Scientific Writing, Integrity and Ethics VII

Privacy

Guoqiang Li
School of Software

SHANGHAI JIAO TONG UNIVERSITY

Reference

Sara Baase, Timothy Henry. Gift of Fire, A: Social, Legal, and Ethical Issues for Computing Technology (5th edition), Pearson, 2017

Michael Quinn. Ethics for the Information Age (8th edition), Pearson, 2019

Privacy Risks and Principles

Key Aspects of Privacy

Freedom from intrusion (being left alone).

Control of information about oneself.

Freedom from surveillance (from being tracked, followed, watched).

No Complete Privacy

We cannot expect complete privacy.

Many know what you look like, whether you are a nice person.

Need not get your permission to observe and talk about you.

If people know nothing about you, they might not rent you a place to live, hire you, and so on.

We give up some privacy for the benefits of dealing with strangers.

Privacy Threats

Intentional, institutional uses of personal information.

Unauthorized use or release by “insiders”.

Theft of information.

Inadvertent leakage of information.

Our own actions.

Philosophical Perspectives

The man who is compelled to live every minute of his life among others and whose every need, thought, desire, fancy or gratification is subject to public scrutiny, has been deprived of his individuality and human dignity. [He] merges with the mass. . . Such a being, although sentient, is fungible; he is not an individual.

Edward J. Bloustein

Personal Perspectives

It's important to realize that privacy preserves not personal secrets, but a sense of safety within a circle of friends so that the individual can be more candid, more expressive, more open with "secrets".

Robert Ellis Smith

New Technology, New Risks

Government and private databases.

Sophisticated tools for surveillance and data analysis.

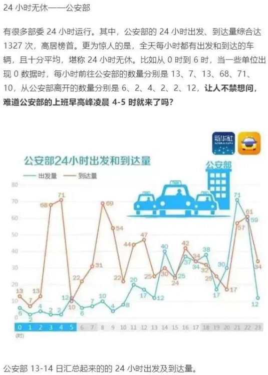
Vulnerability of data.

Examples: Search Query Data

Search engines collect many terabytes of data daily.

Data is analyzed to target advertising and develop new services.

Who gets to see this data? Why should we care?

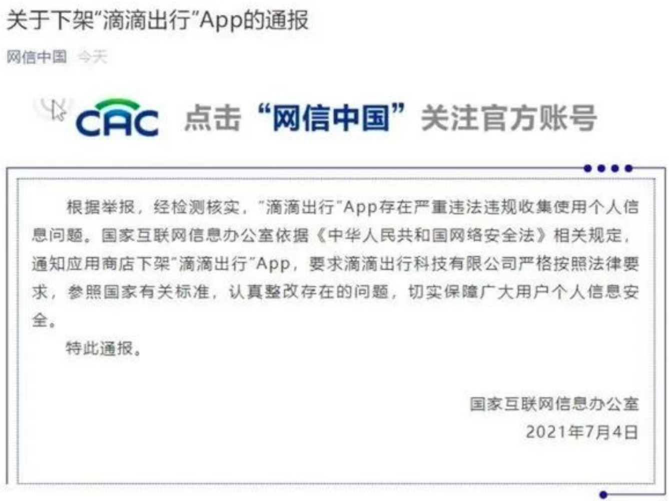


24 小时无休——公安部

有很多部委 24 小时运行。其中，公安部的 24 小时出发、到达量综合达 1327 次，高居榜首。更为惊人的是，全天每小时都有出发和到达的车辆，且十分平均，堪称 24 小时无休。比如从 0 时到 6 时，当一些单位出现 0 数据时，每小时前往公安部的数量分别是 13、7、13、68、71、10，从公安部离开的数量分别是 6、2、4、2、2、12，**让人不禁想问，难道公安部的上班早高峰凌晨 4-5 时就来了吗？**

公安部 13-14 日汇总起来的的 24 小时出发及到达量。

Examples: Search Query Data



关于下架“滴滴出行”App的通报

网信中国 今天

CAC 点击“网信中国”关注官方账号

根据举报，经检测核实，“滴滴出行”App存在严重违法违规收集使用个人信息问题。国家互联网信息办公室依据《中华人民共和国网络安全法》相关规定，通知应用商店下架“滴滴出行”App，要求滴滴出行科技有限公司严格按照法律要求，参照国家有关标准，认真整改存在的问题，切实保障广大用户个人信息安全。

特此通报。

国家互联网信息办公室
2021年7月4日

Examples: Smartphone

Location apps.

Data sometimes stored and sent without user's knowledge.

