Compliance with data protection laws using Hippocratic Database active enforcement and auditing

C. M. Johnson
T. W. A. Grandison

Governments worldwide are enacting data protection laws that restrict the disclosure and processing of personal information. These laws impose administrative and financial burdens on companies that manage personal information and may hinder the legitimate and valuable sharing and analysis of this information. In this paper we describe an integrated set of technologies, known as the Hippocratic Database (HDB), which enables compliance with security and privacy regulations without impeding the legitimate flow of information. HDB's Control Center allows companies to specify fine-grained disclosure policies based on the role of the user, the purpose of the access, the intended recipient, and other disclosure conditions. Its Active Enforcement component transparently enforces these policies by transforming user queries in a middleware layer to ensure that the database returns only policy-compliant information. HDB's Compliance Auditing system efficiently tracks all database accesses and allows auditors to formulate precise audit queries to monitor compliance with privacy and security policies. In this paper, we outline the basic architecture of the HDB solution, discuss the advantages of our approach, and illustrate the features of each component with practical compliance scenarios from the financial services industry.

INTRODUCTION

As private organizations and governments collect, store, and analyze massive amounts of personal information, individuals are increasingly vulnerable to privacy threats. In particular, individuals risk exposure to identity theft, reputation damage, and loss of personal privacy if their information is disseminated in violation of data protection laws or their own privacy preferences. Failure to safeguard this information, in addition to exposing companies to potential liability, may inhibit valuable information sharing or chill free expression, as many individuals may be reluctant to express controver-

sial opinions or reveal sensitive information about themselves without sufficient assurances of anonymity. Innovations in the gathering and analysis of personal information, such as identity resolution and data-mining technologies, while offering great benefits, also create significant challenges for privacy protection.[1]

Many countries have responded to these challenges by enacting laws that limit the processing and disclosure of personal information. Nevertheless, privacy breaches and identity theft continue to increase due to weak or ineffective enforcement of these data protection laws as well as discrepancies and conflicts in legal protections. Enforcement problems stem from: (1) the administrative costs of implementing privacy regulations, (2) the expense of acquiring new technology and reconfiguring applications to impose automated controls, and (3) the inability of existing information systems to enforce fine-grained disclosure policies reliably and efficiently. At the same time, varying constitutional standards and cultural attitudes toward privacy have resulted in conflicting data protection laws among different countries, posing impediments to the free flow of information in the global economy.

Technology can address many of these challenges by limiting the access and disclosure of sensitive personal information stored in automated systems. Such privacy solutions must be able to accommodate the intricacies of various data protection laws and unique individual preferences by discretely managing each information item. However, they must not unduly constrain the legitimate access or disclosure of information. Effective privacy solutions must also be economically and computationally efficient so that they can be incorporated into existing information systems without significant burden or disruption.

In this paper, we describe an integrated set of privacy technologies, known as the Hippocratic Database[2] (HDB), that seamlessly integrate fine-grained privacy controls at the data level without impeding legitimate information flow or significantly impacting system performance.

HDB's Active Enforcement[3] component transparently rewrites user queries in a layer above the database to enforce item-level access and disclosure policies. Most existing privacy solutions operate at the application level, requiring significant recoding for initial implementation and subsequent policy changes. Current database-level solutions typically restrict access to specific tables, rows, or columns, and do not accommodate the item-level semantics necessary to comply with many data protection laws. In contrast, HDB's fine-grained policy controls limit access and disclosure of specific cells in the database, accounting for user authorization privileges, the purpose of access, the intended recipients of the information, and the opt-in or opt-out choices of the individual.

HDB Compliance Auditing[4] provides accountability for policy compliance by tracking the circumstances of all past database access and ascertaining whether a particular query may have disclosed sensitive information. Other commercial auditing systems consume excessive storage overhead and degrade system performance, leading administrators to either turn off the auditing feature or regularly purge the audit logs.[5] HDB resolves these issues by logging only the queries and the updates to the database and deferring all computation until the time of audit. This contrasts with systems that log audit results and create indexes during normal query processing. HDB also provides a simple audit interface and a flexible audit query language, allowing auditors to declaratively specify, with item-level granularity, the disclosures they would like to investigate.

In the sections that follow, we briefly outline the requirements of information privacy laws in several countries, describe the basic architecture of the HDB enforcement and auditing components, and discuss how these technologies enable compliance with privacy laws and promote the responsible governance of sensitive information. We illustrate the functionality of HDB with scenarios from the financial services industry.

