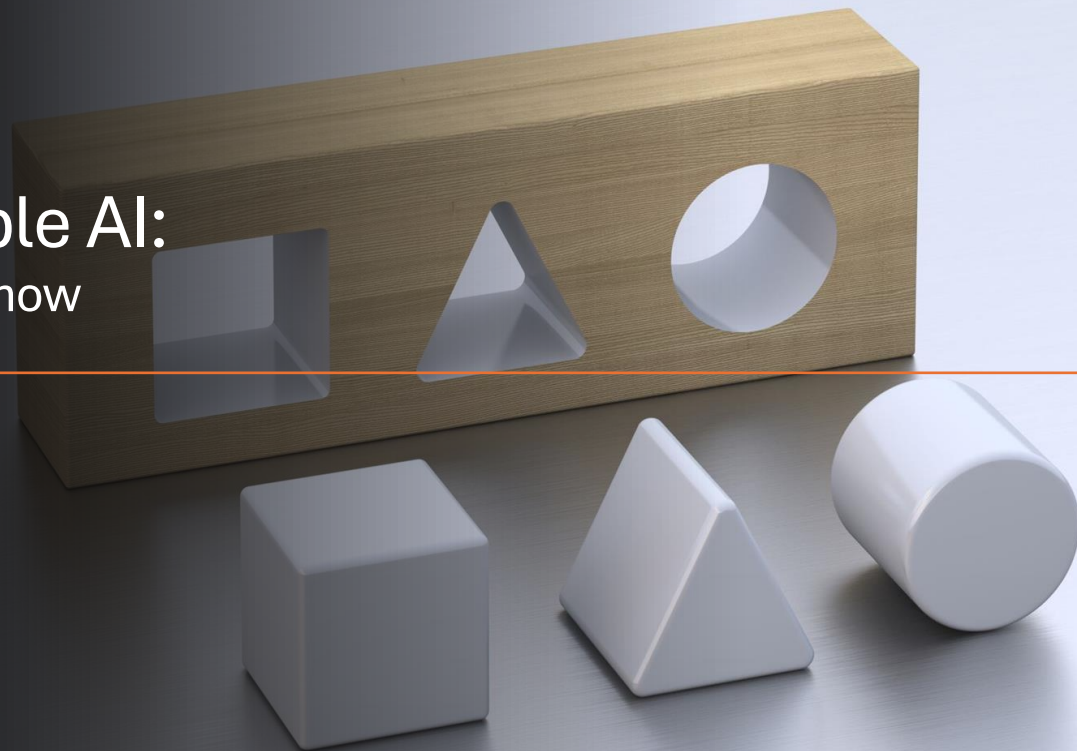


Privacy and Responsible AI:

What Data Scientists Need To Know

Liza Wood
Oct 23, 2023



Nice To Meet You!



Liza Wood
Principal Program Manager



Where I Work

Community Sift is a content moderation platform that classifies, filters, and escalates up to **100 billion** human interactions *per month*, including:

- messages,
- usernames,
- images and videos,

all in **real-time**.

By surfacing **cyberbullying, abuse, hate speech, violent threats**, and **child exploitation**, we enable clients across a variety of social networks to foster safe and healthy online user experiences.

100
BILLION 
lines of chat processed
in one month

6x the monthly
volume of
Twitter

**NOT JUST DATA.
COMPLEX HUMAN
INTERACTIONS.**

Unmatched experience in community management

The deepest experience in the industry

11 YEARS

1 TRILLION+
Human Interactions per Year

22 LANGUAGES

Familiarity with many unique industries



Gaming



Social



Travel



News/Media



Streaming



Education
/ EdTech



Kids



Sports /
Theme Parks



What My Teams Do

Data Rangers

- Data Scientists (AI/ML), Data Engineering, Data Analyst
- Build data features in Community Sift, such as
 - Language detection
 - Predictive Moderation
 - Interactive dashboards
- Increase AI capabilities in Community Sift
- Build tools for Language and Culture team

