Your Data, Your AI

Towards a Decentralised Future

Harshvardhan Pandit | Knowledge and Engineering | TCD 2023 | https://harshp.com/presentations

Harsh(vardhan Pandit)

Assistant Professor, DCU

PhD, Computer Science - TCD (2020)

Data and consent for GDPR compliance

Govt. Of Ireland Postdoctoral Fellowship (2022)
Data protection / privacy impact assessments

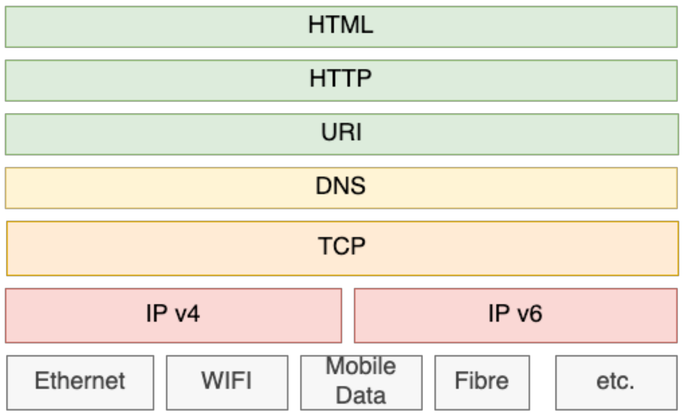
Chair W3C Data Privacy Vocabularies and Controls Community Group (DPVCG)

National Standards Authority Ireland (NSAI)

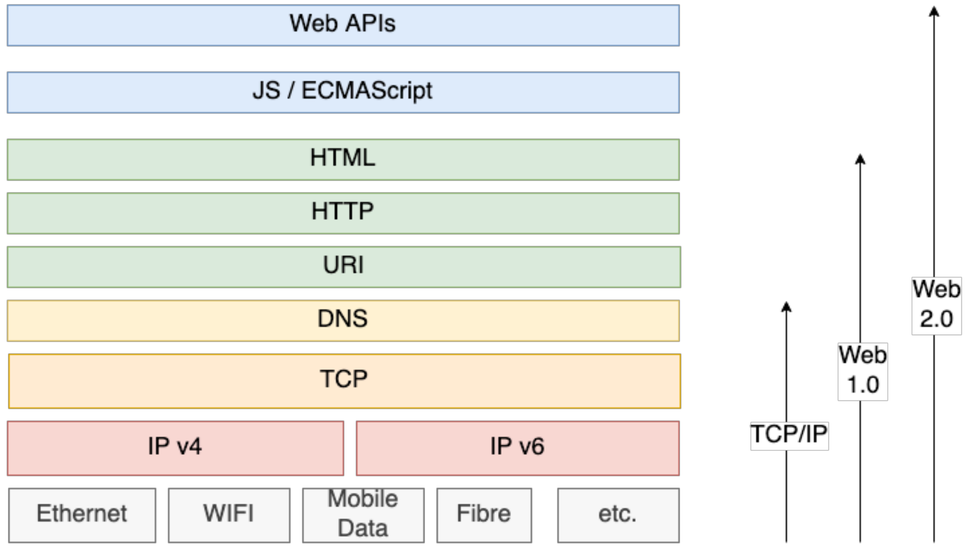
International Standardisation Organisation (ISO)

Conventional Web

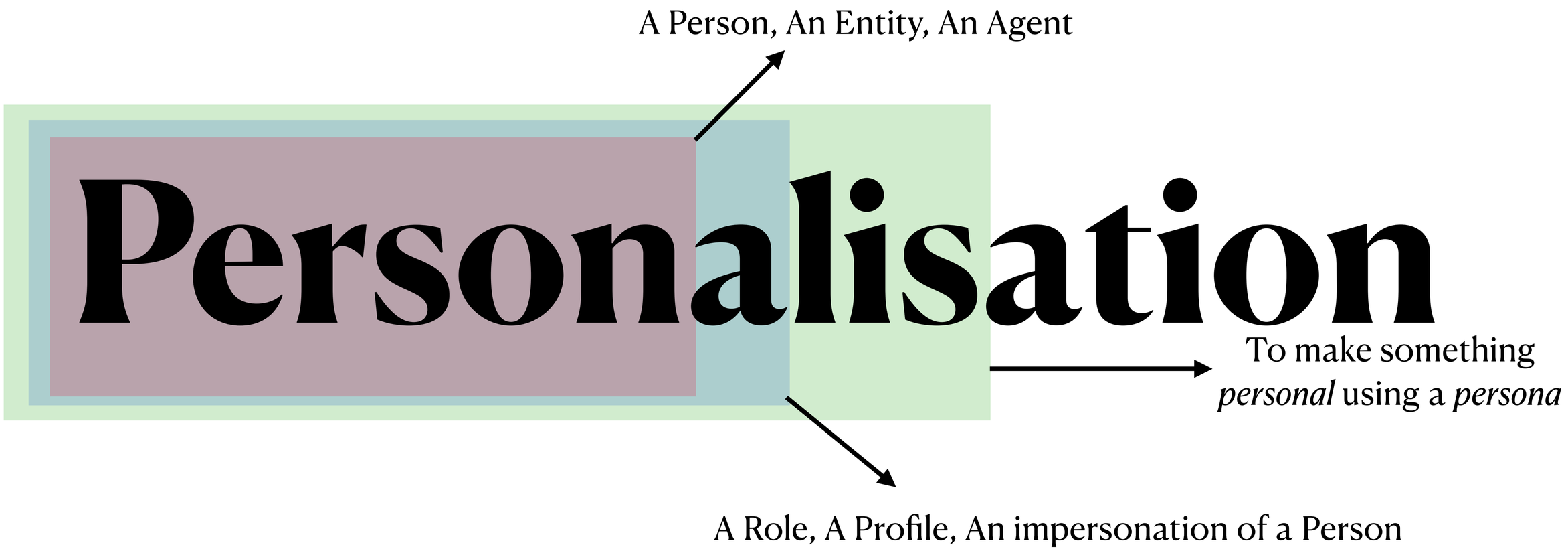
Centralised, Skewed, Linear



Web 1.0



Web 2.0


A Person, An Entity, An Agent
Personalisation
To make something
personal using a *persona*
A Role, A Profile, An impersonation of a Person

Personalisation vs Privacy

Availability of Information Reduces Privacy but Increases potential for Personalisation

