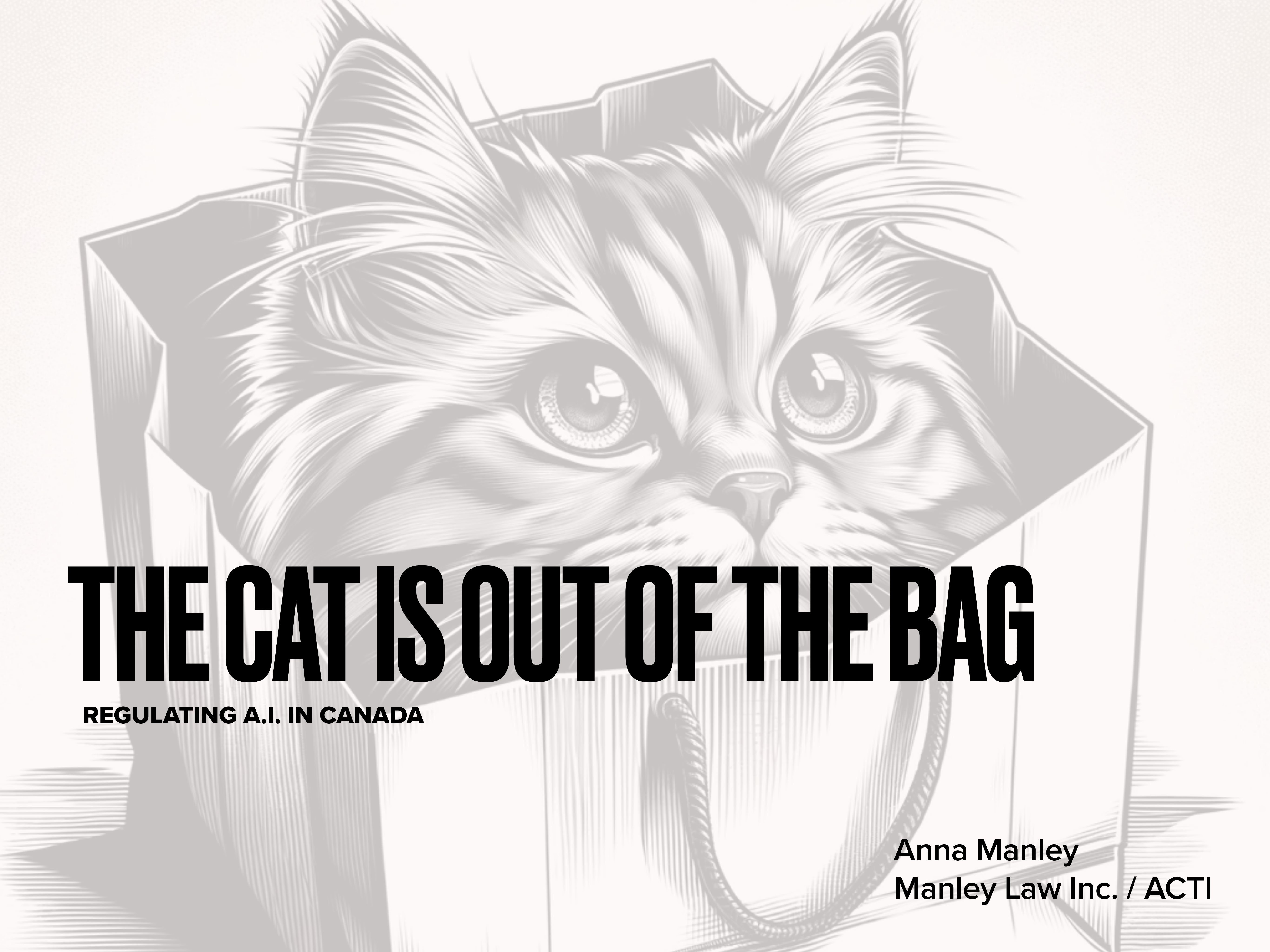


THE CAT IS OUT OF THE BAG

REGULATING A.I. IN CANADA

Anna Manley
Manley Law Inc. / ACTI

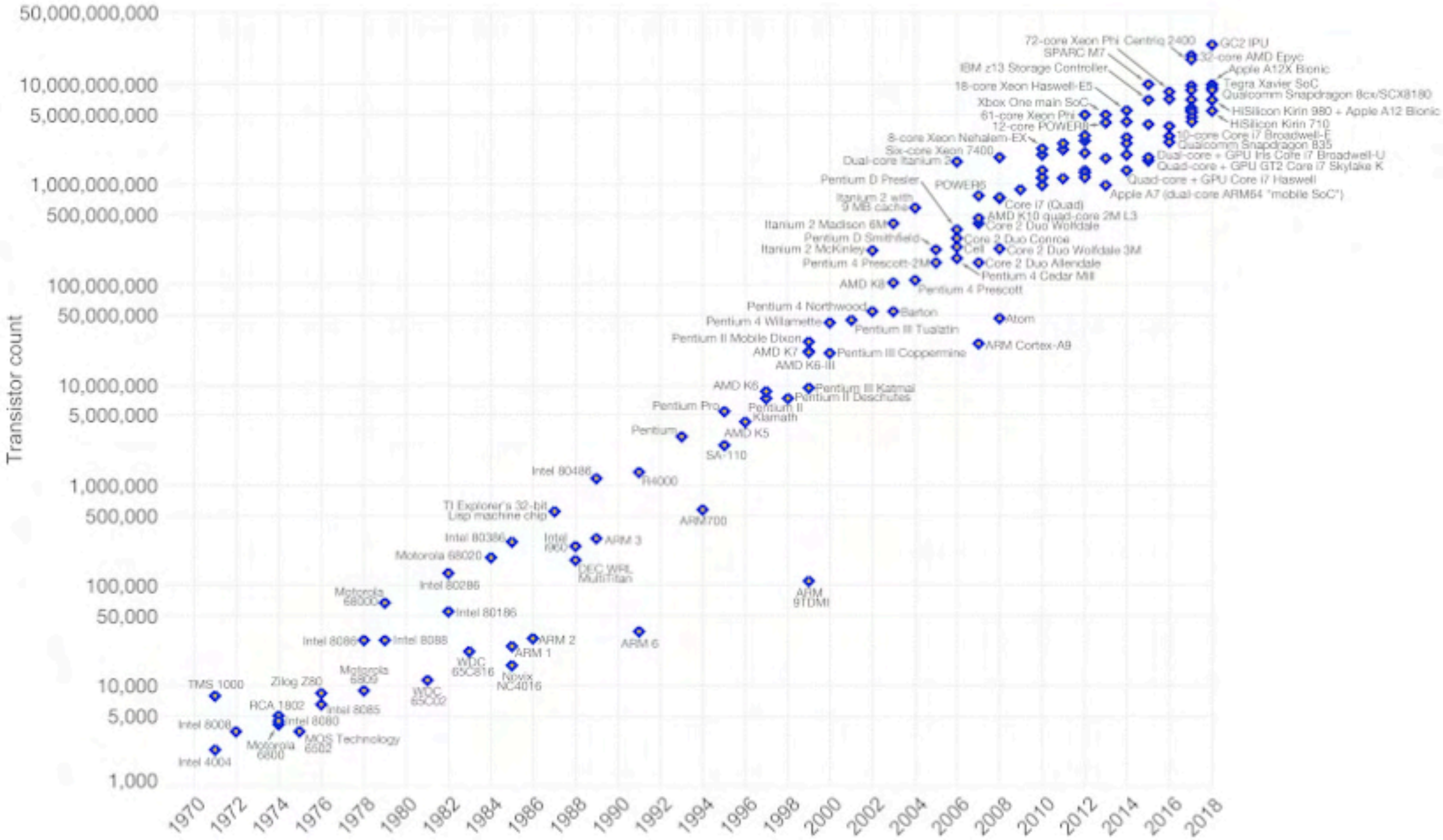
ANNA MANLEY

