**Chapter 10**

**Boundaries**

**They constantly try to escape**

**From the darkness outside and within**

**By dreaming of systems so perfect that no one will need to be good**

– TS Eliot

**Anything your computer can do for you it can potentially do for**

**someone else.**

– Alan Cox

**You have zero privacy anyway. Get over it.**

– SCOTT MCNEALY

**10.1** **Introduction**

When we restrict information ﬂows to protect privacy or conﬁdentiality, a policy  
 goal is usually not to prevent information ﬂowing ‘down’ a hierarchy but to  
 prevent it ﬂowing ‘across’ between smaller groups.

1. If you give the million US Federal employees and contractors with a Top

Secret clearance access to too much Top Secret data, then you get a  
 whistleblower like Ed Snowden if you’re lucky, or a traitor like Aldrich  
 Ames if you’re not.

2. As mobile phones spread round the world, they’ve made wildlife crime

easier. Game rangers and others who ﬁght poaching face organised crime,  
 violence and insider threats at all levels, but unlike in national intelligence  
 there’s no central authority to manage clearances and counterintelligence.

3. If you let too many people in a health service see patient records, you get

scandals where staff look up data on celebrities. And the existence of big  
 central systems can lead to big scandals, such as where a billion English  
 medical records going back a decade were sold to multiple drug companies.

322

*10.1. INTRODUCTION*

4. Similar issues arise in social care and in education. There are frequent

calls for data sharing, yet attempts to do it in practice cause all sorts of  
 problems.

5. If you let everyone in a bank or an accountancy ﬁrm see all the customer

records, then an unscrupulous manager could give really good advice to a  
 client by looking at the conﬁdential ﬁnancial information of that client’s  
 competitors.

The basic problem is that if you centralise systems containing sensitive infor-

mation, you create a more valuable asset and simultaneously give more people  
 access to it. Just as the beneﬁts of networks can scale more than linearly, so  
 can the harms.

A common mitigation is to restrict how much information any individual

sees. In our ﬁve example cases above:

1. Intelligence services put sensitive information into compartments, so that

an analyst working on Argentina might see only the Top Secret reports  
 relating to Argentina and its neighbouring countries;

2. Systems that support game conservation have to do something similar, but

access control has to be a federated effort involving multiple conservancies,  
 researchers, rangers and other actors;

3. Many hospital systems limit staff access to the wards or departments where

they work, to the extent that this is reasonably practical, and patients have  
 a right to forbid the use of their data outside their direct care. Both are  
 becoming more difficult to implement as systems get more complex and  
 their operators lack the incentive to make the effort;

4. In 2010, the UK parliament closed down a system that was supposed to

give doctors, teachers and social workers shared access to all childrens’  
 data, as they realised it was both unsafe and illegal. Yet there’s constant  
 pressure for information sharing, and all sorts of issues with schools and  
 other institutions using dubious cloud services;

5. Financial ﬁrms have ‘Chinese walls’ between different parts of the business,

and bank staff are now often limited to accessing records for which they  
 have a recent customer authorisation, such as by the customer answering  
 security questions over the phone.

We will discuss these kinds of access control in this chapter. There are

several aspects: what sort of technical designs are feasible, the operational costs  
 they impose on the organisation, and – often the critical factor – whether the  
 organisation is motivated to implement and police them properly.

In the last chapter, we discussed multilevel security and saw that it can

be hard to get the mechanisms right. In this chapter, we’ll see that when we  
 go for ﬁne-grained access controls, it’s also hard to get the policy right. Are  
 the groups or roles static or dynamic? Are they set by national policy, by

commercial law, by professional ethics, or – as with your group of Facebook

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 323 | Ross Anderson |

*10.1. INTRODUCTION*

friends – by the system’s users? What happens when people ﬁght over the

rules, or deceive each other? Even where everyone is working for the same boss,  
 different parts of an organisation can have quite different incentives. Some

problems can be technically complex but simple in policy terms (wildlife) while  
 others use standard mechanisms but have wicked policy problems (healthcare).

To start with a simpler case, suppose you’re trying to set security policy at

the tax collection office. Staff have been caught in the past making improper  
 access to the records of celebrities, selling data to outsiders, and leaking income  
 details in alimony cases [188]. How might you go about stopping that?

|  |  |
| --- | --- |
|  | TOP SECRET |

|  |  |
| --- | --- |
|  | SECRET |
|  | CONFIDENTIAL |

|  |  |
| --- | --- |
|  | OPEN |

Fig. 9.1 – multilevel security

Your requirement might be to stop staff looking at tax records belonging to

a different geographical region, or a different industry – except under strict con-  
 trols. Thus instead of the information ﬂow control boundaries being horizontal  
 as we saw in the classic civil service model in Figure 9.1, we actually need the  
 boundaries to be mostly vertical, as shown in Figure 9.2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E |  |
|  | *shared data* | | | | |  |

Fig. 9.2 – multilateral security

Lateral information ﬂow controls may be organizational, as when an in-

telligence agency keeps the names of agents working in one foreign country  
 secret from the department responsible for spying on another. They may be  
 relationship-based, as in a law ﬁrm where different clients’ affairs, and the  
 clients of different partners, must be kept separate. They may be a mixture  
 of the two, as in medicine where patient conﬁdentiality is based in law on the  
 rights of the patient but may be enforced by limiting access to a particular hos-  
 pital department or medical practice. They may be volumetric, as when a game  
 conservancy doesn’t mind declassifying a handful of leopard photos but doesn’t  
 want the poachers to get the whole collection, as that would let them work out  
 the best places to set traps.

Doctors, bankers and spies have all learned that as well as preventing overt

information ﬂows, they also have to prevent information leakage through side-  
 channels such as billing data. The mere fact that patient X paid doctor Y

suggests that X suffered from something in Y’s speciality.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 324 | Ross Anderson |

*10.2. COMPARTMENTATION AND THE LATTICE MODEL*

**10.2** **Compartmentation and the lattice model**

The United States and its allies restrict access to secret information by *codewords*  
 as well as classiﬁcations. These are pre-computer mechanisms for expressing an  
 access control group, such as the codeword *Ultra* in World War 2, which referred  
 to British and American decrypts of messages that had been enciphered using  
 the German Enigma machine. The fact that the Enigma had been broken

was worth protecting at almost any cost. So Ultra clearances were given to  
 only a small group of people – in addition to the cryptologists, translators and  
 analysts, the list included the Allied leaders and their senior generals. No-one  
 who had ever held an Ultra clearance could be placed at risk of capture; and  
 the intelligence could never be used in such a way as to let Hitler suspect that  
 his principal cipher had been broken. So when Ultra told of a target, such as  
 an Italian convoy to North Africa, the Allies would send over a plane to ‘spot’  
 it an hour or so before the attack. This policy was enforced by special handling  
 rules; for example, Churchill got his Ultra summaries in a special dispatch box  
 to which he had a key but his staff did not. (Ultra security is described by  
 David Kahn [1002] and Gordon Welchman [2007].)

Much the same precautions are in place today. Information whose compro-

|  |
| --- |
| mise could expose intelligence sources or methods is marked TS/SCI for ‘Top  Secret – Special Compartmented Intelligence’ and may have one or more code-  words. A classiﬁcation plus a set of codewords gives a *compartment* or security  context. So if you have *N* codewords, you can have 2*N* compartments; some  intelligence agencies have had over a million of them active. This caution was  a reaction to a series of disastrous insider threats. Aldrich Ames, a CIA o�cer  who had accumulated access to a large number of compartments by virtue of  long service and seniority, and because he worked in counterintelligence, was  able to betray almost the entire US agent network in Russia. The KGB’s over-  seas operations were similarly compromised by Vassily Mitrokhin – an o�cer  who’d become disillusioned with communism and who was sent to work in the  archives while waiting for his pension [118]. There was an even earlier precedent  in the Walker spy case. There, an attempt to keep naval vessels in compart-  ments just didn’t work, as a ship could be sent anywhere without notice, and  for a ship to have no local key material was operationally unacceptable. So the  US Navy’s 800 ships all ended up with the same set of cipher keys, which the  Walker family sold to the Russians [876]. You clearly don’t want anybody to  have access to too much, but how can you do that? |

Attempts were made to implement compartments using mandatory access

controls, leading to the *lattice model*. Classiﬁcations together with codewords  
 form a lattice – a mathematical structure in which any two objects *A* and *B*  
 can be in a dominance relation *A > B* or *B > A*. They don’t have to be: *A*  
 and *B* could simply be incomparable (but in this case, for the structure to be a  
 lattice, they will have a least upper bound and a greatest lower bound). As an  
 illustration, suppose we have a codeword, say ‘Crypto’. Then someone cleared  
 to ‘Top Secret’ would be entitled to read ﬁles classiﬁed ‘Top Secret’ and ‘Secret’,  
 but would have no access to ﬁles classiﬁed ‘Secret Crypto’ unless he also had a  
 crypto clearance. This can be expressed as shown in Figure 10.3.

As it happens, the Bell-LaPadula model can work more or less unchanged.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 325 | Ross Anderson |

*10.2. COMPARTMENTATION AND THE LATTICE MODEL*

(TOP SECRET, {CRYPTO, FOREIGN})

(TOP SECRET, {CRYPTO})

(TOP SECRET, {})

(SECRET, {CRYPTO, FOREIGN})

(SECRET, {CRYPTO})

(SECRET, {})

(UNCLASSIFIED, {})

Figure 10.3: – a lattice of security labels

We still have information ﬂows between High and Low as before, where High is  
 a compartment that dominates Low. If two nodes in a lattice are incompatible  
 — as with ‘Top Secret’ and ‘Secret Crypto’ in ﬁgure 10.3 – then there should be  
 no information ﬂow between them at all. In fact, the lattice and Bell-LaPadula  
 models are essentially equivalent, and were developed in parallel. Most products  
 built in the 20th century for the multilevel secure market could be used in  
 compartmented mode. For a fuller history, see the second edition of this book.