PRIVACY

PERSONALISATION

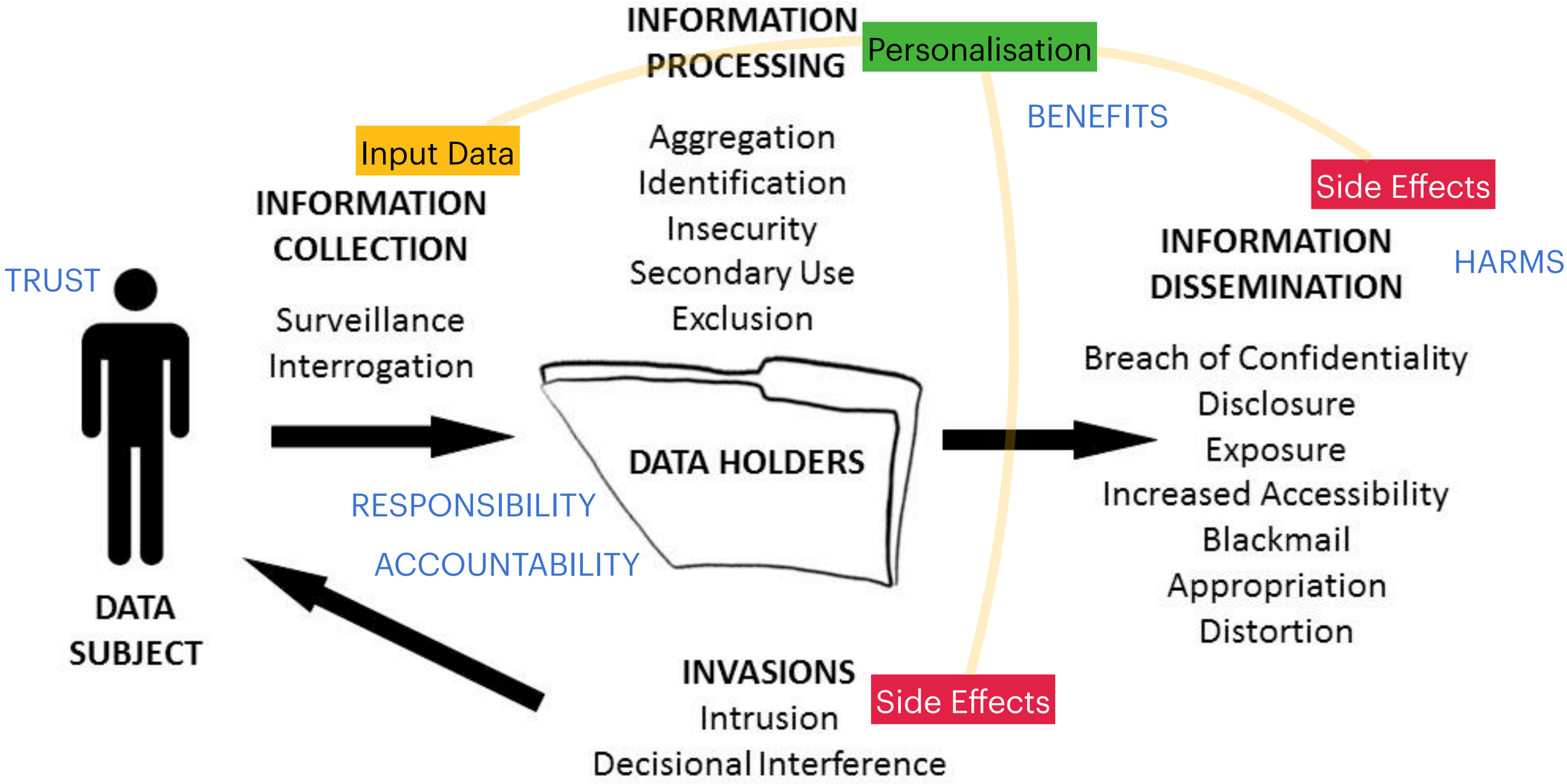
SECURITY

Privacy as Confidentiality
- Data Minimisation
- Reduce data required
- Identifiability of 'Persons'
- Necessity of Data Required

Privacy as Control
- Data Protection, Notices
- Control Flows, Algorithmic
- Involvement, Agency, Ability
- Rights, Norms

Privacy as Practice
- Contextual, Mediation
- Feedback, Negotiation
- Dynamic, Interactive
- Norms, Guidelines

Can you engineer privacy? On the potentials and challenges of applying privacy research in engineering practice - Seda Gurses
https://www.esat.kuleuven.be/cosic/publications/article-2465.pdf



Taxonomy of Privacy - Daniel Solve https://ssrn.com/abstract=667622

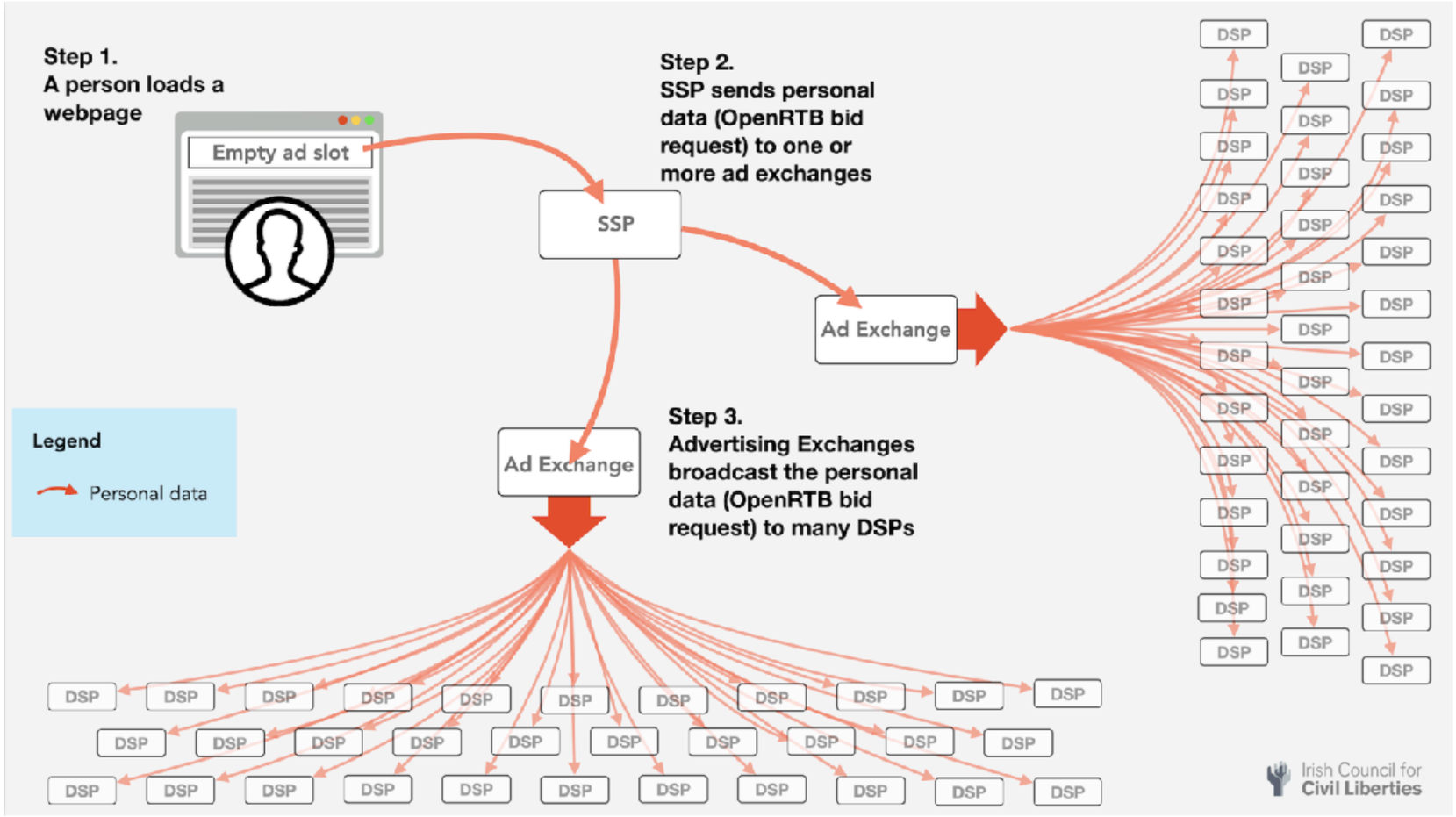
Overview of Personalisation Issues

Key takeaways

- What data is ‘used’ ??? —> Transparency
- What data is ‘needed’? What is ‘necessary’? —> Data Minimisation
- What are the sources of ‘data’ ? —> Transparency
- Is any data ‘sensitive’ ? Is it ‘special’ ? —> Ethical Concerns
- Is data (input/output) ‘accurate’ —> Accountability
- Is the output configurable ? —> Privacy by Design / Default
- Understand distinctions between *Privacy* vs *Security* vs *Identifiability* vs *Control*