Language and Culture

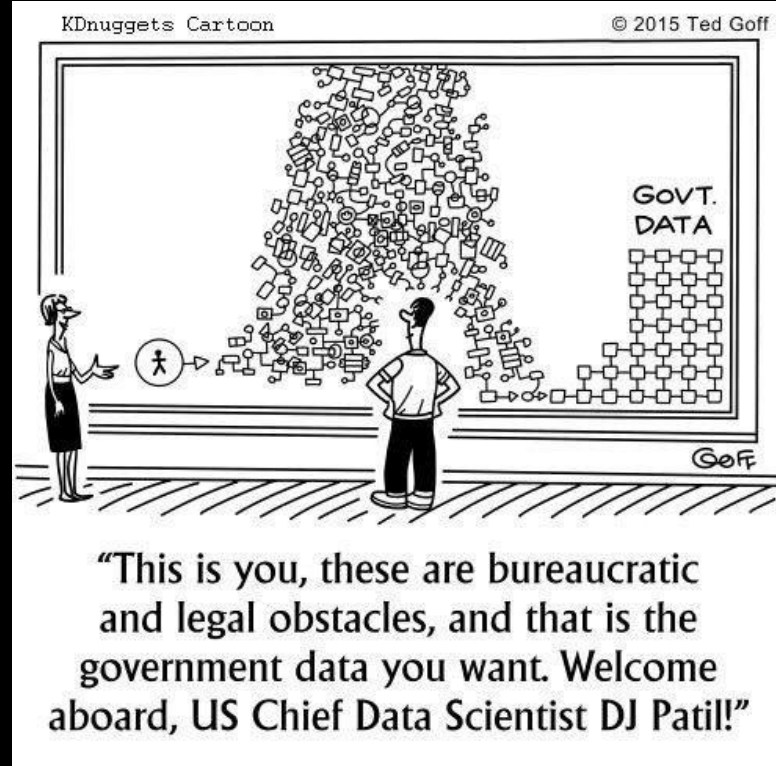
- Team of native-language speakers behind Community Sift's 22 languages
- Ensure text classifications are culturally relevant
- How we communicate, especially online constantly changes
 - Proactively monitor trends and update Community Sift

Disclaimer from Legal

The information contained in this slide deck is provided for informational purposes only and should not be construed as legal advice on any subject matter. You should not act or refrain from acting on the basis of any content included herein without seeking legal or other professional advice. The contents of this slide deck contain general information and may not reflect current legal developments or address your specific needs or situation.

In short: All opinions are my own and based on my own experience with privacy and RAI matters.

Importance of Privacy Regulations



Why Should You Care?

- It's not just about *data*, it's about *people*
 - What if it was data about **you** and something went wrong?
 - What decisions are being made about you? How are they being made?
- Strive to meet the highest standard. For example:
 - COPPA (children)
 - GDPR (general data)
- Know your data pipeline
 - Geography matters

© MARK ANDERSON

WWW.ANDERSTOONS.COM



"Before I write my name on the board, I'll need to know how you're planning to use that data."

COPPA Regulation

US, 2000

COPPA = Children's Online Privacy Protection Act

- Response to a growing awareness of Internet marketing techniques that targeted children and collected their personal information from websites without any parental notification.
 - Intent was to increase parental involvement in children's online activities, ensure children's safety during their participation in online activities, and most importantly, protect children's personal information.
- Specifically protects the privacy of children under the age of 13 by requesting parental consent.
- Applies to commercial websites and online services that are directed at children.

COPPA Regulation - Main Requirements

- Detailed **privacy policy** that describes the information collected from its users.
- **Limited collection of personally identifiable information (PII)** when a child participates in online games and contests.
- **Disclosure to parents** of any information collected on their children by the website.
- Verifiable **parental consent** prior to collecting any personal information from a child under the age of 13.
- **Right to revoke consent** and have information deleted.
- **A general requirement to protect** the confidentiality, security, and integrity of any personal information that is collected online from children.

Meet John Ty



John Ty

Age: 12

Country: USA

**Address: 1133 Lovely Street,
Kansas City**

Phone: 809-201-1233

Instagram: @johnnyBoy

TikTok: @johnny133444

Interests: Basketball

Hates: Birds

**Comment on Instagram:
“Yo! I’m Johnny. S’up?”**



John's PII



John Ty

Age: 12

Country: USA

**Address: 1133 Lovely Street,
Kansas City**

Phone: 809-201-1233

Instagram: @johnnyBoy

TikTok: @johnny133444

Interests: Basketball

Hates: Birds

**Comment on Instagram:
“Yo! I’m Johnny. S’up?”**



Proposed Updates: COPPA 2.0 and KOSA

Senate Commerce Committee Approves 11 Bills, Including Bipartisan Children's Online Privacy Legislation

July 27, 2023

During an [Executive Session](#) today, the U.S. Senate Committee on Commerce, Science and Transportation approved 11 bipartisan bills, including legislation aimed at protecting children's online privacy: the Children and Teens' Online Privacy Protection Act (COPPA 2.0) and the Kids Online Safety Act (KOSA).