In practice, mandatory access control products turned out to be not that

effective for compartmentation. It is easy to use such a system to keep data in  
 different compartments separate – just give them incompatible labels (‘Secret  
 Tulip’, ‘Secret Daffodil’, ‘Secret Crocus’, ...). But the operating system has now  
 become an isolation mechanism, rather than a sharing mechanism; and the real  
 problems facing users of intelligence systems have to do with combining data in  
 different compartments, and downgrading it after sanitization. Lattice security  
 models offer little help here.

There was a sea change in the US intelligence community after 9/11. Leaders

claimed that the millions of compartments had got in the way of the war on ter-  
 ror, and that better information sharing might have enabled the community to  
 forestall the attack, so President Bush ordered more information sharing within  
 the intelligence community. There was a drive by NSA Director Keith Alexan-  
 der to ‘collect it all’, and rather than minimising data collection to maximise it  
 instead and make everything searchable. So nowadays, government systems use  
 mandatory access control to keep the Secret systems apart from the unclassiﬁed  
 stuff, and the Top Secret systems from both, using data diodes and other mech-  
 anisms that we discussed in the previous chapter. The stuff above Top Secret  
 now appears to be mostly managed using discretionary access controls.

The Snowden revelations have told us all about search systems such as

XKeyscore, which search over systems that used to have many compartments. If  
 a search can throw up results with many codewords attached, then reading that  
 result would require all those clearances. In such a world, local labels just get  
 in the way; but without them, as I asked in the second edition of this book, how  
 do you forestall a future Aldrich Ames? Perhaps the US intelligence community

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 326 | Ross Anderson |

*10.3. PRIVACY FOR TIGERS*

was lucky that the failure mode was Ed Snowden instead. As a system admin-  
 istrator he was in a position to circumvent the discretionary access controls and  
 access a large number of compartments.

We later learned that at the CIA, too, compartmentation was not always

effective. In 2017, its hacking tools were leaked in the Vault 7 incident, and a  
 redacted version of the internal report into that was published in 2020 after the  
 trial of the alleged leaker. It revealed that most sensitive cyberweapons were not  
 compartmented, users shared sysadmin passwords, there was no user activity  
 monitoring and historical data were available indeﬁnitely. They did not notice  
 the loss until the tools ended up on Wikileaks a year later. In fact, the Joint  
 worldwide Intel Communications System (JWICS), which the intel community  
 uses for Top Secret data, did not yet use two-factor authentication [2051].

There are a few compartments Ed Snowden didn’t get to, such as the de-

tails of which cryptographic systems the NSA can exploit and how – this was  
 marked ‘extremely compartmented information’ (ECI). Commercial ﬁrms may  
 also have special mechanisms for protecting material such as unpublished ﬁnan-  
 cial results; at my university we compile exam papers on machines that are not  
 even attached to the network. In such cases, what’s happening may be not so  
 much a compartment as a whole new level above Top Secret.

**10.3** **Privacy for Tigers**

People involved in ﬁghting wildlife crime face a fascinating range of problems.  
 The threats range from habitat encroachment through small-scale poaching for  
 bushmeat to organised crime gangs harvesting ivory, rhino horn and tiger body  
 parts on an industrial scale. The gangs may be protected by disaffected com-  
 munities; even heads of government can be a threat, whether by undermining  
 environmental laws or even by protecting poaching gangs. And often the best  
 poacher is a former ranger.

Even where sovereign threats are absent, public-sector defenders often work

for mutually suspicious governments; protecting the snow leopard from poach-  
 ers involves rangers in India, Pakistan, China, Nepal and Tajikistan, while the  
 illegal ivory trade in East Africa spills over borders from Kenya down to South  
 Africa. And technology is making matters worse; as mobile phone masts have  
 gone up in less developed countries, so has poaching. Its military, insider-threat  
 and political aspects are thus similar in many ways to traditional security and  
 intelligence work. The critical difference is that the defenders are a loose coali-  
 tion of NGOs, park rangers and law-enforcement agencies. There isn’t a central  
 bureaucracy to manage classiﬁcations, clearances and counterintelligence.

We had a project with Tanya Berger-Wolf, the leader of Wildbook, an eco-

logical information management system that uses image recognition to match  
 and analyse data collected on animals via tourist photos, camera traps, drones  
 and other data sources [92]. Her idea was that if we could link up the many pho-  
 tographs taken of individual wild animals, we could dramatically improve the  
 science of ecology and population biology, together with the resource manage-  
 ment, biodiversity, and conservation decisions that depend on them. Modern

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 327 | Ross Anderson |

*10.3. PRIVACY FOR TIGERS*

image-recognition software makes this feasible, particularly for large animals  
 with distinctive markings, such as elephants, giraffes and zebras. Wildbook is  
 now deployed for over a dozen species at over a dozen locations.

In 2015, two Spanish citizens were arrested in Namibia’s Knersvlagte nature

reserve with 49 small succulent plants; a search of their hotel room revealed 2000  
 more, of which hundreds were threatened species. It turned out that they sold  
 these plants through a website, had made numerous collecting trips, and found  
 rare specimens via botanical listservs and social networks. They pleaded guilty,  
 paid a $160,000 ﬁne and were banned from the country for life. It turned out that  
 they had also used another citizen-science website, iSpot [2009]. Incidents like  
 this showed that wildlife aggregators need access control, and are also leading  
 to a rethink among botanists, zoologists and others about open data [1166]. So  
 what should the policy be?

What one needs to protect varies by species and location. With rare plants,

|  |
| --- |
| we don’t want thieves to learn the GPS location of even a single specimen. With  endangered Coahuilan box tortoises, we don’t want thieves stealing them from  the wild and selling them as pets with false documents claiming they were bred  in captivity. There, the goal is a public database of all known tortoises, and  conservators are busy photographing all the wild specimens in their range, a  360 km2 region of Mexico. This will enable the US Fish and Wildlife Service to  check shipments. With the snow leopard, Wildbook had three years of camera-  trap data from one Nepal conservancy, and wanted a security policy to help this  scale to ﬁve locations in Nepal, India and Pakistan. This is a Red List species  with only a few hundred individuals in each of these three countries. In Africa  the picture is similar; Wildbook started out by tracking zebras, of which the  Gr´evy’s zebra is endangered. Animals cross borders between mutually suspicious  countries, and tourists post tagged photos despite leaﬂets and warnings that  they should not geotag [2074]. Some tourists simply don’t know how to turn  o↵ tagging; some are so dumb they get out of their cars and get eaten. The  protection requirements also vary by country; in Namibia the authorities are  keen to stop tourists posting tagged photos of rhino, while in Kenya the rhinos  all have their own armed guards and the authorities are less bothered. |

The new wildlife aggregation sites can use image recognition to identify in-

dividual animals and link up sightings into location histories; other machine-  
 learning techniques then aggregate these histories into movement models. We  
 rapidly ﬁnd sensitive outputs, such as which waterhole has lots of leopards, or  
 which island has lots of breeding whales. This is one of the ways animal privacy  
 differs from the human variety: highly abstracted data are often more sensitive  
 rather than less. In effect, our machine-learning models acquire the ‘lore’ that  
 an individual ranger might learn after a decade working at a conservancy. As  
 such individuals make the best poachers if they go over to the dark side, we  
 need to keep models that learn their skills out of the poachers’ hands. And we  
 need to be smart about sensitivity: it’s not enough to protect only the data and  
 movement models of snow leopards, if a poacher can also track them by tracking  
 the mountain goats that they eat.

Our primary protection goal is to not give wildlife criminals actionable in-

telligence, such as “an animal of species A is more likely to be at location X at  
 time T”. In particular, we don’t want the citizen-science data platforms we build

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 328 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

to make the situation worse. Our starting point is to use an operations-research  
 model as a guide to derive access rules for (a) recent geotagged photos, (b) pre-  
 dictive models and (c) photo collections. And we need to be able to tweak the  
 rules by species and location.

There are four levels of access. The core Wildbook team maintains the soft-

ware and has operational access to almost everything; we might call this level  
 zero. At level one are the admins of whom there might be maybe 20 per species;  
 as access control is delegated there will be further admins per conservancy or per  
 reserve. At level two are hundreds of people who work for conservancies collect-  
 ing and contributing data, and who at present are sort-of known to Wildbook;  
 as the system scales up, we need to cope with delegated administration. At  
 level three there are thousands of random citizens who contribute photos and  
 are rewarded with access to non-sensitive outputs. Our threat model is that  
 the set of citizen scientists at level 3 will always include poachers; the set of  
 conservancy staff at level 2 will include a minority who are careless or disloyal;  
 and we hope that the level 1 admins usually won’t be in cahoots with poachers.

The focus of our insider threat mitigation is conservancy staff who may be

tempted to defect. Given that conservancies often operate in weak states, the  
 threat of eventual detection and imprisonment can seem remote. The most pow-  
 erful deterrent available is the social pressure from conservancy peers: loyalty  
 to colleagues, a sense of teamwork and a sense of mission. The task is to ﬁnd  
 a technical means of supporting group cohesion and loyalty. The civil-service  
 approach of having a departmental security officer who looks over everyone’s  
 shoulder all the time is not feasible anyway in a ﬁnancially-stretched conser-  
 vancy employing ten or twenty people on low wages in less-developed country  
 (LDC) conditions.