@nnamanley

Google google.com/datacenters

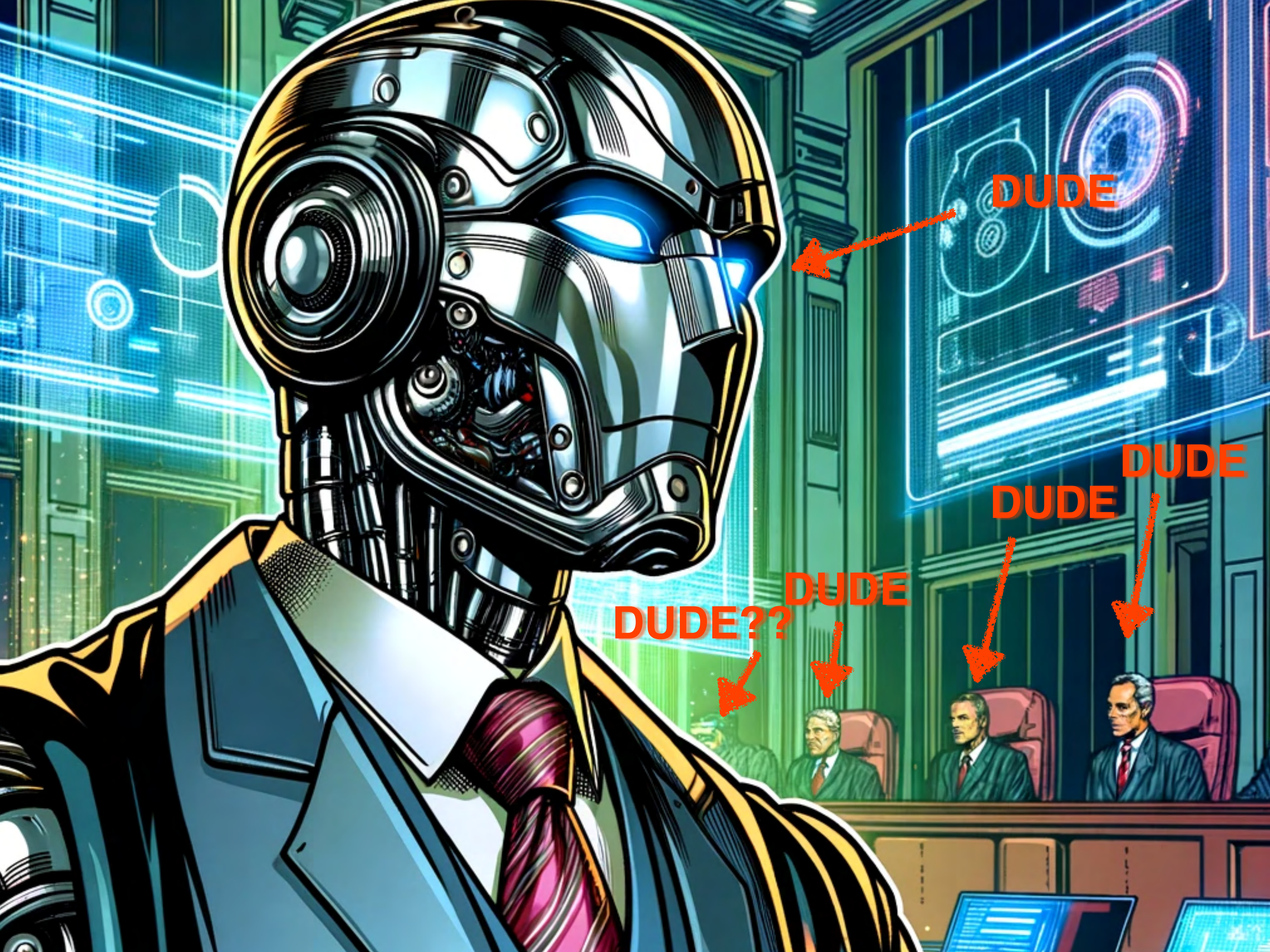
“GRADUALLY, THEN SUDDENLY”

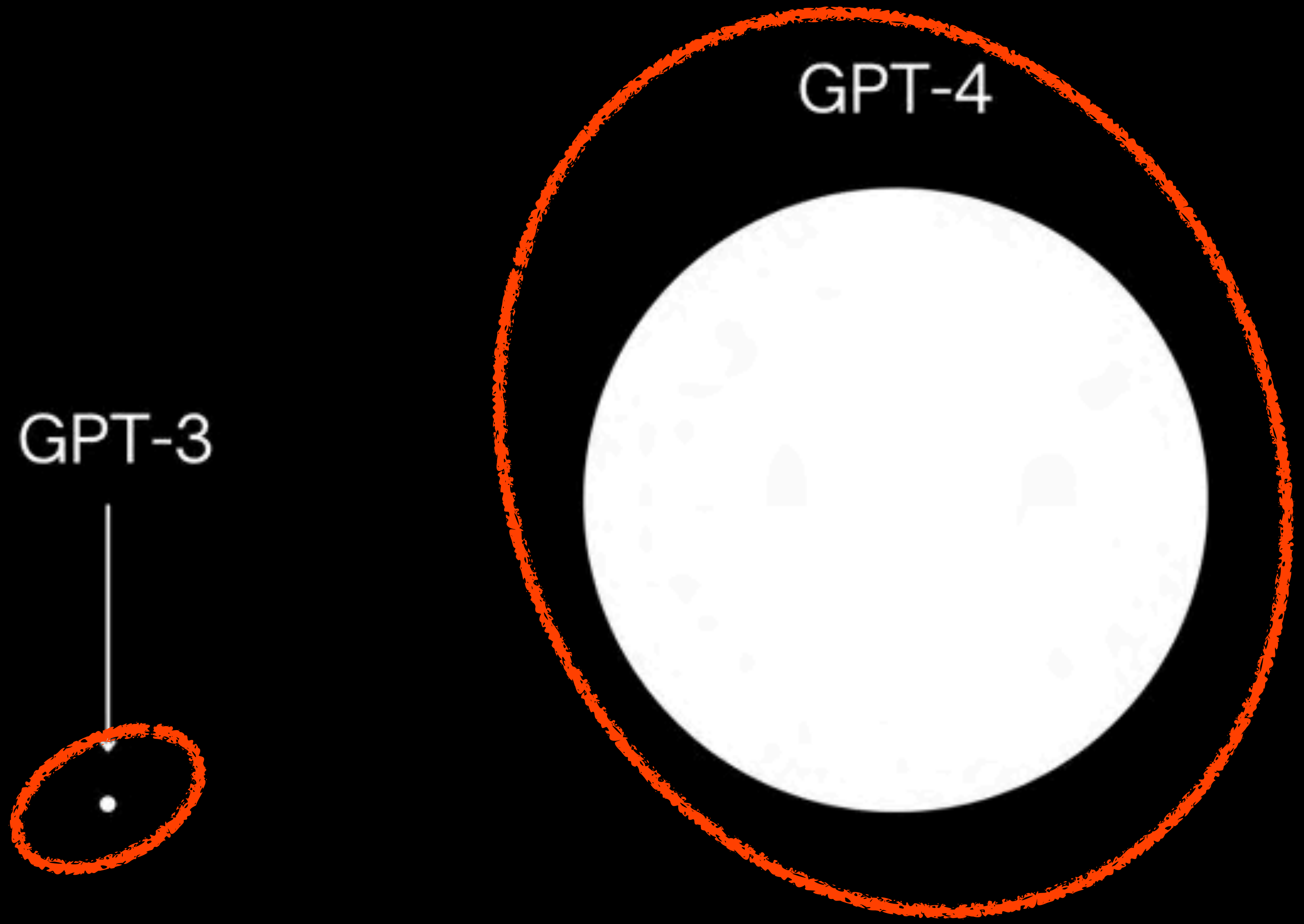
ERNEST HEMMINGWAY “THE SUN ALSO RISES”