Brief survey of global information privacy laws

There have been a number of recent legal efforts around the world to protect the privacy and security of personal information. The European Union Data Protection Directive[6] is the broadest of these laws, setting forth cross-industry rules for data protection that must be followed by European Union member states. The Directive is significant in that it limits the processing, rather than simply the disclosure, of personal data. Generally, it requires data collectors to (1) notify individuals whose data they are

collecting, (2) allow the individuals to access and correct their personal data, (3) obtain opt-in consent from individuals before transferring any of their personal data to third parties, (4) obtain unambiguous opt-in consent from individuals before transferring any sensitive personal data (such as data that reveals race, ethnic origin, political opinions, religion, or beliefs), (5) safeguard any personal information under their control, (6) limit processing of personal data to the purposes for which it was collected or subsequently authorized, subject to limited exceptions, and (7) be accountable for all disclosures of personal data. The Directive also obligates member states to provide enforcement procedures and remedies for violations of data protection laws. Similar cross-industry privacy laws have been enacted in Canada,[7] Japan,[8] Australia,[9] and Argentina[10] to govern the collection, processing, and disclosure of personal information. These laws are inspired by the OECD (Organization for Economic Cooperation and Development) Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,[11] which recommend eight principles of privacy protection to advance the free flow of information and remove social and economic barriers among member countries.

The United States has taken a much different approach, regulating privacy protection in specific industry sectors such as health care and financial services. The Health Insurance Portability and Accountability Act[12] (HIPAA) requires health plans, health-care providers, and health-care clearinghouses (collectively known as "covered entities") to ensure the privacy and security of personally identifiable information about their patients. It obligates covered entities to notify patients of their privacy policies and provide them with an opportunity to opt out of certain disclosures of their personal information to third parties. In the absence of patient consent, HIPAA limits the recipients to whom personal information may be disclosed, the purposes for which it may be disclosed, and the conditions of disclosure. It also requires covered entities to reasonably safeguard personal information and to account for any unlawful disclosures upon the request of a patient or authorized government agency. Any patient information that has been de-identified to a statistically and scientifically acceptable level is not subject to HIPAA and may be processed and disclosed without restriction.

The privacy of financial data in the United States is governed by the Gramm-Leach-Bliley Act (GLBA),[13] which requires financial services companies to provide customers with notice of their privacy policies and an opportunity to opt out of any disclosures of their nonpublic personal information to unaffiliated third parties. However, GLBA does not protect information that is otherwise publicly available. It also allows companies to share any information with "affiliates" and requires customers to affirmatively exercise opt-out rights to avoid disclosure to nonaffiliated parties. In addition, the Fair Credit Reporting Act (FCRA)[14] requires credit reporting agencies to follow reasonable procedures to protect the confidentiality, accuracy, and relevance of credit information. Among other things, the FCRA allows consumers to opt out of disclosing their personal information for credit prescreening, marketing offers from affiliates, and other types of affiliate sharing. Some individual states have enacted more stringent financial privacy laws,[15,16] requiring companies to obtain opt-in consent from customers before disclosing their information to third parties. Other state statutes require notification to individuals in the event of a security breach. For example, California Civil Code Section 1798.82 requires any person or business doing business within the state to disclose any security breach of unencrypted personal information to any California resident whose information was (or is reasonably believed to have been) acquired by an unauthorized person.[17]

The demands of the global economy and the value of free information flow are encouraging countries to resolve legal differences regarding privacy and agree on common elements of data protection. For example, Article 25 of the European Union Data Protection Directive requires member states to restrict transfers of personal data to countries that guarantee an adequate level of privacy protection. To resolve conflicts of law between the United States and the European Union regarding the legality of cross-border data flow among companies, the United States Department of Commerce and the European Commission negotiated a Safe Harbor Agreement[18] that closely resembles the OECD Guidelines. It outlines seven privacy principles that a United States organization must follow in order to accept personal data transfers from the European Union. An organization must provide: (1) notice of the purpose for which it collects and uses informa-

tion, the types of third parties to whom it discloses the information, and the means by which individuals can limit its use and disclosure; (2) the opportunity to opt out of having information disclosed to third parties or used for purposes other than those for which it was collected; (3) assurance that transfer of personal information is limited to third parties that comply with the Safe Harbor Agreement privacy principles; (4) reasonable security to protect the information from loss, misuse, unauthorized access, disclosure, alteration, and destruction; (5) reasonable steps to assure that data integrity is maintained such that the data is reliable for its intended use, accurate, complete, and current; (6) access to an individual's own personal information and the ability to correct, amend, or delete any information that is inaccurate; and (7) enforcement mechanisms to provide recourse to individuals and penalties to organizations for violations of the principles. As the management of information becomes increasingly important to global commerce, it is likely that many industrialized countries will insist on minimum common privacy standards.