Summary of Risks

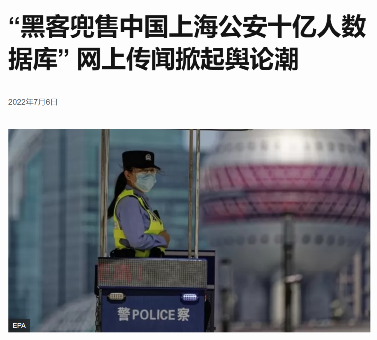
Anything we do in cyberspace is recorded.

Huge amounts of data are stored.

People are not aware of collection of data.

Software is complex.

Leaks happen.



“黑客兜售中国上海公安十亿人数据库” 网上传闻掀起舆论潮

2022年7月6日

EPA

Summary of Risks

A collection of small items can provide a detailed picture.

Re-identification has become much easier due to the quantity of information and power of data search and analysis tools.

If information is on a public Web site, it is available to everyone.

Summary of Risks

Information on the Internet seems to last forever.

Data collected for one purpose will find other uses.

Government can request sensitive personal data held by businesses or organizations.

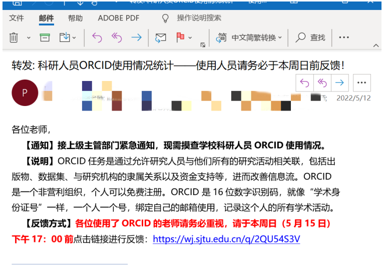
We cannot directly protect information about ourselves. We depend upon businesses and organizations to protect it.

Terminology

Personal information: any information relating to an individual person.

Informed consent: users being aware of what information is collected and how it is used.

Invisible information gathering: collection of personal information about a user without the user's knowledge.



转发: 科研人员ORCID使用情况统计——使用人员请务必于本周日前反馈!

P [redacted] 2022/5/12

各位老师,

【通知】接上级主管部门紧急通知，现需摸查学校科研人员 ORCID 使用情况。

【说明】ORCID 任务是通过允许研究人员与他们所有的研究活动相关联，包括出版物、数据集、与研究机构的隶属关系以及资金支持等，进而改善信息流。ORCID 是一个非营利组织，个人可以免费注册。ORCID 是 16 位数字识别码，就像“学术身份证号”一样，一个人一个号，绑定自己的邮箱使用，记录这个人的所有学术活动。

【反馈方式】各位使用了 ORCID 的老师请务必重视，请于本周日（5 月 15 日）下午 17：00 前点击链接进行反馈：https://wj.sjtu.edu.cn/q/2QU54S3V

Terminology

Cookies: Files a Web site stores on a visitor's computer.

Secondary use: Use of personal information for a purpose other than the purpose for which it was provided.

Data mining: Searching and analyzing masses of data to find patterns and develop new information or knowledge.

Terminology

Computer matching: Combining and comparing information from different databases (using social security number, for example) to match records.

Computer profiling: Analyzing data to determine characteristics of people most likely to engage in a certain behavior.

Examples

Sale of consumer information to marketers or other businesses.

Use of information in various databases to deny someone a job or to tailor a political pitch.

The IRS searching vehicle registration records for people who own expensive cars and boats (to find people with high incomes).

Use of a person's text messages by police to prosecute that person for a crime.

Two Common Forms

opt out: Person must request (usually by checking a box) that an organization not use information.

opt in: The collector of the information may use information only if person explicitly permits use (usually by checking a box).

Fair Information Principles

Inform people when you collect information.

Collect only the data needed.

Offer a way for people to opt out.

Keep data only as long as needed.

Maintain accuracy of data.

Protect security of data.

Develop policies for responding to law enforcement requests for data.

Discussion Questions

Have you seen opt-in and opt-out choices? Where? How were they worded?

Were any of them deceptive?

What are some common elements of privacy policies you have read?

The Business and Social Sectors

Marketing and Personalization

Marketing is an essential task for most businesses and organizations.

Through most of the 20th century, businesses sent out catalogs and advertisements based on a small number of criteria.

Computers and increased storage capacity generated a revolution in targeted marketing.

Marketers argue that targeting ads via personal consumer information reduces the number of ads overall that people will see, provides ads that people are more likely to want, and reduces overhead and, ultimately, the cost of products.

Targeting is so popular with some people that Google advertised that its Gmail displays no untargeted banner ads.

Less Obvious Personalization

The displays, ads, prices, and discounts you see when shopping online might be different from those that other people see.

A clothing site does not display winter parkas on its home page for a shopper from Sanya.