Source:

<https://www.commerce.senate.gov/2023/7/senate-commerce-committee-approves-11-bills-including-bipartisan-children-s-online-privacy-legislation>

Proposed Updates: COPPA 2.0 and KOSA

- COPPA 2.0 covers **privacy** for minors under 16
- KOSA covers **safety** for minors under 18
- Highlights of the bills include:
 - Creating a “duty of care” requiring online providers to prevent serious harms to children on their platforms, including self-harm, eating disorders, sexual exploitation, cyberbullying and social media addiction. [KOSA]
 - Banning manipulative data driven advertising to children and teens. [COPPA 2.0]
 - Requiring platforms to have the strongest, most protective settings on by default for minors. [Both]
 - Giving parents more tools to protect their children’s privacy, restrict purchases, and track their time on a platform. [Both]

Sources:

<https://fairplayforkids.org/kids-online-safety-act-and-children-and-teens-online-privacy-protection-act-2-0-have-advanced/>
<https://www.jdsupra.com/legalnews/kids-privacy-and-safety-redux-amended-1385016/>

GDPR Regulation (2018)

- The **toughest** privacy and security law in the world.
- It imposes obligations onto organizations **anywhere**, so long as they target or collect data **related to people in the EU**.
- From gdpr.eu: “*The regulation itself is large, far-reaching, and fairly light on specifics, making GDPR compliance a daunting prospect...*”
- Basis in the European Convention for Human Rights: “Everyone has the **right to respect** for his private and family life, his home and his correspondence.”

GDPR: Data Subject Rights (DSRs)

The GDPR recognizes **privacy rights for data subjects**, which aim to give individuals more control over the data they loan to organizations:

- The right to be informed
- The right of access
- The right to rectification
- The right to erasure
- The right to restrict processing
- The right to data portability
- The right to object
- Rights in relation to automated decision making and profiling.*

GDPR: Data Protection Principles

Data processing must be done according to **seven protection and accountability principles** outlined in [Article 5.1-2](#):



Lawfulness, fairness and transparency

Processing must be lawful, fair, and transparent to the data subject.



Purpose limitation

You must process data for the legitimate purposes specified explicitly to the data subject when you collected it.



Data minimization

You should collect and process only as much data as absolutely necessary for the purposes specified.

GDPR: Data Protection Principles

Data processing must be done according to **seven protection and accountability principles** outlined in [Article 5.1-2](#):



Accuracy

You must keep personal data accurate and up to date.



Storage limitation

You may only store personally identifying data for as long as necessary for the specified purpose.



Integrity and confidentiality

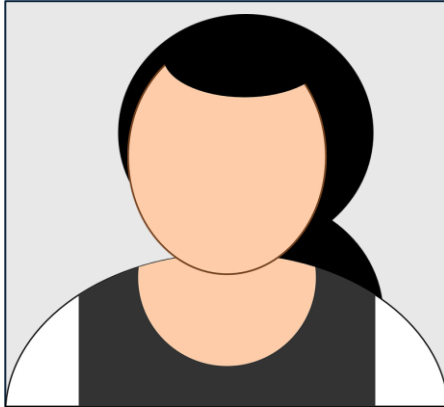
Processing must be done in such a way as to ensure appropriate security, integrity, and confidentiality (e.g. by using encryption).



Accountability

The data controller is responsible for being able to demonstrate GDPR compliance with all of these principles.

Meet Jane Toerg



Jane Toerg

Age: 15

Country: France

Address: 344 Really Nice, Nice

Phone: 00 +809-201-1233

Instagram: @janeTgreat

TikTok:

@janejanejanejaneeeee

Interests: Hiking

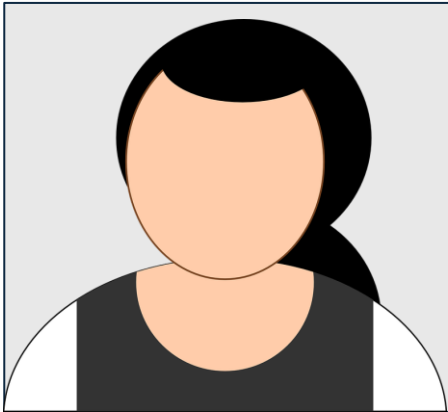
Hates: Birds

Comment on TikTok:

“Awesome dance moves!”



Jane's PII (GDPR)



Jane Toerg

Age: 15

Country: France

Address: 344 Really Nice, Nice

Phone: 00 +809-201-1233

Instagram: @janeTgreat

TikTok:

@janejanejanejaneeeee

Interests: Hiking

Hates: Birds

Comment on TikTok:

“Awesome dance moves!”



GDPR

June 2021 Update

New Standard Contractual Clauses (SCCs):

- additional representations and warranties, **onward transfer obligations**,
- Options and even some requirements for **multi-party use**,
- notification and **recordkeeping requirements**,
- new **sensitive data and accuracy** obligations, and
- **expanded security** and data breach requirements
- **more direct liability**; and
- more explicit requirements on both parties regarding the potential for overly intrusive foreign government access programs.

The background is a dark blue gradient with abstract digital elements. It features a grid of small, light blue dots forming a shape that resembles a map of the world. Overlaid on this are various red and orange circular markers of different sizes, some connected by thin vertical lines. There are also vertical streaks of light blue and orange, giving a sense of data flow or connectivity.

Data Locality

The Question of Data Locality



Since GDPR applies to European citizens, does that mean the data has to stay in Europe?