The common privacy principles enumerated in the OECD Guidelines and the Safe Harbor Agreement are the foundation of the Hippocratic Database.[2] The following sections describe how HDB technology protects information privacy without impacting legitimate business operations.

Financial services scenario

We describe the features of the Hippocratic Database with an example from financial services. Adam is a customer of BankCo in the United States. Upon opening his checking and savings accounts, Adam provides BankCo with various items of personal information necessary to confirm his identity. He also opens a credit card account with BankCo and provides additional information concerning his employment and income. BankCo continues to acquire information about Adam in the course of various banking and credit transactions.

HDB ACTIVE ENFORCEMENT

BankCo uses HDB Active Enforcement,[3] which operates as a middleware layer above the database to enforce fine-grained policies concerning the disclosure of information. It includes a policy creation interface, a preference negotiation mechanism, and a policy enforcement engine. The basic

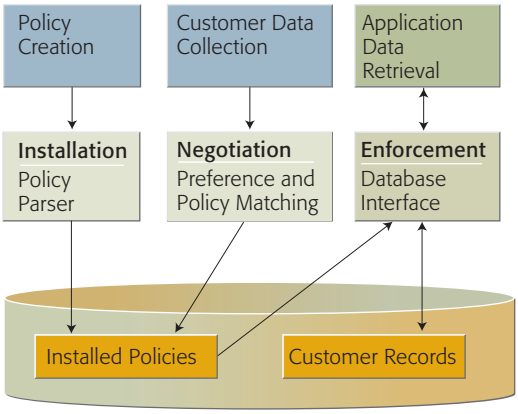

Figure 1
HDB Active Enforcement architecture

Active Enforcement architecture is depicted in *Figure 1*.

Policy creation

In accordance with GLBA,[13] BankCo has adopted a privacy policy detailing its collection, use, and disclosure of customer information. It also has strict internal policies governing each employee's specific authorization to access customer information. In the first stage of Active Enforcement, the HDB Control Center[19] allows BankCo to install its disclosure policies, including user or role authorizations, allowed purposes of access, and permissible recipients of information. For instance, a policy rule could specify that compliance officers may access customer Social Security numbers (SSNs) for identity verification and disclose this information to law enforcement agencies. Organizations express their policies through the Control Center, which installs the policies in the database in a form amenable to symbolic manipulation. HDB accommodates any subsequent changes to the policies and maintains copies of each version of the policies in database tables.[19] This allows BankCo to determine which policy was effective at the time of a particular disclosure or to "roll back" (i.e., revert) to previous versions of the policy.

Many privacy laws require organizations to consider purpose in determining whether data disclosure or use is appropriate. Under the European Union Data Protection Directive (and member state laws adopted thereunder), a European bank would be prohibited from processing Adam's personal data for any

Table 1 Query results showing HDB item-level enforcement

Customer Number	Name	Date of Birth	Address	Phone Number	Social Security Number	Income
1	Adam		Orinda, CA			
2	Brenda		Berkeley, CA	333-3333		
3	Charlotte					
4	David	01-09-70	Palo Alto, CA	111-1111		

purpose other than the purpose for which it was originally collected. Therefore, if a European bank collected Adam's annual income data for the purpose of determining his eligibility for financing, it would be unable to process that information for other purposes without his consent. HDB considers the purpose of each disclosure in determining whether a query complies with the applicable privacy policies. The current operational purpose can be inferred from the application requesting the data or directly specified by the user submitting the query.

