Statistics for Social and Behavioral Sciences

Statistical Confidentiality

Principles and Practice

Springer

Statistics for Social and Behavioral Sciences

For other titles published in this series, go to
http://www.springer.com/series/3463

George T. Duncan · Mark Elliot ·
Juan-José Salazar-González

Statistical Confidentiality

Principles and Practice

Springer

George T. Duncan
Carnegie Mellon University
Santa Fe, NM 87505, USA
gtduncan@gmail.com

Mark Elliot
University of Manchester
Manchester, UK
mark.elliot@manchester.ac.uk

Juan-José Salazar-González
University of La Laguna
La Laguna, 38271 Tenerife, Spain
jjsalaza@ull.es

ISBN 978-1-4419-7801-1 e-ISBN 978-1-4419-7802-8
DOI 10.1007/978-1-4419-7802-8
Springer New York Dordrecht Heidelberg London

Preface

Get together with statisticians and you may see a T-shirt emblazoned, "In God we trust, all others bring data." And, beyond doubt, statisticians are bent on getting data. Indeed they are fully employed in the full gamut of sample surveys, government censuses, observational studies, and clinical trials. Observe another T-shirt bannered, "Top 10 Reasons to be a Statistician," listing, "Estimating parameters is easier than dealing with real life." Surely this self-effacing humor makes the contrary point: that, while statisticians do deal with probability models and their parameters, what really fascinates them is how it all relates to life, and especially the real uncertainties of life. They are intrigued with how data can yield information to reduce these uncertainties, and so enable better decisions to be made. Our concern in this book is that data relevant to issues in the public interest, such as a person's medical record or criminal history, are often highly sensitive. Obtaining and using such data forces us to realize that there is tension between, on the one hand, the desire of the individual for a full and free private life and, on the other hand, the needs of the broader community for information that might, say, improve health care or reduce crime. Statistical confidentiality is pivotal in resolving this tension.

This book on statistical confidentiality is written for all involved with personal and proprietary data from empirical studies. Various roles require an understanding of statistical confidentiality. Here are some instances drawn from issues concerning assessment of a drug rehabilitation program:

- You are a researcher. Your undertaking is to seek rich and convincing evidence to assess benefits of the program, both to addicts and to the community.
- You are a statistician. You design a survey of addicts and link the results to administrative data from the program.
- You are a data steward. Your responsibility is to take the statistician's data and build a database useful to researchers *and* acceptable to privacy advocates.
- You are a privacy advocate. You express qualms about the researcher and statistician matching drug-use records to the addict's personal finances and any criminal behavior.
- You are a respondent to the survey. In principle you are happy to provide data about yourself for the good of society, but you are also concerned about what happens to that data once you have handed it over: will your privacy be respected; will your confidentiality be maintained?

So what exactly is statistical confidentiality? Simply put, it is the stewardship of data to be used for statistical purposes. Stewardship, as expressed in statistical confidentiality, is an active embrace of responsibility for both protecting data and ensuring its beneficial use. Explicitly it requires proper practices for both providing and restricting access to data products.

Getting and using data just to support public policy analysis costs a lot of money. Lane (2003) makes this point: "Billions of taxpayer dollars are spent in supporting the collection and dissemination of federal, state and local data, billions of dollars are spent in data analysis, and this, in turn, both informs scientific understanding of core social science issues and guides decision in how to allocate billions of dollars in social programs." Privacy and confidentiality concerns do not come in dollars, pounds or Euros, but quantifying these concerns via Google search on May 19, 2010, yielded more than 1.42 billion hits on "privacy", more than 19.7 million hits on "confidentiality", and more than 55,000 hits on "statistical confidentiality" (including the quotation marks in the query). Just looking at the first few of these hits on "statistical confidentiality" leads us to information on a United Nations work session in Geneva, Switzerland, to a discussion of pertaining laws in France from the National Institute for Statistics and Economic Studies, to the US Federal Register discussion of the Statistical Confidentiality Order, to a research site of the Computational Aspects of Statistical Confidentiality Project based in the Netherlands, to testimony before the US Congress regarding confidentiality and coordination among statistical agencies, and to a discussion of confidentiality in the Japanese 2000 Census of Population.