The problem is not just one of providing analytics so that we can alarm if a

member of staff starts looking at lots of records of rhino, or lots of records at  
 a Serengeti waterhole. We already have admins per species and per location.  
 The problem is motivating people to pay attention and take action. Our core  
 strategy is local public auditability for situational awareness and deterrence,  
 based on two-dimensional transparency. All conservancy staff are in at least  
 one group, relating to the species of interest to them or the park where they  
 work. Staff in the rhino group therefore see who’s been looking at rhino records  
 – including individual sighting records and models – while staff working in the  
 Serengeti see who’s interested in data and models there. In effect it’s a matrix  
 system for level 2 staff; you get to see Serengeti rhinos if you’re there or if you’re  
 a rhino expert, and in either case you share ﬁrst-line responsibility for vigilance.  
 Level 1 staff can enrol level 2 staff and make peering arrangements with other  
 conservancies, but their relevant actions are visible to level 2 colleagues. We  
 will have to see how this works in the ﬁeld.

**10.4** **Health record privacy**

Perhaps the most complex and instructive example of security policies where  
 access control supports privacy is found in clinical information systems. The  
 healthcare sector spends a much larger share of national income than the mili-

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 329 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

tary in all developed countries, and although hospitals are still less automated,  
 they are catching up fast. The protection of medical information is thus an  
 important case study for us all, with many rich and complex tradeoffs.

Many countries have laws regulating healthcare safety and privacy, which

help shape the health IT sector. In the USA, the Health Insurance Portability  
 and Accountability Act (HIPAA) was passed by Congress in 1996 following  
 a number of privacy failures. In one notorious case, a convicted child rapist  
 working as an orthopedic technician at Newton-Wellesley Hospital in Newton,  
 Massachusetts, was caught using a former employee’s password to go through  
 the records of 954 patients (mostly young females) to get the phone numbers  
 of girls to whom he then made obscene phone calls [317]. He ended up doing  
 jail time, and the Massachusetts senator Edward Kennedy was one of HIPAA’s  
 sponsors.

The HIPAA regulations have changed over time. The ﬁrst set, issued by the

Clinton administration in December 2000, were moderately robust, and based  
 on assessment of the harm done to people who were too afraid to seek treat-  
 ment in time because of privacy concerns. In the run-up to the rulemaking,  
 HHS estimated that privacy concerns led 586,000 Americans to delay seeking  
 cancer treatment, and over 2 million to delay seeking mental health treatment.  
 Meanwhile, over 1 million simply did not seek treatment for sexually transmit-  
 ted infections [873]. In 2002, President Bush rewrote and relaxed them to the  
 ‘Privacy Rule’; this requires *covered entities* such as hospitals and insurers to  
 maintain certain security standards and procedures for *protected health informa-*  
 *tion* (PHI), with both civil and criminal penalties for violations (although very  
 few penalties were imposed in the ﬁrst few years). The rule also gave patients  
 the right to demand copies of their records. Covered entities can disclose infor-  
 mation to support treatment or payment, but other disclosures require patient  
 consent; this led to complaints by researchers. The privacy rule was followed  
 by further ‘administrative simpliﬁcation’ rules in 2006 to promote healthcare  
 systems interoperability. This got a further boost when President Obama’s

stimulus bill allocated billions of dollars to health IT, and slightly increased  
 the penalties for privacy violations; in 2013 his administration extended the  
 rules to the business associates of covered entities. But grumbling continues.  
 Health privacy advocates note that the regime empowered health data holders  
 to freely and secretly aggregate and broker protected health information, while  
 hospitals complain that it adds to their costs and patient advocates have been  
 complaining for over a decade that it’s often used by hospital staff as an excuse  
 to be unhelpful – such as by preventing people tracing injured relatives [827].  
 Although HIPAA regulation gives much less privacy than in Europe, it is still  
 the main driver for information security in healthcare, which accounts for over  
 10% of the U.S. economy. Another driver is local market effects: in the USA,  
 for example, systems are driven to some extent by the need to generate billing  
 records, and the market is also concentrated with Epic having a 29% market  
 share for electronic medical record systems in 2019 while Cerner had 26% [1351].

In Europe, data-protection law sets real boundaries. In 1995, the UK govern-

ment attempted to centralise all medical records, which led to a confrontation  
 with the doctors’ professional body, the British Medical Association (BMA).  
 The BMA hired me to devise a policy for safety and privacy of clinical informa-

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 330 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

tion, which I’ll discuss later in this chapter. The evolution of medical privacy  
 over the 25 years since is a valuable case study; it’s remarkable how little the  
 issues have changed despite the huge changes in technology.

Debates about the safety and privacy tradeoffs involved with medical infor-

mation started around this time in other European countries too. The Germans  
 put summary data such as current prescriptions and allergies on the medical  
 insurance card that residents carry; other countries held back, reasoning that  
 if emergency data are moved from a human-readable MedAlert bracelet to a  
 smartcard, this could endanger patients who fall ill on an airplane or a foreign  
 holiday. There was a series of scandals in which early centralised systems were  
 used to get information on celebrities. There were also sharp debates about  
 whether people could stop their records being used in research, whether out of  
 privacy concerns or for religious reasons – for example, a Catholic woman might  
 want to forbid her gynaecological records being sold to a drug company doing  
 research on abortion pills.

European law around consent and access to records was clariﬁed in 2010 by

the European Court of Human Rights in the case I v Finland. The complainant  
 was a nurse at a Finnish hospital, and also HIV-positive. Word of her condition  
 spread among colleagues, and her contract was not renewed. The hospital’s

access controls were not sufficient to prevent colleagues accessing her record,  
 and its audit trail was not sufficient to determine who had compromised her  
 privacy. The court’s view was that health care staff who are not involved in  
 the care of a patient must be unable to access that patient’s electronic medical  
 record: “What is required in this connection is practical and effective protection  
 to exclude any possibility of unauthorised access occurring in the ﬁrst place.”  
 This judgment became ﬁnal in 2010, and since then health providers have been  
 supposed to design their systems so that patients can opt out effectively from  
 secondary uses of their data.

**10.4.1** **The threat model**

The appropriate context to study health IT threats is not privacy alone, but  
 safety and privacy together. The main objective is safety, and privacy is often  
 subordinate. The two are also intertwined, though in many ways.

There are various hazards with medical systems, most notably safety us-

ability failures, which are reckoned to kill about as many people as road traffic  
 accidents. I will discuss these issues in Part 3 in the chapter on System Eval-  
 uation and Assurance. They interact directly with security; vulnerabilities are  
 particularly likely to result in the FDA mandating recalls of products such as  
 infusion pumps. The public are much more sensitive to safety issues if they have  
 a security angle; we have much less tolerance of hostile action than of impersonal  
 risk.

A second hazard is that loss of conﬁdence in medical privacy causes people

to avoid treatment, or to seek it too late.

1. The most comprehensive data were collected by the US Department of

Health and Human Services prior to the HIPAA rulemaking under Pres-

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 331 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

ident Clinton. HHS estimated that privacy concerns led 586,000 Ameri-  
 cans to delay seeking cancer treatment, and over 2 million to delay seeking  
 mental health treatment. Meanwhile, over 1 million simply did not seek  
 treatment for sexually transmitted infections [873];

2. The Rand corporation found that over 150,000 soldiers who served in Iraq

and Afghanistan failed to seek treatment for post-traumatic stress disorder  
 (PTSD), which is believed to contribute to the suicide rate among veterans  
 being about double that of comparable civilians – a signiﬁcant barrier  
 being access to conﬁdential treatment [1861];

3. The most authoritative literature review concluded that many patients,

particularly teenagers, gay men and prostitutes, withheld information  
 or simply failed to seek treatment because of conﬁdentiality concerns.  
 Anonymised HIV testing more than doubled the testing rate among gay  
 men [1650].

So poor privacy is a safety issue, as well as a critical factor in providing equal

healthcare access to a range of citizens, from veterans to at-risk and marginalised  
 groups. The main privacy threat comes from insiders, with a mix of negligence  
 and malice, in roughly three categories:

1. There are targeted attacks on speciﬁc individuals, ranging from creepy

doctors looking up the records of a date on a hospital computer, to jour-  
 nalists stalking a politician or celebrity. These cause harm to individuals  
 directly;

2. There are bulk attacks, as where governments or hospitals sell millions

of records to a drug company, sometimes covertly and sometimes with  
 the claim that the records have been ‘anonymised’ and are thus no longer  
 personal health information;

3. Most of the reported breaches are accidents, for example where a doctor

leaves a laptop on a train, or when a misconﬁgured cloud server leaves mil-  
 lions of people’s records online [767]. These are reported at ﬁve times the  
 rate of breaches at private ﬁrms, as healthcare providers have a reporting  
 duty. Sometimes accidental leaks lead to opportunistic attacks.

The resulting press coverage, which is mostly of bulk attacks and accidents,

causes many to fear for the privacy of their health data, although they may not  
 be directly at risk. The bulk attacks also offend many people’s sense of justice,  
 violate their autonomy and agency, and undermine trust in the system.

So how big is the direct risk? And how much of the risk is due to technology?

As things get centralised, we hit a fundamental scaling problem. The likelihood  
 that a resource will be abused depends on its value and on the number of  
 people with access to it. Aggregating personal information into large databases  
 increases both these risk factors at the same time. Over the past 25 years,

we’ve moved from a world in which each doctor’s receptionist had access to  
 maybe 5,000 patients’ records in a paper library or on the practice PC, to one  
 in which the records of thousands of medical practices are hosted on common

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 332 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

platforms. Some shared systems give access to data on many patients and have  
 been abused. This was already a concern 25 years ago as people started building  
 centralised systems to support emergency care, billing and research, and it has  
 become a reality since. Even local systems can expose data at scale: a large  
 district hospital is likely to have records on over a million former patients. And  
 privacy issues aren’t limited to organizations that treat patients directly: some  
 of the largest collections of personal health information are in the hands of  
 health insurers and research organizations.

