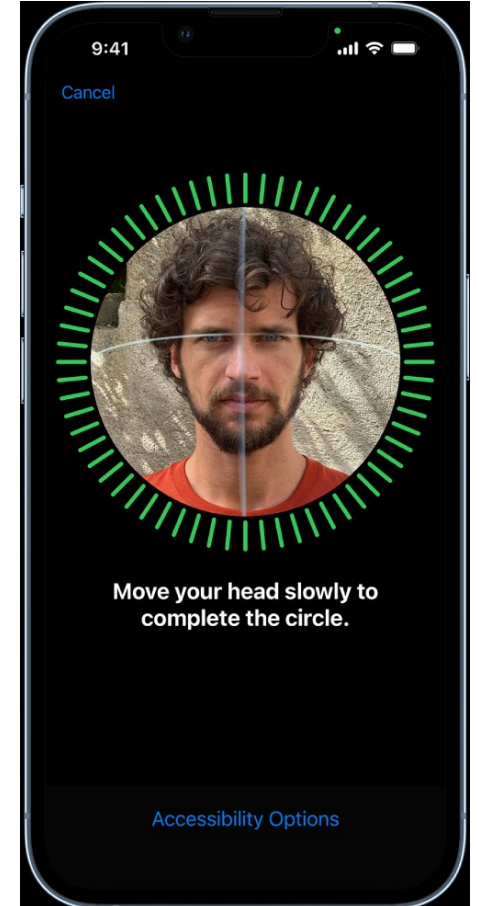
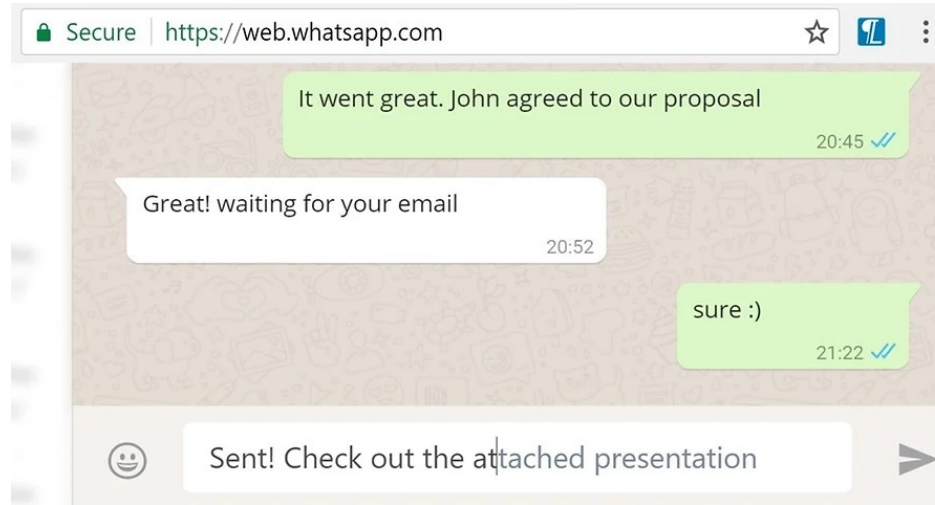
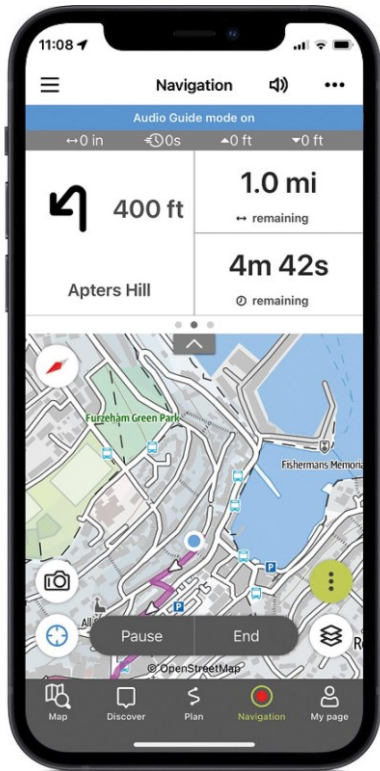


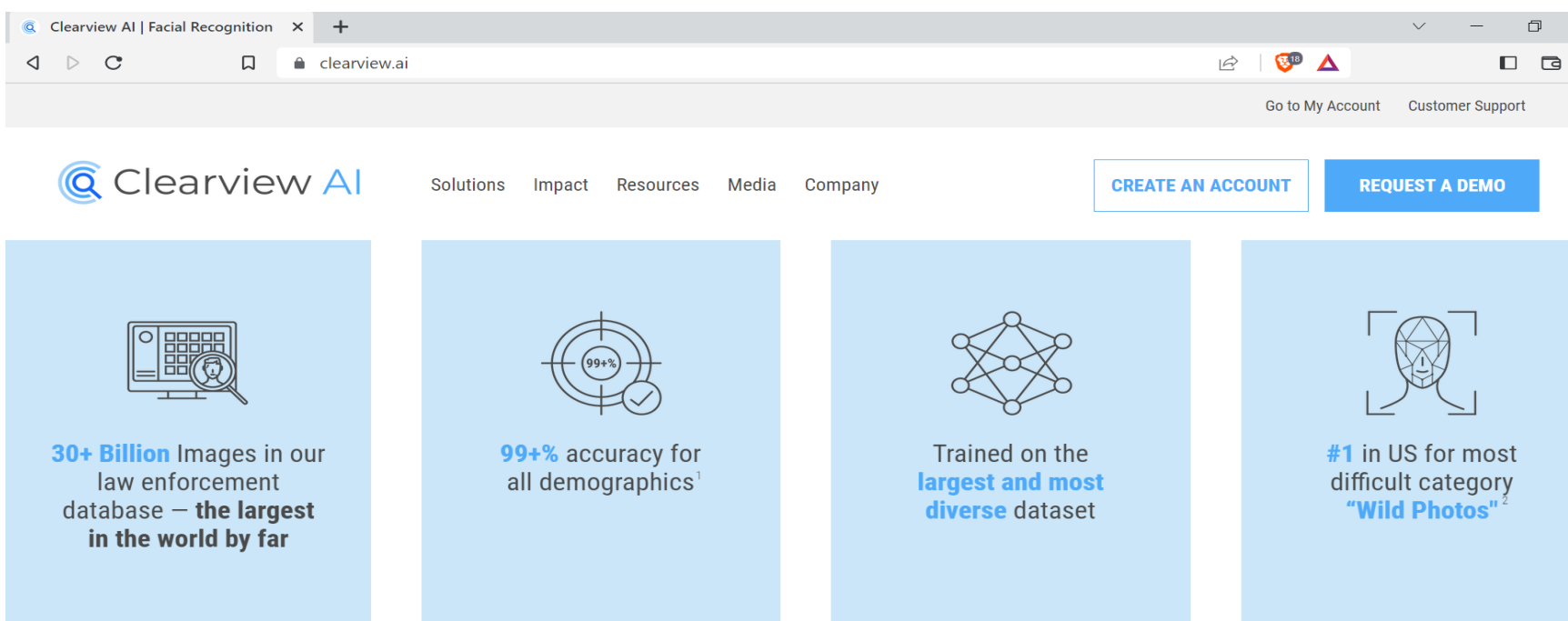
A Decision-Making Process to Implement the 'Right to be Forgotten' in Machine Learning