The issues cited above illustrate the scope of statistical confidentiality. By absorbing the ideas in this book, you will gain understanding in both the principles and practice of this important field that can benefit your work:

- As a researcher, you will understand why an agency holding statistical data does not respond well to approaching them saying, "Just give me the data; I'm only going to do good things with it," and appreciate why you need to learn about what motivates statistical confidentiality and how it works in practice.
- As a statistician, you will incorporate the requirements of statistical confidentiality into your methodologies for data collection and analysis.
- As a data steward, you are caught between those eager for data and those who worry about confidentiality in its dissemination. Fortunately, using the tools of statistical confidentiality you will progress toward satisfying both groups.
- As a privacy advocate, you will comprehend how confidentiality can be protected even though statistical data are, and should be, made available.
- As a respondent, you will have a better understanding of why your data are needed, how they will be used, and how they will be protected.

We have organized this book into eight chapters. In Chapter 1 we motivate and define the study of statistical confidentiality, laying out the dilemma of data stewardship organizations (we will call them DSOs; important examples are statistical

agencies) in resolving the tension between protecting data from snoopers and providing data to legitimate users. We identify the stakeholders in the statistical process, show why statistical data are so useful and in such demand, and explain why DSOs are so concerned about confidentiality. We explore the concept of disclosure risk in terms of an attack by a data snooper and show the basic ways statistical confidentiality can be protected. In Chapter 2 we lay out the fundamental concepts of statistical confidentiality, develop conceptual models of disclosure risk and data utility, identify ways risk can be assessed and controlled, and explore the types of attack that a data snooper might mount. From a rational decision-making perspective, Chapter 3 presents the methodologies of disclosure risk assessment, including a variety of useful metrics for risk assessment. Chapter 4 gives techniques for statistical disclosure limitation of aggregate data, specifically data in tabular form. We present an appropriate definition of disclosure-limited tabular output. We develop deterministic methods, especially through mathematical programming, and stochastic methods, such as cyclic perturbation, for statistical disclosure limitation of tabular data. Chapter 5 gives techniques for disclosure limitation of microdata, that is, data in original record form. We affirm the value of microdata and clarify what users need from such data. We identify the concerns that a DSO has in satisfying the ethical, pragmatic, and legal considerations that motivate their confidentiality promises to data providers. We point out the characteristics of microdata that make it vulnerable to confidentiality attack, and explore various masking methods. We introduce the idea of synthetic data, that is, data that are stochastically generated from a model inferred from the source data. In Chapter 6 we give measures of the impact of disclosure limitation on data utility, and develop the methodology of R-U confidentiality maps and their empirical analog. Chapter 7 provides restricted access methods as a body of administrative procedures for disclosure control. We explore the issues a DSO must face in deciding who can have access, where can access be obtained, what analysis is permitted, and what modes of access should be allowed. Finally, Chapter 8 explores the future of statistical confidentiality. We address a number of important questions: Will privacy and statistical confidentiality have new meanings? Who will care about statistical data? What new forms of DSO will develop? Will statistical data remain valuable? Will there be new issues for statistical confidentiality? Will there be new forms of data snooping? What new strategies of disclosure limitation should be developed?

Plainly, statistical confidentiality is a large and important issue of concern. Research and work in statistical confidentiality is growing rapidly, as is the number of people working on this problem. Many people are screaming that their privacy and confidentiality are vanishing. It is the job of DSOs to provide as much quality data as possible without violating confidentiality laws and promises. This book provides a reader with a comprehensive understanding of the principles and practice of statistical confidentiality. It balances methods and ideas with specific examples. We have written it to be accessible to those just entering the field. Much of the material requires no specific background in mathematics or statistics. Those sections which are more technical are prefaced with explanations of the general ideas without complicated equations.

Our thanks go to our many colleagues who helped us comprehend the need for new ways of dealing with statistical confidentiality, collaborated in our research on this topic over many years, and inspired this book.

Santa Fe, New Mexico, USA — George T. Duncan
Manchester, UK — Mark Elliot
La Laguna, Spain — Juan-José Salazar-González

Contents

Chapter 1
Why Statistical Confidentiality?

As a society we are judged by how we treat those at the dawn of life, those in the sunset of life and those in the shadows of life.