Preference negotiation

In the second phase of HDB Active Enforcement, Adam is notified of BankCo's privacy policies through a Web interface. He is given the opportunity to specify his own privacy preferences through a simple Web browser plug-in, which communicates his preferences to the BankCo system in a privacy preference language, such as XPref.[20] HDB then compares BankCo's policy with Adam's preferences, advises him of any conflicts, and allows Adam to determine whether he would like to proceed with account registration. HDB then presents to Adam any opt-in or opt-out choices that are contained in BankCo's policy. In compliance with GLBA, BankCo advises him that it may share his nonpublic personal information with other financial companies for the purpose of offering him products and services, but it provides him the opportunity to opt out of these disclosures to unaffiliated third parties. BankCo's policy allows customers to opt out of such disclosures categorically or selectively. In our example, Adam opts out of disclosing his age, telephone number, SSN, and income for this purpose. He does not opt out of disclosing his address because he does not mind receiving targeted product and service offers in the mail. BankCo accepts Adam's opt-out choices and stores them in its database for enforcement purposes.

Application data retrieval

In the final stage, the HDB active enforcement engine intercepts and transforms queries to make them comply with company privacy policies and customer opt-in and opt-out choices. It then submits the rewritten queries to the database so that the application retrieves only policy-compliant results. In our financial scenario, suppose that MortageCo would like to purchase a customer list from BankCo to generate leads for its residential lending business. An alliance manager at BankCo submits a query requesting the names, ages, addresses, telephone numbers, and SSNs of all credit card customers with annual incomes over $150,000. Pursuant to GLBA, BankCo may not reveal this information for any customer who has exercised his or her opt-out rights.

BankCo's privacy policy also categorically prohibits revealing SSNs for marketing purposes. The active enforcement engine thus rewrites the manager's query to comply with the privacy policy and the customers' opt-out choices. The database processes the rewritten queries and replaces the prohibited cells with null values. As shown in *Table 1*, Adam's personal information is then filtered from the result, in accordance with his opt-out choices. In the second row, Brenda has opted out of disclosing her age, but agreed to disclose her address and telephone number. Charlotte has opted out of all disclosures of information to unaffiliated third parties, whereas David has not exercised his opt-out rights. All SSNs and incomes are filtered from the result, in compliance with BankCo's policy.

HDB also supports more complex policy conditions. For example, BankCo could implement a policy condition that prohibits anyone other than senior managers from accessing information on accounts with balances over $150,000 for marketing purposes, in case the bank would prefer to have more experienced staff interact with these customers. Such conditions can be installed through the Control Center and enforced with item-level granularity, similar to opt-in and opt-out preferences. In addition, the null values in Table 1 could be replaced by a pseudonym or other de-identified value, if the organization desires.

Advantages of Active Enforcement

HDB Active Enforcement has many distinct advantages over other privacy solutions. First, it operates at the middleware level, transparent to enterprise applications, so that no recoding of applications is necessary to implement policy controls, and all policy changes can be made through a single interface. Second, it handles all query processing (after policy predicates, i.e., clauses added to an SQL [Structured Query Language] query to apply the appropriate policy rules, are added) in the database, taking advantage of the performance optimizations of the database system. Thus, HDB enforcement is scalable and causes a minimal performance impact.[3] In fact, depending on the selectivity of the application (i.e., the ratio of records retrieved to the total number of records) and the choice selectivity of the customers (i.e., the ratio of subjects having not opted out to the total number of subjects), active enforcement may actually improve query processing speed by an order of magnitude.[3,19] In most circumstances, it performs better than application-level solutions, which typically retrieve all records sought by a query and then remove prohibited values from the result.[3] Third, HDB enforces disclosure policies down to the item (cell) level in the database, allowing an organization to enforce opt-in and opt-out choices, as shown in Table 1, without restricting access to an entire row or column. Fourth, HDB supports complex policy constructs, including user role, purpose, recipient, and other arbitrary conditions. For example, a user may be prohibited from accessing Adam's telephone number for marketing purposes, but allowed to access it for account service. Other systems make binary access privileges conditional on only the identity or role of the user. Finally, HDB can be seamlessly integrated[19] into any relational database environment with an SQL interface, such as ODBC (Open Database Connectivity) or JDBC** (Java** Database Connectivity).

HDB COMPLIANCE AUDITING

HDB Compliance Auditing[4] is a fine-grained database auditing system that enables organizations to monitor compliance with privacy laws and data disclosure policies. The audit application allows an administrator to specify an audit query to determine whether sensitive information has been disclosed. In response, HDB returns a list of suspicious user queries that have accessed the specified information, as well as the issuer, time, purpose, and recipient for each query. It also reconstructs past database states to reveal the exact information returned in response to each query. *Figure 2* illustrates the basic architecture of the system.