Katie Hawkins, Nora Alhuwaish, Sana Belguith, Asma Vranaki, Andrew
Charlesworth

University of Bristol



Personal Data is being collected and analysed ...often without an appropriate lawful basis



The screenshot shows the Clearview AI website. The browser address bar displays 'clearview.ai'. The navigation menu includes 'Solutions', 'Impact', 'Resources', 'Media', and 'Company'. Two prominent buttons are 'CREATE AN ACCOUNT' and 'REQUEST A DEMO'. The main content area features four blue boxes with icons and text:

- 30+ Billion** Images in our law enforcement database — **the largest in the world by far**
- 99+%** accuracy for all demographics¹
- Trained on the **largest and most diverse** dataset
- #1** in US for most difficult category **"Wild Photos"**²

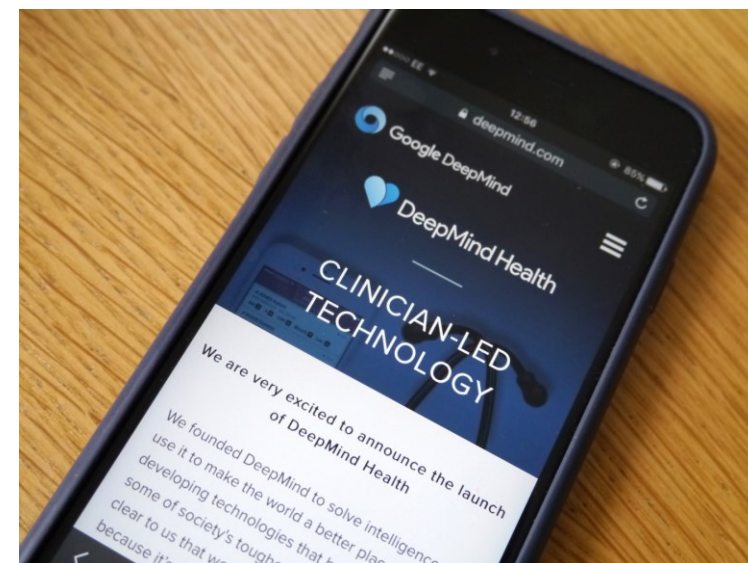
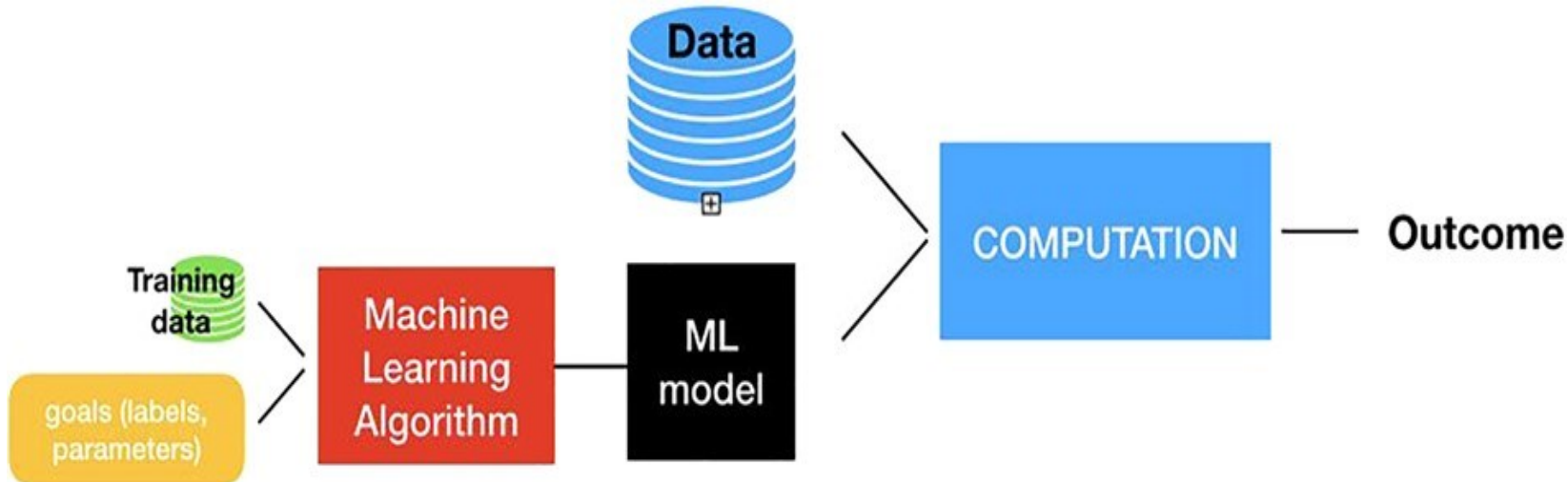
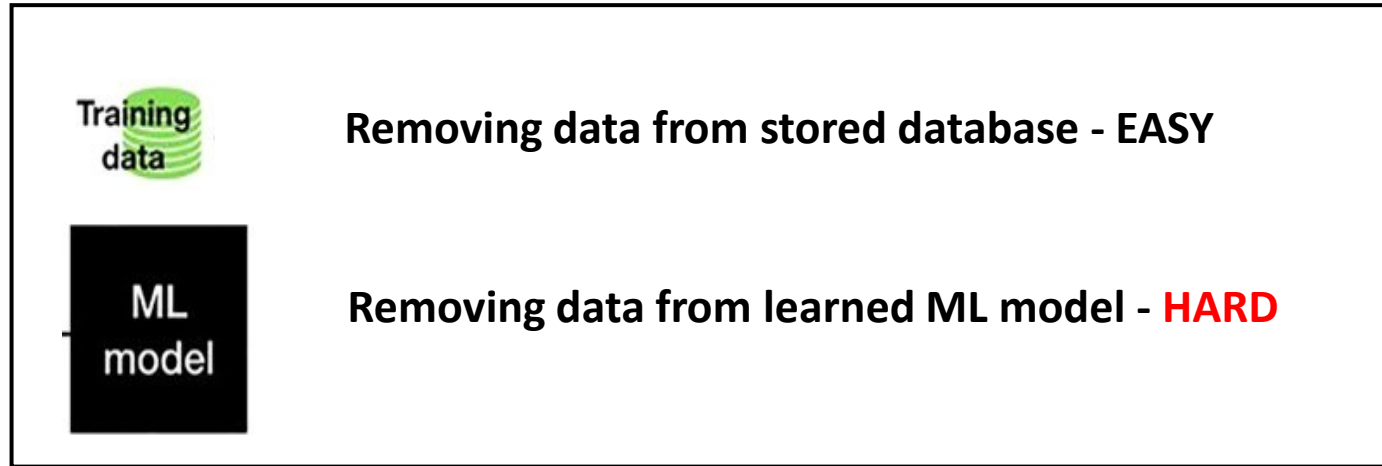


