (i.e, that John Doe may be a terrorist)? Do our friends believe it, too? Is it legal to use the inference for some specific purpose (i.e., to have begun an investigation of John Doe as a possible terrorist but to conclude by arresting him for unpaid child support)? Who's responsible if it's wrong or leads to some damaging action? To explore the more general question of how to instrument transparent reasoning at Web scale, we target an application which uses rule-based applications to combine data from multiple sources along with the application of privacy rules.
Figure 1 below illustrates how sources of information (bottom left) flow through a rules processing and reasoning system, mediated by policies derived from law and regulation (top right) resulting in explanations of inferences drawn for use by analyst/investigators, auditors and data subjects themselves (right side).

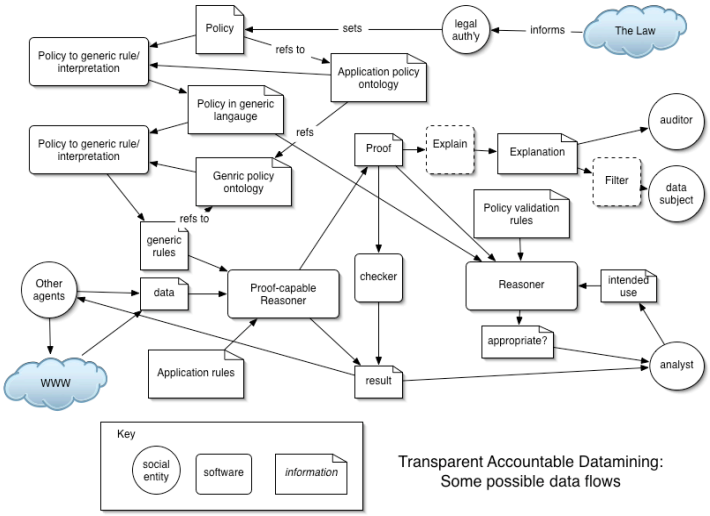

Figure 1. This figure illustrates possible data paths of an exemplar architecture.,

Currently, we are attempting to see if all of the required rules can be expressed using N3. Very preliminary success has been experienced in building the representations of the child support law, the lowest order logic system in our hypothetical. We will explore scenarios that entail expressions in higher order logics as well.

By developing prototypes that address some hypothetical scenarios, we believe we can provide a convincing demonstration that accountability with transparency can be formalized as an basic functional component of inferencing systems used for data mining. We expect to design and deploy the following architectural components to work on top of existing Semantic Web infrastructure. The modules we expect to develop are described below and pictured in figure 2.

Semantic Web Aware Truth Maintenance System

A Truth Maintenance System (TMS) is able to provide a justification for every inference made by the supported inference engine. The cwm reasoning engine attaches a provenance to each triple stored [BL00]. By integrating the cwm reasoning engine with the TMS, the resulting system would be able to respond, in real time, to user-driven hypotheses as to what conclusions could still be drawn if certain inappropriate sources of data were eliminated (i.e., could we still pursue John Doe as a "deadbeat dad" if we couldn't have singled him out for that purpose from out counterterrorism investigation?).

Proof Markup Language

The cwm reasoner will be modified to provide proofs in Stanford's Proof Markup Language (PML). PML provides an interlingua for representing and exchangine justifications. These PML proofs will be loaded into the Inference Web (IW) [McPi04], which is a framework for providing explanations for answers from Semantic Web agents. [McdS01]. It provides tools and services for checking, combining, summarizing, manipulating, and presenting different forms of explanation. As a result of this integration with IW, cwm will be able to generate PML proofs (using an underlying OWL representation) that can be browsed through the IW Browser, thus providing human accessible explanations.

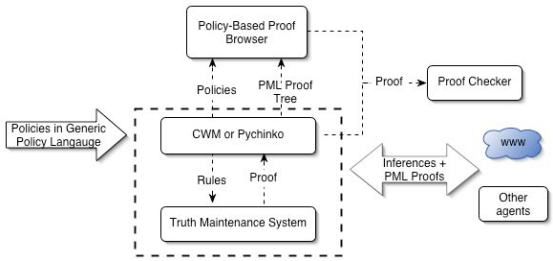

Figure 2. Policy-based Proof Browser.

Policies

Every entity will also include a set of policies that determine how it collects, uses, and reasons over data on the Semantic Web. In our context, we propose to identify laws that are applicable to data collection and usage and represent these laws as policies. We will use a rule-based approach [PrSa02] [Ko92] [JoSe93] [JoSe92] to model these privacy laws as policies. We envisage that in real scenarios, the policy itself is set by the end user and will be defined in a constrained way, from combinations of simple choices. From this, there is, effectively, a local application-specific ontology of policy options. These in turn are expressed in a policy language, for which we will use or adapt Rei [Ka04]. This language, by being constrained more than arbitrary rule languages, allows certain properties of the system to be reified by analysis of the policies. We will investigate the chosen language's suitability for the purposes of accountability and transparency and adapt and extend it as necessary.