Confirming statistical confidentiality as vital to the stewardship of personal data, the United Nations set out Principle 6 of its Fundamental Principles of Official Statistics: *Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.*[1] Data stewardship is active on two fronts:

- protecting entrusted data by providing confidentiality
- assuring its beneficial use by researchers and policy analysts

The US Census Bureau assures, *Data Stewardship is the formal process we use to care for your information—from the beginning, when you answer a survey, to the end, when we release statistical data products. Data Stewardship goes beyond the law to ensure that any decisions we make will fulfill our ethical obligations to respect your privacy and protect the confidentiality of your information . . . The Data Stewardship program ensures that we protect the information you provide, while enabling us to release high quality information about our population and the economy.*[2]

In this chapter, we address the following questions:

1. What is statistical confidentiality? (Section 1.1)
2. Who are the stakeholders in the statistical process and what are the stakes for them? (Section 1.2)
3. What is the dilemma in statistical confidentiality? (Section 1.3)
4. What is the utility of statistical data and why is it in such demand? (Section 1.4)
5. Why are Data Stewardship Organizations (DSOs) so concerned about confidentiality? (Section 1.5)

[1] unstats.un.org/unsd/goodprac/bpabout.asp

[2] www.census.gov/privacy/data_stewardship/partnership_and_trust.html

G.T. Duncan et al., *Statistical Confidentiality*, Statistics for Social and Behavioral Sciences, DOI 10.1007/978-1-4419-7802-8_1,

6. What is it about high-quality statistical data that raises confidentiality concerns? (Section 1.6)
7. What is disclosure risk and who is the data snooper who might want to attack statistical confidentiality? (Section 1.7)
8. How can statistical confidentiality be protected? (Section 1.8)

Building on this base, Chapter 2 lays out the concepts of statistical disclosure limitation. Chapter 3 presents the methodologies of disclosure risk assessment. Chapter 4 gives techniques for statistical disclosure limitation (SDL) of aggregate data, specifically data in tabular form, while Chapter 5 gives techniques for SDL of microdata, that is, data in original record form. Chapter 6 gives measures of the impact of SDL on data utility. Chapter 7 describes restricted access methods as a body of administrative procedures for disclosure control. Finally, Chapter 8 explores the future of statistical confidentiality.

1.1 What Is Statistical Confidentiality?

First, what do we mean by *confidentiality*—for now leaving aside the adjective "statistical"? Broadly, confidentiality reflects a complex amalgam of societal values about information privacy, secrecy of personal information, and autonomy of the individual. But more specifically, confidentiality is a status accorded to information about a person. Under confidentiality, the party holding the information is bound by an implicit or explicit promise that it be protected from unauthorized or inappropriate access and usage. Fienberg (2005) defines confidentiality this way: "Broadly, a quality or condition accorded to information as an obligation not to transmit that information to an unauthorized party."

The term "confidentiality" is used in many different contexts, ranging over religious confessionals, communications from patient to doctor, employment recommendations, and asserted prerogatives of officials to garner input shielded from outside monitoring. Regardless of context, confidentiality is a promise that is made to the provider of the information by the receiver and current holder of the information. There are two overarching principles to confidentiality: information should be (1) reserved exclusively for intended purposes, and (2) used only by authorized individuals.

Promise keeping is not the only aspect of the ethical basis for confidentiality. Other ethical considerations include a respect for the individual's autonomy, a desire not to cause individuals embarrassment, and a view that society functions better when confidentiality is considered to be a human right. Additional considerations, especially pragmatic and legal ones, motivate concern for confidentiality. We will explore this in detail in Section 1.5.

We view statistical confidentiality as a body of principles, concepts, and procedures that permit confidentiality to be afforded to data, while still permitting its use for statistical purposes. This is not an easy task, and so arises the need for this book.

1.2 Stakeholders in the Statistical Process

To understand how statistical confidentiality functions, we must appreciate that the various interests among researchers, statisticians, data users, and data providers rarely align and so must be reconciled by the responsible organization. Data stewardship organizations manage the process of data capture, storage, integration, and dissemination. DSOs include statistical agencies, national statistical offices, data archives, trade associations, unions, credit bureaus, and health information associations. Many DSOs support research and policy analysis based on statistical methods.