HDB Compliance Auditing offers efficient performance and low overhead because it relies largely on existing database infrastructure and defers computation until the time of the audit. The HDB logging system stores all data updates, insertions, and deletions in backlog tables, which are populated by using database triggers (i.e., procedural code that is automatically executed in response to certain events on a particular table in a database). Alternatively, existing replication logs or point-in-time query features[21] can be used in place of the backlog tables. HDB also records all queries and relevant contextual information (i.e., user identity, time, purpose, recipient) in query logs. Upon receiving an audit query, the system performs a static analysis of the query logs and generates a list of suspicious queries. A query is suspicious if it shares an indispensable tuple (i.e., row) with the audit query. A tuple is indispensable for a given query if its omission makes a difference in the results of the query. HDB then combines these suspicious queries into a single SQL audit query, which it runs against the backlog tables. This SQL audit query identifies all queries that accessed the information specified in the audit query. For each of these queries, the audit application outputs the query string, user identity, time, purpose, recipient, and exact information disclosed in response to the query.

Auditing promotes enforcement

HDB Compliance Auditing provides a reliable and efficient tool to aid in the enforcement of data protection laws and to make organizations accountable for their management of personal infor-

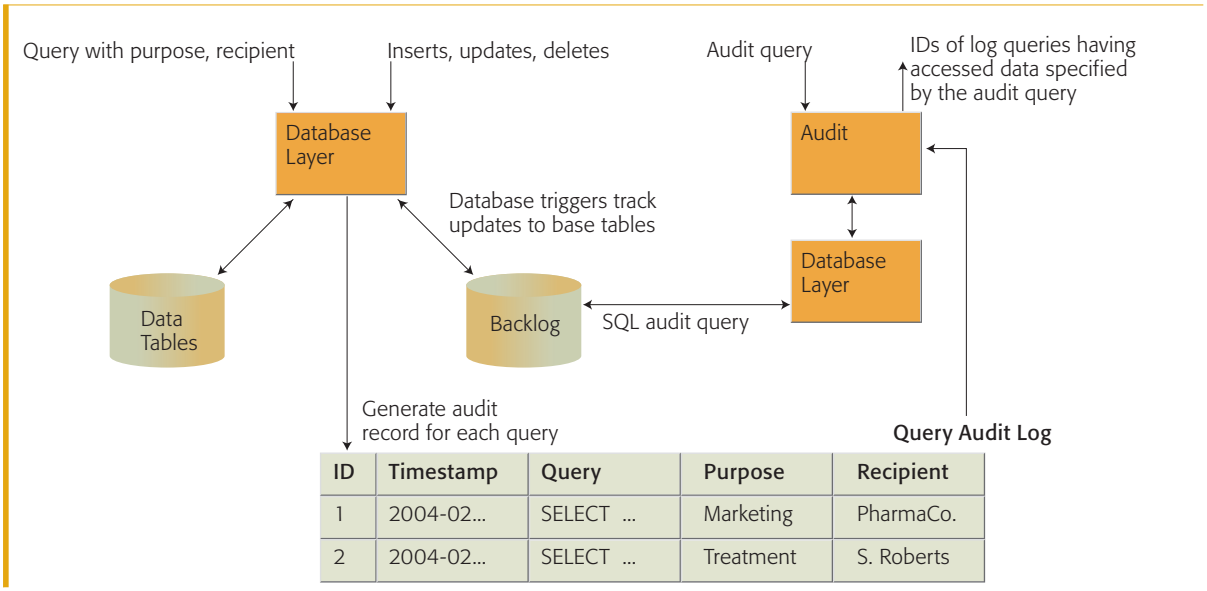

Figure 2
HDB Compliance Auditing architecture

mation. The European Union Data Protection Directive[6] and the Safe Harbor Agreement[18] require governments and organizations to provide individuals with legal recourse in order to ensure compliance with data protection requirements. In the United States, the HIPAA Privacy Rule[22] requires covered entities to provide a written accounting of certain disclosures of personal information upon the request of a patient or authorized government agency. GLBA does not have a specific auditing requirement, but the ability to trace the access and disclosure of personal data is essential in maintaining an effective compliance program. HDB auditing allows auditors to investigate customer complaints and take corrective action, if necessary. It can also serve as a deterrent to unauthorized access and unlawful disclosures if employees are aware that their database usage is subject to review. In addition, organizations can use HDB auditing to proactively investigate past database access to uncover any suspicious users or abusive practices.