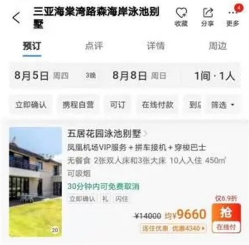
Some sites offer discounts to first-time visitors.

Some sites guess a visitor's gender based on clicking behavior.

If a person hesitates over a product, a site might offer something extra, perhaps free shipping.

A hotel reservation website begins showing more expensive options to visitors who use Macs.

Less Obvious Personalization



三亚海棠湾路森海岸泳池别墅

收藏 分享 更多

预订 点评 详情 周边

8月5日 周四 3晚 8月8日 周日 1间 · 1人

立即确认 携程自营 可订 在线付款 到店 筛选

五居花园泳池别墅

凤凰机场VIP服务＋拼车接机＋穿梭巴士

无餐食 2张双人床和3张大床 10人入住 450㎡

可吸烟

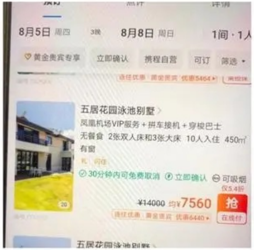
30分钟内可免费取消

立即确认 礼 闪住

仅6.9折

¥14000 均¥9660 抢

连住优惠 优惠4340 在线付



预订 点评 详情

8月5日 周四 3晚 8月8日 周日 1间 · 1人

黄金贵宾专享 立即确认 携程自营 可订 筛选

连住优惠 黄金贵宾 优惠5464

五居花园泳池别墅

凤凰机场VIP服务＋拼车接机＋穿梭巴士

无餐食 2张双人床和3张大床 10人入住 450㎡

有窗

礼 闪住

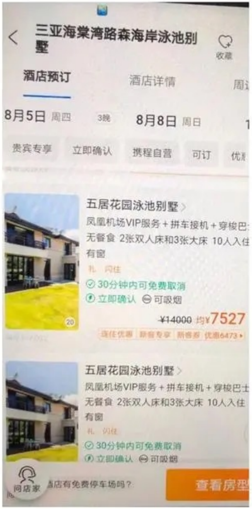
30分钟内可免费取消 立即确认 可吸烟

仅5.4折

¥14000 均¥7560 抢

连住优惠 黄金贵宾 优惠6440 在线付

五居花园泳池别墅



三亚海棠湾路森海岸泳池别墅

收藏

酒店预订 酒店详情 周边

8月5日 周四 3晚 8月8日 周日 1

贵宾专享 立即确认 携程自营 可订 优惠

五居花园泳池别墅

凤凰机场VIP服务＋拼车接机＋穿梭巴士

无餐食 2张双人床和3张大床 10人入住

有窗

礼 闪住

30分钟内可免费取消

立即确认 可吸烟

¥14000 均¥7527

连住优惠 贵宾专享 新客价 优惠6473

五居花园泳池别墅

凤凰机场VIP服务＋拼车接机＋穿梭巴士

无餐食 2张双人床和3张大床 10人入住

有窗

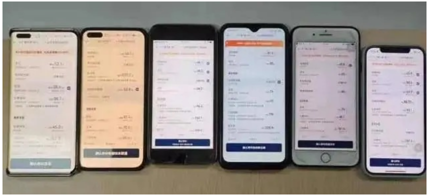
礼 闪住

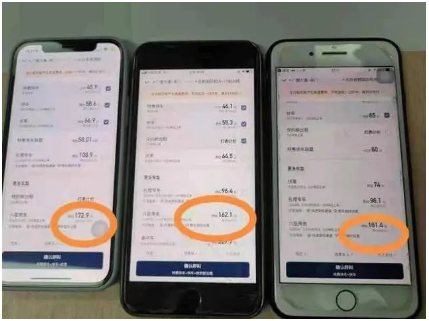
30分钟内可免费取消

立即确认 可吸烟

问店家 酒店有免费停车场吗？

查看房型





Borderline

Salesclerks can see our gender and our approximate age. They can form other conclusions about us from our clothing, conversation, and behavior.

Good salespeople in expensive specialty stores, car dealerships, and flea markets make judgments about how much a potential customer will pay. They modify their price or offer extras accordingly.

Is the complex software that personalizes shopping online merely making up for the loss of information that would be available to sellers if we were shopping in person?

Are some people uneasy mainly because they did not realize that their behavior affects what appears on their screen? Or are there privacy threats lurking in these practices?

A salesclerk in a store does not have a list of our online search queries. Who does? Who should?