To illustrate the DSO's problem let us consider medical data. Here is a listing of just some of the stakeholders in such data:

- Patients
- Physicians
- Family members
- Hospital administration
- Public health officials
- Insurance companies
- Governments at all levels
- Employers
- Health care researchers
- Educators
- Journalists
- Marketers
- Law enforcement
- Litigants

Patients choosing a hospital might want statistical information on hospital mortality rates. Insurance companies seek statistical information on costs per patient. Health care researchers want objective, high-quality statistical data about things such as hospital admissions of teenage drug users and how they relate to socioeconomic conditions of the patients. Marketers want statistical data on trends regarding the willingness of patients to accept generic alternatives to prescription drugs. Law enforcement wants statistical data on admissions to emergency facilities for gunshot wounds. Hospital administrators need medical information for billing purposes and quality assurance purposes. Journalists want to provide analyses to their readers. Family members want information about their loved ones. Given the variety of these expressed needs, a DSO faces a real and growing tension between protecting privacy and satisfying the demand for information.

1.3 The Data Stewardship Organization's Dilemma

Astonishing advances in technology for computing and telecommunications have dramatically altered how we make decisions. In our domain of statistical confidentiality, these advances have also precipitated a dilemma: On the one hand,

technology has boosted the perceived benefits of statistical information. Consider that three-quarters of Americans polled[3] said it was very important or somewhat important for doctors and hospitals to use electronic records instead of paper. On the other hand, people have heightened anxiety about whether confidentiality is, or even can be, provided. The same poll revealed that nearly the same number of respondents said they were not confident that their computerized records would remain safe from prying eyes. What is the meaning of these conflicting developments for the DSO, which must make decisions about confidentiality and data access?

On the one hand, the benefits of statistical information about people became apparent as early as 1830 when Adolphe Quetelet examined factors influencing social phenomena such as crime, marriage, and suicide (see Stigler, 1986). Illustrating the importance of government data in public health, John Snow used data in 1854 from the Registrar of Deaths to show the link between cholera and contaminated water in London. Today the social statistics enterprise is huge. For example, the Organization for Economic Co-Operation and Development (OECD) gathers voluminous statistical data from at least 30 countries on topics ranging from agriculture to health to globalization to transport, and disseminates data products to an estimated 6 million users in over 100 countries.

On the other hand, increasing availability of massive amounts of data about individuals has also fueled confidentiality concerns. Here are two examples:

- In 2006 US Congress members and the Federal Communications Commission raised concerns about the sale on several Internet sites of customers' wireless and landline phone records, including the date, time, and length of calls placed by consumers.
- Controversy ensued when in 1998 the Icelandic parliament granted the biopharmaceutical company deCODE genetics the right to construct an electronic database of the country's health records together with genetic information on some 65% of the Icelandic population and a genealogy that for most stretches back 1000 years.

Because a DSO is a broker between the providers of data and the users of data, it is caught in the colliding paths of information demands and confidentiality concerns. DSOs broadly encompass all organizations that capture, store, integrate, and disseminate information—the CSID (Capture Storage Integration Dissemination) data process (Duncan, 2004). In the statistical realm we are concerned with data capture that involves surveys (for example the American Community Survey), censuses (such as the UK Census 2011), and administrative procedures (such as providing the required information to obtain a medical license in Spain). Today, data storage is ubiquitously electronic and measured in terabytes. Electronic storage facilitates the integration of records across disparate databases, for example a health economics

[3] http://www.npr.org/templates/story/story.php?storyId=103362165

study might integrate an individual's medical record with their employment history. Data dissemination makes a data product such as a collection of tables or a microdata record file available to a data user.

The largest and most prominent of DSOs are those whose primary function is statistical. These include government statistical agencies (e.g., the US Bureau of the Census), national statistical offices (e.g., Statistics Canada), and cross-national statistical offices (most prominently, Eurostat—the Statistical Office of the European Community).

Another grouping of DSOs is the data libraries and data archives (e.g., the National Data Archive for Child Abuse and Neglect at Cornell University[4]), which typically are not heavily involved in primary data collection, but do assemble statistical databases and make them available.

A third grouping of DSOs includes those with rather specialized functions. These include trade associations (e.g., National Association of Manufacturers), unions (e.g., in Spain, Unión General de Trabajadores, UGT), credit bureaus (e.g., Experian), marketing data firms (e.g., Nielsen//NetRatings, Marketing Opinion and Research International (MORI), which is a large market research firm in Great Britain), and health insurance information agencies (e.g., the Health Insurance Industry Benchmarking Association).

A fourth grouping of DSOs is organizations for which the collection of data for statistical purposes is not part of their primary business, but which are increasingly been called on to enable the use of such data that they do collect for administrative purposes to be used for statistical research. For example, government departments in the United Kingdom provide data for the Neighbourhood Statistics Service (NeSS).