Image Credits: Natasha Lomas / TechCrunch

How should controllers erase personal data that has already been deployed in ML models?

Erasure in supervised machine learning models



WHEN YOU TRAIN PREDICTIVE MODELS ON INPUT FROM YOUR USERS, IT CAN LEAK INFORMATION IN UNEXPECTED WAYS.

Erasure Challenges



Memorising

What does that mean for erasure?

Models classified as personal data?
Recent paper explains how the GDPR is likely to classify models as personal data due to this memorisation (Veale, M., Binns, R., Edwards, L)



Disconnect between the law and technology

A range of different interpretations from the legal and technical literature.

Technical literature focus on a range of erasure techniques *without* the necessary legal analysis.



Increase in erasure requests

How can we erase requested personal data completely, efficiently, and for many erasure requests?

And without impacting the ML model currently in operation?

Paper contributions



Integrate the multidisciplinary problem space

Consider expertise from both the legal and technical space



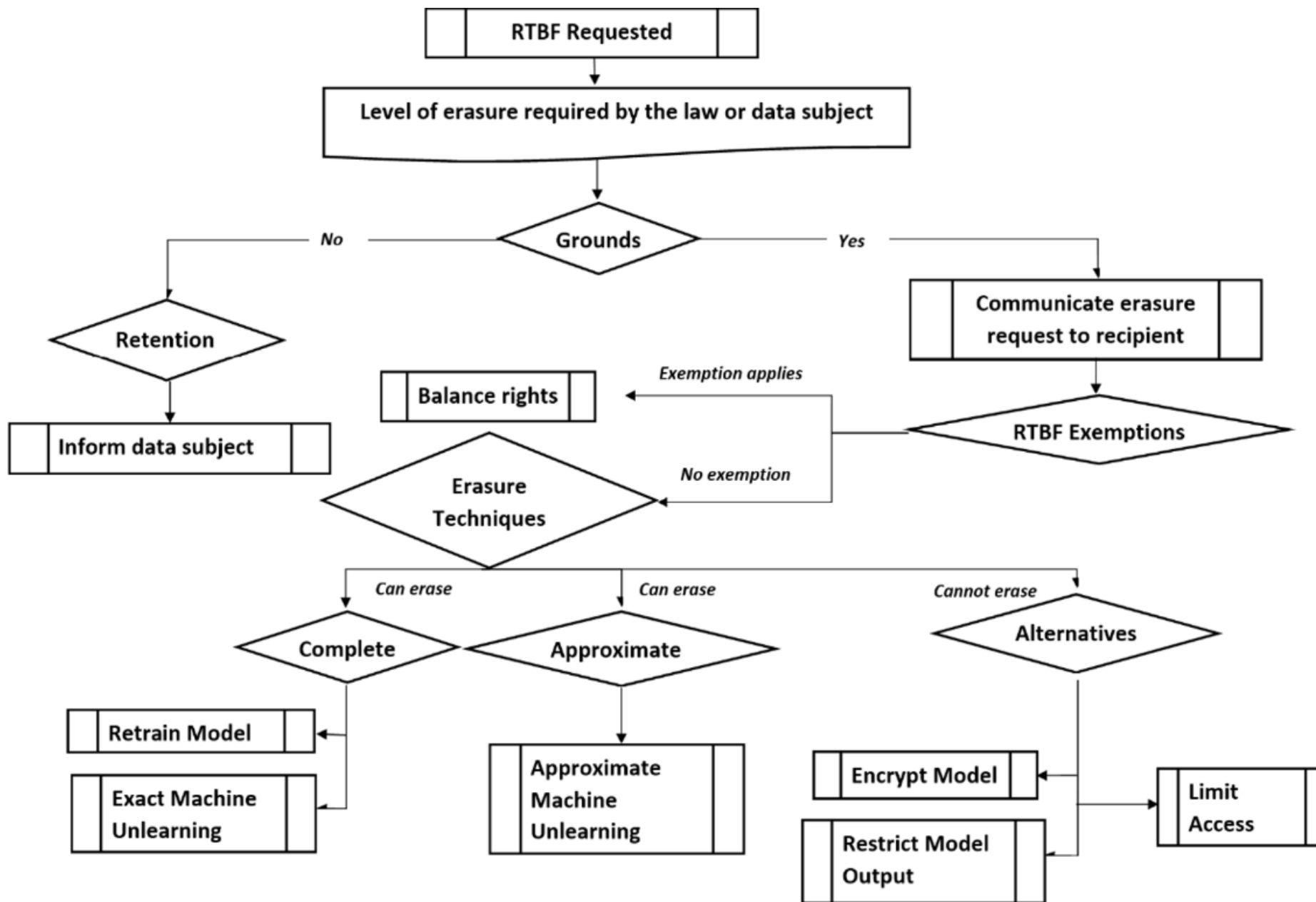
Outline erasure in ML

There are various techniques that lead to differing erasure outcomes



Decision-making flow: a practical outcome for the controller to implement the RTBF in ML models

