HOW WILL THE GDPR IMPACT MACHINE LEARNING?

Steve Touw, Chief Technology Officer
Immuta

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 27 April 2016

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Three points for today:

1. Background on the GDPR.
2. Three major questions about its impact on ML.
3. Some caveats.

Three points for today:

1. **Background on the GDPR.**
2. Three major questions about its impact on ML.
3. Some caveats.

A Deeper Dive Into the GDPR

The EU's General Data Protection Regulation was enacted on 27 April 2016, with a two-year grace period before enforcement, which will began 25 May 2018. With fines of up to 4 percent of global turnover, GDPR represents the biggest regulatory change in the usage of global this century (if not ever).

Key facts:

- The regulation comprises 88 pages, 99 **Articles**, and 173 **Recitals**.
- Applies to EU "personal data"... even extraterritorially.
- Enforced by state-level Data Protection Authorities within the EU, meaning interpretation will likely vary by country.

4.5.2016 EN Official Journal of the European Union L 119/1

I

(Legislative acts)

REGULATIONS

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 27 April 2016

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 16 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee (1),

Having regard to the opinion of the Committee of the Regions (2),

Acting in accordance with the ordinary legislative procedure (3),

Whereas:

(1) The protection of natural persons in relation to the processing of personal data is a fundamental right. Article 8(1) of the Charter of Fundamental Rights of the European Union (the 'Charter') and Article 16(1) of the Treaty on the Functioning of the European Union (TFEU) provide that everyone has the right to the protection of personal data concerning him or her.

(2) The principles of, and rules on the protection of natural persons with regard to the processing of their personal data should, whatever their nationality or residence, respect their fundamental rights and freedoms, in particular their right to the protection of personal data. This Regulation is intended to contribute to the accomplishment of an area of freedom, security and justice and of an economic union, to economic and social progress, to the strengthening and the convergence of the economies within the internal market, and to the well-being of natural persons.

(3) Directive 95/46/EC of the European Parliament and of the Council (4) seeks to harmonise the protection of fundamental rights and freedoms of natural persons in respect of processing activities and to ensure the free flow of personal data between Member States.

(1) OJ C 229, 31.7.2012, p. 90.
(2) OJ C 391, 18.12.2012, p. 127.
(3) Position of the European Parliament of 12 March 2014 (not yet published in the Official Journal) and position of the Council at first reading of 8 April 2016 (not yet published in the Official Journal). Position of the European Parliament of 14 April 2016.
(4) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (OJ L 281, 23.11.1995, p. 31).

TIME TO GDPR!

01:	12 :	40 :	05
DAYS	HOURS	MINUTES	SECONDS

Information Commissioner's Office

Information Rights Strategic Plan 2017-2021

ico.
Information Commissioner's Office

by Max Metzger

Follow @MetzgerSC

March 23, 2017

ICO to take on 200 staff as it help UK businesses to GDPR compliance

The ICO is set to grow by 40 percent over the next two years to help with the mammoth task of making UK businesses compliant with GDPR before its comes into effect next year.

The Information Commissioner's Office (ICO) is expected to grow its staff by 40 percent in the next few years to bear the weight of incoming European regulation.

Information commissioner Elizabeth Denham: needs to hire 200 more staff

The ICO, which governs data protection in the UK, will add 200 people to its staff of 500 who are already said to be buckling under the pressure. The office may battle with skill shortages for as long as two years as it attempts to hire all the lawyers, investigators and specialists it requires.

Elizabeth Denham, the information commissioner, appeared before the House of Lords on 8 March to discuss the implication of the EU's General Data Protection Regulation (GDPR) and the added resources her office would require.

Helping UK firms comply with the GDPR appears to be at the heart of this new employment drive.

Though the UK is set to leave the EU by 2019, the ICO has been focused on ensuring that UK firms are compliant with the regulation. Denham told parliament last week that the ICO's cooperation with EU member data protection authorities will also be ramped up as the office looks towards more long-standing data sharing agreements.

MOST READ ON SC

1. Women of influence in UK cyber security 2017: 20 women to watch
2. Locky ransomware back in huge spam campaign; new variant escapes sandbox
3. Spambot weaponises 711M accounts to spread Ursnif malware
4. 320 m compromised passwords hashes cracked by research 'cracktivists'
5. Key-logging malware, dubbed EHDevel, found intelligence gathering

“Under the EU data protection reform package (GDPR) we will also see an increase in the scale and impact of the sanctions at our disposal. We are committed to using these increased powers in ways which target the most serious areas of non-compliance.”