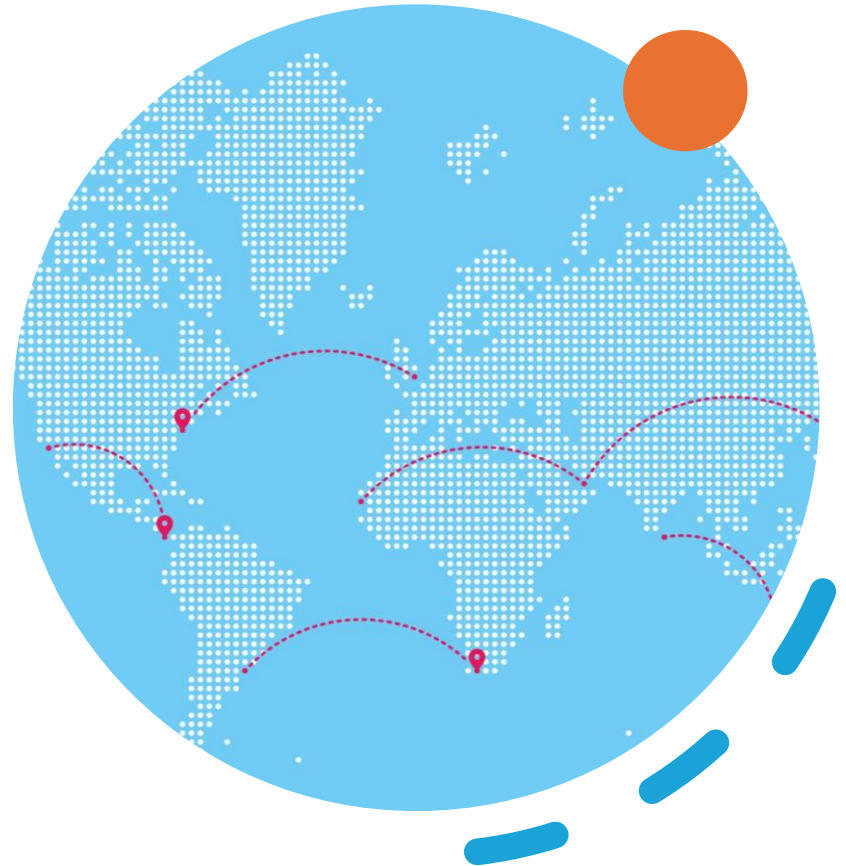
The Question of Data Locality

The **protection offered by GDPR travels with the data**, meaning that the rules protecting personal data continue to apply regardless of where in the world the data lands.

- The transfers to an ‘adequate’ third country must be comparable to a transmission of data within the EU.
- This limits *where* data can be processed and/or stored.

BUT

Europe’s highest court in July 2020 ruled that the main transatlantic data transfer deal hammered out between Brussels and Washington – **Privacy Shield** – was invalid because of concerns about U.S. surveillance. (aka **Schrems II** ruling)



The Trouble With Google Analytics (2022)



UPDATE: Further EU DPA orders stop of Google Analytics

The Italian DPA (GDPD) has joined the consensus shared by the [EDPS](#), as well as the [French](#) and [Austrian DPA](#) and has banned the use of Google Analytics (GA). Following our 101 complaints on data transfers, the GDPD concluded that websites using GA collected user interactions and transferred user data to the US; a country without an adequate level of data protection, making said transfer unlawful.

Source: <https://noyb.eu/en/update-further-eu-dpa-orders-stop-google-analytics>



Google Sustainability Summit
Moving from climate pledges to progress

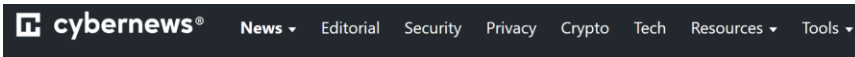
Read more

[Home](#) / [News](#) / [Technology](#) / [Data protection](#) / France joins Austria in finding Google Analytics illegal

France joins Austria in finding Google Analytics illegal

By Mathieu Pollet | EURACTIV France | translated by Daniel Eck | Feb 10, 2022 (updated: Feb 14, 2022)

Source: <https://www.euractiv.com/section/data-protection/news/france-joins-austria-says-google-analytics-data-not-protected-in-us/>



If you purchase via links on our site, we may receive **affiliate commissions**.

[Home](#) » [News](#)

Denmark: use of Google Analytics is illegal



Anna Zhadan, Editor

Updated on: 23 September 2022



Source: <https://cybernews.com/news/denmark-use-of-google-analytics-is-illegal/>

The Trouble With Google Analytics (2023)

Customer Data Management

Swedish Privacy Watchdog Slaps Over \$1 Million in Fines, Warns Companies Against Using Google Analytics

The fines came regarding exporting European users' data via Google Analytics to the US, a breach of the EU's privacy rulebook.