From the steps and decisions once a RTBF has been requested, to the existing techniques for erasure in ML



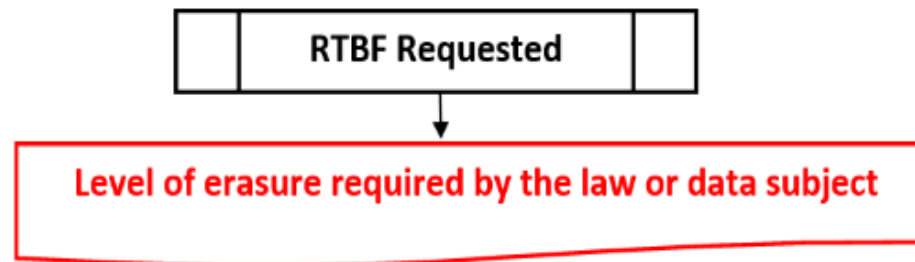
KEY

Process or Action

Decision or Result

Written contract, law or policy

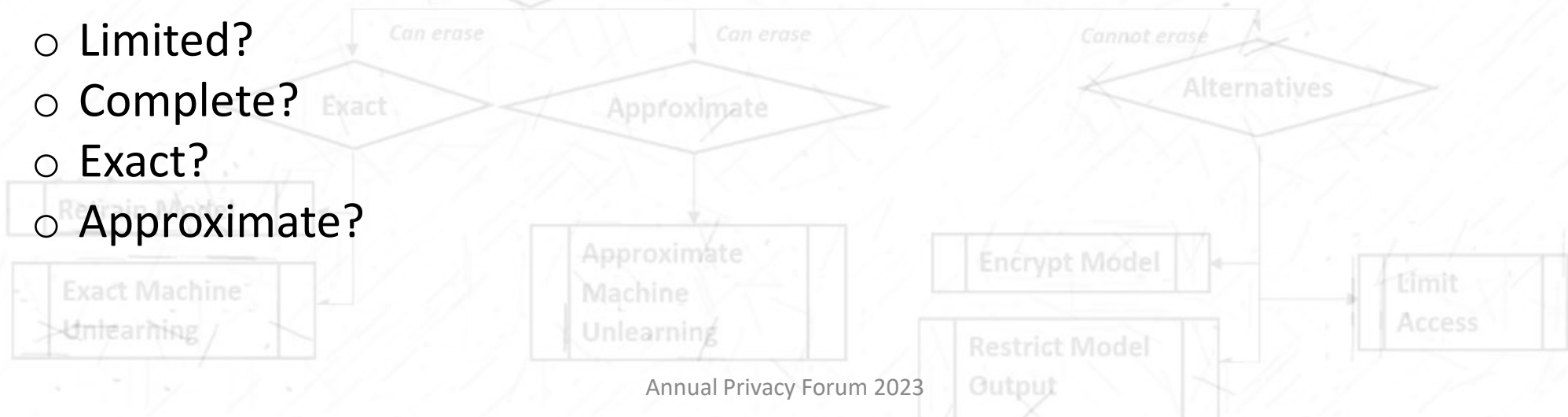
Level of Erasure



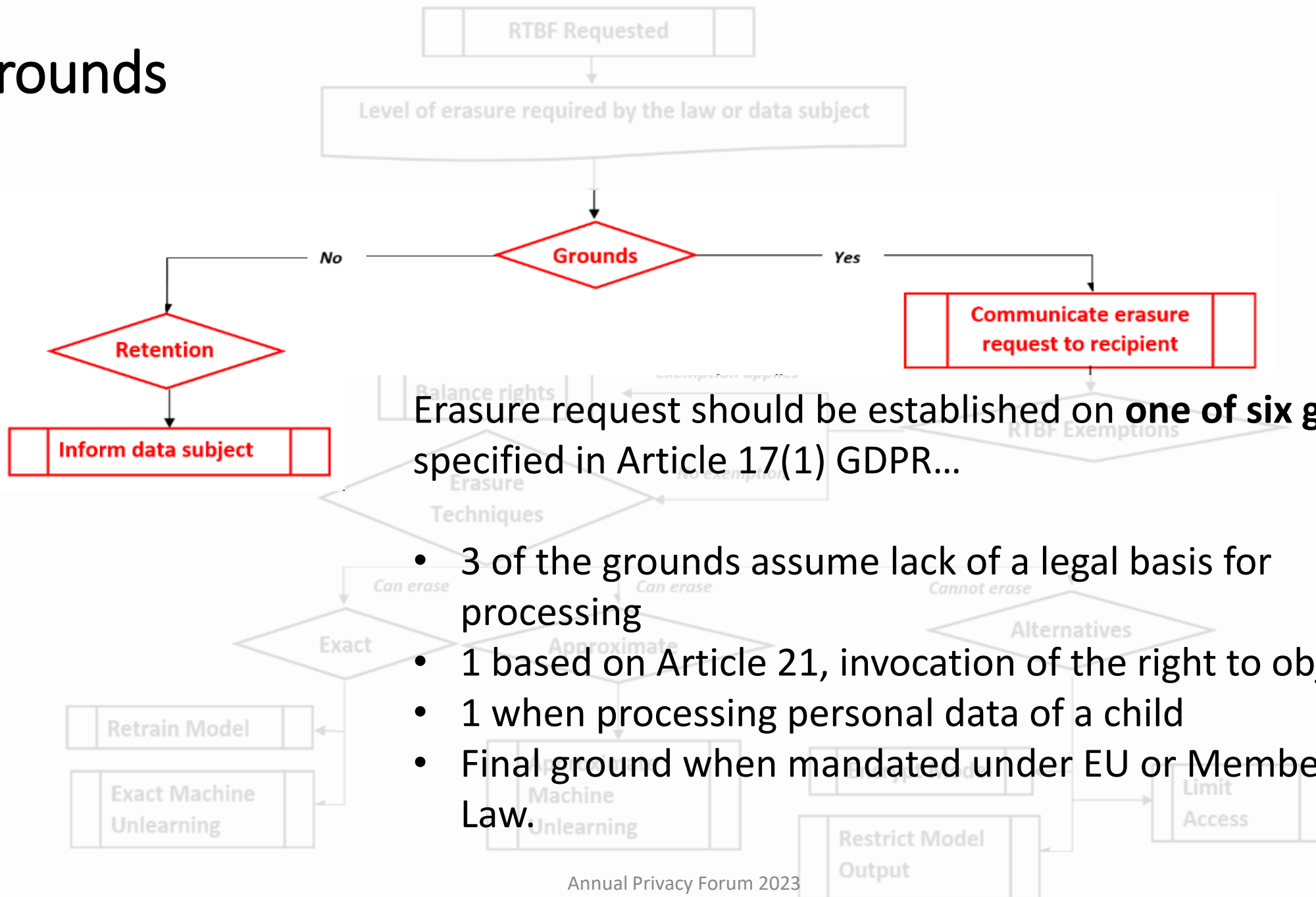
Determined by: The data subject's request & the law