To prevent abuses scaling, lateral information ﬂow controls are needed. Early

hospital systems that gave all staff access to all records led to a number of privacy  
 incidents, of which the most notable was the one that led to the I v Finland  
 judgment of the European court; but there were similar incidents in the UK  
 going back to the mid-1990s. All sorts of ad-hoc privacy mechanisms had been  
 tried, but by the mid-1990s we felt the need for a proper access control policy,  
 thought through from ﬁrst principles and driven by a realistic model of the  
 threats.

**10.4.2** **The BMA security policy**

By 1995, most medical practices had computer systems to keep records; the  
 suppliers were small ﬁrms that had often been started by doctors whose hobby  
 was computing rather than golf or yachting, and they were attuned to doctors’  
 practical needs. Hospitals had central administrative systems to take care of  
 billing, and some were moving records from paper to computers. There was  
 pressure from the government, which pays for about 90% of medical care in  
 Britain through the National Health Service; officials believed that if they had  
 access to all the information, they could manage things better, and this caused  
 tension with doctors who cared about professional autonomy. One of the last  
 things done by Margaret Thatcher’s government, in 1991, had been to create  
 an ‘internal market’ in the health service where regional commissioners act like  
 insurers and hospitals bill them for treatments; implementing this was a work in  
 progress, both messy and contentious. So the Department of Health announced  
 that it wanted to centralise all medical records. The Internet boom had just  
 started, and medics were starting to send information around by private email;  
 enthusiasts were starting to build systems to get test results electronically from  
 hospitals to medical practices. The BMA asked whether personal health infor-  
 mation should be encrypted on networks, but the government refused to even  
 consider this (the crypto wars were getting underway; see 26.2.7.3 for that story).  
 This was the last straw; the BMA realised they’d better get an expert and asked  
 me what their security policy should be. I worked with their staff and members  
 to develop one.

We rapidly hit a problem. The government strategy assumed a single elec-

tronic patient record (EPR) that would follow the patient around from concep-  
 tion to autopsy, rather than the traditional system of having different records  
 on the same patient at different hospitals and doctors’ offices, with information  
 ﬂowing between them in the form of referral and discharge letters. An attempt  
 to devise a security policy for the EPR that would observe existing ethical norms  
 became unmanageably complex [821], with over 60 rules. Different people have

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 333 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

access to your record at different stages of your life; your birth record is also  
 part of your mother’s record, your record while you’re in the army or in jail  
 might belong to the government, and when you get treatment for a sexually  
 transmitted disease you may have the right to keep that completely private.

The Department of Health next proposed a multilevel security policy: sex-

ually transmitted diseases would be at a level corresponding to Secret, normal  
 patient records at Conﬁdential and administrative data such as drug prescrip-  
 tions and invoices at Restricted. But this was obviously a non-starter. For

example, how should a prescription for anti-retroviral drugs be classiﬁed? As  
 it’s a prescription, it should be Restricted; but as it identiﬁes a person as HIV  
 positive, it should be Secret. It was wrong in all sorts of other ways too; some  
 people with HIV are open about their condition while others with minor con-  
 ditions are very sensitive about them. Sensitivity is a matter for the patient  
 to decide, not the Prime Minister. Patient consent is central: records can only  
 be shared with third parties if the patient agrees, or in a limited range of legal  
 exceptions, such as contact tracing for infectious diseases like TB.

Medical colleagues and I realised that we needed a security context with

ﬁner granularity than a lifetime record, so we decided to let existing law and  
 practice set the granularity, then build the policy on that. We deﬁned a record  
 as the maximum set of facts to which the same people have access: patient +  
 doctor, patient + doctor plus surgery staff, patient + patient’s mother + doctor  
 + staff, and so on. So a patient will usually have more than one record, and  
 this offended the EPR advocates.

A really hard problem was the secondary use of records. In the old days,

|  |
| --- |
| this meant a researcher or clinical auditor sitting in the library of a hospital or  medical practice, patiently collecting statistics; consent consisted of a notice in  the waiting room saying something like ‘We use our records in medical research  to improve care for all; if you don’t want your records used in this way, please  speak to your doctor.’ By 1995, we’d already seen one company o↵ering sub-  sidised computers to General Practitioners (GPs)1 in return for allowing remote  queries by drug companies to return supposedly anonymous data. |

The goals of the BMA security policy were therefore to enforce the principle

of consent, and to prevent too many people getting access to too many records.  
 It did not try to do anything new, but merely to codify existing best practice,  
 and to boil it down into a page of text that everyone – doctor, engineer or  
 administrator – could understand.

Starting from these principles and insights, we proposed a policy of nine

principles.

1. Access control: each identiﬁable clinical record shall be marked with an

access control list naming the people who may read it and append data  
 to it.

2. Record opening: a clinician may open a record with herself and the patient

1Britain’s GPs are the equivalent of family doctors in the USA; they have historically

acted as gatekeepers to the system and as custodians of each patient’s lifetime medical record.  
 They also act as the patient’s advocate and join up care between medical practice, hospital  
 and community. This helps keeps healthcare costs down in the UK, compared with the USA.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 334 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

on the access control list. Where a patient has been referred, she may  
 open a record with herself, the patient and the referring clinician(s) on  
 the access control list.

3. Control: One of the clinicians on the access control list must be marked

as being responsible. Only she may alter the access control list, and she  
 may only add other health care professionals to it.

4. Consent and notiﬁcation: the responsible clinician must notify the patient

of the names on his record’s access control list when it is opened, of all  
 subsequent additions, and whenever responsibility is transferred. His con-  
 sent must also be obtained, except in emergency or in the case of statutory  
 exemptions.

5. Persistence: no-one shall have the ability to delete clinical information

until the appropriate time period has expired.

6. Attribution: all accesses to clinical records shall be marked on the record

with the subject’s name, as well as the date and time. An audit trail must  
 also be kept of all deletions.

7. Information ﬂow: Information derived from record A may be appended to

record B if and only if B’s access control list is contained in A’s.

8. Aggregation control: there shall be effective measures to prevent the ag-

gregation of personal health information. In particular, patients must

receive special notiﬁcation if any person whom it is proposed to add to  
 their access control list already has access to personal health information  
 on a large number of people.

9. Trusted computing base: computer systems that handle personal health

information shall have a subsystem that enforces the above principles in  
 an effective way. Its effectiveness shall be subject to evaluation by inde-  
 pendent experts.

From the technical viewpoint, this policy is strictly more expressive than

the Bell-LaPadula model of the last chapter, as it contains an information ﬂow  
 control mechanism in principle 7, but also contains state. In fact, it takes

compartmentation to the logical limit, as there are more compartments than  
 patients. A discussion for a technical audience can be found at [59]. The full  
 policy dealt with a lot more issues, such as access to records by vulnerable  
 patients who might be coerced [58].

Similar policies were developed by other medical bodies including the Swedish

and German medical associations; the Health Informatics Association of Canada,  
 and an EU project (these are surveyed in [1077]). The BMA model was adopted  
 by the Union of European Medical Organisations (UEMO) in 1996, and feedback  
 from public consultation on the policy can be found in [60].

**10.4.3** **First practical steps**

Feedback from the ﬁeld came from a pilot implementation in a medical prac-  
 tice [870], which was positive, and from a hospital system developed in Hastings,

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 335 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

which controlled access using a mixture of roles and capabilities, rather than the  
 ACLs in which the BMA model was expressed. It turned out that the practical  
 way to do access control at hospital scale was by rules such as ‘a ward nurse  
 can see the records of all patients who have within the previous 90 days been  
 on her ward’, ‘a junior doctor can see the records of all patients who have been  
 treated in her department’, and ‘a senior doctor can see the records of all pa-  
 tients, but if she accesses the record of a patient who has never been treated in  
 her department, then the senior doctor responsible for that patient’s care will  
 be notiﬁed’2.

The technical lessons learned are discussed in [535, 536, 870]. With hindsight,

the BMA model was a lossless compression of what doctors said they did while  
 the role-based model was a slightly lossy version but which implemented what  
 hospitals do in practice and worked well in that context. One of the BMA

rules, though, created difficulty in both contexts: the desire for a small trusted  
 computing base. GPs ended up having to trust all the application code that  
 they got from their suppliers, and while they could inﬂuence its evolution, there  
 was no useful trusted subset. The hospital records system was much worse: it  
 had to rely on the patient administrative system (PAS) to tell it which patients,  
 and which nurses, are on which ward. The PAS was ﬂaky and often down,

so it wasn’t acceptable to make a safety-critical system depend on it. The

next iteration was to give each hospital staff member a smartcard containing  
 credentials for their departments or wards.

The policy response from the Department of Health was to set up a com-

mittee of inquiry under Dame Fiona Caldicott. She acknowledged that some  
 60 established ﬂows of information within the NHS were unlawful, and rec-  
 ommended the appointment of a responsible privacy officer in each healthcare  
 organisation [367]. This was at least a start, but it created a moral hazard:  
 while the privacy officer, typically a senior nurse, was blamed when things went  
 wrong, the actual policy was set by ministers – leading to the classic security-  
 economics gotcha we discussed in chapter 8, of Bob guarding the system while  
 Alice pays the cost of failure. Anyway, the government changed, and the new  
 administration of Tony Blair went for a legal rather than a technical ﬁx – with  
 a data-protection law that allowed data controllers to pretend that data were  
 anonymous so long as they themselves could not re-identify them, even if others  
 could re-identify them by matching them with other data3. We will discuss the  
 limits of anonymisation in the following chapter.