— UK ICO, 2017-2021 Strategic Plan

MAR 29, 2018 @ 03:41 PM 2,858

Rise Of Les Machines: France's Macron Pledges $1.5 Billion To Boost AI

French President Emmanuel Macron delivers a speech during the 'Artificial Intelligence for Humanity' event in Paris on March 29, [+]

HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

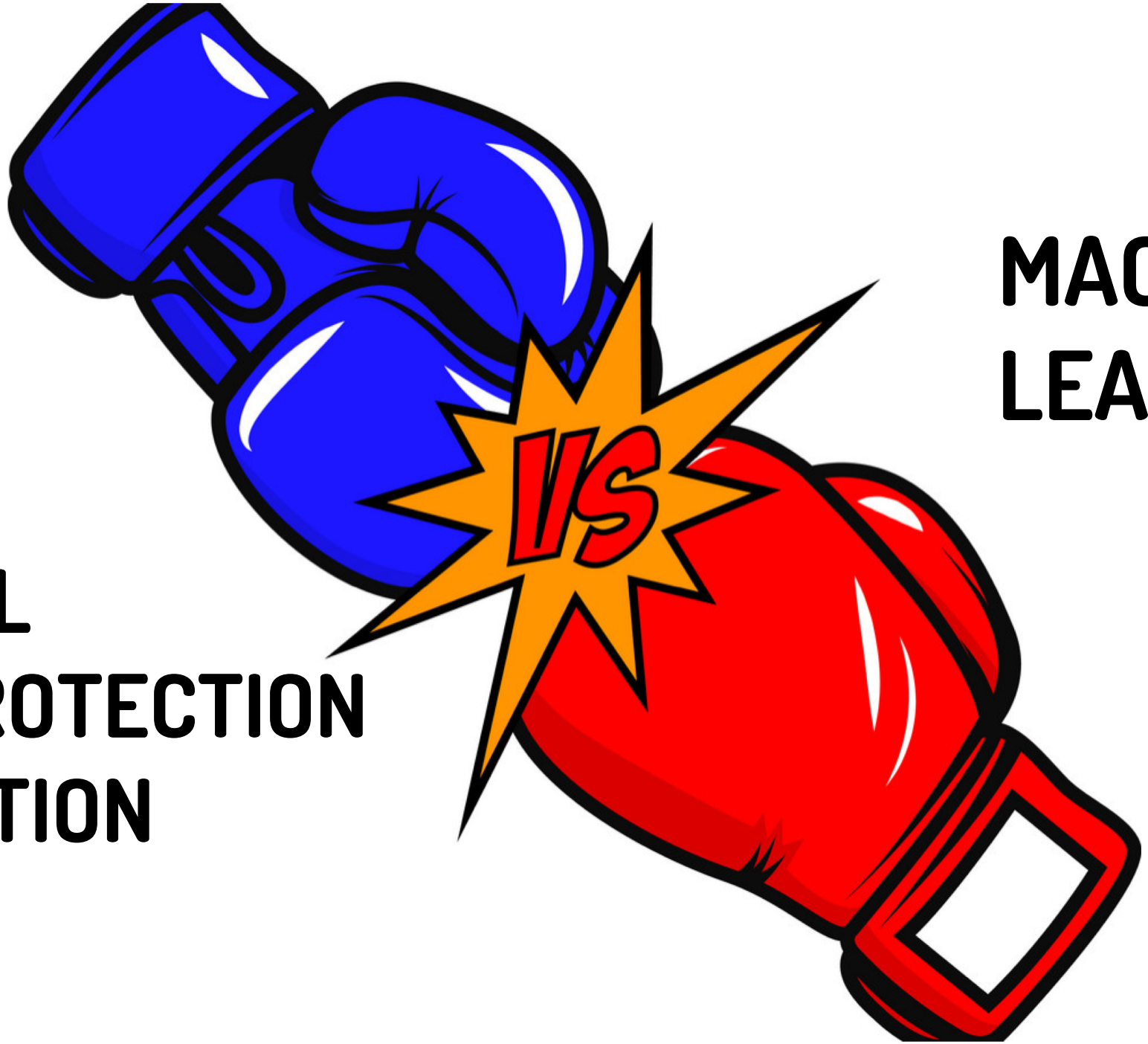
AI in the UK: ready, willing and able?

