Background

v.0.0.2

1 Intro

We're going to talk about topics that cut across a wide range of disciplines. That means we need to be familiar with a bunch of different concepts. Here's a very rough and very quick overview of some of them, just so we can all be on the same page. I'll then outline the questions we will seek to answer.

2 Data

Much of what we'll discuss revolves around the use of data. So, what's data?

Very roughly, a piece of data (a datum) is a piece of information or a representation of a (purported) fact. If we're being careful, we'll say that data is information or a representation of a fact in the context of some use.[1] The temperature of the room you are in now is not data. It's just part of the way the world is. It would be data if we were adjusting the HVAC or running an experiment about how temperature affects the ability to read boring material.

This probably won't matter too much for our purposes, but just in case let me be clear about the relations between the world, facts, and data.

2.1 Facts

Data represent facts. What are facts?

Let's start with the world. A <u>state-of-affairs</u> is the way part of the world could be at a time. Consider three ways the world could be right now:

> At 6 PM on 10 May 2019, there is a taco on Mars.
> At 6 PM on 10 May 2019, there is a taco in Adam's hand.
> At 6 PM on 10 May 2019, the dog needs attention.

All three are possible states of affairs. There could be a taco on Mars (perhaps via one of Elon Musk's schemes). I could have a taco (don't remind me). The dog could be needy.

While all 3 are possible. Only 1 is the way the world actually is right now; it is a state of affairs which <u>obtains</u>. There are no tacos on Mars; I am (sadly) taco-less. The dog keeps nudging me, bringing over toys, and otherwise acting cute in wanton disregard of my need to write this. Be right back.

So far so good? A fact is a states-of-affairs which obtains.[2] Since the dog is actually (still!) in need of attention, it is a fact about how the world is right now.

2.2 Representations of facts

A fact is thus a way the world is at a given time. That has absolutely nothing to do with us. The universe is full of facts and was full of facts for billions of years before we

[1] Some writers invert this formula and define information as data plus use; we needn't wade into this fight here.

[2] As you might guess, there are plenty of alternative views. I'm describing one influential picture which comes from D.M. Armstrong.

arrived on the scene. Once creatures, especially ones with language, get on the scene, things get more complicated. That's because a fact can be represented in different ways.

Suppose over time erosion has uncovered a big diamond on a beach. It is a fact that there is a diamond on that beach. If a turtle wanders in range, the fact that there is a diamond on the beach will be represented by certain patterns of neural activity in its brain. An autonomous drone's camera could capture an image representing the fact. I could see it and squeal "Holy crap! There's a big diamond on the beach!" thus representing it with language; my Japanese speaking friend could do the same with "海に金剛石があります!"

Thus a piece of data is a representation of a way the world is (i.e., a fact).[3]

Why not just say that data are facts and vice-versa? Well, one of the first questions we will ask concerns ownership of data. What would it even mean to say that someone owns a fact? Facts are abstract things.[4] The fact that I weigh 190 lbs[5] can be expressed with different sentences. For example, 'if you make a big pile of all the things which weigh 190 lbs and look through it, you will find Adam' (hopefully near the top). Thus no one can own a fact. Even if you owned me, you would not own the fact about my weight.

Since data are representations of facts, they are the right sort of thing. Data plausibly can be owned (just like a copyright holder can own an image). If my weight is recorded in a database and someone hacks into the database and erases the record, presumably the company can sue her for destroying its property.

3 Personal data

Data comes in myriad flavors. Weather stations capture meteorological data. Stock markets create financial data. We will mostly be concerned with personal data. So, what's personal data? Data of a personal nature. Moving on….

[3] NB, on this definition, there is no bad data —data which fails to represent a fact. Those sentences/numbers would fail to be data. Try not to worry about that; we're deep enough in the weeds as it is.

[4] At least on some definitions of 'abstract', if you're getting hung up on this, congratulations, you have a promising career as a metaphysician.

[5] Sigh. That was the pre-pandemic Adam….