Current Personalised Advertising Model

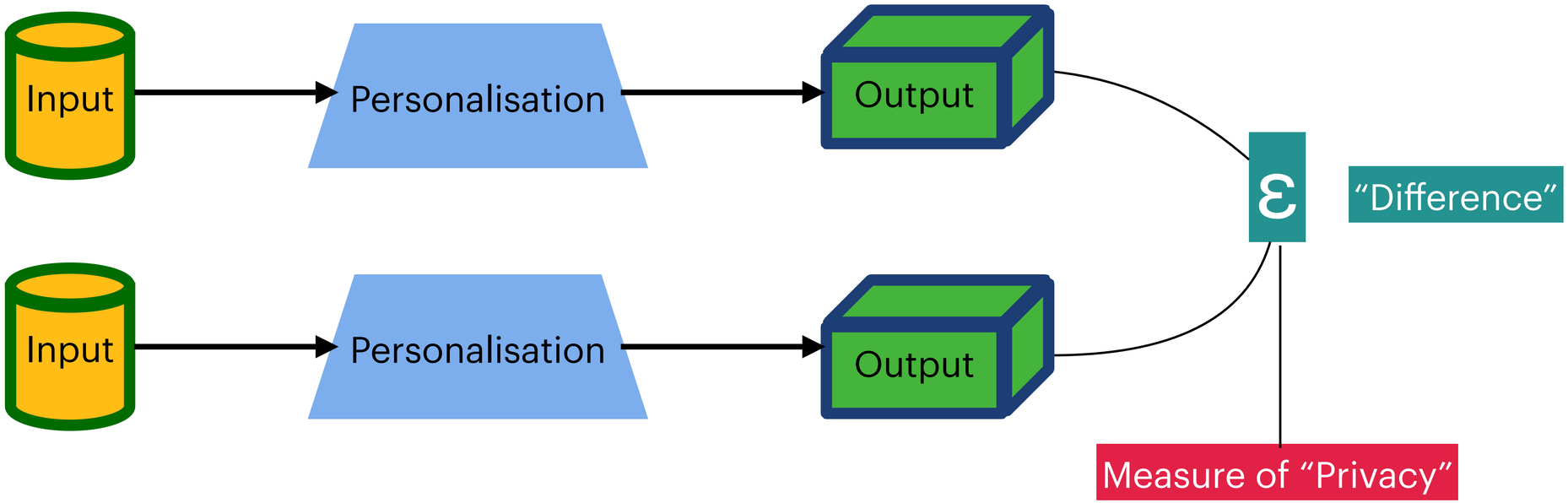
Surveillance-based Targeted Advertising



https://www.iccl.ie/digital-data/iab-europe-cant-audit-what-1000-companies-that-use-its-tcf-system-do-with-our-personal-data/

Differential Privacy

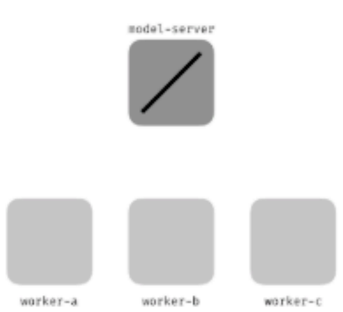
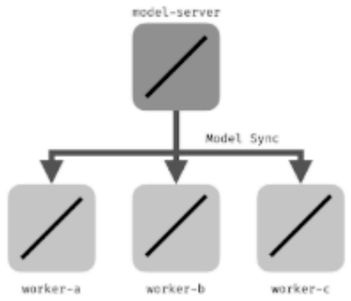
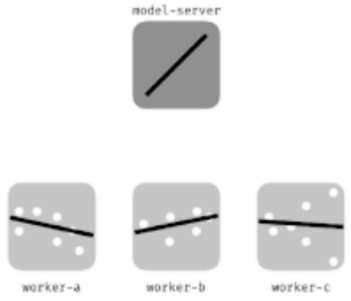
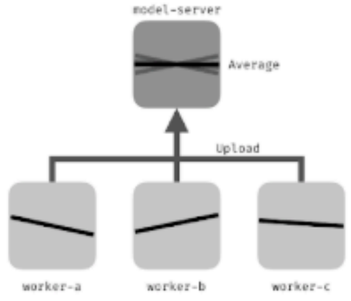
Performing Personalisation with lesser loss of Privacy



Differential Privacy: A Primer for a Non-Technical Audience - Wood et al. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3338027

Federated Learning

Do ML locally and pool models globally

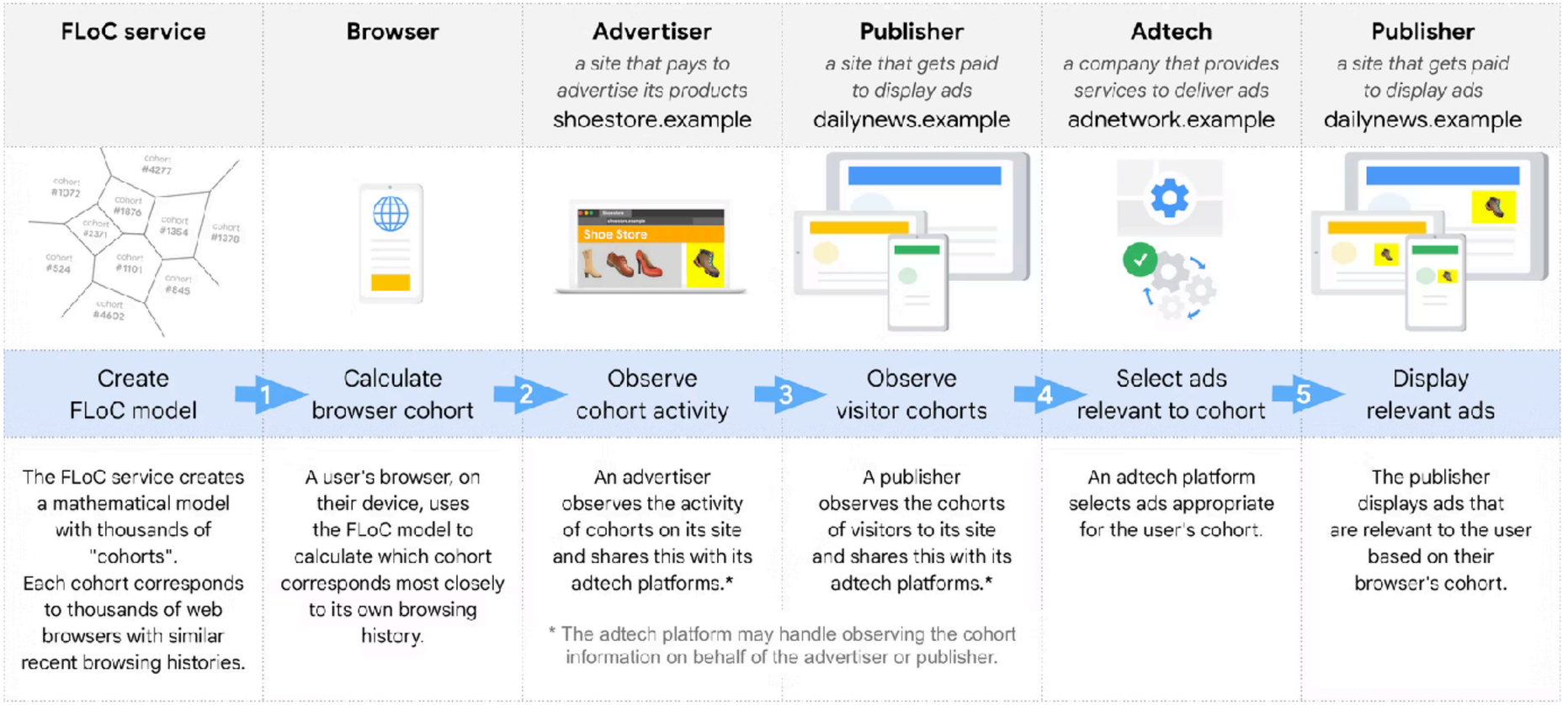
Step 1	Step 2	Step 3	Step 4
			
Central server chooses a statistical model to be trained	Central server transmits the initial model to several nodes	Nodes train the model locally with their own data	Central server pools model results and generate one global mode without accessing any data

