Data Science Ethics

College of Science

Al-Mustaqbal University

Stage 2 , Semester 1
@ Department of Intelligent Medical Systems

Prepared by
Dr. Samir Badrawi

Anonymity

Anonymity

"On the Internet No One Knows You're a Dog"

Cartoon in the
New Yorker
July 5, 1993

In 1993, The New Yorker magazine published a famous cartoon with the caption, "*On the Internet, No One Knows You're a Dog*". This cartoon has become a cultural icon simply because it encapsulated the idea that you can say whoever you are, you can say whatever you are, you can make up your personality on the internet. And indeed, many people do.

However, as our experience has grown with the internet, we find that this is less and less true. So let's think a little bit about when anonymous transactions are possible on the internet.

Services like **Tor***, using which, one can post anonymously.

There are things like **Bitcoin** where one can use to pay anonymously.

(*Tor (The Onion Router): a search browser is free and enabling anonymous communication.

But many transactions that you might wish to do, even if you're doing them on the internet, need ID. **So**;

*if you're ordering goods that are to be shipped to you, **you must** provide an address,*

*if you're booking an airline ticket or a hotel room, **you'll need** to provide your name and other identification,*

*if you get cellular service, **you've got** to provide your location,*

*if you want effective medical care, **you have to** disclose confidential details of your health and lifestyle.*

And so there are some things that you might be able to do without saying who you are, but for most things, you're going to have to say something about you, at least reveal some aspects of yourself.

A thing that people don't realize is how much they can reveal without meaning to.

So, there's this idea called (*egosurfing)* which is searching for yourself on a search engine.

Can Reveal ID Without Meaning To

- **Egosurfing = Searching for yourself on a search engine**
  - *Important for reputation management, etc.*
  - *Most people do it.*
- **Most people egosurf periodically.**
- **Most people do not repeatedly look for the same person.**

=> Search pattern for person can reveal your identity.

This is something that is actually recommended by reputation management consultants. We've got to see what somebody will find about you if they google you on the web. Not even knowing that means you can't manage your reputation on the web. And so, *a lot of people egosurf periodically.*

So, if a web search company can observe your web search history and they see a particular person has been looked for repeatedly over a number of years, *that is most likely you*

So, even without your saying anything about who you are, simply based on your searching for yourself and the fact that a *lot of us egosearch*, *the web search company knows your identity.*

And this is all not theoretical. *This actually happened*.

So AOL* in 2006 released three months' worth of search logs for a chunk of users, and they actually had good intentions in mind releasing this.

They released it for research purposes, this is new technology, they wanted to get the research community, the computer scientists, to help improve the way in which AOL and other companies responded to search requests.

*(*AOL : American On Line; a Web search company)*

However, once the search logs were there , two New York Times journalists used the data to identify several users, and one of them was **_Thelma Arnold_** of Lilburn, from Georgia (GA), and *her searches included the following*:

landscapers in Lilburn, Georgia;

property values in her specific subdivision;

pet care issues for her dog.

These kinds of searches told the journalists quite a lot about who she might be. *And so even though she didn't actually search for herself by name, they were very quickly able to zero-in on who this person was.*