Data Mining and Clever Marketing

Example (Target retail chain)

The Target retail chain had its data miners analyze purchases by women who signed up for baby registries. Target discovered that pregnant women tend to increase their purchases of a group of 25 products. So when a woman began to purchase more of those products (e.g., unscented lotions and mineral supplements), Target sent coupons and ads for pregnancy and baby products—even timing the coupons for stages of the pregnancy.

Data Mining and Clever Marketing

Example (Tesco)

Customers of the British retailing firm Tesco, permit the company to collect information on their buying habits in exchange for discounts. The company identifies young adult males who buy diapers and sends them coupons for beer— assuming that, with a new baby, they have less time to go to a pub.

Data Mining and Clever Marketing

Example (Tesco)

To compete with Walmart, Tesco aimed to identify customers who were most price conscious and hence most likely to be attracted to Walmart's low prices. By analyzing purchase data, the company determined which customers regularly buy the cheapest version of products that are available at more than one price level. Then the company determined what products those customers buy most often, and set prices on those products below Walmart's.

Borderline

Companies can use face recognition systems in video game consoles and televisions to target ads to the individual person who is playing a game or watching TV.

What risks to privacy does this entail?

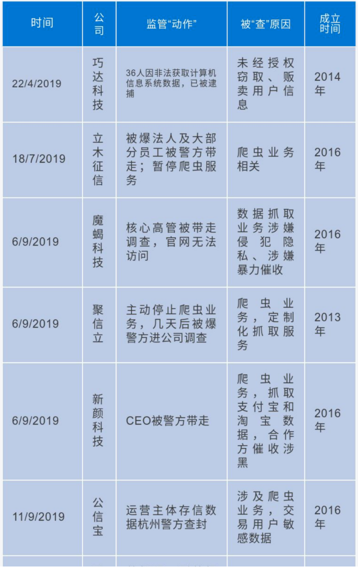
Is it unethical to include such features?

Will most people like the customization?

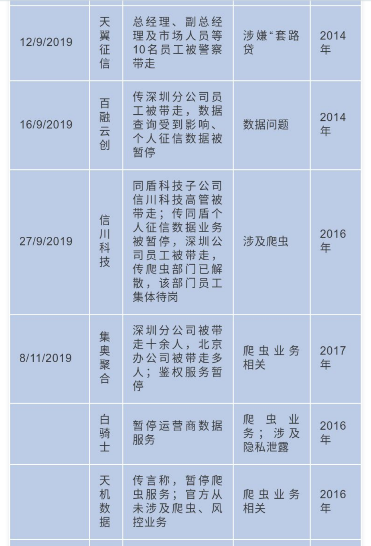
Do they understand that if they see ads targeted to their interests, someone somewhere is storing information about them?

Does it matter if a human ever views the data or if it is processed and acted on only by software?

Example: Crawler



时间	公司	监管“动作”	被“查”原因	成立时间
22/4/2019	巧达科技	36人因非法获取计算机信息系统数据，已被逮捕	未经授权窃取、贩卖用户信息	2014年
18/7/2019	立木征信	被爆法人及大部分员工被警方带走；暂停爬虫服务	爬虫业务相关	2016年
6/9/2019	魔蝎科技	核心高管被带走调查，官网无法访问	数据抓取业务涉嫌侵犯隐私、涉嫌暴力催收	2016年
6/9/2019	聚信立	主动停止爬虫业务，几天后被爆警方进公司调查	爬虫业务，定制化抓取服务	2013年
6/9/2019	新颜科技	CEO被警方带走	爬虫业务，抓取支付宝和淘宝数据，合作方催收涉黑	2016年
11/9/2019	公信宝	运营主体存信数据杭州警方查封	涉及爬虫业务，交易用户敏感数据	2016年



12/9/2019	天翼征信	总经理、副总经理及市场人员等10名员工被警察带走	涉嫌“套路贷	2014年
16/9/2019	百融云创	传深圳分公司员工被带走，数据查询受到影响、个人征信数据被暂停	数据问题	2014年
27/9/2019	信川科技	同盾科技子公司信川科技高管被带走；传同盾个人征信数据业务被暂停，深圳公司员工被带走，传爬虫部门已解散，该部门员工集体待岗	涉及爬虫	2016年
8/11/2019	集奥聚合	深圳分公司被带走十余人，北京办公司被带走多人；鉴权服务暂停	爬虫业务相关	2017年
	白骑士	暂停运营商数据服务	爬虫业务；涉及隐私泄露	2016年
	天机数据	传言称，暂停爬虫服务；官方从未涉及爬虫、风控业务	爬虫业务相关	2016年

Key Issue

The issue is informed consent.