Proof Checker

The Semantic Web is made up of entities that may or may not trust each other. When information is exchanged between trusting entities, it may be the case that no justification is required. However, if entities do not trust each other, or if they need to inspect assumptions or sources, all inferences made by an entity may need to be re-verified by entities with which it is interacting. Our infrastructure will support both trusting and distrusting entities. In order to support mutually distrusting entities, all information exchanges will include a proof tree to justify the information with respect to list of logical rules that were used (to the extent that they are shared between the agents or publicly available). The proof checker's function is simply to check the steps in the proof, including access to shared information, calculations on the data, and cryptographic signature verification. We may exploit a trust infrastructure such as IWTrust[ZPM05] that facilitates encodings and combinations of trust values.

Policy-based Proof Browser

The Policy-based Proof Browser (PPB) provides complete configurable transparency into how the system works. It will not be part of a deployed system and is only used to demonstrate the system. IW has a browser that displays the PML, provides different kinds of explanations for the proof tree, and allows level of magnification to be changed. PPB will provide additional functionality: it will (i) associate policies with different nodes of a proof tree; (ii) provide policy-based obscurity or transparency of a proof of a certain inference (i.e., show or hide appropriate nodes in the tree based on the policy); (iii) allow the backtracking through a proof tree to verify if the inference is legal; and (iv) display the proof tree and show sufficiency nodes (i.e., the nodes or premises that need to be proved false in order for the inference to be invalidated).

Conclusion

We believe that transparent reasoning will be important for a variety of applications on the Web of the future, including compliance with laws and assessing the trustworthiness of conclusions presented by reasoning agents such as search engines. Our particular focus is on using transparent inferencing to increase accountability for compliance with privacy laws. We also expect that this technical research will provide important guidance to policy makers who are considering how to fashion laws to address privacy challenges raised by data mining in homeland security contexts.

References

[BL00] T. Berners-Lee, "Cwm", http://www.w3.org/2000/10/swap/doc/cwm.html

[CS01] 18 U.S.C. § 288 ("Failure to pay legal child support obligations").

[JoSe92] A.J.I. Jones and M.J. Sergot. "Deontic logic in the representation of law: Towards a methodology." Artificial Intelligence and Law, 1:45--64, 1992.

[JoSe93] A.J.I. Jones and M.J. Sergot. "On the Characterization of Law and Computer Systems: The Normative Systems Perspective." In John-Jules Ch. Meyer and Roel J. Wieringa, editors, Deontic Logic in Computer Science: Normative System Specification, chapter 12, pages 275--307. John Wiley & Sons, Chichester, England, 1993.

[JT01] A Joint Terrorism Task Force is an FBI-led multi-agency effort , usually including state and federal officers and analysts; there are more than 100. Although not explicitly created by law, their authority appears to derive from 28 CFR Section 0.85 in which the Attorney General designates the FBI as the lead agency to investigate terrorism (see, http://www.fbi.gov/page2/dec04/jttf120114.htm).

[Ka04] L. Kagal, "A Policy-Based Approach to Governing Autonomous Behavior in Distributed Environments", Phd Thesis, University of Maryland Baltimore County, November 2004.

[Ko92] R. Kowalski, "Legislation as Logic Programs," In: Logic Programming in Action (eds. G. Comyn , N. E. Fuchs, M. J. Ratcliffe), Springer- Verlag, pages 203-230 (1992).

[McPi04] D. L. McGuinness and P. Pinheiro da Silva. Explaining Answers from the Semantic Web: The Inference Web Approach. Web Semantics: Science, Services and Agents on the World Wide Web Special issue: International Semantic Web Conference 2003 - Edited by K.Sycara and J.Mylopoulis. Volume 1, Issue 4. 2004.

[NY01 http://appsext4.dos.state.ny.us/csewarrants_public/cse_search

[PMF05] Paulo Pinheiro da Silva, Deborah L. McGuinness and Richard Fikes. A Proof Markup Language for Semantic Web Services. Information Systems, to appear.

[PrSa02] H.Prakken and G. Sartor, "The role of logic in computational models of legal argument: a critical survey. In A. Kakas and F. Sadri (eds.), Computational Logic: Logic Programming and Beyond. Essays In Honour of Robert A. Kowalski, Part II. Springer Lecture Notes in Computer Science 2048, Berlin 2002, 342-380.

[SF01] Federal Register: June 22, 2005 (Volume 70, Number 119), p. 3619, 3621 (System of Records Notice for Secure Flight, "Categories of Records in the System," subsection (a), describing the acquisition of Passenger Name Records (PNRs) in response to the Transportation Security Administration Order issued November 15, 2004 (69FR)65625)).

[WHBC05] Weitzner, Hendler, Berners-Lee, Connolly, Creating the Policy-Aware Web: Discretionary, Rules-based Access for the World Wide Web in Elena Ferrari and Bhavani Thuraisingham, editors, Web and Information Security. IOS Press, forthcoming.

[ZPM05] Ilya Zaihrayeu, Paulo Pinheiro da Silva and Deborah L. McGuinness. IWTrust: Improving User Trust in Answers from the Web. Proceedings of 3rd International Conference on Trust Management (iTrust2005), Springer, Rocquencourt, France, 2005.