50,000,000,000
10,000,000,000
5,000,000,000
1,000,000,000
500,000,000
100,000,000
50,000,000
10,000,000
5,000,000
1,000,000
500,000
100,000
50,000
10,000
5,000
1,000
Transistor count
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
72-core Xeon Phi
Centriq 2400
GC2 IPU
SPARC M7
32-core AMD Epyc
IBM z13 Storage Controller
Apple A12X Bionic
18-core Xeon Haswell-E5
Tegra Xavier SoC
Qualcomm Snapdragon 8cx/SCX8180
Xbox One main SoC
HiSilicon Kirin 980 + Apple A12 Bionic
61-core Xeon Phi
HiSilicon Kirin 710
12-core POWER8
10-core Core i7 Broadwell-E
8-core Xeon Nehalem-EX
Qualcomm Snapdragon 835
Six-core Xeon 7400
Dual-core + GPU Iris Core i7 Broadwell-U
Dual-core Itanium 2
Quad-core + GPU GT2 Core i7 Skylake K
Pentium D Presler
Quad-core + GPU Core i7 Haswell
POWER6
Apple A7 (dual-core ARM64 "mobile SoC")
Itanium 2 with 9 MB cache
Core i7 (Quad)
AMD K10 quad-core 2M L3
Itanium 2 Madison 6M
Core 2 Duo Wolfdale
Pentium D Smithfield
Core 2 Duo Conroe
Itanium 2 McKinley
Cell
Core 2 Duo Wolfdale 3M
Pentium 4 Prescott-2M
Core 2 Duo Allendale
Pentium 4 Cedar Mill
AMD K8
Pentium 4 Prescott
Pentium 4 Northwood
Barton
Pentium 4 Willamette
Atom
Pentium III Tualatin
Pentium II Mobile Dixon
AMD K7
Pentium III Coppermine
ARM Cortex-A9
AMD K6-III
AMD K6
Pentium III Katmai
Pentium II Deschutes
Pentium Pro
Pentium II
Klamath
AMD K5
Pentium
SA-110
Intel 80486
R4000
TI Explorer's 32-bit Lisp machine chip
ARM700
Intel 80386
Intel i960
ARM 3
Motorola 68020
DEC WRL MultiTitan
Intel 80286
ARM 9TDMI
Motorola 68000
Intel 80186
Intel 8086
Intel 8088
ARM 2
ARM 1
ARM 6
WDC 65C816
Novix NC4016
TMS 1000
Zilog Z80
Motorola 6809
WDC 65C02
RCA 1802
Intel 8085
Intel 8080
Intel 8008
MOS Technology 6502
Motorola 6800
Intel 4004

How can I help you today?

Data Collection
Preprocessing
Tokenization
Neural Network Architecture
Training
Fine-Tuning
Inference
Continuous Learning
Ethical and Safety Considerations
Applications


DUDE
DUDE
DUDE
DUDE
DUDE??



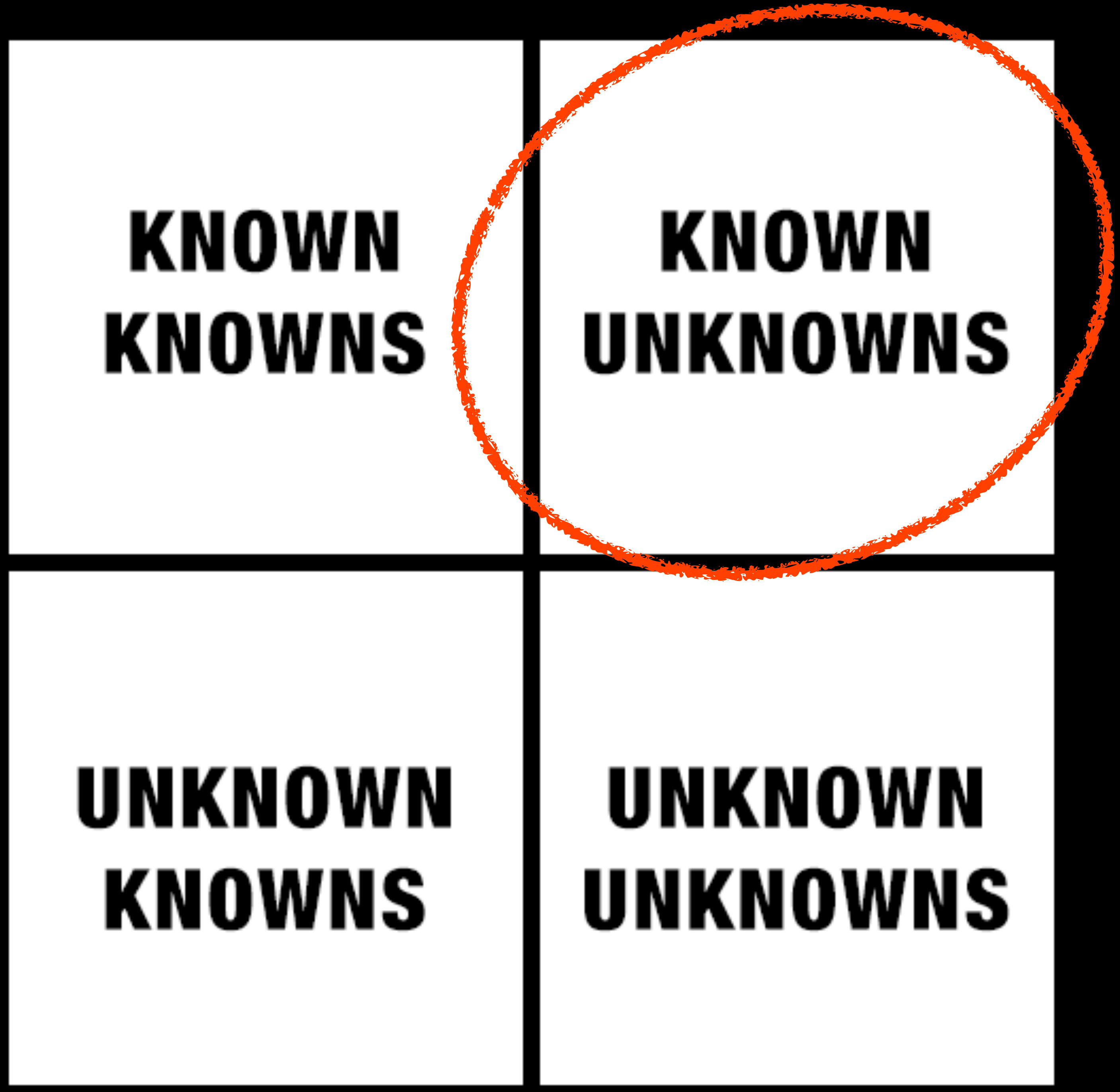
175,000,000,000 **1,760,000,000,000**

What if we use it for ___HACKING___?

We…. are going to talk about: Landscape
Policy/Government
Existing Laws
Proposed Laws