Informed Consent

Technological and social changes make people uncomfortable, but that does not mean the changes are unethical.

Informed Consent

Technological and social changes make people uncomfortable, but that does not mean the changes are unethical.

Collection of consumer data for marketing without informing people or obtaining their consent used to be widespread, essentially standard practice.

Informed Consent

Technological and social changes make people uncomfortable, but that does not mean the changes are unethical.

Collection of consumer data for marketing without informing people or obtaining their consent used to be widespread, essentially standard practice.

- Opt-out and opt-in options matter.

Informed Consent

Technological and social changes make people uncomfortable, but that does not mean the changes are unethical.

Collection of consumer data for marketing without informing people or obtaining their consent used to be widespread, essentially standard practice.

- Opt-out and opt-in options matter.

Awareness of online tracking varies among consumers.

WE need to help to educate consumers and encourage responsible choices.

Social Networks - What We Do

Polls show that people care about privacy.

Why don't they act that way?

Ian Kerr

Social Networks - What We Do

You may enjoy the feature on a social network site that told you which members read your profile, but you may be surprised and upset to find that people whose profiles you read knew that you read them.

Social Networks - What We Do

You may enjoy the feature on a social network site that told you which members read your profile, but you may be surprised and upset to find that people whose profiles you read knew that you read them.

People often want information about others, but they do not want others to have access to the same kinds of information about themselves.

Social Networks - What We Do

You may enjoy the feature on a social network site that told you which members read your profile, but you may be surprised and upset to find that people whose profiles you read knew that you read them.

People often want information about others, but they do not want others to have access to the same kinds of information about themselves.

Some people do not know or understand or think enough about information sharing policies to make good decisions about what to do in cyberspace.

Social Networks - What We Do

Many young people post opinions, gossip, and pictures that their friends enjoy.

Their posts might cause trouble if parents, potential employers, law enforcement agents, or various others see them.

People who try to clean up their online personas before starting a job search find that it is hard to eliminate embarrassing material.

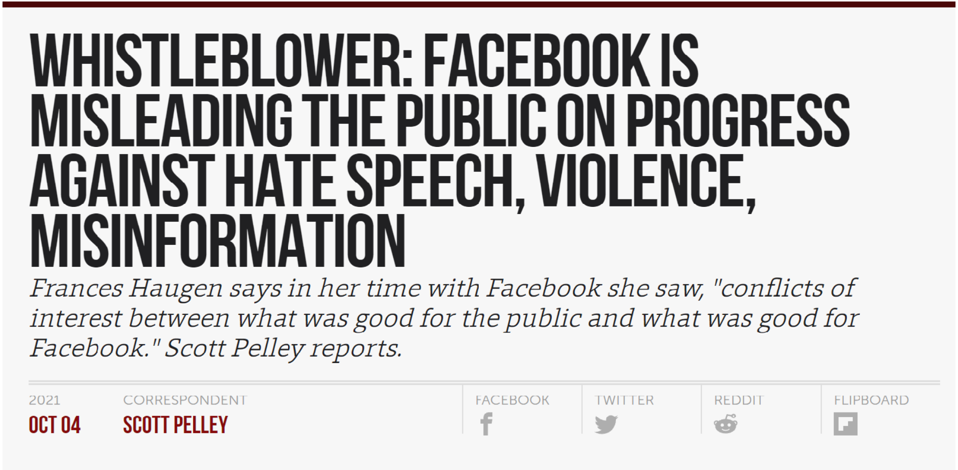
Social Networks - What We Do

Why was it for so long standard practice to stop mail and newspaper delivery when going away on a trip?

This one detail about location ("away from home") was important to protect from potential burglars.

Yet, now, a great many people post their location (and that of their friends) to social networks. Is this less risky?

Social Networks - What They Do



WHISTLEBLOWER: FACEBOOK IS MISLEADING THE PUBLIC ON PROGRESS AGAINST HATE SPEECH, VIOLENCE, MISINFORMATION

Frances Haugen says in her time with Facebook she saw, "conflicts of interest between what was good for the public and what was good for Facebook." Scott Pelley reports.

2021
OCT 04

CORRESPONDENT
SCOTT PELLEY

Discussion Questions

Is there information that you have posted to the Web that you later removed? Why did you remove it? Were there consequences to posting the information?

Have you seen information that others have posted about themselves that you would not reveal about yourself?