**10.4.4** **What actually goes wrong**

In his second term as Prime Minister, Tony Blair announced a £6bn plan to  
 modernise health service computing in England. The National Programme for  
 IT (NPfIT), as it came to be known, turned out to be the world’s most expensive

2The Hastings system was initially designed independently of the BMA project. When we

learned of each other we were surprised at how much our approaches coincided, and reassured  
 that we had captured the profession’s expectations in a reasonably consistent way.

3The UK law was supposed to transpose the EU Data Protection Directive (95/46/EC)

into UK law to provide a level playing ﬁeld on privacy; this loophole was one of several that  
 allowed UK ﬁrms a lot of wriggle room, annoying the French and Germans [597]. The EU  
 eventually pushed through the stricter General Data Protection Regulation (2016/679).

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 336 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

civilian IT disaster. After David Cameron came to power in 2010, an inquiry  
 from the National Audit Office noted of a total expenditure of about £10bn,  
 some £2bn spent on broadband networking and digital X-ray imaging resulted  
 in largely working systems, while the rest didn’t give value for money, and the  
 core aim that every patient should have an electronic care record would not be  
 achieved [1390]. Cameron formally killed the project, but its effects continued  
 for years because of entrenched supplier contracts, and health IT was held up  
 for a decade [1559].

NPfIT had called for all hospital systems to be replaced during 2004–2010

with standard ones, to give each NHS patient a single electronic care record.  
 The security policy had three main mechanisms.

1. There are role-based access controls like those pioneered at Hastings.

2. In order to access patient data, a staff member also needs a *legitimate*

*relationship*. This abstracts the Hastings idea of ‘her department’.

3. There was a plan that patients would be able to seal certain parts of their

records, making them visible only to a particular care team. However, the  
 providers never got round to implementing this. It wasn’t consistent with  
 the doctrine of a single electronic health record, which had been repeated  
 so often by ministers that it had become an article of religious faith. As  
 late as 2007, Parliament’s Health Committee noted that suppliers hadn’t  
 even got a speciﬁcation yet [925].

As a result, patients receiving outpatient psychiatric care at a hospital found

that the receptionist could see their case notes. Formerly, the notes were kept  
 in paper in the psychiatrist’s ﬁling cabinet; all the receptionist got to know was  
 that Mrs Smith was seen once a month by Dr Jones. But now the reception-  
 ist role had to be given access to patient records so that they could see and  
 amend administrative data such as appointment times; and everyone working  
 reception in the hospital wing where Dr Jones had his office had a legitimate  
 relationship. So they all got access to everything. This illustrates why the doc-  
 trine of a single record with a single security context per patient was a bad idea.  
 Thanks to project mismanagement, less than ten percent of England’s hospitals  
 actually installed these systems, though the doctrine of ‘RBAC + relationship’  
 has affected others since. It now looks like the failure to support multiple secu-  
 rity contexts per patient is about to become an issue in the USA as ﬁrms start  
 pushing health apps supported by the FIHR standard, to which I’ll return in  
 section 10.4.5.

**10.4.4.1** **Emergency care**

The next thing to go wrong was emergency medical records. One of the stories  
 used by politicians to sell NPfIT had been ‘Suppose you fall ill in Aberdeen and  
 the hospital wants access to your records in London ...’. This was, and remains,  
 bogus. Paramedics and emergency-room physicians are trained to treat what  
 they see, and assume nothing; the idea that they’d rely on a computer to tell the  
 blood group of an unconscious patient is simply daft. But policy was policy, and

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 337 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

in Scotland the government created an ‘emergency care record’ of prescriptions  
 and allergies that is kept on a central database for use by emergency room clin-  
 icians, paramedics and the operators of out-of-hours medical helpline services.  
 Sensitive information about 2.5 million people was made available to tens of  
 thousands of people, and the inevitable happened; one doctor of Queen Mar-  
 garet Hospital in Dunfermline was arrested and charged for browsing the health  
 records of then Prime Minister Gordon Brown, First Minister Alex Salmond and  
 various sports and TV personalities. The case was eventually dropped as ‘not  
 in the public interest’ to prosecute [1741]. Patients had been offered the right  
 to opt out of this system, but it was a very odd opt-out: if you did nothing,  
 your data were collected from your GP and made available to the Department of  
 Health in Edinburgh and also to the ambulance service. If you opted out, your  
 data were still collected from your GP and made available to the Department  
 of Health; they just weren’t shared with the ambulance crew.

This was also policy in England where it was called ‘consent-to-view’: the

state would collect everything and show users only what they were allowed to  
 see. Everybody’s records would be online, and doctors would only be allowed  
 to look at them if they claimed the patient had consented. Officials assured  
 Parliament that this was the only practical way to build NPfIT; they described  
 this as ‘an electronic version of the status quo’ [925]. The English emergency  
 system, the Summary Care Record (SCR), also has sensitive data on most cit-  
 izens, is widely accessible, but is little used; if you end up in an ambulance,  
 they’ll take a medical history from you en route to hospital, just as they always  
 have4. Something similar also happened in the Netherlands, where a database  
 of citizens’ medical insurance details ended up being accessible not just by doc-  
 tors and pharmacists but alternative healers and even taxi ﬁrms, with entirely  
 predictable results [186].

**10.4.4.2** **Resilience**

The move to centralised systems typically makes failures rarer but larger, and  
 health systems are no exception. The NPfIT’s only real achievement was to  
 standardise all X-ray imaging in England using digital machines and cloud stor-  
 age. An early warning of fragility came on 11th December 2005, when a leak  
 of 250,000 litres of petrol at the Bunceﬁeld oil storage depot formed a vapour  
 cloud and detonated – the largest peacetime explosion in Europe. Oil compa-  
 nies were later ﬁned millions of pounds for safety breaches. Our local hospital  
 lost X-ray service as both the primary and backup network connections to the  
 cloud service passed nearby. A further warning came when the Wannacry worm  
 infected machines at another nearby hospital in 2017; managers foolishly closed  
 down the network, in the hope of preventing further infection, and then found  
 that they had to close the emergency room and send patients elsewhere. With  
 no network they could do no X-rays (and get no pathology test results either,  
 even from the hospital’s own lab). There have been further incidents of hospitals  
 closed by ransomware since, particularly in the USA.

4In the coronavirus crisis, the SCR was ‘enriched’ by adding a lot of data from the GP

record, making it available to planners, and making it opt-out by default. It’s still not clear  
 that any worthwhile use has been made of it.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 338 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

**10.4.4.3** **Secondary uses**

Databases relating to payment usually don’t allow a real opt-out, and the UK  
 example is the Hospital Episode Statistics (HES) database, which collects bills  
 sent by hospitals to the commissioning bodies that pay them, and has exten-  
 sive information on every state-funded hospital visit and test in England and  
 Wales since 1998 – about a billion records in total5. These records have proved  
 impossible to protect, not just because anonymisation of complete records is im-  
 practical but because of the intense political pressure for access by researchers.  
 More and more people had got access under the 1997–2010 Labour government;  
 and after David Cameron became Prime Minister in 2010, the ﬂoodgates opened.  
 Cameron hired a ‘transparency tsar’ who’d previously run a health IT business,  
 and announced ‘Open Data measures’ in 2011 which the goal that every NHS  
 patient would be a research patient, in order to make Britain a world leader in  
 pharmaceutical research. Officials claimed that ‘All necessary safeguards would  
 be in place to ensure protection of patients’ details – the data will be anonymised  
 and the process will be carefully and robustly regulated’ [1807]. Anonymisation  
 meant that your personal details were redacted down to your postcode and date  
 of birth; this is quite inadequate, as we’ll discuss in the next chapter.

In 2013 the government announced that records would also be harvested

from GP systems; GPs were given eight weeks to inform their patients of  
 the impending upload. This caused enough disquiet that privacy campaign-

ers, GPs and others got together to set up a medical privacy campaign group,  
 medConfidential.org. The initial impetus was consent, and in particular that  
 patients who tried to exercise their European-law rights to opt out of such  
 systems have ended up being ignored or even de-registered from the health  
 service. Campaigners pushed for the government to obey the newly clariﬁed  
 European law on consent; the government wriggled and evaded. How could

doctors’ bonuses be calculated if some of their records could not be uploaded?

In January 2014, some digging revealed that the HES data had been sold to

over 1000 drug companies, universities and others round the world – often in the  
 form of a set of DVDs containing a billion episodes going back to 1998. A medic  
 revealed that the data had appeared online; it was quickly taken down [1800].  
 This ‘care.data’ scandal, as it became known after the proposal to collect all  
 the GP data, went mainstream. Surveys show that most people are prepared  
 to let their data be used in academic research, so long as they’re asked; but  
 most are not prepared to share it with for-proﬁt researchers, and most object  
 to having it simply taken. On inspection, it turned out to be easy to re-identify  
 patients, even if their postcode and date of birth had not been included in the  
 dataset; we’ll discuss the technical details in the following chapter. There was  
 a ﬁnancial scandal: despite ministers talking of the huge value of research data  
 to the health service, the data had been sold on a cost-recovery basis, for a

5HES is advertised as ‘a data warehouse containing details of all admissions, outpatient

appointments and A and E attendances at NHS hospitals in England’ including private and  
 foreign patients treated at NHS hospitals, and treatments at private hospitals for which the  
 NHS pays. It is now claimed that ‘We apply a strict statistical disclosure control in accordance  
 with the NHS Digital protocol, to all published HES data. This suppresses small numbers to  
 stop people identifying themselves and others, to ensure that patient conﬁdentiality is main-  
 tained.’. See https://digital.nhs.uk/data-and-information/data-tools-and-services/

data-services/hospital-episode-statistics.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 339 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

few thousand dollars a set. There was also an issue of jurisdiction: it turned  
 out that PA Consulting had loaded the HES data to a Google cloud system for  
 resale to its clients, as at 20Gb it was too big for Excel.