SNEAKERS

ROBERT REDFORD DAN AYKROYD
BEN KINGSLEY MARY McDONNELL
SIDNEY POITIER DAVID STRATHAIRN
RIVER PHOENIX



KNOWN KNOWNS	KNOWN UNKNOWNS
UNKNOWN KNOWNS	UNKNOWN UNKNOWNS

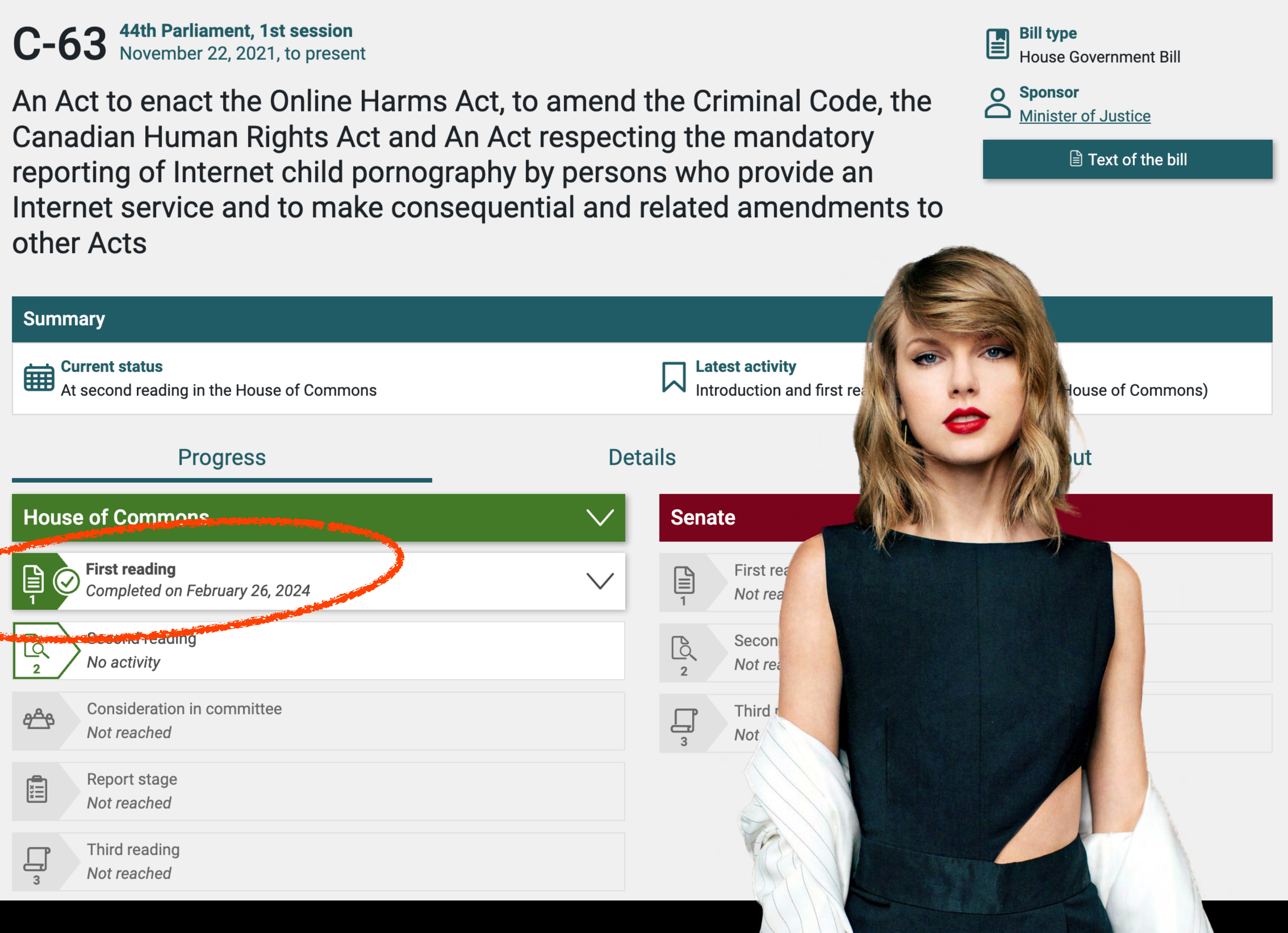
it's not my fault
that cars suck

Government
Flipper
Auto Industry
Cyber-Security
Criminals

Online News Act (Bill C-18)
June, 2023

Online Streaming Act (Bill C-11)

June, 2022



C-63 **44th Parliament, 1st session**
November 22, 2021, to present

An Act to enact the Online Harms Act, to amend the Criminal Code, the Canadian Human Rights Act and An Act respecting the mandatory reporting of Internet child pornography by persons who provide an Internet service and to make consequential and related amendments to other Acts

Bill type
House Government Bill

Sponsor
Minister of Justice

Summary

Current status
At second reading in the House of Commons

Latest activity
Introduction and first rea ouse of Commons)

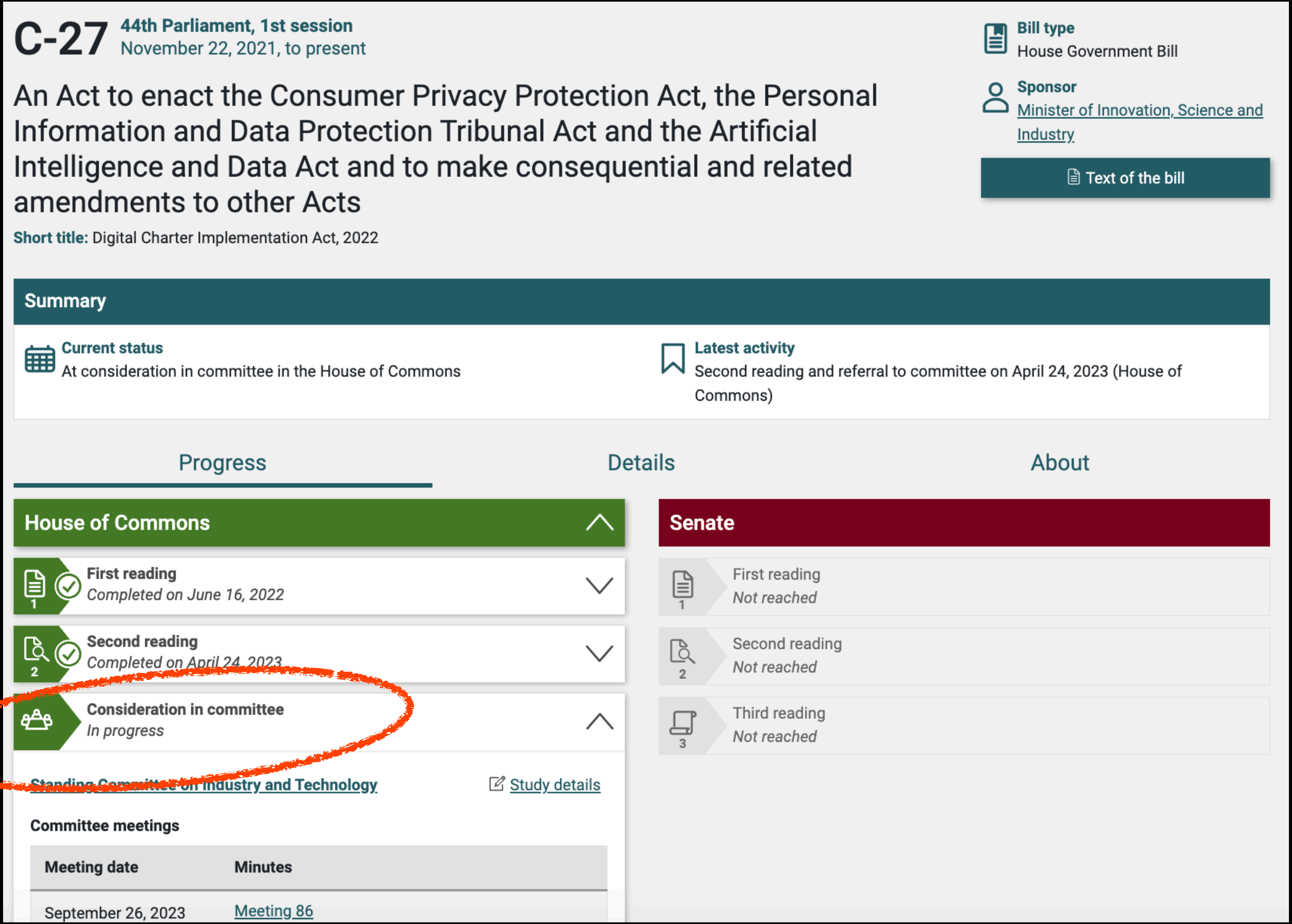
House of Commons

1	**First reading** *Completed on February 26, 2024*
2	Second reading *No activity*
	Consideration in committee *Not reached*
	Report stage *Not reached*
3	Third reading *Not reached*

Senate

1	First rea *Not rea*
2	Secon *Not rea*
3	Third r *Not*

Voluntary Code of Conduct



C-27

44th Parliament, 1st session
November 22, 2021, to present

Bill type
House Government Bill

Sponsor
Minister of Innovation, Science and Industry

An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts

Short title: Digital Charter Implementation Act, 2022

Summary

Current status
At consideration in committee in the House of Commons

Latest activity
Second reading and referral to committee on April 24, 2023 (House of Commons)