Ordered to be printed 13 March 2018 and published 16 April 2018

Published by the Authority of the House of Lords

HL Paper 100

MACHINE LEARNING



GENERAL DATA PROTECTION REGULATION

Three points for today:

1. Background on the GDPR.
2. Three major questions about the GDPR's impact on ML.
3. Some caveats.

Three points for today:

1. Background on the GDPR.
2. **Three major questions about the GDPR's impact on ML.**
3. Some caveats.

1. Does the GDPR prohibit machine learning?

Yes! (in theory)

No! (in practice)

Key terms:

Automated decision-making: decisions made about data subjects without direct human intervention.

Profiling: “any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.” Art. 4 (4).

Key terms:

Legal or significant affects: threshold for when the GDPR regulates ML. Examples include “as automatic refusal of an online credit application or e-recruiting practices without any human intervention.” Very wide scope!

Let's Walk Through Some Examples

Example of profiling:

> "A data broker collects data from different public and private sources, either on behalf of its clients or for its own purposes. The data broker compiles the data to develop profiles on the individuals and places them into segments. It sells this information to companies who wish to improve the targeting of their goods and services. The data broker carries out profiling by placing a person into a certain category according to their interests."

From the Article 29 Data Protection Working Party.

Example of automated decision-making:

"Imposing speeding fines purely on the basis of evidence from speed cameras is an automated decision making process that does not necessarily involve profiling.

It would, however, become a decision based on profiling if the driving habits of the individual were monitored over time, and, for example, the amount of fine imposed is the outcome of an assessment involving other factors, such as whether the speeding is a repeat offence or whether the driver has had other recent traffic violations."

From the Article 29 Data Protection Working Party.

Example of NON automated decision-making:

"An automated process produces what is in effect a recommendation concerning a data subject.

If a human being reviews and takes account of other factors in making the final decision, that decision would not be 'based solely' on automated processing.

From the Article 29 Data Protection Working Party.

But... no tricks!

"The controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing.

To qualify as human involvement, the controller must ensure that any **oversight of the decision is meaningful**, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision"

From the Article 29 Data Protection Working Party.

Example of “legal” or “significant” effects:

“Hypothetically, a credit card company might reduce a customer’s card limit, based not on that customer’s own repayment history, but on non-traditional credit criteria, such as an analysis of other customers living in the same area who shop at the same stores.

This could mean that someone is deprived of opportunities based on the actions of others. In a different context using these types of characteristics might have the advantage of extending credit to those without a conventional credit history, who would otherwise have been denied.”

From the Article 29 Data Protection Working Party.

ARTICLE 29 DATA PROTECTION WORKING PARTY

17/EN

WP251rev.01

Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679

Adopted on 3 October 2017

As last Revised and Adopted on 6 February 2018

"[A]s a rule, there is a general prohibition on fully **automated** individual decision-making, including **profiling** that has a **legal or** similarly **significant effect**.

From the Article 29 Data Protection Working Party.

“[A]s a rule, there is a general prohibition on fully automated individual decision-making, including profiling that has a legal or similarly significant effect.

[But] there are **exceptions** to the rule.”

From the Article 29 Data Protection Working Party.

The three exceptions:

(a) necessary for the performance of or entering into a contract;
(b) authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
(c) based on the data subject's explicit consent.

GDPR Art. 22.

The contractual exception:

- For contractual purposes to achieve an appropriate objective that routine human involvement would make impractical or impossible.
- The controller must be able to demonstrate that this type of processing is necessary.
- If less privacy-intrusive methods exist and could achieve the same goals, it's not "necessary"!

Example of contractual necessity for automated decisions:

> "A business advertises an open position. As working for the business in question is popular, the business receives tens of thousands of applications. Due to the exceptionally high volume of applications, the business may find that it is not practically possible to identify fitting candidates without first using fully automated means to sift out irrelevant applications. In this case, automated decision-making may be necessary in order to make a short list of possible candidates, with the intention of entering into a contract with a data subject."

From the Article 29 Data Protection Working Party.

The member state law exception:

- Another law explicitly authorizes the processing.
- That law must also lay down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests.
- GDPR Recital 71 gives a few examples: monitoring and preventing fraud and tax-evasion, or to ensure the security and reliability of a service provided by the controller.

The consent exception:

- “Explicit consent” required
- “Explicit consent” happens to not be defined in the GDPR!
- For now, think of this as clear and unambiguous agreement to have a data subject’s data used for this specific purpose

Example of “consent”:

“By clicking ‘**submit**,’ I agree to allow the data I have provided, along with other data about me possessed by Acme Corp., to assess the likelihood of me adhering to this payment plan.