But hang on, said members of parliament, how can that be legal? Google

didn’t have any data centres in the UK, and there are all sorts of regulations  
 against taking NHS data overseas [1573]. Also, officials had promised that UK  
 data wouldn’t be sold overseas, yet they were advertised in the USA; and it  
 turned out that even the regulator, the Medicines and Healthcare Products  
 Regulatory Agency (MHRA)6, had been selling personal data [1645]. Ministers  
 went into damage-containment mode; the privacy regulator was persuaded to  
 believe that the exported data were anonymous enough, and the UK website  
 of a ﬁrm claiming to be able to identify patients from these records was taken  
 offline [1574]. Ministers talked of lessons being learned, and a review of all data  
 releases was commissioned; but when this appeared, it only investigated whether  
 internal guidelines had been followed, not whether they were legal [1496].

UK health privacy scandals have continued at the rate of about once a year

since then:

*•* In 2015, Google Deepmind obtained a copy of all the 1.6m patient records  
 velop an app to detect acute kidney injury (it took all the records, not  
 just those of kidney patients). Patient consent was not sought, the deal  
 was later found to be unlawful, and when the app was developed using  
 US data obtained from the VA instead, it was unimpressive [1541]. The  
 Information Commissioner reprimanded the hospital but failed to order  
 Google Deepmind to delete the data. Eventually Deepmind transferred  
 the records to Google, contrary to previous assurances [1281].

*•* Also in 2015, a tabloid newspaper discovered the online pharmacy Phar-  
 cluding lottery fraudsters who targeted unwell elderly men and a health-  
 care supplement vendor that had already been sanctioned for mislead-  
 ing advertising and unauthorised health claims [662]. The ﬁrm was ﬁned  
 £130,000 and its commercial director suspended by the General Pharma-  
 ceutical Council. A major backer, the UK’s largest GP software supplier  
 EMIS, sold its shareholding.

*•* SCR data were also sold to Boots, a high-street pharmacy chain that pres-

*•* In 2017, leading GP software supplier TPP which has 6,000 customers  
 records on 26 million patients – switched on ‘enhanced data sharing’ so  
 that records could be seen by doctors at local hospitals. It was soon noticed  
 that records could be seen at all other practices that were TPP customers;  
 GPs had not been aware of this [577]. The records were also visible to

6The MHRA had also been a lot less keen about making data about adverse clinical trial

results available to medics who wanted it. The essence of the complaint against it was that it  
 acted more in the interests of the drug companies and medical device makers rather than in  
 the interest of patients, becoming in effect a captured regulator.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 340 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

TPP customers in care homes, prisons and immigration detention centres.  
 TPP failed to answer questions about whether any of its customers in  
 India, China and the UAE had access.

*•* In 2018, the records of all 180,000 lung cancer patients diagnosed in Eng-  
 England, which had claimed that cancer registry data would only be sold  
 for a ‘medical purpose’.

Standard central systems do have real advantages. In the USA, the Veterans’

Administration runs such systems for its hospital network; after Hurricane Ka-  
 trina, veterans from Louisiana who’d ended up as refugees in Texas or Florida,  
 or even Minnesota, could go straight to local VA hospitals and ﬁnd their notes  
 there at the doctor’s ﬁngertips, when patients of many other hospitals in New  
 Orleans lost their notes altogether.

But there have also been controversies in the USA. In November 2019, it

emerged that Google had done an outsourcing deal to process the medical  
 records of 50 million Americans on behalf of Ascension, and a whistleblower  
 revealed that the data were not even being lightly de-identiﬁed; staff at both  
 Google and Ascension had full access to patient data. A federal inquiry was  
 started into whether the arrangement was HIPAA compliant [121].

Google also got VA data from the USA, which it used in place of the Lon-

don data once the ICO ruled against it there. With a few such exceptions in  
 egregious cases, policymakers ﬁnd it hard to resist lobbying from marketers and  
 researchers for access. The EU General Data Protection Regulation has a con-  
 venient exemption for ‘research’, put there by the pharma lobby, which doesn’t  
 exclude market research. And, of course, law enforcement and intelligence agen-  
 cies demand access. This started off in the 1990s with the collection of opiate  
 prescribing records and has greatly expanded.

**10.4.5** **Conﬁdentiality – the future**

What can we say about healthcare privacy now, almost a quarter of a century  
 after the BMA policy? Well, some things change, but a surprising number of  
 things stay the same. We noted in chapter 2 that the cybercrime ecosystem had  
 not been changed much by the huge technological changes of the past decade;  
 much the same holds for the health privacy ecosystem. The move to cloud-  
 based medical records is hard to resist as it saves individual care providers the  
 trouble and expense of maintaining servers and backups. The move to ever more  
 complex outsourcing also seems inexorable; we can expect that specialist ﬁrms  
 will handle X-ray images, pathology tests and the like, while subject specialists  
 will support care for speciﬁc diseases such as diabetes.

Since 2014, there has emerged a draft standard for Fast Healthcare Interop-

erability Resources (FHIR, pronounced ‘ﬁre’) which describes how two systems  
 talk to each other, once you’ve allowed them to do so. The security engineering  
 is outside this standard; Deepmind’s smartphone apps, for example, use OAuth  
 2. FIHR has been mandated in the NHS from 2021. In America, new federal  
 information-sharing rules may require providers to send your record to third-

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 341 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

party apps, like Apple’s Health Records, after you have authorized the data  
 exchange. The details alarm doctors who note that once you do that you’ll be  
 open to serious abuse, as the data will fall outside HIPAA and the apps can sell  
 it off as they please. Data such as substance abuse could not only limit access  
 to insurance but even be demanded by employers and others. The government  
 responds that opening up health data will enable people to manage their care  
 better and understand costs, while opening the sector up to competitive inno-  
 vation [1815]. Quite apart from whether people would trust Microsoft, Amazon  
 and Google with their health data, you have to share it all or not at all; there  
 is no provision for ﬁner-grained access control than your whole lifetime record.  
 The last 25 years’ experience suggests that this will not be satisfactory.

In the UK, the medical professors and drug companies are having another

push to collect all the GP data, talking about three big new health industries,  
 based on medical records, AI and genomics. Research policy is that while R&D  
 should be 2% of GDP, only a third of that should be from the state and the rest  
 from industry. It was announced in 2019 that ﬁve hospitals had done deals with  
 a pharmaceutical company run by a former minister: they supply ‘anonymised’  
 data for research in return for an equity stake [500]. On the other hand, the UK’s  
 biggest medical-research charity, the Wellcome Trust, is predicting that as many  
 as 40% of patients might opt out of having their data being used in research if  
 there’s another scandal on the scale of care.data. Certainly the data show that  
 while about 80% of people trust doctors with their health data, this falls to just  
 over 50% for health insurers and pharmacies, around 40% for researchers, 20%  
 for drug companies and 10% for tech [1100]. How can we navigate this thicket?

The view of the UK campaign group medConﬁdential is that three things

are needed.

1. First, to enable us to enforce our rights under European law, there must

be real patient consent. This means a single opt-out from secondary uses,  
 rather than the current Facebook-like approach of changing the opt-out  
 mechanisms every year or two and forcing people to opt out all over again.

2. Second, it should not be the patient’s job to defend their data, so both

the privacy architecture and the security engineering must be safe by de-  
 fault. People must not be quietly opted in to secondary data uses that  
 are misdescribed or not mentioned at all; and there must be appropriate  
 security mechanisms about which patients are told the truth, particularly  
 when they fail.

3. Third, there must still be real transparency. At present my GP can see

who has had access to my record, but I want to see too. If tens of millions  
 of patients can audit access, then even if only a few hundred thousand  
 actually do so, this should deter most of the abuse.

History should have taught us that it’s best to be honest with patients. In

the UK we’ve wasted 20 years: a decade with NPfIT and a further decade trying  
 to sell data while pretending not to. Yet hospitals that set out to get positive  
 consent for the use of data in research get it 70–80% of the time, and we have  
 had large collaborative research projects such as UK Biobank where 500,000

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 342 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

people not only consented in 2006–10 to lifetime monitoring but also provided  
 blood samples, so that researchers could sequence their DNA and correlate that  
 with health outcomes. There’s a further research database of 100,000 genomes  
 collected from other patients who consented.

Another development is the OpenSAFELY collaboration, which has been

pioneering rapid analysis of the Covid-19 epidemic by working in situ with the  
 live medical records held by TPP, a large provider of cloud electronic health  
 record services which supports about 40% of GPs in England. They imported a  
 list of death notiﬁcations and were able to analyse mortality not just by age and  
 sex, as in official statistics, but by social deprivation, race, smoking history, body  
 mass index and speciﬁc comorbidities, establishing risk factors over more than  
 17 million patients and over 6,000 deaths over February to April 2020 [2025].  
 They were ﬁrst to establish, for example, that the excess mortality observed in  
 black and Asian patients was signiﬁcantly greater than could be explained by  
 social deprivation alone, The speed and scale of this study were unprecedented  
 and make the case for taking ethically-approved queries directly to the live data  
 and taking away only statistics, rather than abstracting anonymised subsets for  
 offsite use that still carry privacy hazards (as we’ll discuss at length in the next  
 chapter). The privacy risks may be more controllable as there are fewer copies  
 of the data and as patient opt-outs can be enforced. And although this might  
 be seen as a ‘new’ research technique, enabled by the emergence of cloud-based  
 medical records, it’s actually a very old technique. In the days before computers,  
 observational epidemiology meant sitting in the library of a hospital or surgery,  
 sifting through thousands of paper records, looking for diagnoses of interest,  
 and departing after weeks or months of work with statistical tables rather than  
 with identiﬁable personal information.