Audit scenario

In a continuation of our scenario, Adam receives an unsolicited telephone call from MortgageCo offering a competitive rate to refinance his home. The loan officer advises Adam that this interest rate is available only to "high net worth" individuals with outstanding credit records. Because his number is unlisted and the loan officer seems overly familiar with his finances, Adam suspects that BankCo has disclosed his nonpublic personal information to MortgageCo, in violation of his opt-out choice. Thus, he contacts BankCo to complain that his private information was disclosed. Upon receiving Adam's complaint, BankCo's privacy officer determines whether BankCo may have been the source of an illegal disclosure.

First, using the HDB auditing system, the privacy officer requests an audit of all accesses to Adam's records since the date of his BankCo credit card application. The Control Center automatically converts this request to an audit query in the following form:

```
AUDIT*
FROM bankcodb.customers
WHERE customer = 'Adam'
DURING '2004-1-17 00:00:00.0' AND CURRENT_DATE
```

In response to this audit query, HDB returns an audit trail of all queries that accessed Adam's records within the specified date range, including user, date, recipient, purpose, and the actual query string. The results are displayed in *Table 2*.

Table 2 Results of general audit of access to customer records

User	Date	Query	Customer	Purpose	Recipient
D. Allen	2004-10-12	SELECT...	Adam	Marketing	BankCo Auto Finance
S. Roberts	2005-04-25	SELECT...	Adam	Accounting service	S. Roberts
...	...	...	...	...	...
B. Jones	2005-11-10	SELECT...	Adam	Marketing	MortgageCo

Next, the privacy officer narrows her request to seek only the results of those queries that accessed Adam's address, telephone, number, SSN, or annual income, including aggregate queries. The Control Center submits the following audit query:

```
AUDIT address, phone, ssn, income
FROM bankcodb.customers
WHERE customer = 'Adam'
DURING '2005-1-17 00:00:00.0' AND CURRENT_DATE
```

The HDB audit application then provides the results displayed in *Table 3*. As HDB reconstructs past states of the database, the audit results show the exact information that was accessed at the time of each query.

The privacy officer determines that two queries issued by S. Roberts accessed Adam's address, telephone number, SSN, and income, but the results confirm that S. Roberts did not obtain Adam's current address (in Orinda rather than Oakland). Because the MortgageCo solicitation referenced his new address, it is unlikely that either of these queries improperly disclosed Adam's information. The audit results also reveal that B. Jones accessed Adam's name and current address and disclosed them to MortgageCo. However, this disclosure did not violate Adam's preferences as it did not reveal his telephone number, SSN, or income. It is therefore unlikely that BankCo was responsible for an unauthorized disclosure.

Advantages of HDB Compliance Auditing

HDB Compliance Auditing offers several important advantages over conventional auditing solutions. First, its method of logging offers significant performance and storage advantages over systems that log the results of every database query. HDB stores only the updates, which are usually maintained by the existing database infrastructure. Thus, it does not incur a cost for read accesses (i.e., those in which the user reads the information, but does not insert, update, or delete any information), which may constitute the bulk of database queries. Second, HDB does not burden normal query processing, as it requires only that the query string and a few annotations be logged during normal query processing. Third, HDB provides a flexible audit query language that allows users to declaratively specify the precise information that they would like to audit through the Control Center interface. Fourth, HDB's modest storage overhead allows it to maintain audit logs for long periods of time and trace data access for years in the past. Because the Active Enforcement component maintains multiple policy versions, HDB's audit system can determine whether past disclosures were made in accordance with the

Table 3 Result of audit of access to customer's address, phone number, SSN, and income

User	Date	Customer	Address	Phone Number	SSN	Income
S. Roberts	2005-04-25	Adam	Oakland, CA	444-4444	012-34-5678	$200,000
S. Roberts	2005-09-17	Adam	Oakland, CA	444-4444	012-34-5678	$200,000
B. Jones	2005-11-10	Adam	Oakland, CA			

policy that was effective at the time.[23] Finally, because HDB relies mostly on the existing database infrastructure and has a minimal impact on normal query processing, it does not disrupt existing production systems.