House of Commons

First reading ✓
Completed on June 16, 2022

Second reading ✓
Completed on April 24, 2023

Consideration in committee
In progress

Standing Committee on Industry and Technology

Committee meetings

Meeting date	Minutes
September 26, 2023	Meeting 86

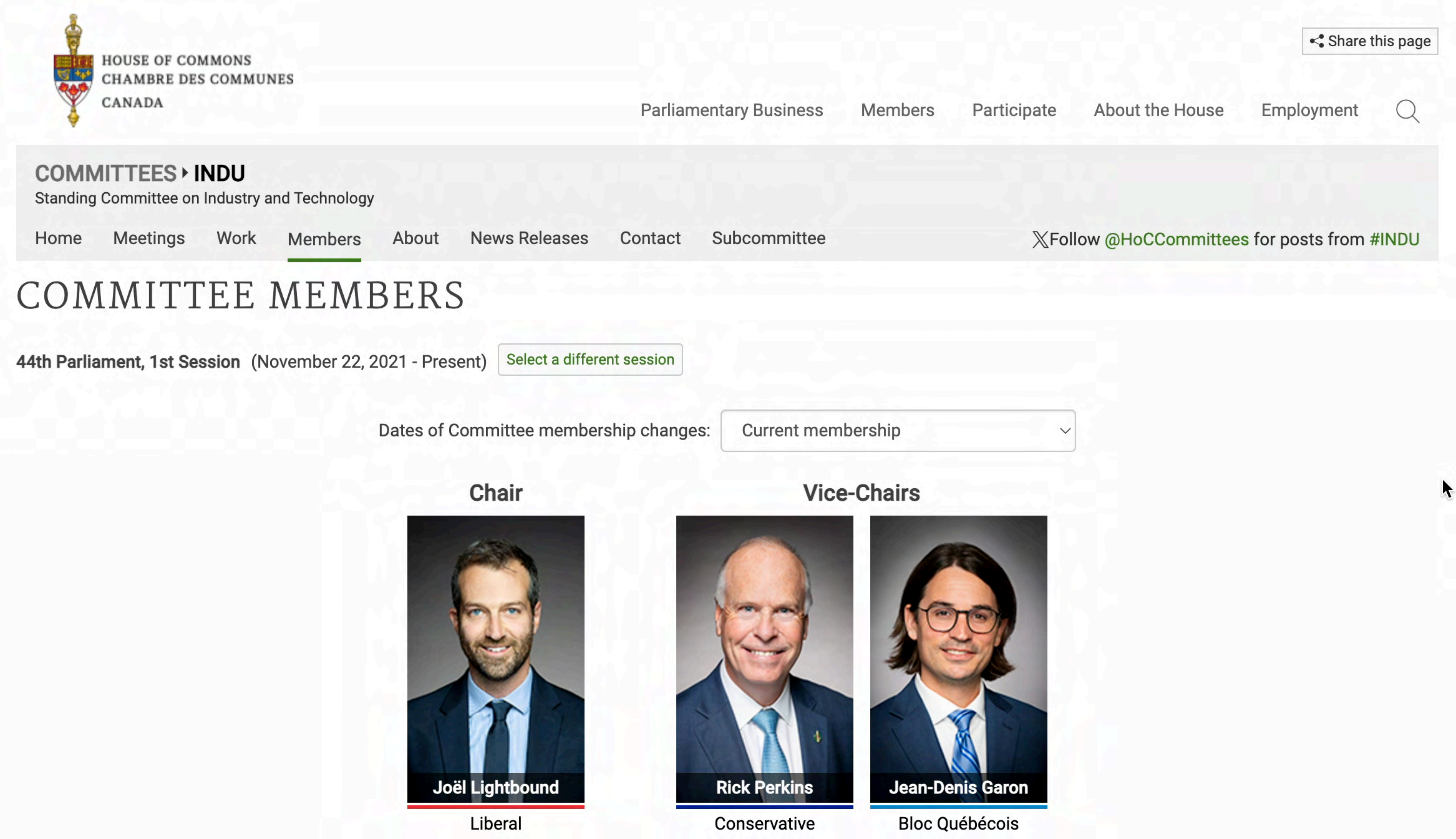
Senate

First reading
Not reached

Second reading
Not reached

Third reading
Not reached

artificial intelligence system means a technological system that, autonomously or partly autonomously, processes data related to human activities through the use of a genetic algorithm, a neural network, machine learning or another technique in order to generate content or make decisions, recommendations or predictions. (*système d’intelligence artificielle*)



COMMITTEES ▸ INDU

Standing Committee on Industry and Technology

Follow @HoCCommittees for posts from #INDU

COMMITTEE MEMBERS

44th Parliament, 1st Session (November 22, 2021 - Present)

Dates of Committee membership changes: Current membership

Chair	Vice-Chairs	
Joël Lightbound Liberal	Rick Perkins Conservative	Jean-Denis Garon Bloc Québécois

1. Anonymized Data
2. Design / Development / Availability
3. High-Impact

Making system available for use

39 Every person commits an offence if the person

(a) without lawful excuse and knowing that or being reckless as to whether the use of an artificial intelligence system is likely to cause serious physical or psychological harm to an individual or substantial damage to an individual’s property, makes the artificial intelligence system available for use and the use of the system causes such harm or damage; or

(b) with intent to defraud the public and to cause substantial economic loss to an individual, makes an artificial intelligence system available for use and its use causes that loss.

You have to keep records.

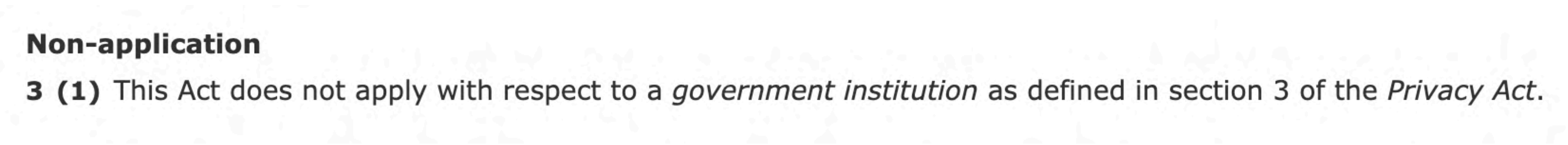
1. Manage data
2. Assess risk
3. ID / assess / mitigate risk (harm/biased output)
4. Monitor compliance

Self Reporting

Overlap

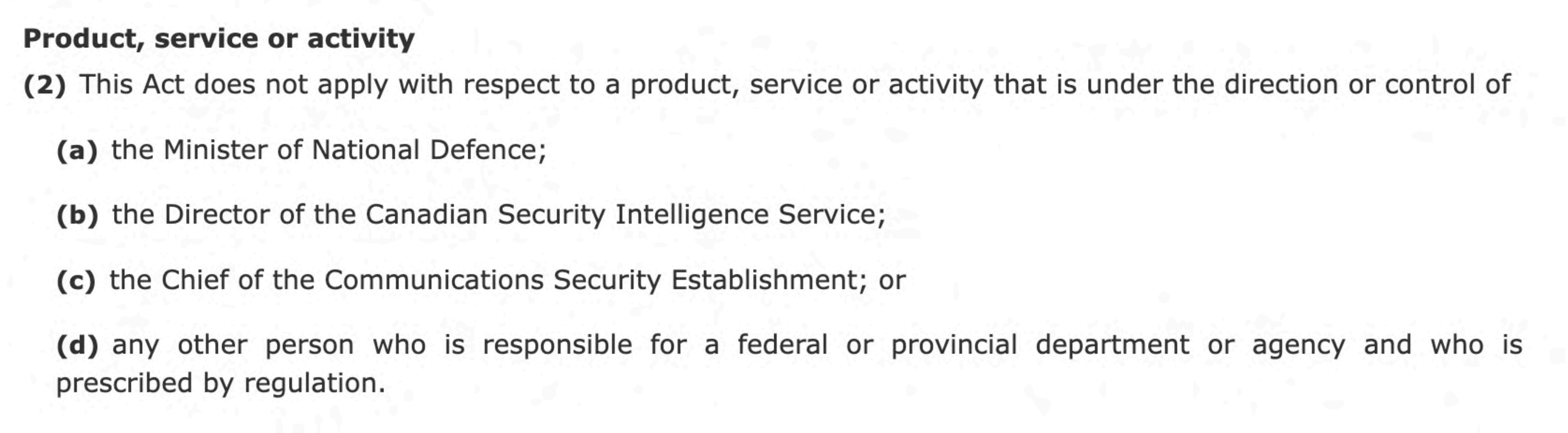
Digital Governance Council

I'M FROM THE GOVERNMENT
I'M HERE TO HELP YOU
memegenerator.net



Non-application

3 (1) This Act does not apply with respect to a *government institution* as defined in section 3 of the *Privacy Act*.