Just kidding (although not by much). Personal data is a proper subset[6] of data. It is data about a natural person. Since we will be concerned with data that can be used in ways which harm individuals, we'll use the somewhat narrower definition is the found in the European Union's Data Protection Directive, namely

> Any information relating to an identified or identifiable natural person[7]

Here's a summary of what we are and aren't interested in from the Stanford Encyclopedia of Philosophy entry on privacy and information technology

> Personal information or data is information or data that is linked or can be linked to individual persons. Examples include date of birth, sexual preference, whereabouts, religion, but also the IP address of your computer or metadata pertaining to these kinds of information. Personal data can be contrasted with data that is considered sensitive, valuable or important for other reasons, such as secret recipes, financial data, or military intelligence. Data that is used to secure other information, such as passwords, are not considered here. Although such security measures may contribute to privacy, their protection is only instrumental to the protection of other information, and the quality of such security measures is therefore out of the scope of our considerations here.[8]

We thus need to talk a bit about natural persons and the something being linked to a person.

3.1 Natural persons

You, me, and all the people in your life except for your imaginary friends are natural persons. Trees, rocks, and tacos are non-persons. Corporations, states, and other entities whose existence depends on things like acts of law are unnatural persons.

There may be borderline cases. If a robot turns out to be sufficiently like a human being that similar moral considerations should be extended to them, then perhaps the robot would be a natural person. For now, natural persons are limited to human beings.

3.2 Linkability

[6] A fancy way of saying all personal data is data, but not all data is personal data

[7] EU Data Protection Directive (95/46/EC) Article 2(a)

[8] https://plato.stanford.edu/entries/it-privacy/

What it is for data to be 'about' or 'linked to' a natural person? Since data represent facts, part of the answer is that a person is the subject of the fact. For example, Adam weighs 190 lbs is[9] a fact. I am the subject of the sentence 'Adam weighs 190 lbs'. Thus the entry in the database which records this fact is about me.

But our definition of personal data requires that the person in question be 'identified or identifiable'. The ability to link a piece of data to a natural person is the decisive feature.

The tricky bit comes from the fact that we can talk about the properties of things whose identity we don't know. Suppose I say "The murderer of Tupac is a horrible person". Should you assume I know the jerk? Not if my reason for saying this is that the murderer deprived the world of some good music.

To see why this matters, suppose a medical researcher is collecting data to study the relationship between obesity and existing heart disease. To do this, she hands out forms requesting 3 pieces of information from each person in the sample:

Height
Weight
Whether they have heart disease.

Once she records these three variables on her computer, she shreds the forms (then burns them, then separates the ashes into several small bags, which she disposes in trashcans spread out across the country).[10]

Is the data on her computer personal data? Probably not. If it is impossible to tell whose height, weight, and disease status each entry represents to, then it is not linkable to a natural person.

I say 'probably not' because the question turns on a bunch of contingent matters. For example, if the data was collected in an area where everyone is very tall and skinny, the 1 short and pudgy guy might be identifiable. In that case, the (his?) data would be personal data. This sort of exception may seem silly, but it will prove extremely important later on (e.g., Technical problems). It's surprisingly difficult to reliably anonymize data sets, especially when there are multiple data points for individuals.

[9] Sigh. Was a fact.

[10] HIPAA is no joke.

4 Algorithms

If you ask a computer scientist or mathematician what an algorithm is, you'll get something like

> An ordered set of unambiguous steps that produces a result and terminates in a finite time.

That's way more formal than we need. We'll just say that an algorithm is a stepwise computational procedure for doing something.

In the articles we'll read, writers sometimes use the term in narrower or more loaded ways. For example, some talk about algorithms as computational processes used to make decisions. That's fine when such algorithms are the focus of the discussion. But, technically, decision-making is a subset of the things we might do with algorithms.[11]

4.1 Adam's attendance algorithm

For a simple example, here's the attendance algorithm I follow[12] at the beginning of every class:

1. For each student, call out their name. If they answer, mark them present
2. When done, ask if anyone was missed
3. If anyone answers 'yes', mark each answering student as present
4. Stop

Following these steps produces a list of who is present. Assuming I don't have infinite students[13], this is an algorithm.

If wanted a computer to do this, we'd (of course) write it in python. It might look something like (lines starting with '#' or enclosed between triple quotes """this is a

[11] For a discussion of this see *Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency* (2015) (pp. 1–16).

[12] Sigh. Stupid pandemic.

[13] Class sizes are creeping up, but I don't think we have to worry about that.

comment””” are comments for humans to understand what’s going on and not part of the instructions the computer follows):

```
def take_attendance(list_of_students):
        “””Use at beginning of class to record which students are present”””

    for student in list_of_students:
        # Do the following for each student in the list
        answer = call_student_name(student)

        if answer:
            mark_present(student)

        # Do nothing if no answer

    # Now we're done calling the initial list
    missed_students = ask_if_anyone_was_missed()

    if len(missed_students) > 0:
        for student in missed_students:
            mark_present(student)
```

Note that `mark_present` and `call_student_name` are two other functions which do what they’re named.