- ❖ Data subject can either request limited erasure or complete erasure:
- ❖ If the data subject requests complete erasure, to what extent does the GDPR require the controller to erase the data in complicated systems such as ML?

- Limited?
- Complete?
- Exact?
- Approximate?



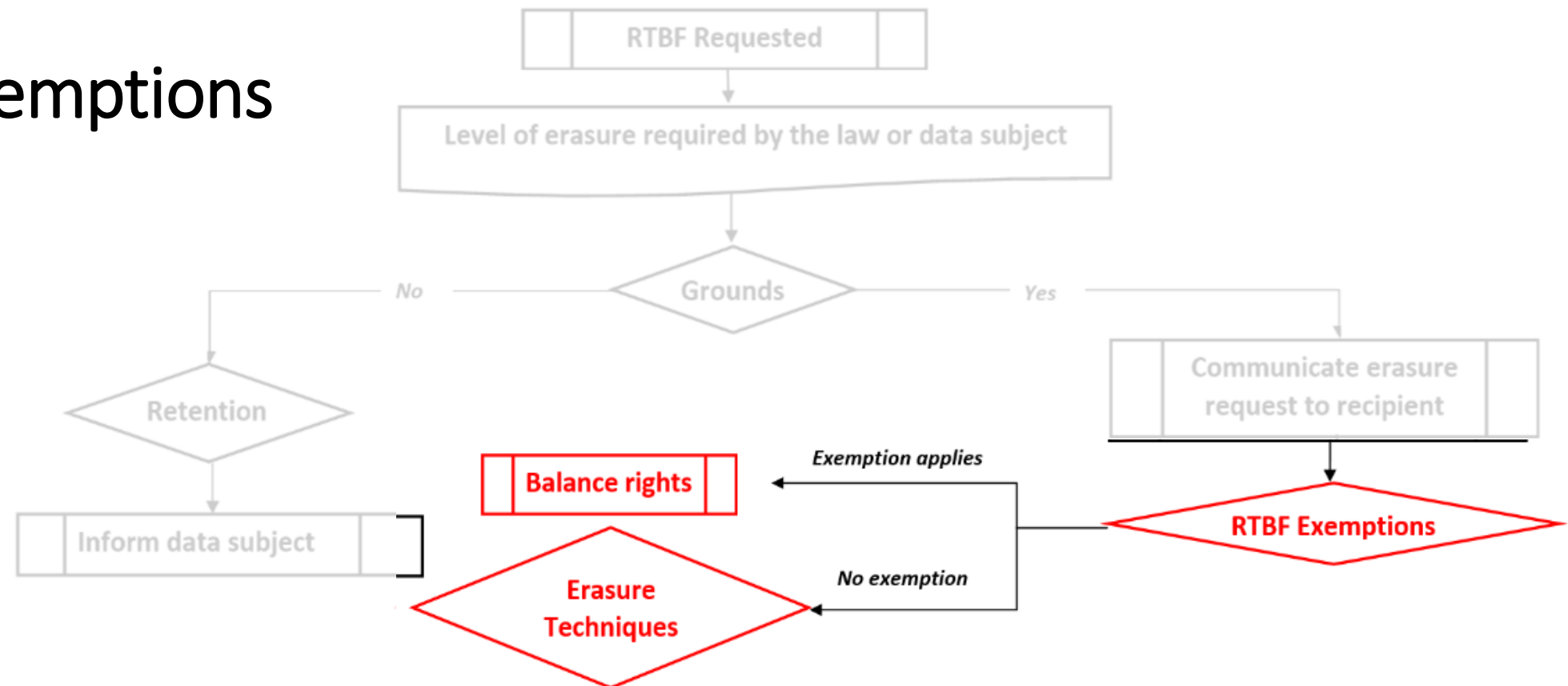
Grounds



Erasure request should be established on **one of six grounds** specified in Article 17(1) GDPR...