https://en.wikipedia.org/wiki/Federated_learning

Google's FLoC Proposal

Topics

Federated Learning of Cohorts uses 'cohorts' to target advertisements



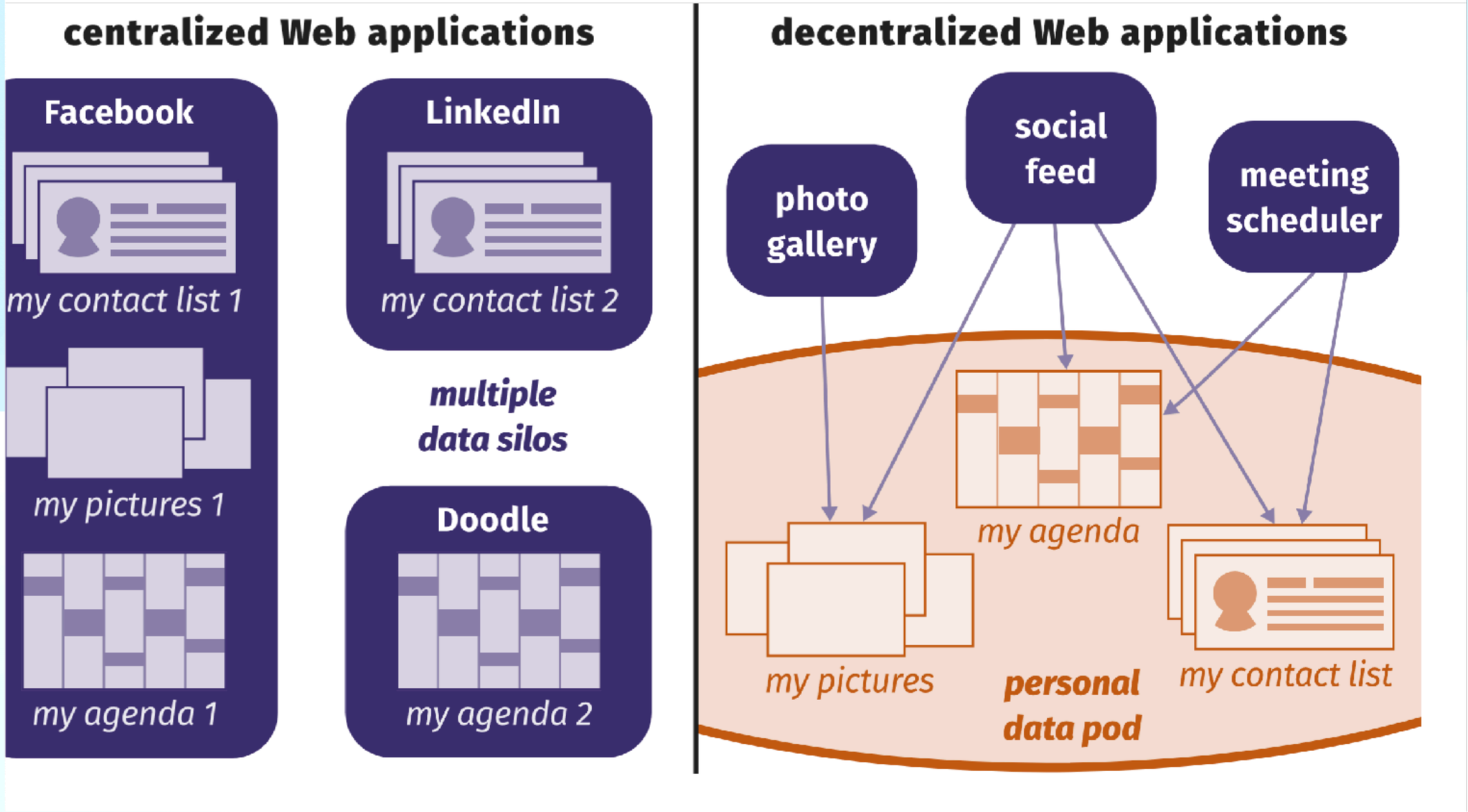
https://developer.chrome.com/docs/privacy-sandbox/floc/

Data Value

How to maximise the value of data?

How to minimise the “risks” associated with data?

Centralisation vs Decentralisation



SOLID: A Decentralised Web

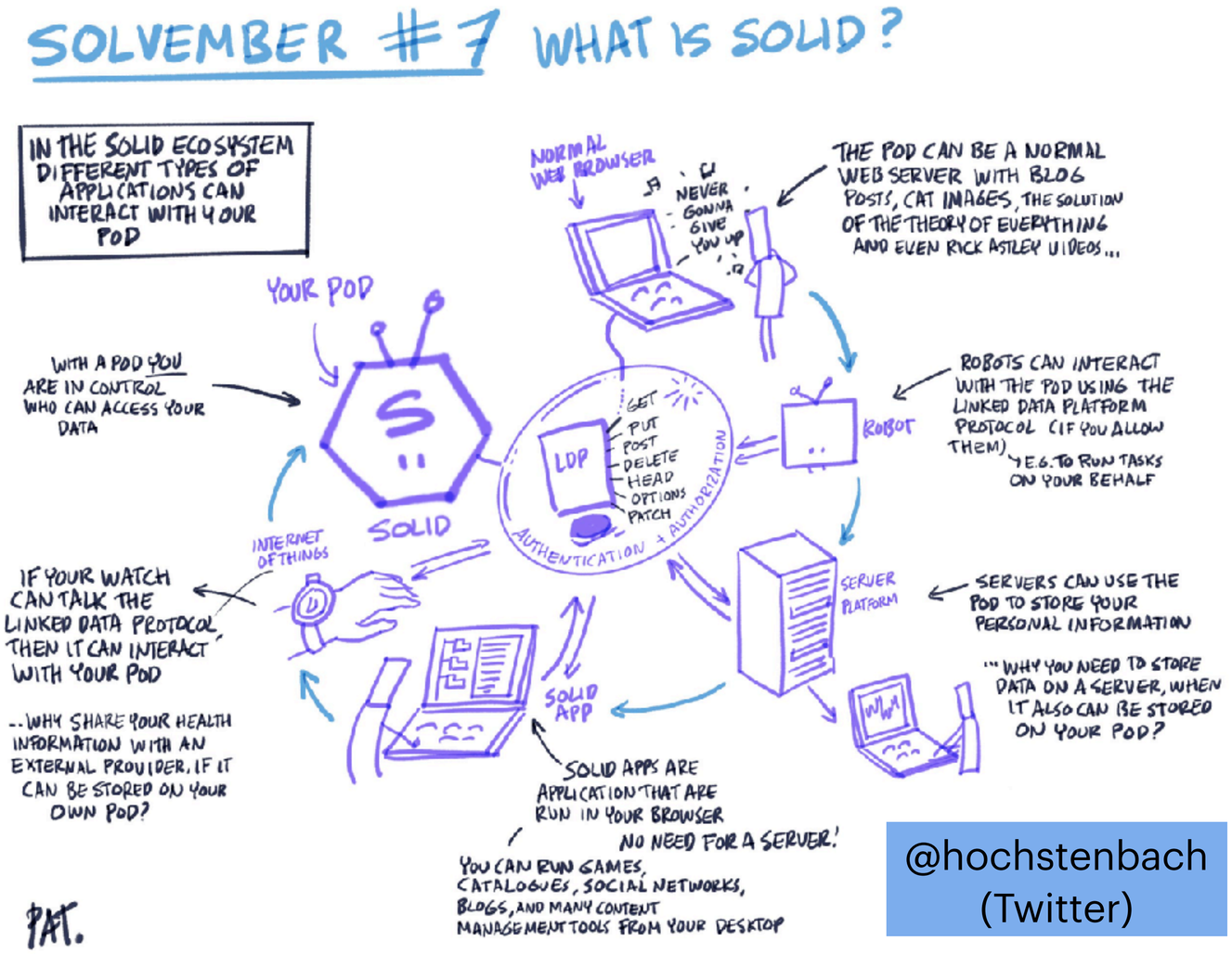
https://solidproject.org/

Centralised

- Companies decide how to collect, store data
- Companies decide how/where to use it
- Companies offer you choices and controls