**10.4.5.1** **Ethics**

So researchers working with health data had better pay attention to ethics.  
 In 2014–5, the Nuffield Bioethics Council commissioned a dozen of us from a  
 variety of backgrounds in tech, genetics, medicine, insurance and ethics to write  
 a detailed report on what happens to medical ethics in a world of cloud-based  
 medical records and pervasive genomics [1600]. Historically, it was a series of  
 ethical abuses in medical research that drove the development of research ethics  
 more generally.

*•* In the Tuskegee syphilis experiment, US doctors studied the progression of  
 they were getting free healthcare. The experiment ran from 1932 to 1972,  
 but even after effective antibiotic treatments became available in 1947,  
 infected men were not treated.

*•* Dr Karl Brandt was Hitler’s personal physician, and ran a euthanasia  
 and the civilians of occupied countries without their consent, as did his  
 colleague Dr Josef Mengele who experimented on twins at Birkenau from  
 1943–5; subjects were often killed and dissected afterwards. Brandt was  
 convicted at the Nuremberg trials and hanged in 1948.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 343 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

*•* In the UK Alder Hey scandal, the press discovered that pathologists were  
 without any kind of consent. Parents discovered that body parts of their  
 dead children had been kept without their knowledge. This did serious  
 damage to public trust and the consequences impaired research in pathol-  
 ogy in the UK. There was a similar scandal in Ireland.

The Nazi doctors’ trial led to the Nuremberg code in 1948, under which

the voluntary and informed consent of subjects is essential. The subject must  
 have the freedom to choose, without deceit or duress, and must be able to  
 exit from the experiment at any time. This led later to the Declaration of

Helsinki on ethics in medical research in 1964, which was revised in 1975 after  
 Tuskegee to incorporate the need for an independent institutional review board  
 or ethics committee, and subsequently in 1983, 1989, 1996, 2000 and 2008.  
 The Declaration is managed by the World Medical Association and is ethically  
 binding on physicians. The Declaration upholds the right of patients to make  
 informed decisions about participation in research, both initially and afterwards.

Until about the mid-1990s, the main ethical debates were related to drug

trials: was it wrong to give placebos to HIV sufferers once effective anti-retroviral  
 drugs existed? And was it ethical to test drugs in less developed countries if  
 their citizens or health services could not afford them? Since then, the growing  
 issues have been informational: is it ethical to use whole populations as subjects  
 in observational epidemiology and research, without giving them a right to opt  
 out? And what are the ethical issues arising from low-cost sequencing of the  
 human genome?

After spending a year considering in detail the history and issues I’ve sum-

marised in this section, we concluded that, when working in such a complex  
 and fast-moving ethical ﬁeld, that holds a lot of promise but is also riven with  
 vested interests and political chicanery, it’s not enough for researchers to hide  
 behind the law or just act in accordance with this year’s government guidelines.  
 A morally reasonable set of expectations should embody four principles. To  
 quote the report:

1. The set of expectations about how data will be used in a data initiative

should be grounded in the principle of respect for persons. This includes  
 recognition of a person’s profound moral interest in controlling others’ ac-  
 cess to and disclosure of information relating to them held in circumstances  
 they regard as conﬁdential.

2. The set of expectations about how data will be used in a data initiative

should be determined with regard to established human rights. This will  
 include limitations on the power of states and others to interfere with the  
 privacy of individual citizens in the public interest (including to protect  
 the interests of others).

3. The set of expectations about how data will be used (or re-used) in a data

initiative, and the appropriate measures and procedures for ensuring that  
 those expectations are met, should be determined with the participation  
 of people with morally relevant interests. This participation should in-  
 volve giving and receiving public account of the reasons for establishing,

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 344 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

conducting and participating in the initiative in a form that is accepted as  
 reasonable by all. Where it is not feasible to engage all those with relevant  
 interests – which will often be the case in practice – the full range of values  
 and interests should be fairly represented.

4. A data initiative should be subject to effective systems of governance and

accountability that are themselves morally justiﬁed. This should include  
 both structures of accountability that invoke legitimate judicial and politi-  
 cal authority, and social accountability arising from engagement of people  
 in a society. Maintaining effective accountability must include effective  
 measures for communicating expectations and failures of governance, ex-  
 ecution and control to people affected and to the society more widely.

In short, you have to treat people as ends rather than means, and not just

treat their data as an industrial raw material; you have to tell people in advance  
 what you’re doing, and if you can’t tell everyone you must tell a good sample,  
 not just some friends on your ethics committee; you have to obey the law,  
 including the difficult bits of human-rights law; and you have to tell people  
 what you’ve done afterwards – which includes public breach disclosure [1600].  
 Beware, though, that there is a lot of moral hazard around ethics processes; big  
 ﬁrms who abuse data routinely set up ethics bodies to excuse what they do. I’ll  
 return to this ethics washing in section 11.4.4.

Since then we have used this model to guide our own research in cybercrime,

which is similar in a number of ways. For example, we may sometimes use

data that may be of questionable origin and from which it may be possible  
 to draw inferences about living people who did not give consent. However, in  
 many cases, an ethical case for an investigation can be made but the processes  
 for taking and recording such decisions need careful thought. Transparency is  
 vital; we put all the papers we write on our website, so everyone can see what’s  
 been done with the data.

The same principles may be a good starting point for thinking about the

ethics of machine learning. Many if not most of the AI ethics controversies in  
 the real world so far have been around health data.

**10.4.6** **Social care and education**

The same issues have spilled over into education and social care. While building  
 the NHS national programme for IT, the UK government also started to build  
 a national database of all children, for child-protection and welfare purposes,  
 containing a list of all professionals with which each child has contact. In 2006,  
 the UK Information Commissioner asked a group of us to study the safety and  
 privacy aspects of this. Now the fact that child X is registered with family doctor  
 Y may be innocuous, but a child’s registration with a social work department  
 is different; teachers have lower expectations of children whom they know to  
 have been in contact with social workers. And a record of contact with drug-  
 addiction services or prostitution services is highly stigmatizing. We concluded  
 that the failure to keep such metadata private is both unsafe and unlawful [101].

This became an even hotter political issue in November 2007, when the tax

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 345 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

authorities lost two DVDs containing the UK’s entire child beneﬁt database –  
 personal information on every family in Britain with children. A charity asso-  
 ciated with the Liberal Democrat party commissioned a further report entitled  
 ‘Database State’ on the safety, privacy and legality of a range of public-sector  
 systems [102]; the coalition government of which the Liberal Democrats were  
 part after the 2010 election killed the children’s database as well as discontinu-  
 ing NPfIT, repealing the previous Labour government’s legislation to make ID  
 cards compulsory, and destroying the data and hardware associated with that  
 project. After a further review, it also abandoned a plan for a new ‘eCaf’ system  
 to organise social workers involved in child protection. There the issue was not  
 just privacy but also poor design, as eCaf demanded so much information that  
 social workers were starting to spend more time ‘feeding the beast’ than they  
 did actually talking to children and their families [1354].

Attempts to share data between medicine and social care by direct electronic

access threw up issues of integrity as well as privacy. As an example, when social  
 workers in Oxford were given access to GP records, a social worker could enter  
 ‘diabetic?’ directly into a GP system – which would interpret this as a diagnosis  
 and start trying to schedule all the rest of the diabetes care machinery. The  
 GP would have their work cut out stopping this, as medical records are append-  
 only; and they might start failing to meet their targets for scheduling eye tests  
 for diabetics, which would cut their income. There are also problems with

automating exchanges between care services and schools; in fact, any automated  
 interaction between different types of professional practice needs to be designed  
 with extensive consultation and exploration of a lot of edge cases.

The ‘Database State’ report also highlighted privacy in education. In Eng-

land, the Department for Education had set up a National Pupil Database that  
 initially held census data but gradually accreted test results, behaviour and at-  
 tendance data, whether the child was poor enough to get free school meals and  
 whether they were in care. In addition, schools started adding further surveil-  
 lance ranging from ﬁngerprint scanners to record attendance and library book  
 loans, to CCTV recording the classroom continuously (with the sales pitch that  
 teachers could defend themselves against false accusations by children).

In Scotland, the government proposed a ‘named person’ scheme in 2014,

whereby each child would be allocated one public-sector worker (typically a  
 teacher or health visitor) to promote and safeguard their wellbeing. Rather than  
 stigmatising the poor children who have a social worker, why not give everybody  
 one? This aroused widespread opposition, was defeated in the Supreme Court  
 in 2016, and ﬁnally abandoned in 2019 after ministers couldn’t ﬁgure out a way  
 to do it that was both legal and politically acceptable. A body set up to devise a  
 statutory code of practice decided it ‘would not be desirable as the complexity of  
 this would mean it would not be easy to understand or apply in practice’ [520].

Following sporadic protests by parents, there is now at least one NGO work-

ing for children’s rights7. Concerns range from biometrics to the widespread  
 adoption of cloud services in education, with numerous small providers selling a  
 huge range of teaching support and other services, and children’s data getting  
 everywhere. Even the privacy regulator, the Information Commissioner, has

7https://www.defenddigitalme.org

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 346 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

been criticised for being blind to children’s issues, for example using Vimeo to  
 make instructional videos available on her website, when its terms of service  
 prohibit use by under-13s. If even the regulator can’t manage her own web-  
 site, what chance does the average school have? More fundamentally, should a  
 school treat each pupil as a citizen/customer – responsible and in control – or as  
 a suspect/recidivist to be tracked, scanned and ﬁngerprinted? The temptation  
 with young people is the latter.