5 Databases

Databases store data. Databases come in many different formats. Index cards in a shoebox can be a database. A simple spreadsheet is a database. A complex relational databases with hundreds of tables is one; I mean it’s right there in the name. Massively overhyped[14] and more-trouble-than-they’re-worth NoSQL stores are databases.

The type of database won’t really matter for the questions we are interested in. Though the type of database involved may affect how concerned we are about certain types of privacy violations. If your personal information is written on an index card, randomly tossed in a shoebox, and that shoebox randomly tossed into a storage unit with thousands of similar shoeboxes, it’s going to take an awful lot of luck or work to violate

[14] That’s right. I’m looking at you MongoDB

your privacy. In a well-designed relational database, the information will be broken up across a whole series of tables. The most time consuming step in finding your information can be typing in the query.

Database type will also affect how powerful data mining techniques will be. Again, a well-designed relational database abstracts out all the data which allows us to connect things in novel ways. For example, if we have one table that holds people's names, one table that holds weights, and a third that holds pet preferences, we can easily look for correlations in weights and pet ownership. Try that with a shoebox full of index cards.

6 Data mining

[ToDo: brief overviews of how some data mining techniques work]

7 Main Questions

Companies have collected data on their customers forever. Many of the hottest data-mining algorithms have roots in statistical techniques that have been around awhile. Why are we so worried about this stuff now? What's changed?

One major set of changes involves the drastically declining costs of storage and computation. It used to be that if you wanted to keep data on something, the benefit of that data needed to outweigh the costs. Now, the marginal cost of storing and processing it is basically trivial. For many companies there is now very little financial reason not to capture all the data you can and store it for an unlimited amount of time.[15]

Another factor is the availability of extremely individualized data. The internet enabled this to some degree. But the rise of social media and the ubiquity of smart phones makes it possible to gather a detailed profile of every potential customer.

Indeed, the wealth of very granular data has given rise to companies who specialize in what used to be called Knowledge Discovery in Databases (KDD). These companies aim at

> discovering non-trivial new insights in existing datasets, insights that cannot simply be observed in datasets or follow automatically from datasets, but

[15] Check out how cheap Amazon Web Services is: https://aws.amazon.com/pricing/

> insights that have to be extracted or generated since they do not 'lie at the surface' [Sax 27]

Nowadays, we call this big data.

To paraphrase an expert on the B.I.G., big data, big problems.

We're going to ask 3 questions.

7.1 (Q1) When may an ethical company profit from the use of personal data?

First, as Sax notes

> big data's entrepreneurial potential resides in the fact that advanced mining techniques can extract/generate unanticipated, non-trivial, new, and (commercially) interesting insights. [Sax 27]

If that's true, then as he writes

> Big data's entrepreneurial potential is equally dependent on the legitimacy of the appropriation of these newly extracted/generated insights by commercial parties [Sax 27]

This is roughly our first main question:

> (Q1) When may an ethical company profit from the use of personal data?

This extends to insights derived through machine learning and other analytical techniques done on users data.

We will consider 2 possible answers to Q1:

Q1 Answer 1: When they own it

Q1 Answer 2: When we let them

7.2 (Q2) Harms of informational privacy violation

Second, people can be hurt by the use of their data. These injuries may be tiny —the annoyance of ads following you around the web. They may be large —your location data being available to a stalker or outing you to a bigoted boss. We want to draw a line between what's annoying and what's wrong. We want to know

> (Q2) When is a misuse of personal data significant enough to warrant moral condemnation, regulation, criminalization, or other forms of coercion to prevent?

Q2 Harms of informational privacy violation

7.3 (Q3) Responsibility

Third, when the widespread use of sophisticated machine learning algorithms meets the widespread gathering and sharing of personal data, it is possible that some of the harms discussed under Q2 don't have any human beings in the loop.

If you are hurt by a process that had no human involvement, who is to blame? Sure the company may be legally liable. But who in the company may be blame? It will take a lot of work later on to make this question make sense, but for now let's just note that a lot of harm can be done to people without any human beings being involved in causing the harm. This puts pressure on our ordinary understandings of blame, punishment, and other notions of moral responsibility. Thus our question will be

> (Q3) How should we assign responsibility when people are harmed by algorithmic uses of personal data?

Q3: Responsibility