Decentralised

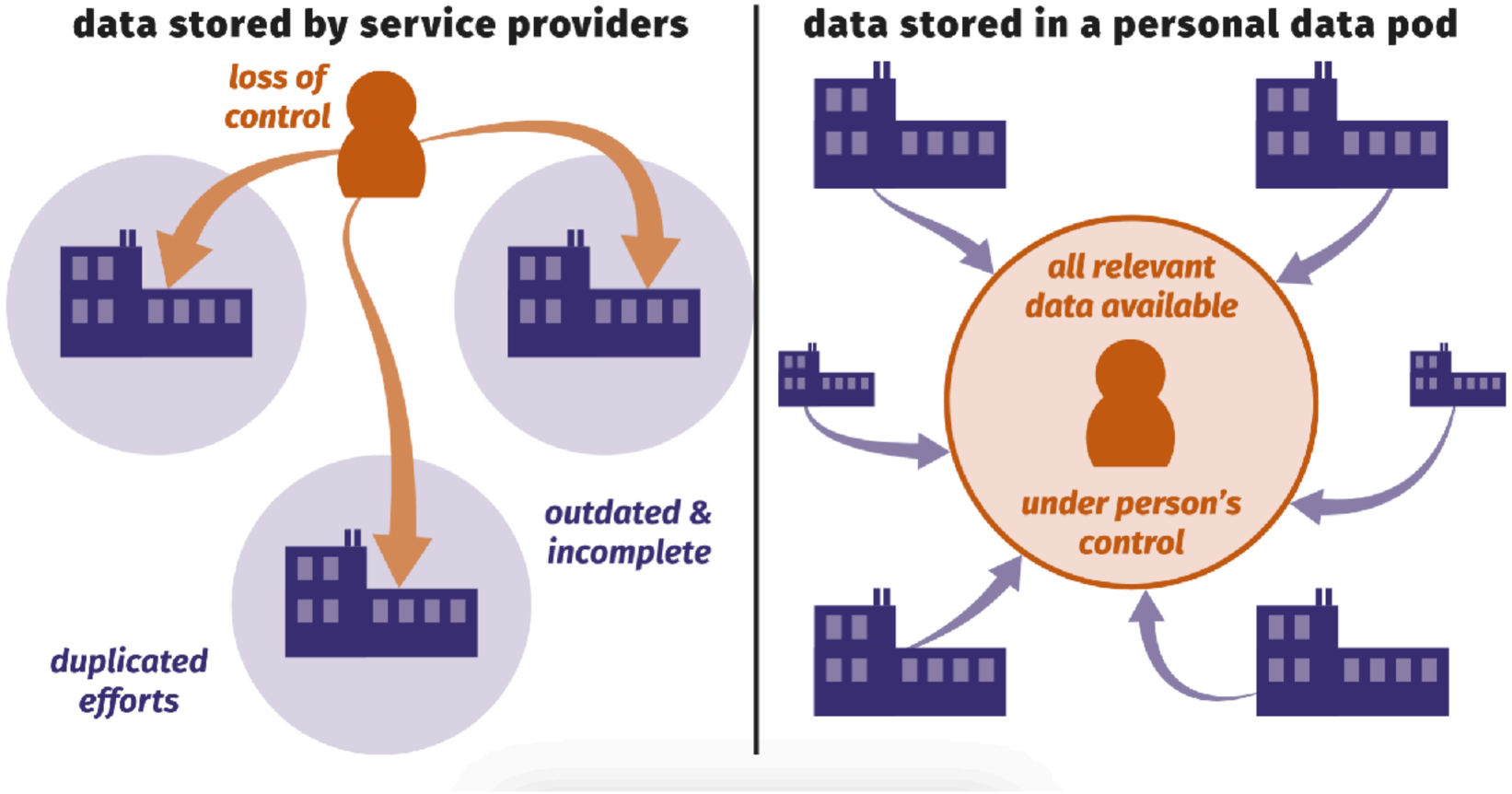
- You "control" where your data is stored
- You "control" how it is used by apps/services
- You offer choices and controls



@hochstenbach (Twitter)

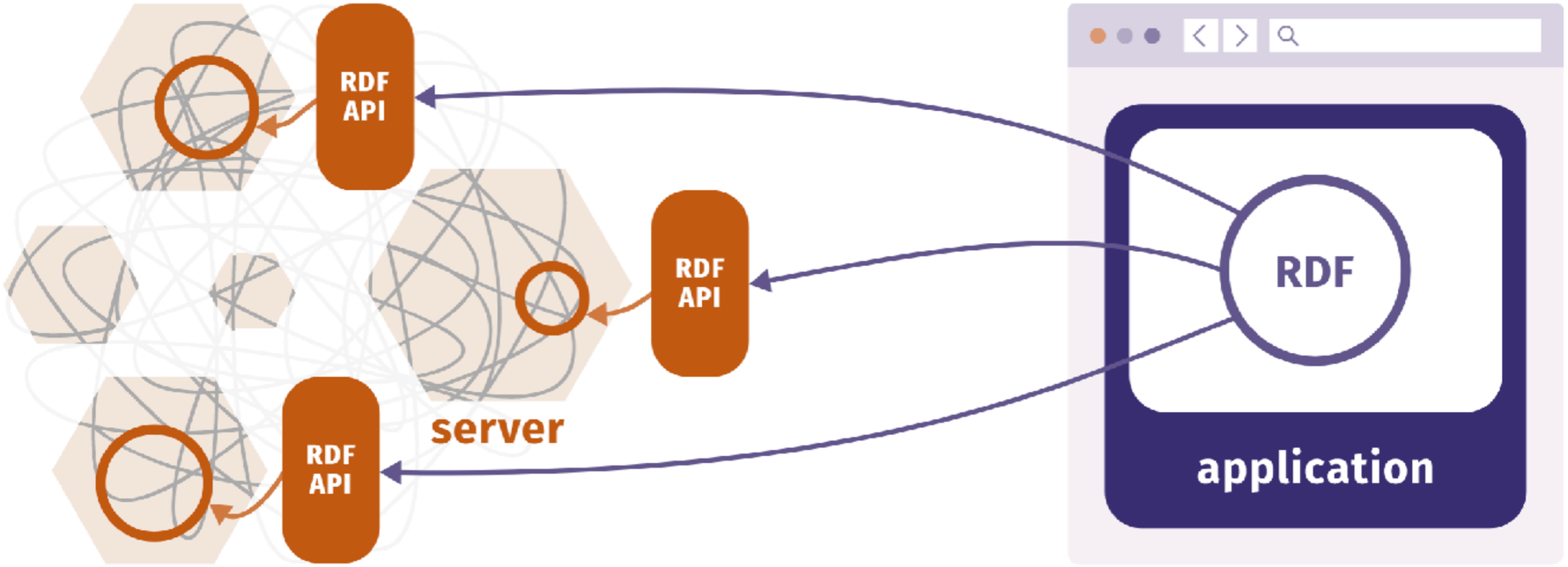
Solid - Inversion of Control

Every piece of data created *by* a person or *about* them, is stored in a data pod.