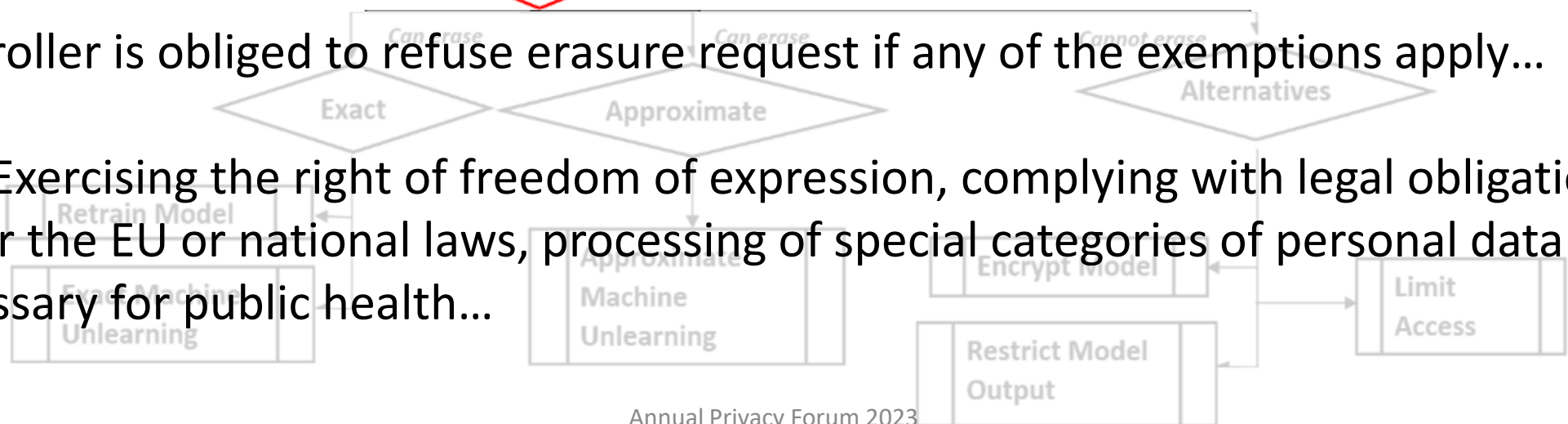
- 3 of the grounds assume lack of a legal basis for processing
- 1 based on Article 21, invocation of the right to object
- 1 when processing personal data of a child
- Final ground when mandated under EU or Member State Law.

Exemptions



Controller is obliged to refuse erasure request if any of the exemptions apply...

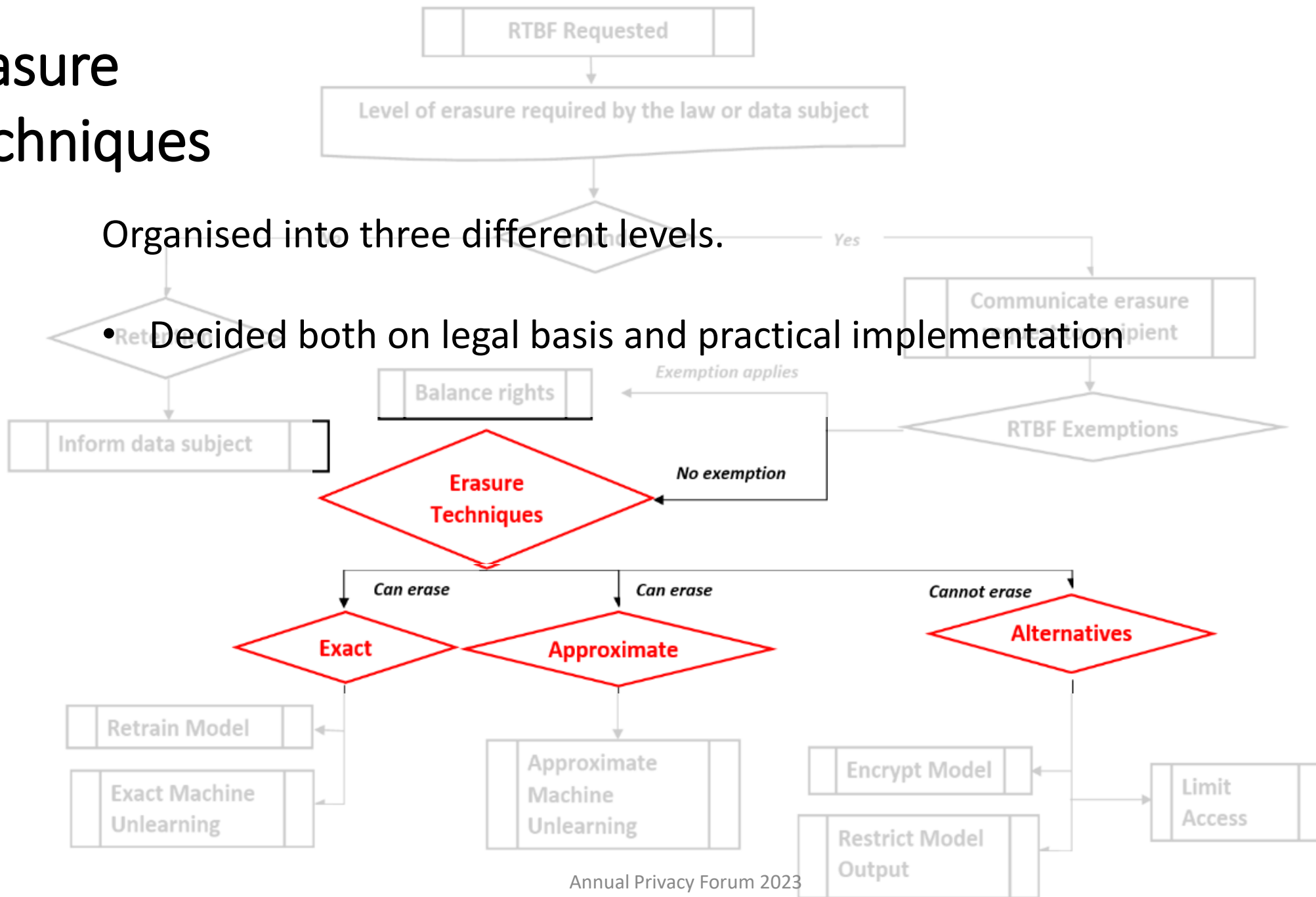
e.g. Exercising the right of freedom of expression, complying with legal obligation under the EU or national laws, processing of special categories of personal data is necessary for public health...