I consent to allow statistical methods to automatically determine this likelihood without direct human intervention . . .”

So does the GDPR *prohibit* ML?

- In practice, no.
- But a high compliance burden, especially regarding consent.
- You can't just deploy an ML model without **VERY careful legal planning** if it uses a data subject's data, **even if the model doesn't interact with the subject directly**.

2. Is there a “right to explainability” from ML?

No!*

*Kind of.

Intelligent Machines

The Dark Secret at the Heart of AI

No one really knows how the most advanced algorithms do what they do. That could be a problem.

by Will Knight April 11, 2017

Where does the GDPR talk about explainability?

- Articles 13-15
- Article 22
- Recital 71

Articles 13-15

the controller shall, at the time when personal data are obtained, provide the data subject . . . [in the case of] automated decision-making, including profiling, . . . **meaningful information about the logic** involved, as well as the **significance and the envisaged consequences** of such processing for the data subject.

Article 22

The data subject shall have **the right not to be subject to a decision based solely on automated processing**, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

Recital 71

[Automated decision making] should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, **to obtain an explanation of the decision reached after such assessment** and to challenge the decision.

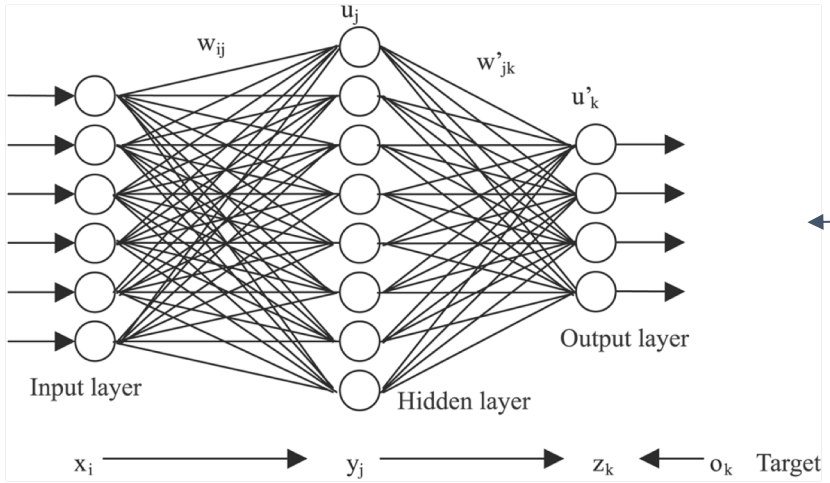
Is Explaining models possible?

Maybe?
...Probably not...

One School of Thought

The field of Artificial Intelligence will eventually be able to solve this problem

Examples:
CAMEL: Causal Models to Explain Learning (DARPA research)
LIME: Local Interpretable Model-Agnostic Explanations



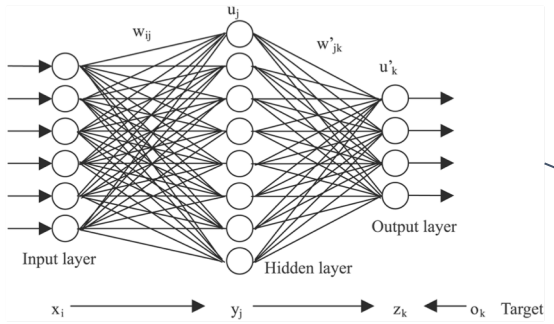
Boil this mess into understandable concepts

The Other School of Thought

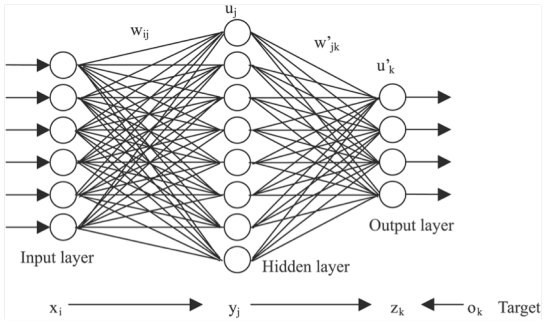
We'll never understand, and that's the point

Think about "Multiplicity of Good Models"

input x, y, z



input a, b, c



Same conclusion

"We are embarking on the age of the impossible-to-understand reason, when marketers will know which style of shoe to advertise to us online **based on the type of fruit we most often eat for breakfast**, or when the police know which group in a public park is most likely to do mischief **based on the way they do their hair or how far from one another they walk**"

-Paul Ohm, THE INTUITIVE APPEAL OF EXPLAINABLE MACHINES

Let’s Walk Through Some Working Party 29 Examples

Example of "meaningful information about the logic":

This may include, for example:

- the information provided by the data subject on the application form;
- information about previous account conduct , including any payment arrears;
- and official public records information such as fraud record information and insolvency records.