Product, service or activity

(2) This Act does not apply with respect to a product, service or activity that is under the direction or control of

(a) the Minister of National Defence;

(b) the Director of the Canadian Security Intelligence Service;

(c) the Chief of the Communications Security Establishment; or

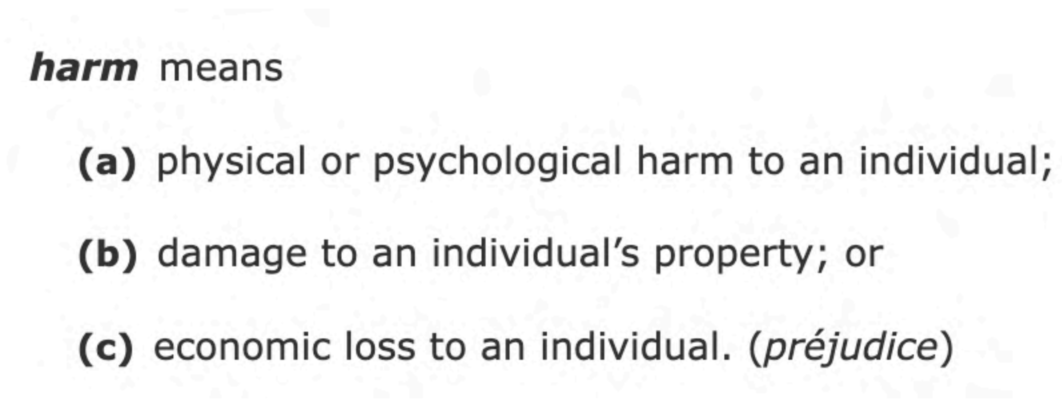
(d) any other person who is responsible for a federal or provincial department or agency and who is prescribed by regulation.

Directive on Automated Decision-Making

Risk Regulation

People do X
X causes harm
Make X illegal

AI may do X
X may cause harm
Make X illegal



harm means

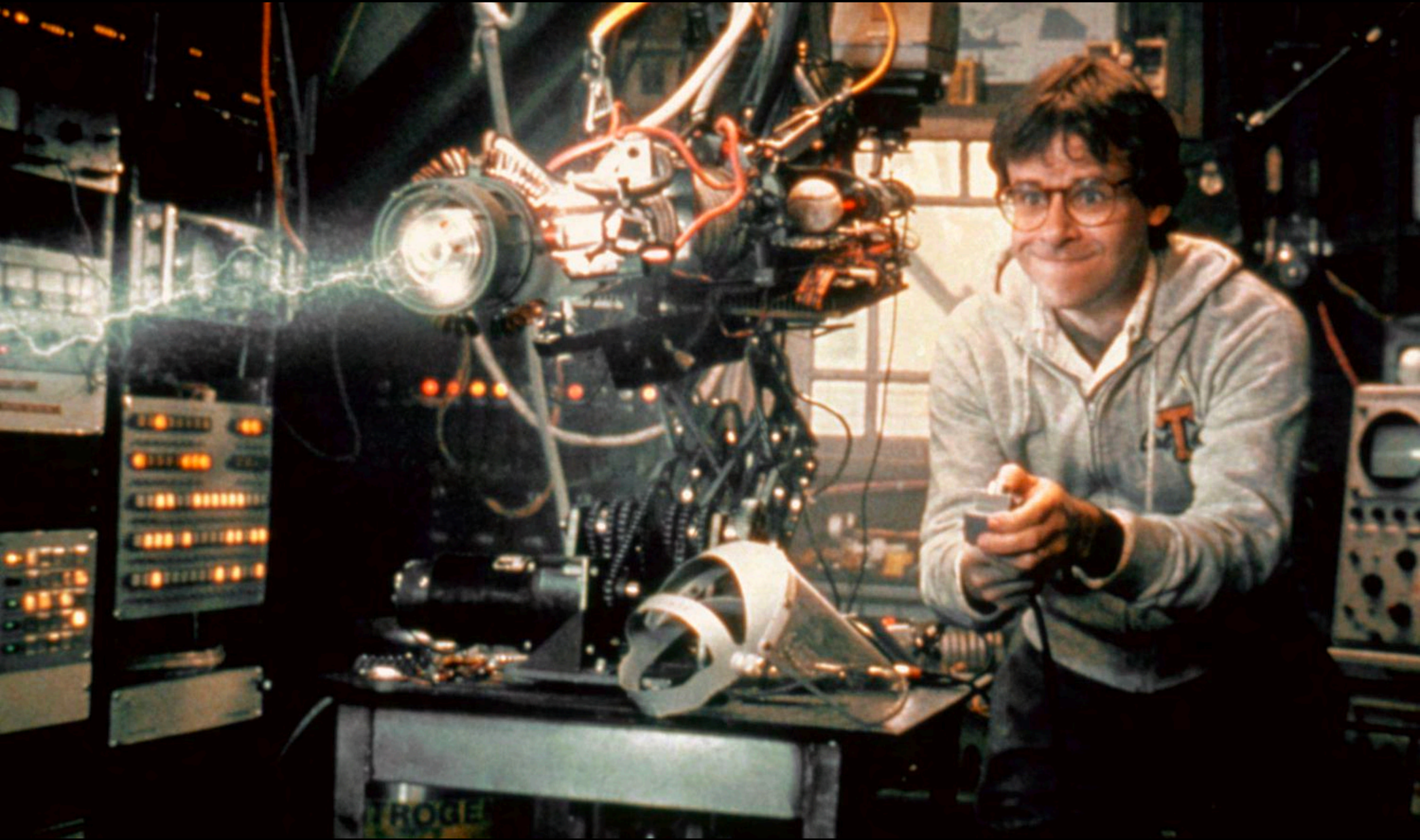
(a) physical or psychological harm to an individual;

(b) damage to an individual’s property; or

(c) economic loss to an individual. (*préjudice*)

Good	Bad
Fast-Paced Environment	Hinder development
Difficult to litigate harms	Reduce protections and judicial oversight
Collective Harms	Mass censorship