Erasure Techniques

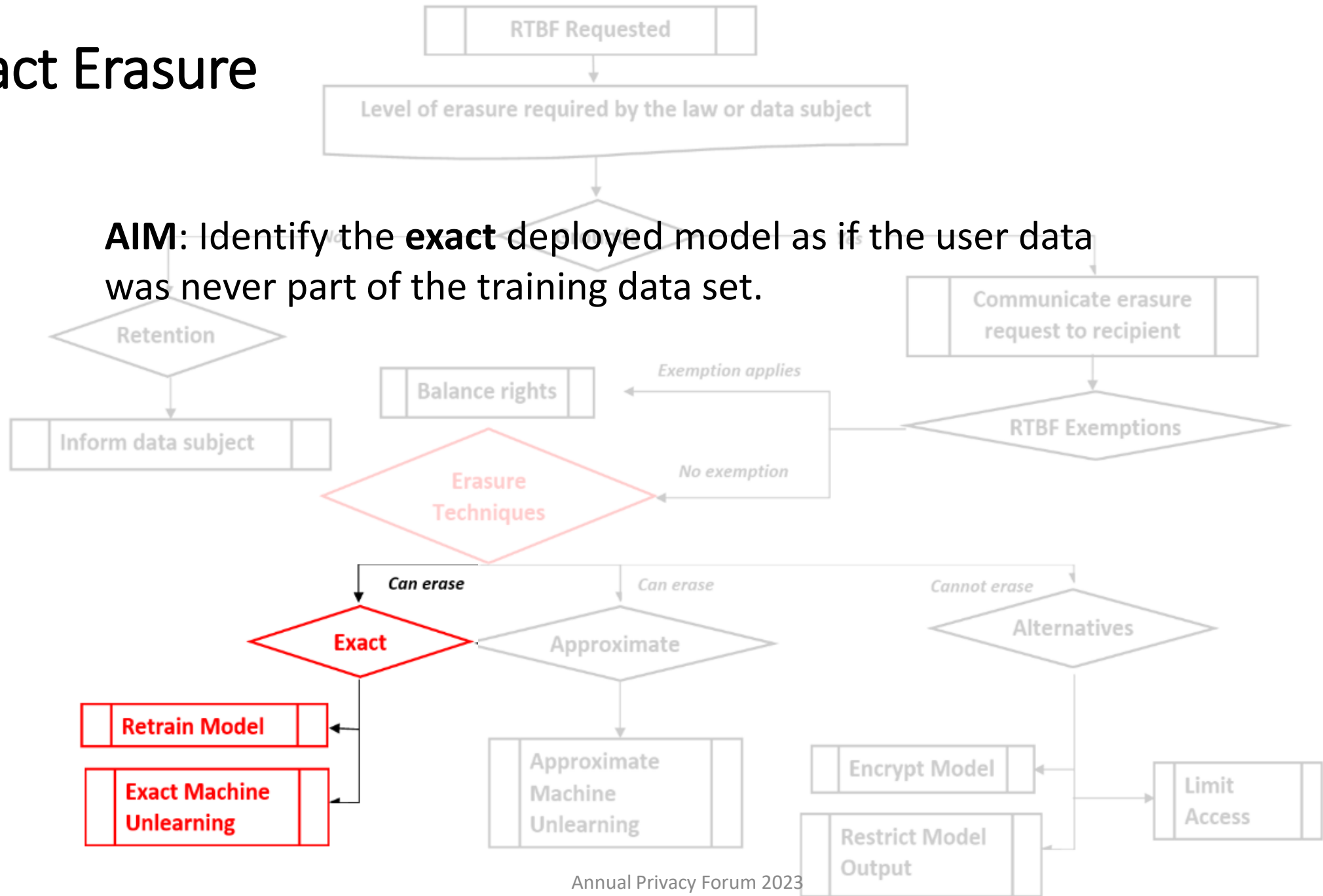
Organised into three different levels.

- Decided both on legal basis and practical implementation



Exact Erasure

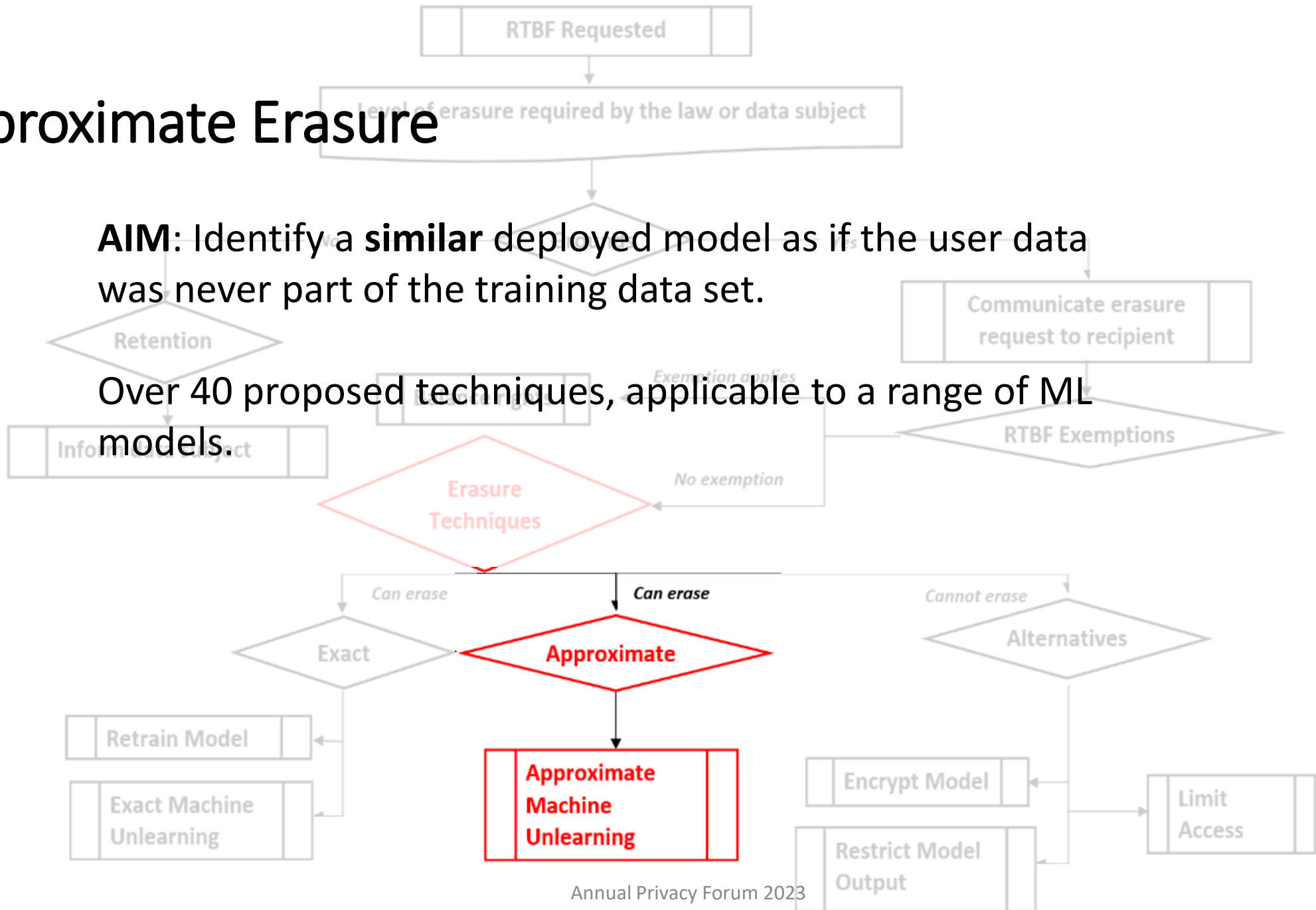
AIM: Identify the **exact** deployed model as if the user data was never part of the training data set.