Many DSOs are experts in what is needed to acquire high-quality data from individuals, households, and establishments. A DSO serves the operational and decision needs of the community at large in the case of government statistical agencies, specialized researchers and policy analysts in the case of data libraries and data archives, and various client constituencies in the case of the specialized DSOs. Regardless of their mission, the challenge to each DSO is to manage the critical and sensitive data under its care so that they can be used to their fullest capacity.

Also, whatever the DSO and its particular function, the DSO's professional staff has the responsibility of developing and implementing confidentiality and data access policies. The DSO has to be concerned with what data should be afforded confidential status and how it can be protected—specifically how the risk of disclosure, which is access by unauthorized parties, can be limited (see, for example, Duncan, 2001).

The Secretary General of the European Commission affirmed the value of statistical confidentiality to the European Parliament: "Official statistics must be

[4] www.ndacan.cornell.edu/

produced and disseminated according to common standards guaranteeing compliance with the principles of impartiality, reliability, objectivity, scientific independence, cost-effectiveness and statistical confidentiality."[5] According to this statement, the DSO must fulfill conflicting standards: one standard that requires an organization to *provide* quality data products at low cost, and a second mandating that the organization maintain statistical confidentiality. This second standard is essentially an imperative to *not provide* information that can be tied to individuals. Having one mandate to provide data and a second to not provide data emphasizes the vexing nature of the DSO's predicament. How can one have a duty to reveal and yet a duty to hide? Nonetheless, the DSO has an opportunity to develop creative ways to resolve this inherent tension and so provide the factual basis for decision making—a critical component that benefits the common good—while simultaneously demonstrating that the individual's concerns about confidentiality can be respected.

To summarize, given that providing complete access to statistical data and maintaining confidentiality are ineluctably opposing goals, key staff within DSOs face a predicament whose resolution requires a grasp of the principles and practice of statistical confidentiality. In Section 1.4 we discuss why statistical data are of such value and in Section 1.5 we discuss why DSOs are so concerned about statistical confidentiality.

1.4 The Value of Statistical Data

Statistical data are the foundation for empirical analysis, providing the factual information needed to guide policy and decision making and to enable a better scientific understanding of how our world works. Empirical analysis addresses questions about human conditions and needs. For example, empirical analysis using statistical data can address such questions as the following:

1. Can middle-class neighborhoods absorb low-income families under a housing mobility policy without leading to neighborhood decline through out-migration of affluent families?
2. How can chest pain be diagnosed as myocardial ischemia in women?
3. What combination of pharmacological and psychosocial interventions works best for bipolar mood disorder?
4. Will lowering tax rates increase tax revenue?
5. Will restricting immigration of the technically skilled raise employment opportunities for poor citizens?

To illustrate the need for statistical data, let us consider issues akin to Question 1 above. In 1999, the Urban Institute sponsored the Symposium on Section 8 Mobility

[5] eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52005DC0217:EN:NOT

and Neighborhood Health (Section 8 is a tenant-based housing assistance program in the United States). The report on this symposium[6] identified the need for the following statistical data:

1. Data to map Section 8 locations, identify clusters of Section 8 recipients, and assess market conditions in neighborhoods where these clusters are located.
2. Local data to understand and address potential behavior problems and neighborhood consequences.
3. Data for determining the availability of below-Fair Market Rent housing units in neighborhoods of different types regionwide, and relating clusters of Section 8 housing to the distribution of affordable housing.

Those who work to bring such statistical data to bear on research questions include policy analysts, sociologists, epidemiologists, psychiatric researchers, and economists. They are employed variously by universities, think tanks, trade associations, unions, marketing and consulting firms, corporations, and government departments and agencies.

Answering such questions requires statistical data of high quality, which necessitates that it be accurate, detailed, and comprehensive. Often it needs to be at the level of the individual person. Further, it often must identify what is happening in both time and space. So the data need to be geographically specific and longitudinal. These criteria are explored in the next section in conjunction with their implications for statistical confidentiality. Here we note that the demand for statistical data has led to the development of a variety of accessible data sets that are widely useful. Here are two examples that illustrate what is readily available on the web and suggest how useful they can be:

1. *US Census Bureau PUMS files—Public Use Microdata Sample*[7] Recognizing that "Because of the rapid advances in computer technology and the increased accessibility of census data to the user community, the Census Bureau has had to adopt more stringent measures to protect the confidentiality of public use microdata through disclosure-limitation techniques," this sample provides individual and household-level data from the 2000 Census long form. These data provide comprehensive information on social and economic characteristics of the people as well as physical and financial aspects of housing.
2. *The WHO Statistical Information System*[8] Here the World Health Organization presents the most recent statistics since 1997 of 50 health indicators for WHO's 192 Member States. A highlight: "While some countries are making

[6] www.urban.org/url.cfm?ID=309465

[7] www.census.gov/population/www/cen2000/pums.html

[8] www.who.int/whosis/en/

progress and achieve greater equality in child survival chances within the country, the general picture is that little progress has been made during the last decade."