The controller also includes information to advise the data subject that the credit scoring methods used are regularly tested to ensure they remain fair, effective and unbiased. The controller provides contact details for the data subject to request that any declined decision is reconsidered. "

From the Article 29 Data Protection Working Party.

Example of “significance” and “envisaged consequences”:

An insurance company uses an automated decision making process to set motor insurance premiums based on monitoring customers’ driving behaviour.

To illustrate the significance and envisaged consequences of the processing it explains that dangerous driving may result in higher insurance payments and provides an app comparing fictional drivers, including one with dangerous driving habits such as fast acceleration and last-minute braking. It uses graphics to give tips on how to improve these habits and consequently how to lower insurance premiums.

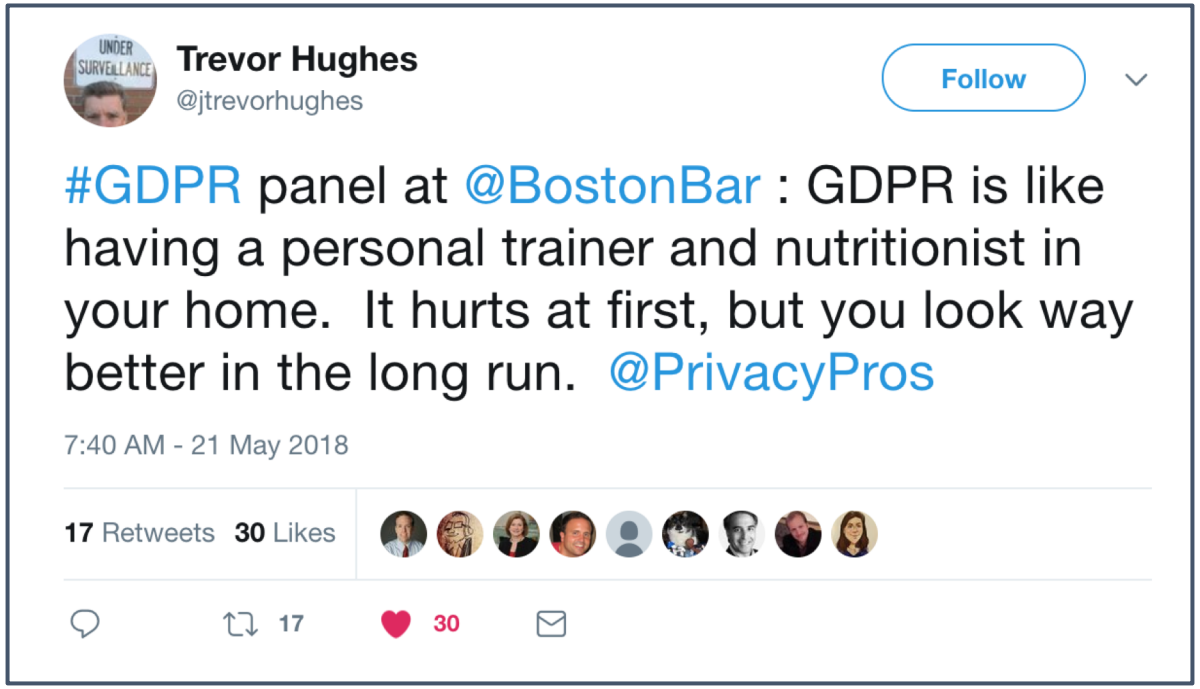
From the Article 29 Data Protection Working Party.

What to make of Recital 71?

- Non-binding!
- Great in theory, and something that DPA's will want organizations to aspire to.
- But far from mandatory and extremely hard to achieve and implement.

So is there a right to explainability in the GDPR?

- In practice, not *technically*.
- But data subjects must have a basic understanding of how ML models are using their data and what that means.
- This is about EMPOWERMENT, which is a good thing!



3. Do data subjects have the ability to demand that models be retrained without their data? e.g. “the right to be forgotten”

No!*

*Maybe.

When could withdrawal of consent force the retraining of a model?