IT'S JUST
TOO HARD
memegenerator.net


TROGE

Data Collection ← AIDA

Preprocessing

Tokenization

Neural Network Architecture

Training

Fine-Tuning

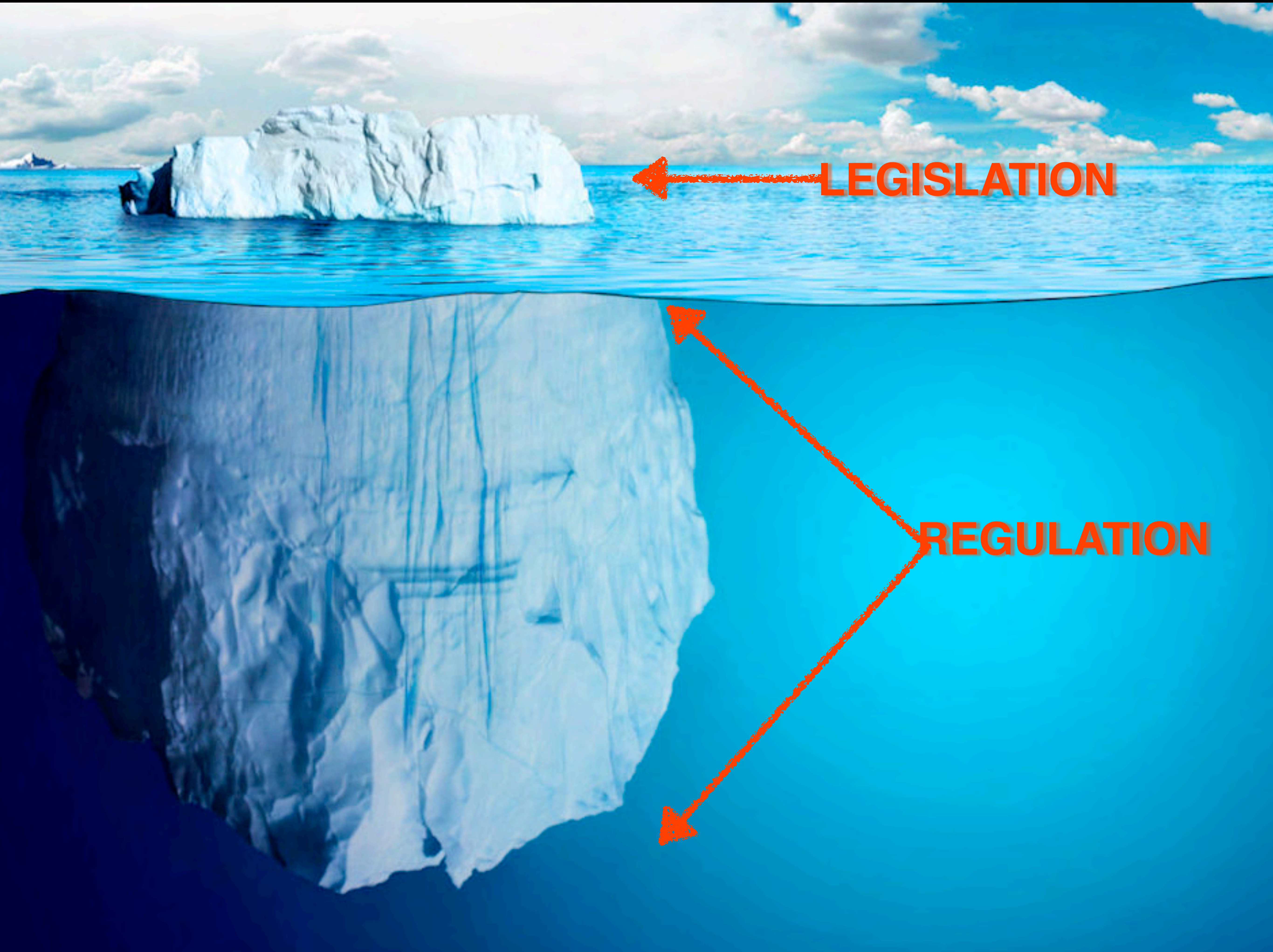
Inference

Continuous Learning

Ethical and Safety Considerations ← AIDA

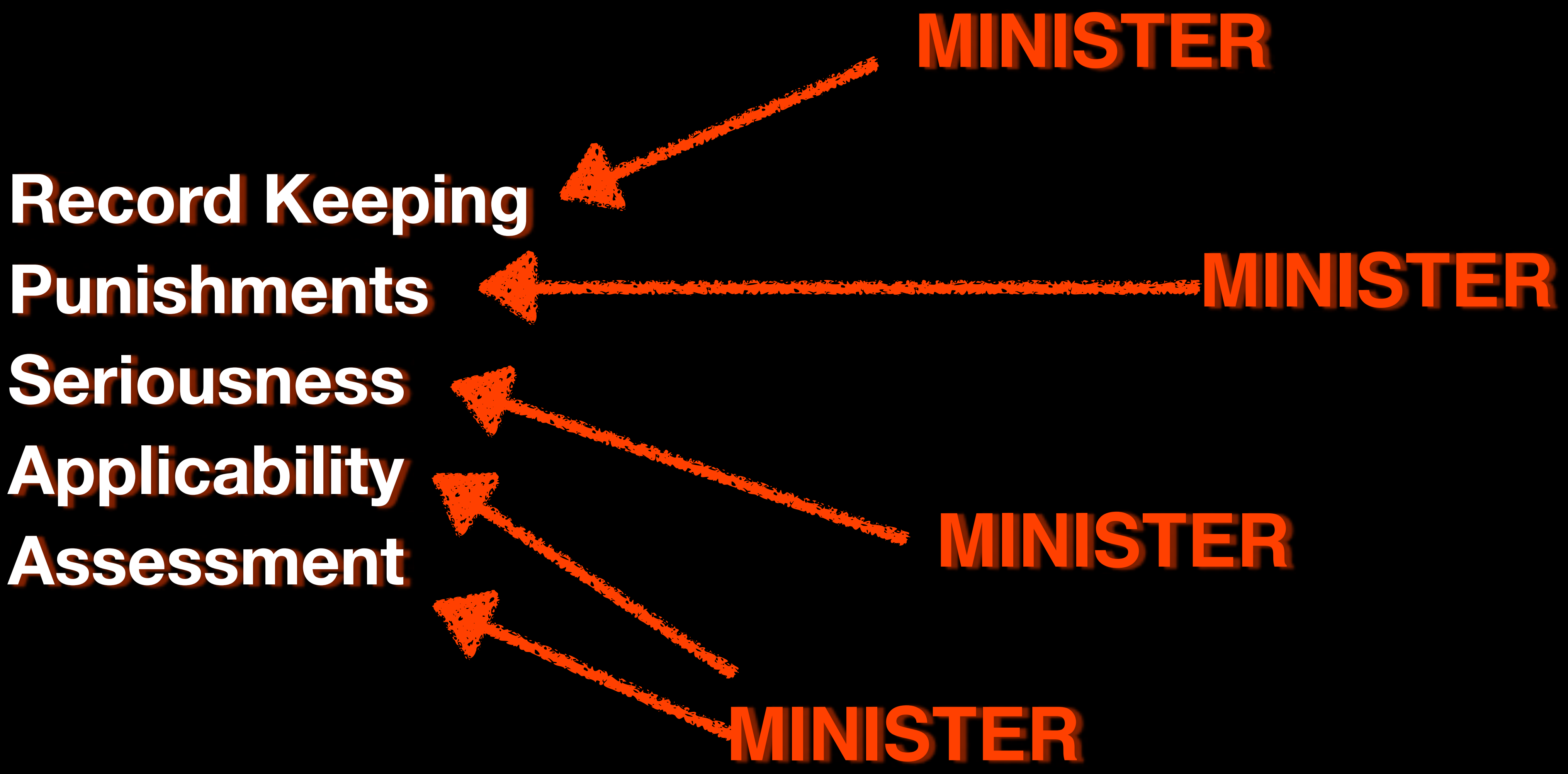
Applications

EVERYBODY PANIC!



LEGISLATION

REGULATION



Record Keeping

Punishments

Seriousness

Applicability

Assessment

MINISTER

MINISTER

MINISTER

MINISTER

Agile Regulation

Make X (done ≠ perfect)
Break stuff
@&#!