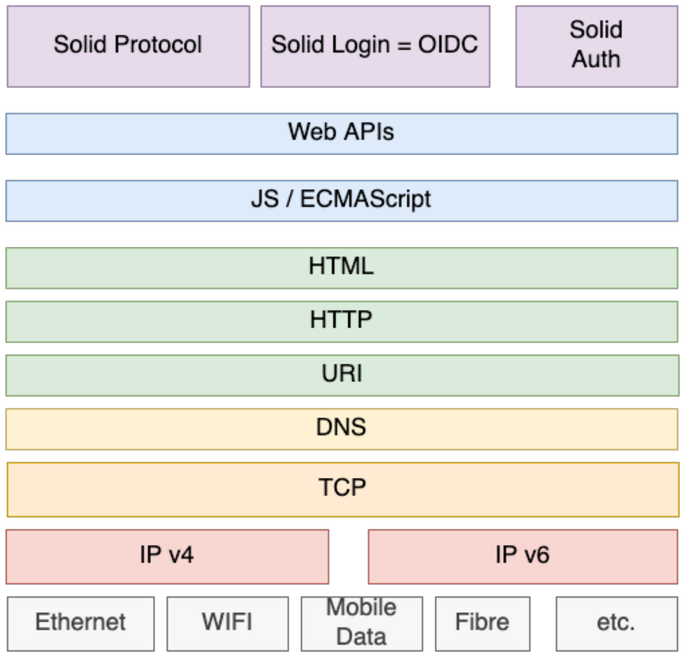


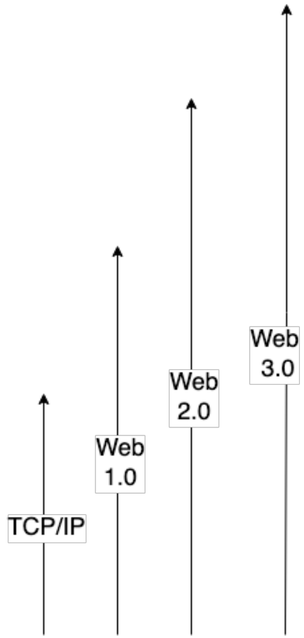
Solid - RDF ‘data standard’

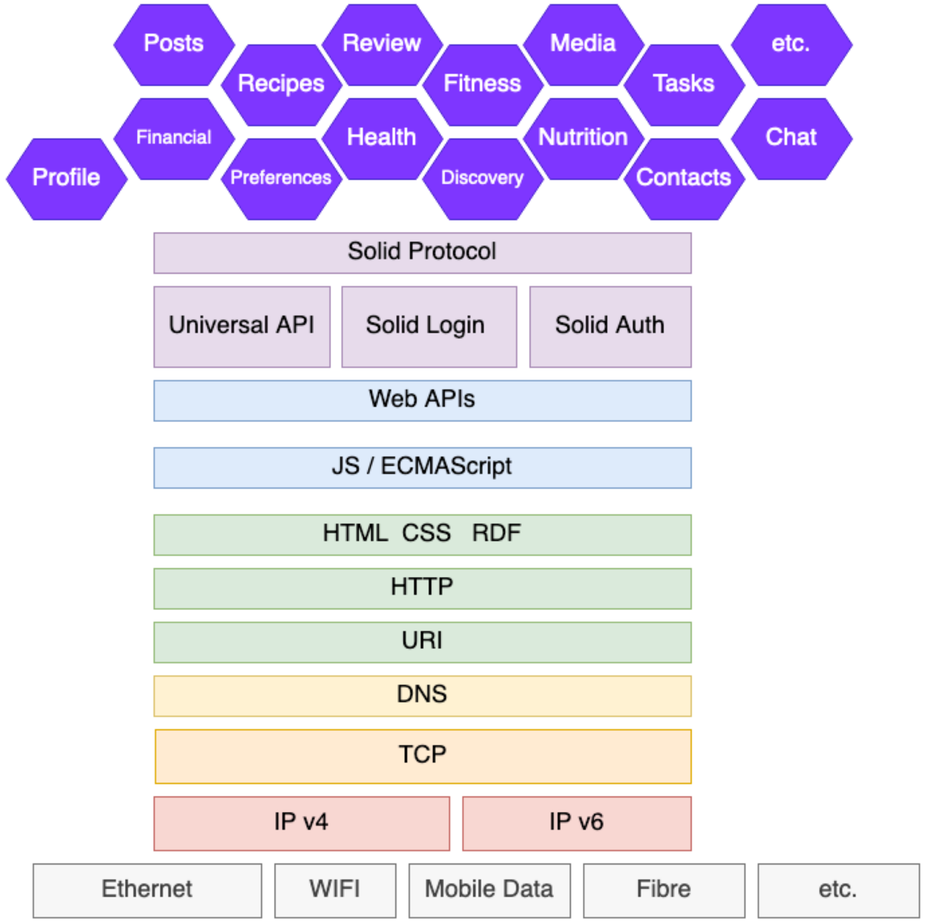
Each data pod exposes its part as RDF through a Web API.



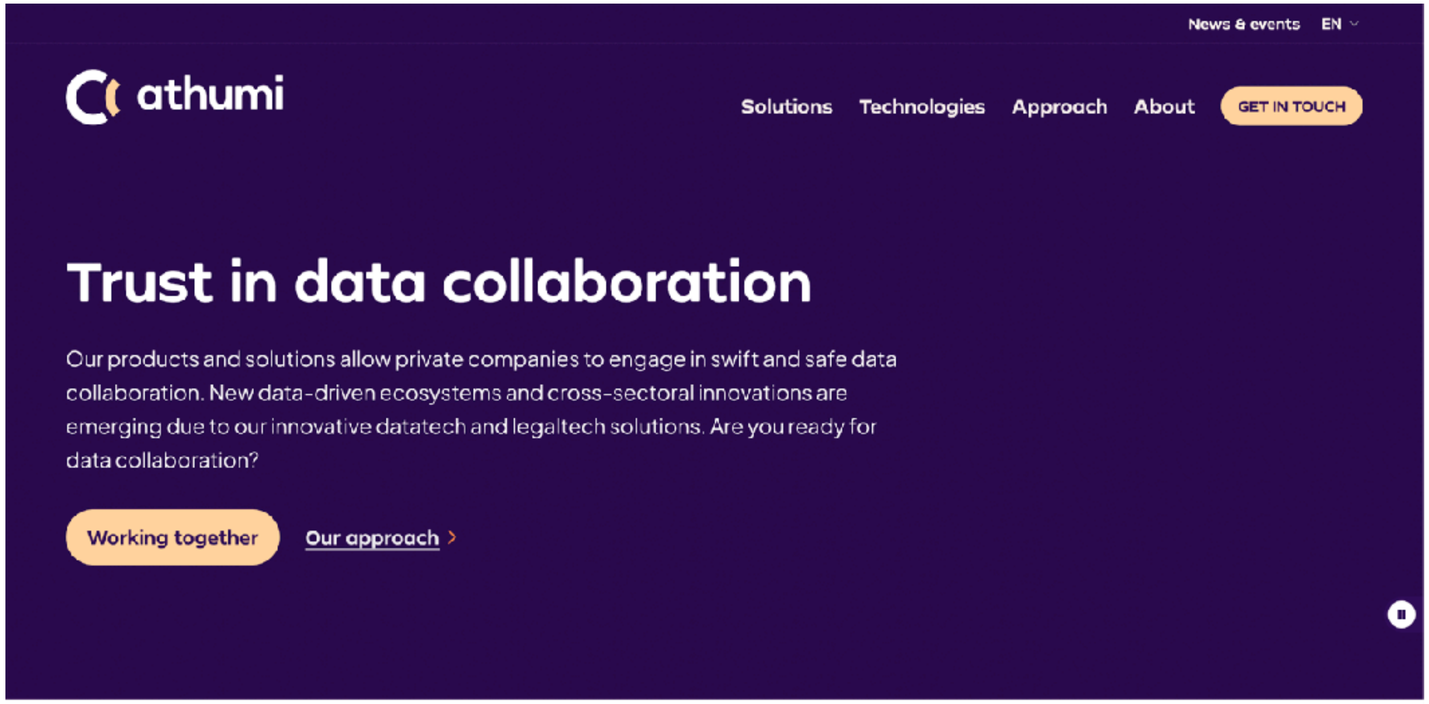
Web 3.0







Flanders (Belgium) - a Pod for every Citizen



Relevant Questions for Decentralised Data Governance

- Anelia Kurteva, Harshvardhan Pandit (2023)