Only if the data is still being *used*.

Is it possible to consider the model's training data to still be being used in a production model?

- You really have to split hairs on the word "used"
- Is the data still in the model? Yes, but in a complex way

Model Inversion Attack

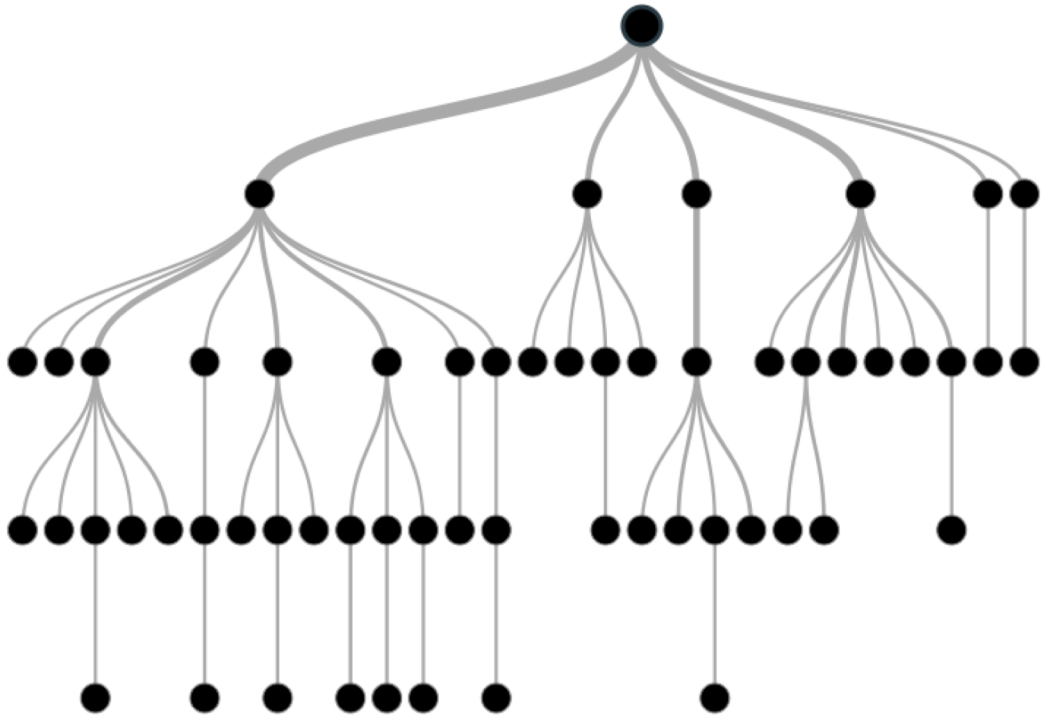
Fredrikson et al. (2015) Model Inversion Attacks

Public model, you provide image, it spits out name

But, you could invert this. Knowing a name, you can iteratively un-fuzz an image until you get the “real” image based on the confidence scores of the model

Model Inversion Attack 2

Fredrikson et al. (2015) Model Inversion Attacks



Lifestyle survey to predict x

But, you can predict how someone will answer a sensitive question. In this case if they cheated on their spouse!

This is done by using the context of the other answers in the decision tree model

Model Inversion Attack - Brought to you by Explainability!!

Fredrikson et al. (2015) Model Inversion Attacks

- In both cases, details about the models inner-workings enabled the attack!
- If the confidence scores for the images were rounded, the attack was not nearly as effective
- If you didn't know the structure of the decision tree, the attack would be impossible

Conundrum: explainability could result in privacy attacks...

Is this likely to occur in practice?

- For the moment, this is theoretical.
- Don't expect regulators to force model retraining under this argument.
  - But there are techniques to enable privacy preserving models to be built
- Technically, though, this is a possibility in the future!

Three points for today:

1. Background on the GDPR.
2. Three major questions about the GDPR's impact on ML.
3. Some caveats.

Three points for today:

1. Background on the GDPR.
2. Three major questions about the GDPR's impact on ML.
3. **Some caveats.**

There's a lot we don't know:

- A huge amount of nuance to all these questions - and future nuances will surely arise.
- With 99 Articles and 173 Recitals, the GDPR is likely to get more complex over time as it's enforced by regulators.
- What we do know: the most important component of your ML programs could be your lawyers and privacy engineers!

QUESTIONS?

Contact Me:

steve@immuta.com

@steve_touw

VISIT OUR BOOTH

IMMUTA

booth #215