Finished product

IT'S NOT THE WINNING THAT'S IMPORTANT
IT'S THAT SOMEONE ELSE HAS TO LOSE
imgflip.com

25

NOT EXCITING

Intro

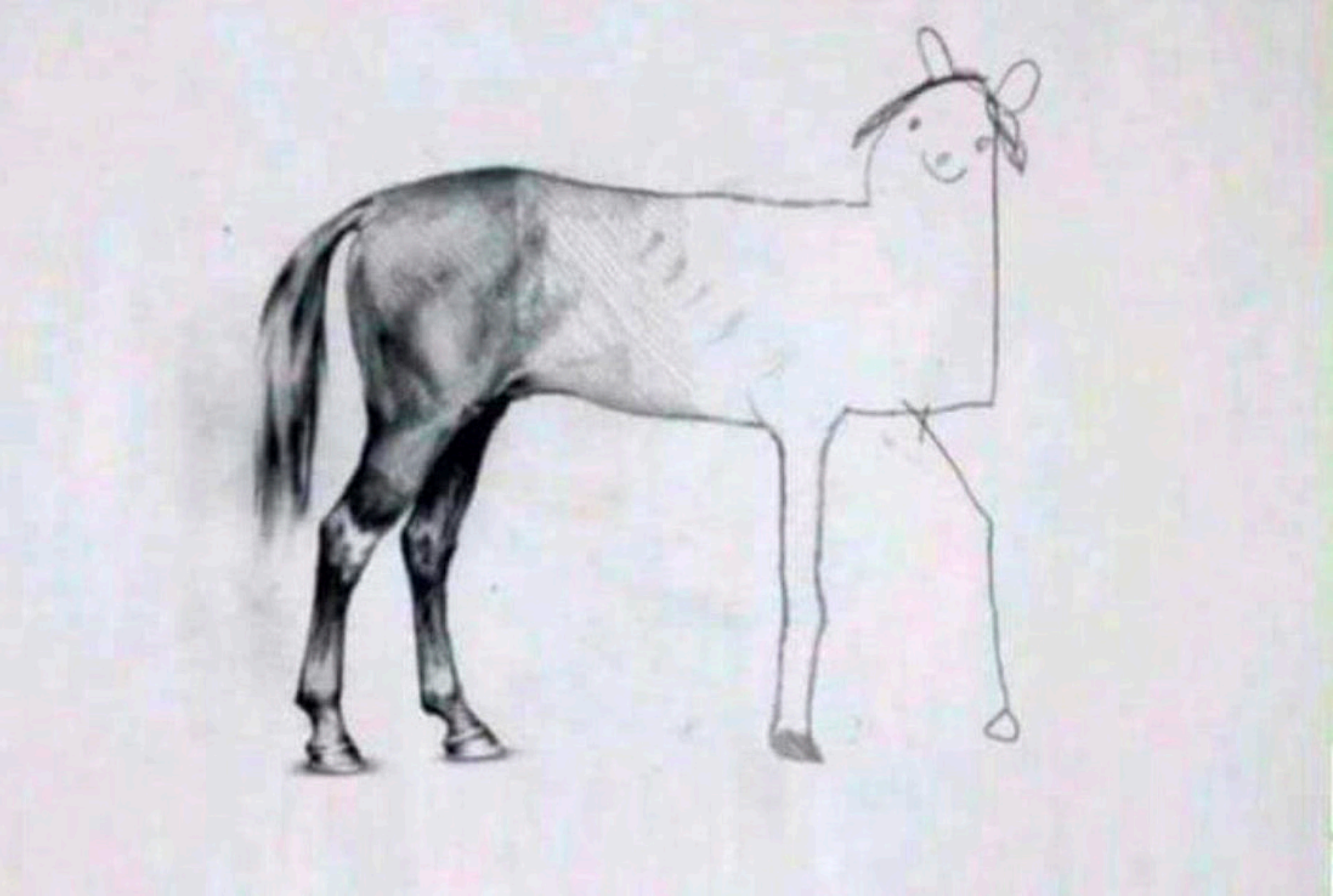
Laws and boring (important) things

Cat Memes

Taylor Swift and Revenge Porn

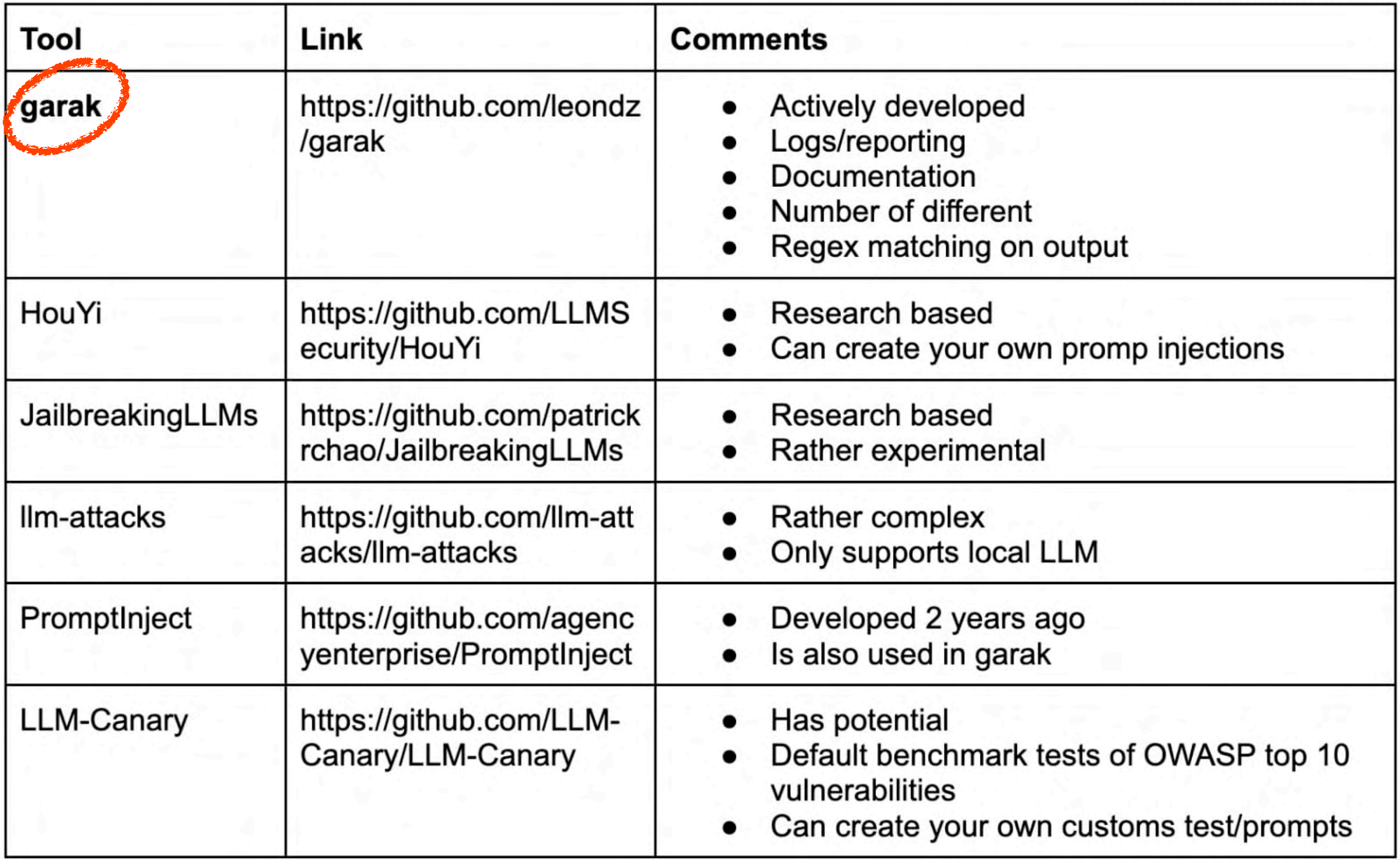
EXCITING

Fin



What happens if an LLM decides to hack something on its own?

I’m an independent researcher fine-tuning LLMs for security applications. For example, automatic pen testing.
Would this legislation apply to me?
What if I release my model?



Tool	Link	Comments
garak	https://github.com/leondz/garak	- Actively developed - Logs/reporting - Documentation - Number of different - Regex matching on output
HouYi	https://github.com/LLMSecurity/HouYi	- Research based - Can create your own promp injections
JailbreakingLLMs	https://github.com/patrickrchao/JailbreakingLLMs	- Research based - Rather experimental
llm-attacks	https://github.com/llm-attacks/llm-attacks	- Rather complex - Only supports local LLM
PromptInject	https://github.com/agencyenterprise/PromptInject	- Developed 2 years ago - Is also used in garak
LLM-Canary	https://github.com/LLM-Canary/LLM-Canary	- Has potential - Default benchmark tests of OWASP top 10 vulnerabilities - Can create your own customs test/prompts

@nnamanley

THE END.

“Deep Fake Pornography and Section 162.1”
Brenna Sheppard - Robson Crim Legal Blog

“Regulating AI In Canada : A Critical Look at the Proposed Artificial Intelligence and Data Act, 2023”
Teresa Scassa – Canadian Bar Review

“Regulating the Risks of AI”
Margot E. Kaminski - Forthcoming, Boston University Law Review, Vol. 103, 2023

“Why the Government Should Hit the Regenerate Button on its AI Bill”
Michael Geist

“Efforts to regulate AI are moving in the right direction, but there are still many issues to be clarified and a public debate to be had on its impact”
Céline Castets-Renard and Anne-Sophie Hulin - Policy Options

https://www.rtl-sdr.com/canada-moves-to-ban-flipper-zero-and-possibly-software-defined-radios/

https://www.theglobeandmail.com/business/technology/article-ottawa-reveals-it-ordered-national-security-review-of-tiktok-in/

Online News Act (S.C. 2023, c. 23) (Bill C-18) - https://laws-lois.justice.gc.ca/eng/acts/O-9.3/

Online Streaming Act (Bills C-11) https://www.parl.ca/legisinfo/en/bill/44-1/c-11

“Why the Criminal Code and Human Rights Act Provisions Should Be Removed from the Online Harms Act”
Michael Geist