Looking back at almost a quarter century of tussles around the safety and

privacy of health IT, and the related subjects of IT in education and social care,  
 one can see the failures conforming to political stereotypes. Britain’s Labour  
 governments from 1997–2010 failed in a typical left-wing way. They were well-  
 meaning but na¨ıve; they could only think in terms of bureaucratic centralism  
 and billion-pound contracts (some with ﬁrms that hired ministers before or after  
 their term of office); they had no idea how to write the speciﬁcations; they lied  
 like mad when things went wrong; and they were suckers for special interests  
 such as medical researchers demanding access to everything. The Conservative  
 governments since 2010 have failed in a typical right-wing way8: talking about  
 rights and freedoms but cynically selling off data to their friends in the drug  
 companies, and for a pittance; lying like mad when things went wrong; while  
 undermining regulators and appointing leaders disposed to turn a blind eye to  
 both safety and privacy failures.

**10.4.7** **The Chinese Wall**

Our ﬁnal ﬂavour of multilateral security is the Chinese Wall model, formalised  
 by David Brewer and Michael Nash [319]. Financial services ﬁrms from invest-  
 ment banks to accountants are required by their regulators to have internal  
 rules designed to prevent conﬂicts of interest wherever two of their clients are  
 competitors, and these controls are called Chinese Walls.

The model’s scope is wider than ﬁnance. There are many service ﬁrms whose

clients may be in competition with each other: advertising agencies are another  
 example. A typical rule is that ‘a partner who has worked recently for one

company may not see the papers of any other company in the same sector’. So  
 once a copywriter has worked on the Shell account, they will not be allowed to  
 work on another oil company’s account for some ﬁxed period of time.

The Chinese Wall model thus mixes free choice and mandatory access con-

trol: a partner can choose which oil company to work for, but once that de-  
 cision is taken their actions in that sector are constrained. It also introduces  
 the concept of *separation of duty* into access control; a given user may perform  
 transaction A or transaction B, but not both. Access controls thus become

stateful.

Part of the attraction of the Chinese Wall model to the security research

community comes from the fact that it’s easy to formalise; in fact, it can be  
 expressed in terms similar to Bell-LaPadula. If we write, for each object *c*, *y*(*c*)  
 for *c*’s company and *x*(*c*) for *c*’s conﬂict-of-interest class, then like BLP it can

8This was despite the fact that the 2010–15 government had Liberal Democrat coalition

partners

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 347 | Ross Anderson |

*10.4. HEALTH RECORD PRIVACY*

be expressed in two properties:

|  |
| --- |
| *•* The *simple security property*: a subject *s* has access to *c* if and only if, for  all *c0* which *s* can read, either *y*(*c*) */2 x*(*c0*) or *y*(*c*) = *y*(*c0*) |
| *•* The *\*-property*: a subject *s* can write to *c* only if *s* cannot read any *c0*  with *x*(*c0*) *6*= *↵* and *y*(*c*) *6*= *y*(*c0*). |

The Chinese Wall model sparked a debate about the extent to which it is

consistent with the BLP tranquility properties, and some work on the formal  
 semantics of such systems9. There are also some interesting new questions

about covert channels. For example, could an oil company ﬁnd out whether  
 a competitor which used the same investment bank was planning a bid for a  
 third oil company, by asking which specialists were available for consultation  
 and noticing that their number had dropped suddenly?

In practice Chinese Walls still get implemented using manual methods. One

large software consultancy has each of its staff maintain an ‘unclassiﬁed’ CV  
 containing entries that have been sanitized and agreed with the customer. A  
 typical entry might be:

**Sep 17 – Apr 18**: consulted on security requirements for a new  
 branch accounting system for a major US retail bank

This is not the only control. A consultant’s manager should be aware of

possible conﬂicts and not forward the CV to the client if in doubt; if this fails,  
 the client can spot potential conﬂicts himself from the CV; and if this also fails  
 then the consultant is duty bound to report any potential conﬂicts as soon as  
 they appear.

There remains the issue of micro-level access. What if a bank manager simply

looks at the bank statements of his best customer’s competitors? Here, modern  
 systems tend to limit access except where the staff member has established  
 a security context for that customer, for example by getting the customer to  
 answer some authentication questions. I’ll discuss this further in the chapter on  
 Banking and Bookkeeping.

One conspicuous failure mode of Chinese walls is where the conﬂict period

is too short. Governments typically have conﬂict rules that prevent a minister  
 working in any sector that they have regulated for six months after leaving office.  
 This is way too little. Someone who was an energy minister six months ago still  
 knows all the top people in the industry, and anyone who’s beneﬁted from their  
 policy may express their gratitude by hiring them. Five years might be more  
 sensible, but if you think you can get your local legislature to pass such a law,  
 good luck.

9See, for example, Foley [700] on the relationship with non-interference. The practical

resolution of tranquility is usually a cooling-off period: having worked for one oil company,  
 you might be forbidden to work for another for two years

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 348 | Ross Anderson |

*10.5. SUMMARY*

**10.5** **Summary**

In this chapter, we looked at the problem of setting boundaries when systems  
 scale up to collect large amounts of sensitive information, to which many people  
 need access in order to do their jobs. This is an issue in many information secu-  
 rity problems, ranging from the protection of national intelligence data and data  
 about wildlife at risk from poaching, through the privacy and conﬁdentiality of  
 medical and social-care information, to professional practice in general.

We looked at medical records in the greatest detail, and found that the easy

problem is setting up access controls in a direct care setting so that access to  
 each record is limited to a sensible number of staff. Such systems can be designed  
 by automating existing working practices, and role-based access controls are a  
 natural way to implement them. However, the incentives in health-care systems  
 are such that the implementation is often poor, and needs regulation to enforce  
 compliance. The traditional approach to privacy, which might be summarised as  
 ‘consent or anonymise’, is being undermined by growing complexity with many  
 outsourced systems that are often opaque even to doctors (let alone patients).  
 The harder problems are the growing number of central systems, particularly  
 those related to payments, from which opt-outs aren’t available; the growing  
 use of genetic data, and the effects of social media from which sensitive personal  
 health information can often be inferred. Here, too, the governance problems  
 are even less tractable than the technical ones. The only realistic solution lies in  
 regulation, and here the USA and the EU are moving ever further apart. Europe  
 gives its citizens the right to restrict their personal health information to the  
 clinicians involved directly in their case; America does not. However it can be  
 hard for Europeans to enforce our rights. Both America and Europe have huge  
 lobbying and ﬁnancial pressures from drug ﬁrms and others who want all our  
 data; politicians tend to side with the industry and undermine the regulators.

Since the 1990s, health providers and services have tried to have their cake

and eat it by building ‘anonymised’ databases of medical records (or school  
 records, or census returns) so as to allow researchers to make statistical enquiries  
 without compromising individuals’ privacy. There are some applications where  
 this is a complete non-starter, such as in ﬁghting wildlife crime; there, the  
 aggregate data are even more valuable to poachers than individual sightings.  
 In the case of medical records, computer scientists have known since the 1980s  
 that anonymising rich data is a lot harder than it looks, and in recent years  
 we’ve acquired a robust theory of this that lets us work out when it can work  
 and when it won’t. I’ll discuss this in the next chapter.

Another takeaway message is this. Just as multilevel security was the ‘hedge-

hog’ approach to information security, where you hope to get a good result by  
 just getting one big thing right, multilateral security requires the ‘fox’ approach;  
 you need to understand your application in detail, learn what’s gone wrong in  
 the past – and also be good at adversarial thinking if you want to anticipate  
 what’s likely to go wrong in future.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 349 | Ross Anderson |

*10.5. SUMMARY*

**Research Problems**

The coronavirus pandemic is likely to make health surveillance much more per-  
 vasive so personal health information will become more widespread and the con-  
 ﬂicts discussed here will spread way beyond the healthcare sector. What will  
 that entail, and how should technical and policy mechanisms evolve to cope?

Also, in the near future, more and more medical treatment will involve ge-

netic information. Is there any sensible way in which privacy models can be  
 extended to deal with multiple individuals? For example, in many countries  
 you have the right not to know the outcome of a DNA test that a relative  
 has for an inheritable disease such as Huntington’s Chorea, as it may affect  
 the odds that you have it too. Your relative does have a right to know, and  
 may tell others – so unwelcome news might reach you indirectly. As I write,  
 there are cases going through the courts in the UK and Germany that push in  
 different directions on the rights of the children of people diagnosed with Hunt-  
 ington’s [606]. Such tensions over information rights long predate the Internet  
 and cannot be managed purely by technological mechanisms. But social media  
 change the scaling factors in such a way as to make them more widespread and  
 acute. The long-term solutions may well involve some mix of laws, social norms  
 and technology support; but they are likely to take years to work out, and we  
 may well end up with different solutions in different cultures. For example,

East Asian countries have tolerated much more intrusive surveillance, and have  
 suffered far fewer deaths in the pandemic, at least so far. Might that change  
 attitudes elsewhere?

**Further Reading**

The literature on compartmented-mode security is scattered: most of the public-  
 domain papers are in the proceedings of the NCSC/NISSC and ACSAC con-  
 ferences, while Amoroso [47] and Gollmann [779] cover the basics of the lattice  
 and Chinese-wall models. For a survey of privacy failures in health, social care  
 and education in the UK in 2009, see *‘Database State’* [102]. For a case study of  
 the NHS National Programme for IT, see [379], and for a later report on total  
 costs by the UK Parliament’s Public Accounts Committee, see [1559]. For the  
 BMA model see the policy itself [58], the Oakland version [59], the proceedings  
 of a conference on the policy [63], and the papers on the pilot system at Hast-  
 ings [535, 536]. For a National Research Council study of medical privacy in the  
 USA, see [1412]; there is also an HHS report on the use of de-identiﬁed data in  
 research at [1191]. But the best sources for up-to-date news on medical privacy  
 issues are the websites of the relevant lobby groups: medConﬁdential for the  
 UK, and Patient Privacy Rights for the USA.

|  |  |  |
| --- | --- | --- |
| **Security Engineering** | 350 | Ross Anderson |