Karthik Kashyap

July 5, 2023

Source: <https://www.spiceworks.com/marketing/customer-data/news/swedish-watchdog-slaps-1-million-warns-against-using-google-analytics/>

EU-US Data Privacy Framework (2023)

Data Protection Report

Data protection legal insight at the speed of technology

[Home](#) > [Fintech](#) > European Commission adopts its adequacy decision for the EU-US Data Privacy Framework

European Commission adopts its adequacy decision for the EU-US Data Privacy Framework



By **Lara White (UK)**, **Marcus Evans (UK)**, **Susan Ross (US)**, **Jurriaan Jansen**, **Christoph**

Ritzer (DE) and **Nadège Martin (FR)** on July 11, 2023
Posted in **Fintech**

On 10 July, the European Commission adopted its **adequacy decision for the EU-US Data Privacy Framework** (the **DP Framework**). It thereby declared that the United States (the **US**) ensures an adequate level of protection for personal data transferred from the EU to US companies that have self-certified their adherence to the DP Framework Principles. More broadly, the DP Framework will significantly reduce the challenges faced by EU organisations transferring personal data to the US, at least for the time being.

Source: <https://www.dataprotectionreport.com/2023/07/european-commission-adopts-its-adequacy-decision-for-the-eu-us-data-privacy-framework/>

Cross-Border Data Transfer News (2023)



cross border data transfer regulations



 International Association of Privacy Professionals

[ANPD releases draft regulation for cross-border transfers under SCCs](#)

Brazil's data protection authority, the Autoridade Nacional de Proteção de Dados, unveiled its proposed regulations for international data...

Aug 15, 2023



 Lexology

[Significant Changes to China's CBDT Regulatory Framework -A Brief Analysis on CAC's new Draft Regulation](#)

On September 28, 2023, the Cyberspace Administration of China ("CAC") issued the Notice to Seek Public Consultation on the Provisions on...

2 days ago



 Reuters

[China to waive security checks for some data exports](#)

China on Thursday proposed relaxing some rules concerning the transfer of data abroad in a move likely to bring some relief to foreign and...

3 weeks ago



 The National Law Review

[China Releases Draft Regulation to Significantly Ease Cross-border Data Transfers](#)

Cyberspace Administration of China releases Regulations on Regulating Facilitating Crossborder Data Flow to improve free data flows,...

3 weeks ago



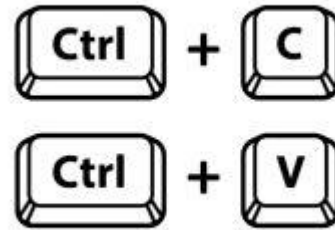
The Question of Data Locality



If the data is stored in Europe but is viewable in Canada, is that considered a cross-region transfer?

Viewing Data and Data Locality

*Yes - unless the method of viewing **prevents** copy/paste or screen capture.*



The Question of Data Locality

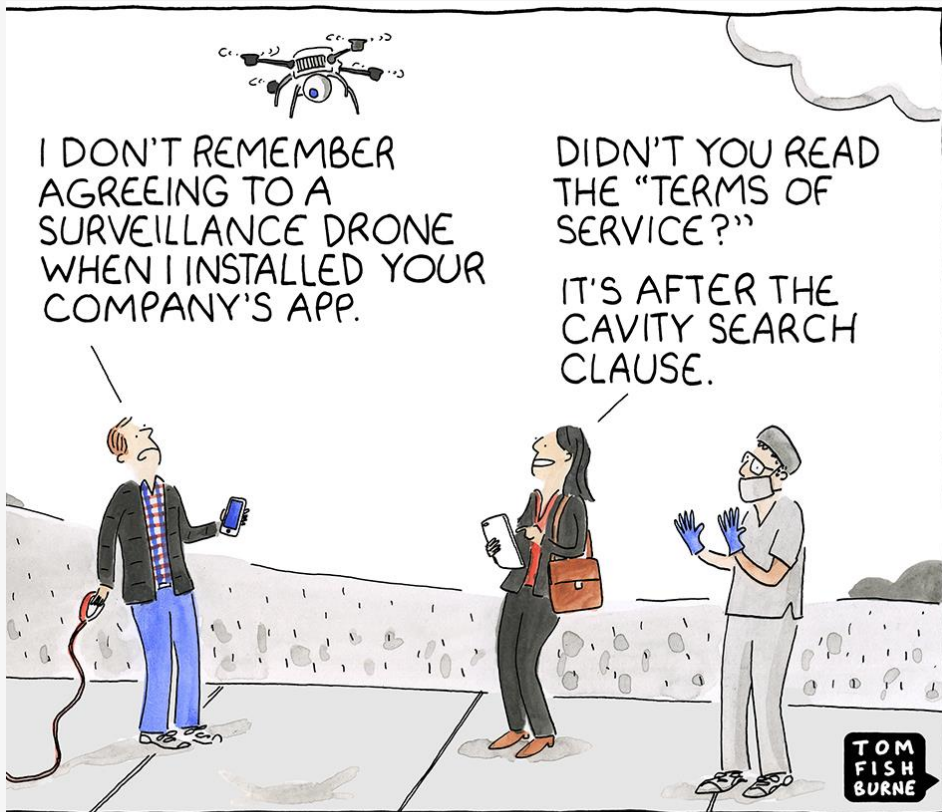


When you run `head()` to view the first few lines of data in Jupyter/R, are you storing that data in a new location?