1.5 Why Are DSOs Concerned About Statistical Confidentiality?

Concerns about confidentiality have presumably been an issue ever since one person learned something about another. Indeed, the 4th century BC Hippocratic Oath mandated for physicians, "All that may come to my knowledge in the exercise of my profession or in daily commerce with men, which ought not to be spread abroad, I will keep secret and will never reveal." In the Hebrew Oath of Asaph, the practitioner is admonished: "Ye shall not disclose secrets confided unto you."[9] The Bible acknowledges there is "a time to keep silence and a time to speak" (Ecclesiastes 3:7). Notwithstanding this long history of concern, many of today's information privacy worries are new, and are driven by both technological advances that have lowered costs and societal changes, especially those involving mobility and population growth.

A large, mobile population has raised the demand for personal information. Health care is no longer the sole province of a local physician providing continuing care to a patient over dozens of years. Instead, health care must be provided wherever the patient may happen to live, work, and play—whether in Madrid or Adelaide—and so benefits from databases of electronic records. Commerce is not just the purview of the local merchant, but extends globally and works through instant credit checks against electronic databases and electronic tracking of shipping.

1.5.1 A Difficult Context for a DSO

Over the years, statisticians and other professionals in DSOs have contributed to methodologies and practices that enable the collection of quality statistical data. Yet today the analysis that is required to answer many questions of societal importance is threatened because the faucet to useful statistical data may become so restricted that nothing but a trickle emerges. The statistical community is in this crisis of data access for four inter-related reasons: privacy worries, confidentiality concerns, a changing legal and social context, and an increased sensitivity to social impact. We now examine each of these reasons, which all essentially relate to whether a data provider can trust a DSO.

[9] http://www1.umn.edu/phrm/oaths/oath5.html

1.5.1.1 Privacy Worries

Privacy is closely aligned with confidentiality, but distinct from it. Privacy is linked to the desire of individuals to control how they are presented to others. Also, as those others see it, privacy is not being intrusive into the lives of people. Information privacy has been defined to encompass "an individual's freedom from excessive intrusion in the quest for information and an individual's ability to choose the extent and circumstances under which his or her beliefs, behaviors, opinions, and attitudes will be shared or withheld from others"(Duncan et al., 1993).

Not surprisingly, the appropriate scope of privacy is hotly debated. Much of this debate is outside the scope of this book. For example, we will not discuss how much privacy employees should have in using e-mail in the workplace, or how much authority government should have to search your laptop's files. Yet, much of the debate about privacy *is* within the scope of this book. For example, we will address privacy issues in areas such as research use of personal medical information, the sharing of statistical data in the United States between the Social Security Administration and the National Institute on Aging, and providing public-use data files from the Korean National Fertility Survey.

A quick search of the web reveals the breadth of public concern about privacy in almost every area where information is collected:

1. The Electronic Frontier Foundation shows how secret codes inserted in the output of color printers can be used to trace their source.[10]
2. BBC News raises concerns about Radio Frequency Identification (RFID) tags on consumer goods.[11]
3. CBS News reports on personal data on airline passengers being given to the government for testing a computerized background-check project.[12]

1.5.1.2 Confidentiality Concerns

As with privacy worries, confidentiality concerns are ubiquitous, appearing wherever information is shared. Think of these problems:

1. Duke University Medical Center reports that HIV patients in rural areas say they're afraid to seek treatment because they fear breaches of confidentiality by their medical providers.
2. In the United Kingdom the Equal Opportunities Commission seeks pay records to check for sex bias, but does this require employee permission?

---

[10] http://hardware.slashdot.org/article.pl?sid=08/02/15/1612226

[11] http://news.bbc.co.uk/2/hi/technology/6691139.stm

[12] http://ssc.sagepub.com/cgi/content/abstract/23/4/401