HDB also has security advantages over auditing systems that log only query results, as such systems may not track all information actually disclosed in response to certain queries. For example, users can formulate queries to return nonsense data whenever the presence of sensitive data is detected. Such a query could be written as: `Select [1] if customer [Adam] has an annual income > $150,000`. Although the output reflects only that "1" was disclosed, the user determined that Adam's income exceeds $150,000. In addition, these result logging systems may allow users to determine information about particular customers by running a series of aggregate queries without tracking that any sensitive information was disclosed. In the HDB system, auditors can analyze the query logs to reveal suspicious sets of aggregate queries. Cryptographic techniques,[24] serialization,[25] and distributed log storage[25] can be used to ensure that audit logs are resistant to administrator tampering.

CONCLUSIONS AND FUTURE WORK

We have shown how Hippocratic Database enforcement and auditing technologies can implement many of the common principles of various data protection laws around the world. Organizations must be able to enforce automated security and privacy controls that conform to these common principles to facilitate cross-border information transfers. HDB provides a scalable architecture for enforcing such controls with minimal impact on performance, storage overhead, and the functions of existing applications.

In the following subsections, we describe several important research topics that could extend HDB enforcement and auditing technologies.

Distributed policy enforcement

Many laws require companies to assure that third parties are also in compliance with data protection obligations before transferring personal information to those parties. This is usually accomplished by contractually obligating the third parties to provide relevant data protection and to accept at least joint responsibility for any breaches of information privacy. We are currently researching extensions of HDB that transfer disclosure policies as metadata with sensitive information to assure that third-party recipients enforce the appropriate policies (of the source and recipient) subsequent to data transfer. This would require the remote systems to maintain compatible enforcement or auditing mechanisms.

Intrusion detection

There are many instances in which a query appears benign in isolation, but is suspicious if considered in the context of other queries or background information available to the query issuer. It would be desirable to have an intrusion detection capability to identify these suspicious queries and either block them or alert the system administrator. A database instance-independent approach is proposed by Miklau and Suciu[26] to deny access based upon the circumstances of the query, regardless of the information contained in the database. However, determining whether a particular query violates a disclosure rule for all possible database instances is an intractable problem.[26] Further research is needed to develop practical and efficient methods for identifying suspicious queries in light of other queries and contextual data.

Querying encrypted data

HDB Active Enforcement could be used in combination with techniques that allow querying of encrypted data without significantly degrading performance. This would ensure that the system guards against attacks to the storage media that occur outside of the database system. Research is necessary in the development of efficient and useful methods of querying encrypted numeric and categorical data.

Data integrity

The current HDB auditing system maintains a query log of all `SELECT` statements to support data-access tracking. This could be extended to log other types of database updates and contextual information, so that an audit could reveal any modifications made to a specified data item and the relevant circumstances of the update. Such a capability would help to uncover the sources of improper modifications and assure the accuracy of personal data.

**Trademark, service mark, or registered trademark of Sun Microsystems, Inc. in the United States, other countries, or both.