Viewing Data and Data Locality

*Yes.
Remember to clear your
notebooks.*





© marketoonist.com

Notification And Consent

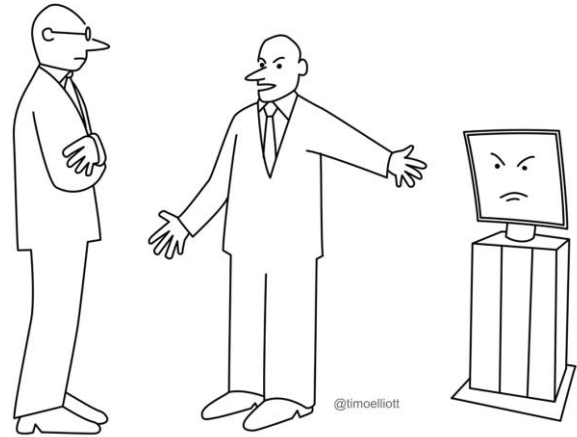
- Consumers and businesses need to know
 - What data you're collecting
 - What you're using it for
 - How long you're keeping it
 - Risks/consequences
- Provide opportunity for consent
- Have the ability to control level of detail and opt-out *at any time*
- Clear and simple terms



Constantly Evolving Privacy Regulations

- Even with GDPR, there are *additional* regulations which vary from country to country
 - From state to state within the US (e.g. California's CCPA)
- Geography matters:
 - Know where your data is coming from
 - Know where it's going
 - Know where it's landing
- Regulations vary on what type of data you're dealing with
 - Need to understand what is personal and/or sensitive data you're handling
 - Details matter!

Privacy and Responsible AI



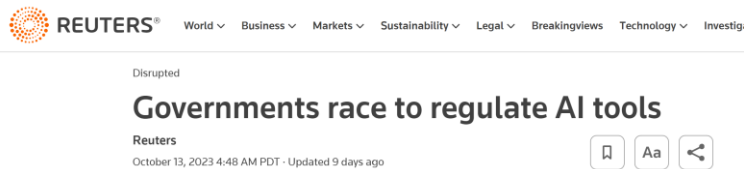
*His decisions aren't any better than yours
— but they're WAY faster..*

The Impact of Generative AI

- The mass-market introduction of text and image generators has accelerated the need for AI regulation.
- Responsible AI policies and practices are how companies are preparing.
- Privacy is an important component.



AI Regulation In The News



Source: <https://www.reuters.com/technology/governments-race-regulate-ai-tools-2023-10-13/>



Source: <https://www.bloomberg.com/news/articles/2023-10-18/eu-plans-stricter-rules-for-most-powerful-generative-ai-models#xj4y7vzkg>

WHO outlines considerations for regulation of artificial intelligence for health

19 October 2023 | News release | Geneva | Reading time: 2 min (610 words)

Source: <https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health>



Source: <https://www.theglobeandmail.com/business/article-champagne-ai-voluntary-code-of-conduct-tech-companies/>

Global AI Legislation Tracker



This map shows which jurisdictions are in focus and covered by this tracker. It does not represent the extent to which jurisdictions around the world are active on AI governance legislation.

Jurisdictions in focus

Australia | Brazil | Canada | China | EU | India | Israel | Japan | New Zealand
Saudi Arabia | Singapore | South Korea | United Arab Emirates | U.K. | U.S.

**Click on the country names above to navigate to its location in the tracker.*

Singapore's Model Framework (2019)

- The Model Framework aims to promote **public understanding and trust** in technologies by:
 - explaining how AI systems work,
 - building good data accountability practices, and
 - creating open and transparent communication
- Ties together the concerns of data privacy, AI ethics, and governance.
- Guiding Principles:
 - Decisions made by AI should be **explainable, transparent and fair**
 - AI systems should be **human-centric**



Singapore's Model Framework (2019)

From Principles to Practice



Internal Governance Structures and Measures

- Clear roles and responsibilities in your organisation
- SOPs to monitor and manage risks
- Staff training



Determining the Level of Human Involvement in AI- augmented Decision- making

- Appropriate degree of human involvement
- Minimise the risk of harm to individuals



Operations Management

- Minimise bias in data and model
- Risk-based approach to measures such as explainability, robustness and regular tuning









Stakeholder Interaction and Communication

- Make AI policies known to users
- Allow users to provide feedback, if possible
- Make communications easy to understand

Common Features of Proposed Law/Policy (2023)

These are early days. Expect changes.