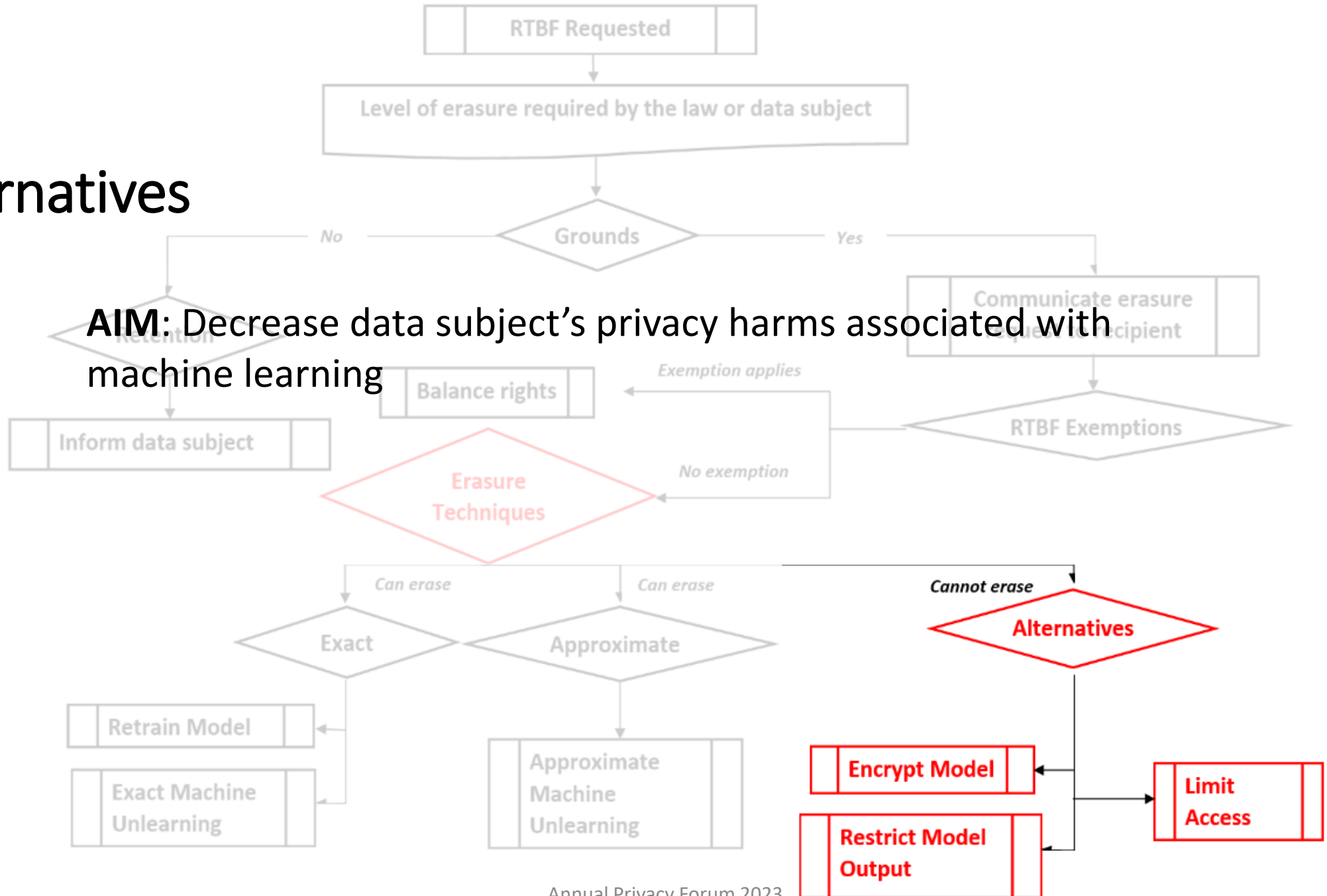
Approximate Erasure

AIM: Identify a **similar** deployed model as if the user data was never part of the training data set.

Over 40 proposed techniques, applicable to a range of ML models.



Alternatives



Summary

The proposed decision-making flow:

- Bridges the gap between law, computer science and industry application...
- Propose ways to efficiently erase personal data...
- Differing levels of erasure based on grounds/exceptions etc.

(all this while) Prioritising the privacy of the data subject.



Future Work

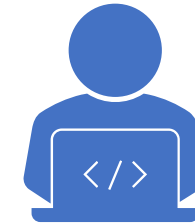
THE AI ACT



How personal data can be inferred from the output of the ML model?



Testing the proposed solution on both hypothetical and real-world scenarios to solidify its validity and feasibility.



Consider other complex ML applications such as federated learning, repurposing, transfer learning and one-shot learning.

Thank you.
Questions?

Email: katie.hawkins@bristol.ac.uk