Q1. Data Discovery
Q2. Identity
Q3. Security in/after Transit
Q4. Minimising End-user Cognitive Overload
Q5. Accountability
Q6. Preventing Legal Obligations from becoming a Hindrance
Q7. Digital Infrastructure
Q8. Automation Potential

Regulating Data

F: Findable
A: Accessible
I: Interoperable
R: Reusable

Harmonising FAIR data sharing with Legal Compliance

Harmonising FAIR data sharing with Legal Compliance

F: Findable
A: Accessible
I: Interoperable
R: Reusable

Harmonising FAIR data sharing with Legal Compliance

- Copyright Law
- Privacy Regulation
- Data Protection Regulation
  - GDPR
- AI Act
- Digital Markets Act
- Data Act
- Digital Services Act

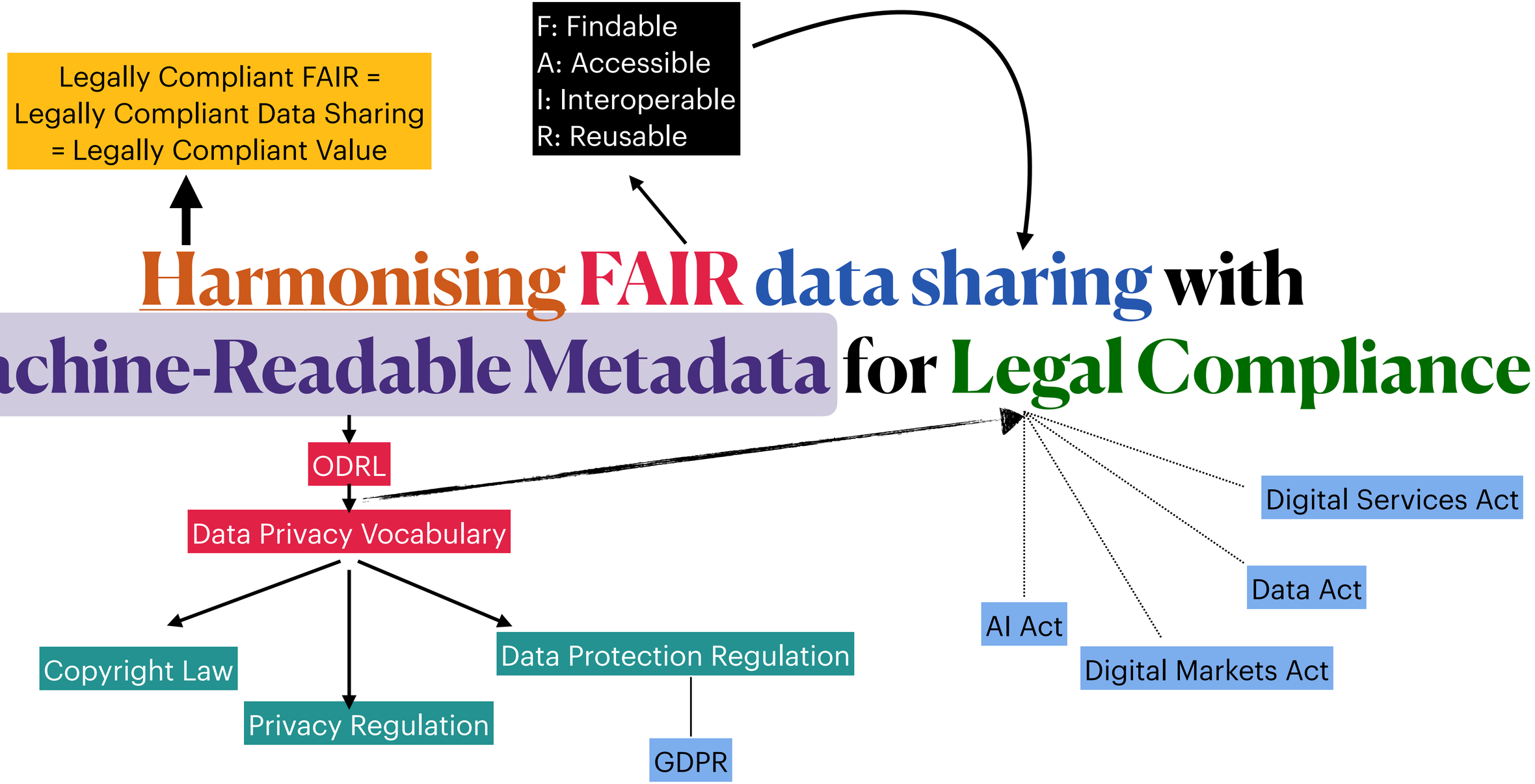
Legally Compliant FAIR =
Legally Compliant Data Sharing
= Legally Compliant Value

F: Findable
A: Accessible
I: Interoperable
R: Reusable

Harmonising FAIR data sharing with Legal Compliance

- Copyright Law
- Privacy Regulation
- Data Protection Regulation
  - GDPR
- AI Act
- Digital Markets Act
- Digital Services Act
- Data Act

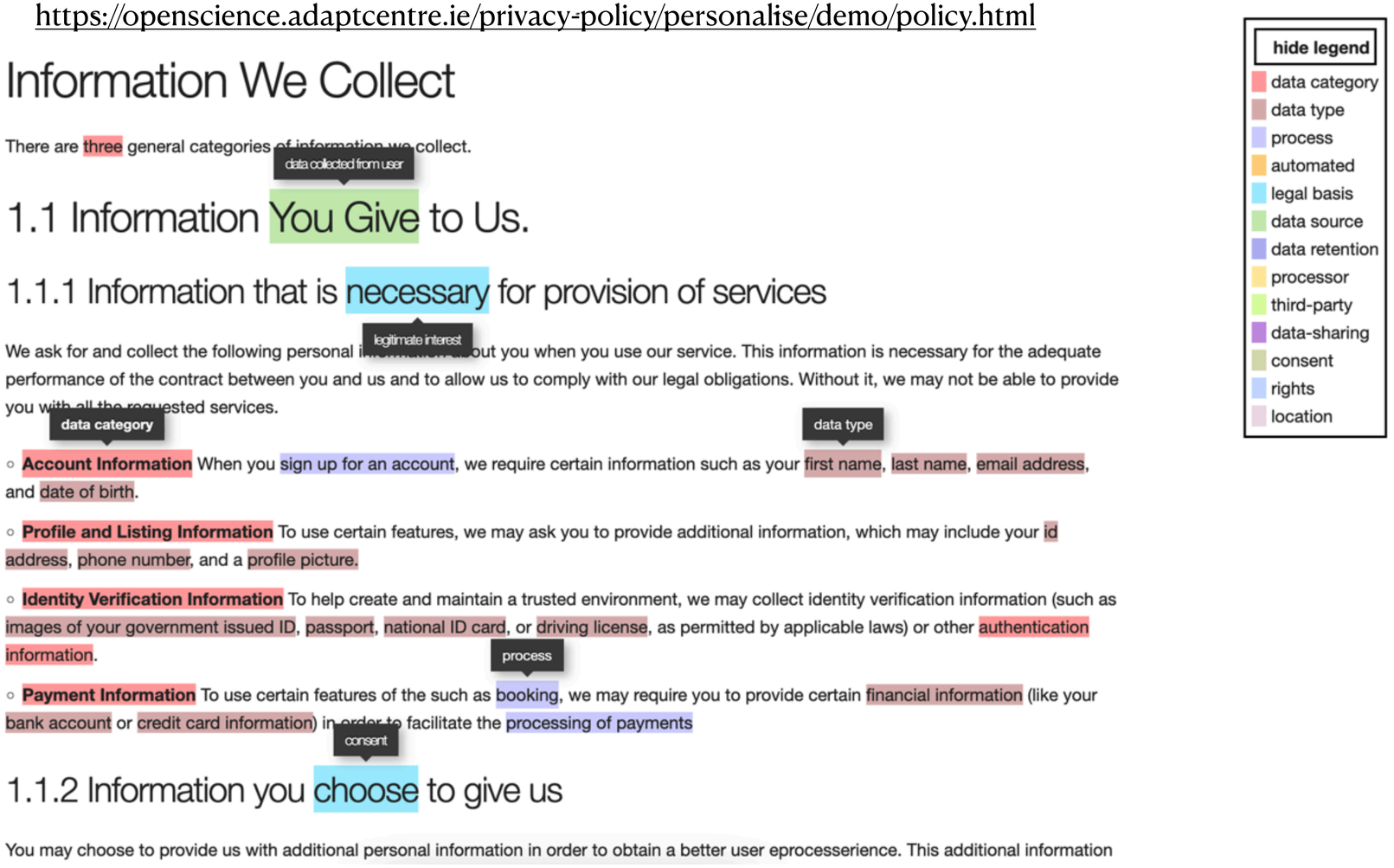
Harmonising FAIR data sharing with Machine-Readable Metadata for Legal Compliance