CITED REFERENCES

1. D. Solove, *The Digital Person*, NYU Press, New York (2004).
2. R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Hippocratic Databases," *Proceedings of the 28th International Conference on Very Large Databases*, Morgan Kaufmann Publishers, San Francisco, CA (2002), pp. 143–154.
3. K. LeFevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan, Y. Xu, and D. DeWitt, "Limiting Disclosure in Hippocratic Databases," *Proceedings of the 30th International Conference on Very Large Databases*, Morgan Kaufmann Publishers, San Francisco, CA (2004), pp. 108–119.
4. R. Agrawal, R. Bayardo, C. Faloutsos, J. Kiernan, R. Rantzau, and R. Srikant, "Auditing Compliance with a Hippocratic Database," *Proceedings of the 30th International Conference on Very Large Databases*, Morgan Kaufmann Publishers, San Francisco, CA (2004), pp. 516–527.
5. *Report to the President, Revolutionizing Health Care Through Information Technology*, President's Information Technology Advisory Committee (June 2004), http://www.nitrd.gov/pitac/meetings/2004/20040617/20040615_hit.pdf.
6. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, *Official Journal of the European Communities of 23 November 1995 No L. 281 p. 31*, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri = CELEX:31995L0046:EN:HTML.
7. Personal Information Protection and Electronic Documents Act, Statutes of Canada, Second Session, Thirty-sixth Parliament, 48–49 Elizabeth II, 1999–2000 (2000), http://laws.justice.gc.ca/en/P-8.6/text.html.
8. Personal Information Protection Act, Law 57 of 2003, Japan, http://www5.cao.go.jp/seikatsu/kojin/foreign/act.pdf.
9. Privacy Act 1988, Act No. 119 of 1988 as amended, Commonwealth of Australia, http://www.privacy.gov.au/publications/privacy88_030706.pdf.
10. Personal Data Protection Act, Act 25,326, enacted October 4, 2000, Argentine Republic, http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-63297.
11. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, adopted September 23, 1980, Organisation for Economic Co-operation and Development, http://www.oecd.org/document/18/0,2340,en_2649_201185_1815186_1_1_1_1,00.html.
12. Health Insurance Portability and Accountability Act of 1996, Public Law 104-191, 104th Congress of the United States of America, http://www.cms.hhs.gov/HIPAAGenInfo/Downloads/HIPAALaw.pdf.
13. Gramm-Leach Bliley Financial Services Modernization Act of 1999, 15 U.S.C. §§ 6801-6809 (Disclosure of Nonpublic Personal Information).
14. Fair Credit Reporting Act, 15 USC § 1681, *et seq.*, http://www.ftc.gov/os/statutes/031224fcra.pdf.
15. *Privacy of Consumer Financial and Health Information Regulation*, State of Vermont, Department of Banking, Insurance, Securities & Health Care Administration, Banking Division, Regulation B-2001-01 (2001), http://www.bishca.state.vt.us/RegsBulls/bnkregs/REG_B2001_01.pdf.
16. California Financial Information Privacy Act, Financial Code §§ 4050–4060, http://www.leginfo.ca.gov/cgi-bin/displaycode?section=fin&group=04001-05000&file=4050-4060.
17. California Civil Code § 1798.82, http://www.leginfo.ca.gov/cgi-bin/displaycode?section=civ&group=01001-02000&file=1798.80-1798.84.
18. Safe Harbor Privacy Principles, United States Department of Commerce, http://www.export.gov/safeharbor/.
19. *IBM Hippocratic Database Active Enforcement User Guide, Version 1.0*, IBM Corporation, http://www.almaden.ibm.com/cs/projects/iis/hdb/Publications/papers/HDBEnforcementUserGuide.pdf.
20. R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "An XPath-Based Preference Language for P3P," *Proceedings of the 12th International World Wide Web Conference*, ACM Press, NY (2003), pp. 629–639.
21. A. Nanda and D. K. Burleson, *Oracle Privacy and Security Auditing*, Rampant TechPress, USA (2003).
22. Standards for Privacy of Individually Identifiable Health Information, United States Department of Health and Human Services, 45 CFR Parts 160 and 164, http://www.hhs.gov/ocr/combinedregtext.pdf.
23. *IBM Hippocratic Database Auditing: User Guide, Version 1.0*, IBM Corporation, http://www.almaden.ibm.com/cs/projects/iis/hdb/Publications/papers/HDBAuditingUserGuide.pdf.
24. R. Snodgrass, S. Yao, and C. Collberg, "Tamper Detection in Audit Logs," *Proceedings of the 30th International Conference on Very Large Databases*, Morgan Kaufmann Publishers, San Francisco, CA (2004), pp. 504–515.
25. *Implementing a Trusted Information Sharing Environment: Using Immutable Audit Logs to Increase Security, Trust and Accountability*, Markle Foundation (February 2006).
26. G. Miklau and D. Suciu, "A Formal Analysis of Information Disclosure in Data Exchange," *Proceedings of the ACM SIGMOD International Conference on Management of Data*, ACM Press, New York (2004), pp. 575–586.

Accepted for publication November 9, 2006.
Published online March 21, 2007.

Christopher M. Johnson
IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (johnsocm@us.ibm.com).

Mr. Johnson is a researcher in the Intelligent Information Systems group of the Computer Science department at the IBM Almaden Research Center. He received B.A. and M.B.A. degrees from the University of California at Berkeley in 1993 and 2003, respectively, and a J.D. degree from the University of Southern California in 1996. At Almaden, his research focuses on privacy and legal compliance technologies.

Tyrone W. A. Grandison
IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (tyroneg@us.ibm.com).

Dr. Grandison manages the Data Disclosure team in the Computer Science department at the Almaden Research Center. He received B.Sc. and M.Sc. degrees from the University of West Indies in Jamaica in 1997 and 1998, respectively, and a Ph.D. degree from the Imperial College of Sciences, Technology, and Medicine of the University of London, United Kingdom in 2003. His research interests include security-sensitive and privacy-aware data disclosure, security and trust management, fundamental data science for new frameworks, models, methodologies, and opportunities for specific application domains. ■