	Prohibit and/or regulate specific “high risk” systems
	Guarantee various individual rights (e.g. fairness, explainability)
	Trust and safety
	Transparency and/or accountability
	Accessibility/inclusivity
	Existing regulations and/or self-regulation

The New York Times

Pressured by Biden, A.I. Companies Agree to Guardrails on New Tools

Amazon, Google and Meta are among the companies that announced the guidelines as they race to outdo each other with versions of artificial intelligence.

Source: <https://www.nytimes.com/2023/07/21/us/politics/ai-regulation-biden.html>



```
mirror_mod = modifier_ob.  
    # Set mirror object to mirror  
    mirror_mod.mirror_object =  
    operation == "MIRROR_X":  
    mirror_mod.use_x = True  
    mirror_mod.use_y = False  
    mirror_mod.use_z = False  
    operation == "MIRROR_Y":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = True  
    mirror_mod.use_z = False  
    operation == "MIRROR_Z":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = False  
    mirror_mod.use_z = True
```

```
selection at the end -add  
mirror_ob.select= 1  
modifier_ob.select=1  
context.scene.objects.active  
("Selected" + str(modifier  
mirror_ob.select = 0  
= bpy.context.selected_obj  
data.objects[one.name].se  
print("please select exactly
```

```
types.Operator):  
    # X mirror to the selected  
    object.mirror_mirror_x"  
    mirror X"
```

```
context):  
    context.active_object is not
```

Corporate Approach: Responsible AI

Microsoft: “We are committed to advancement of AI driven by ethical principles that put people first.”

- Principles include fairness, transparency, inclusiveness, reliability & safety, **privacy** & security, accountability
- Published Responsible AI practices for Azure Open AI models



```
mirror_mod = modifier_ob.  
    mirror object to mirror  
    mirror_mod.mirror_object  
    operation == "MIRROR_X":  
    mirror_mod.use_x = True  
    mirror_mod.use_y = False  
    mirror_mod.use_z = False  
    operation == "MIRROR_Y":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = True  
    mirror_mod.use_z = False  
    operation == "MIRROR_Z":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = False  
    mirror_mod.use_z = True  
    selection at the end -add  
    mirror_ob.select= 1  
    modifier_ob.select=1  
    context.scene.objects.active  
    ("Selected" + str(modifier  
    mirror_ob.select = 0  
    = bpy.context.selected_obj  
    data.objects[one.name].se  
    print("please select exactly  
    --- OPERATOR CLASSES ---  
    types.Operator):  
    X mirror to the selected  
    object.mirror_mirror_x"  
    mirror X"  
    context):  
    context.active_object is not
```

Corporate Approach: Responsible AI

Amazon: “We are committed to developing AI responsibly, taking a people-centric approach that prioritizes education, science, and our customers, to integrate responsible AI across the end-to-end AI lifecycle.”

- Principles include fairness, explainability, privacy & security, robustness, governance, transparency.
- Meta has declared similar principals

```
mirror_mod = modifier_ob.  
    mirror object to mirror  
    mirror_mod.mirror_object
```

```
operation == "MIRROR_X":  
    mirror_mod.use_x = True  
    mirror_mod.use_y = False  
    mirror_mod.use_z = False  
operation == "MIRROR_Y":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = True  
    mirror_mod.use_z = False  
operation == "MIRROR_Z":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = False  
    mirror_mod.use_z = True
```

```
selection at the end -add  
mirror_ob.select= 1  
modifier_ob.select=1  
context.scene.objects.active  
("Selected" + str(modifier  
mirror_ob.select = 0  
= bpy.context.selected_obj  
data.objects[one.name].se  
print("please select exactly
```

```
--- OPERATOR CLASSES ---
```

```
types.Operator):  
    X mirror to the selected  
    object.mirror_mirror_x"  
    mirror X"
```

```
context):  
context.active_object is not
```

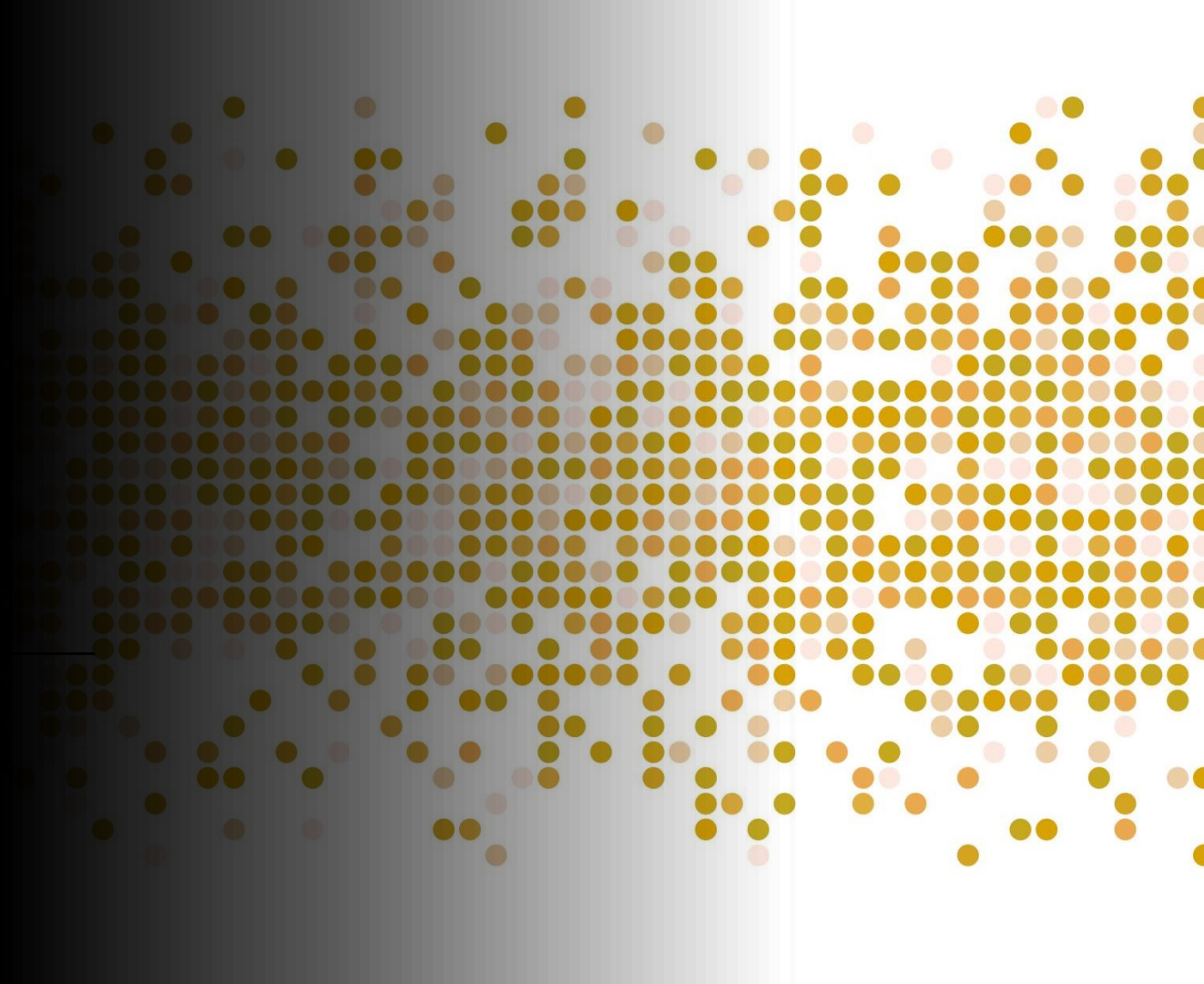
Corporate Approach: Responsible AI

Google: “Reliable, effective user-centered AI systems should be designed following general best practices for software systems together with practices that address considerations unique to machine learning.”

- Practices include human-centered design, multiple metrics, testing, and monitoring
- Added principles and practices for fairness, interpretability, **privacy**, safety and security



Working with your Data Privacy Manager



How To Make Your Privacy Manager's Life Easier

- Get familiar with your company's privacy policy and RAI policy (if available)
- Know your data pipeline
 - What data are you handling and why
 - Where the data came from, where it's going, where it lands
- Initiate privacy and/or RAI reviews **early** in development
 - Changes in data, processing and how the data is being used
 - Consider, and be transparent about, **all** possible risks
- Be disciplined about documentation
 - Data inventory
 - Data flows



"I'm updating my privacy policy."



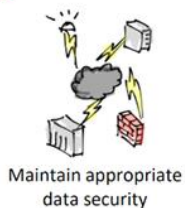
Thank You!

Additional Reading

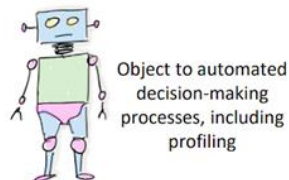
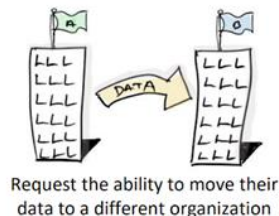
- COPPA: <https://epic.org/privacy/kids/>
- GDPR:
 - https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en
 - <https://gdpr.eu/>
- Canada's Consent Guidelines:
 - https://www.priv.gc.ca/en/privacy-topics/collecting-personal-information/consent/gl_omc_201805/
- Impact of GDPR on AI:
 - [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU\(2020\)641530_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf)
- International Association of Privacy Professionals: <https://iapp.org/>

High level view of the GDPR

What organizations have to do



What individuals can do



What regulators can do