Real-World Use-Cases

Privacy Policy Analysis

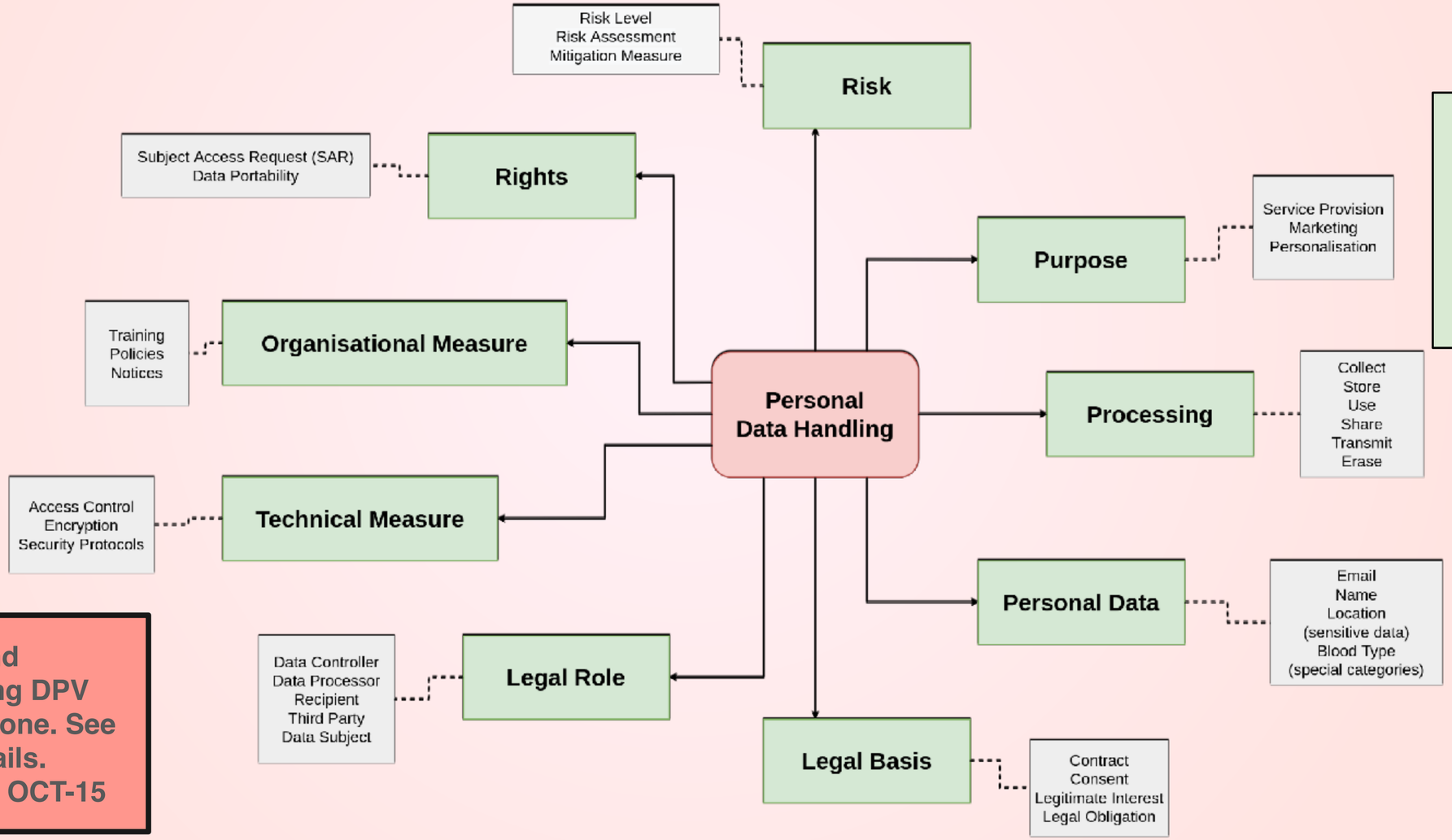
https://openscience.adaptcentre.ie/privacy-policy/personalise/demo/policy.html



Data Privacy Vocabulary (DPV)

Description of Personal Data Processing https://w3id.org/dpv

:Taxonomies:
Purpose
Personal Data
Processing
Legal Basis
Legal Role
Tech/Org Measures
Risk
Rights



Serialisations
RDF (canonical)
CSV, JSON

Semantics
RDFS+SKOS
OWL2

We invite comments and feedbacks for publishing DPV v1 - a significant milestone. See DPV spec for more details. Comment period: up to OCT-15

Artificial Intelligence - Real Risks

automate

machine-readable

interoperable format

queryable

verifiable

reuse standards & guidelines

RQ: How to assist with this?

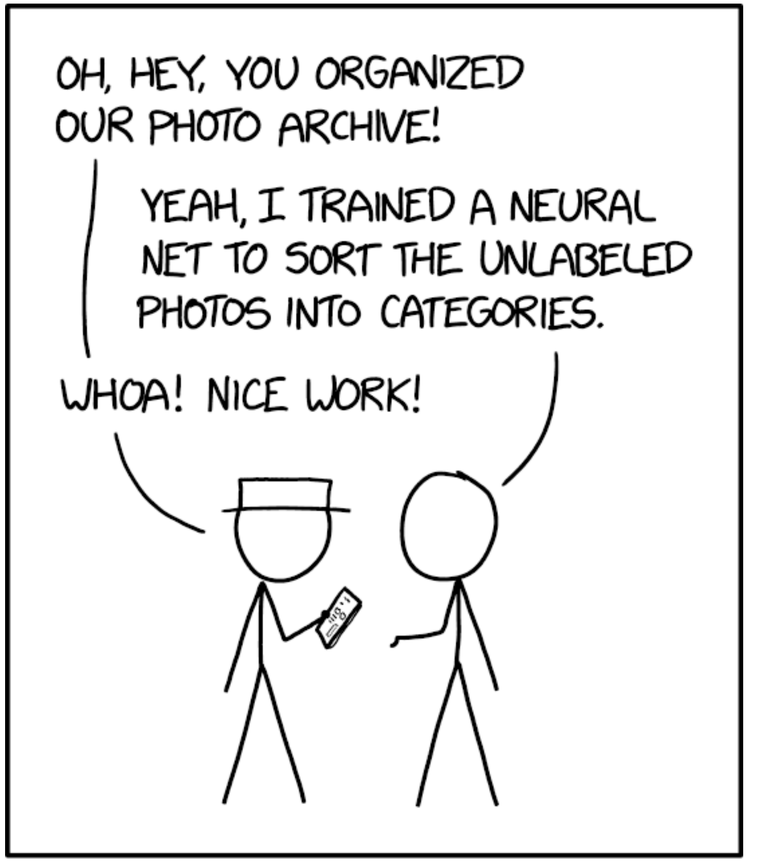
share information

simplify risk assessment

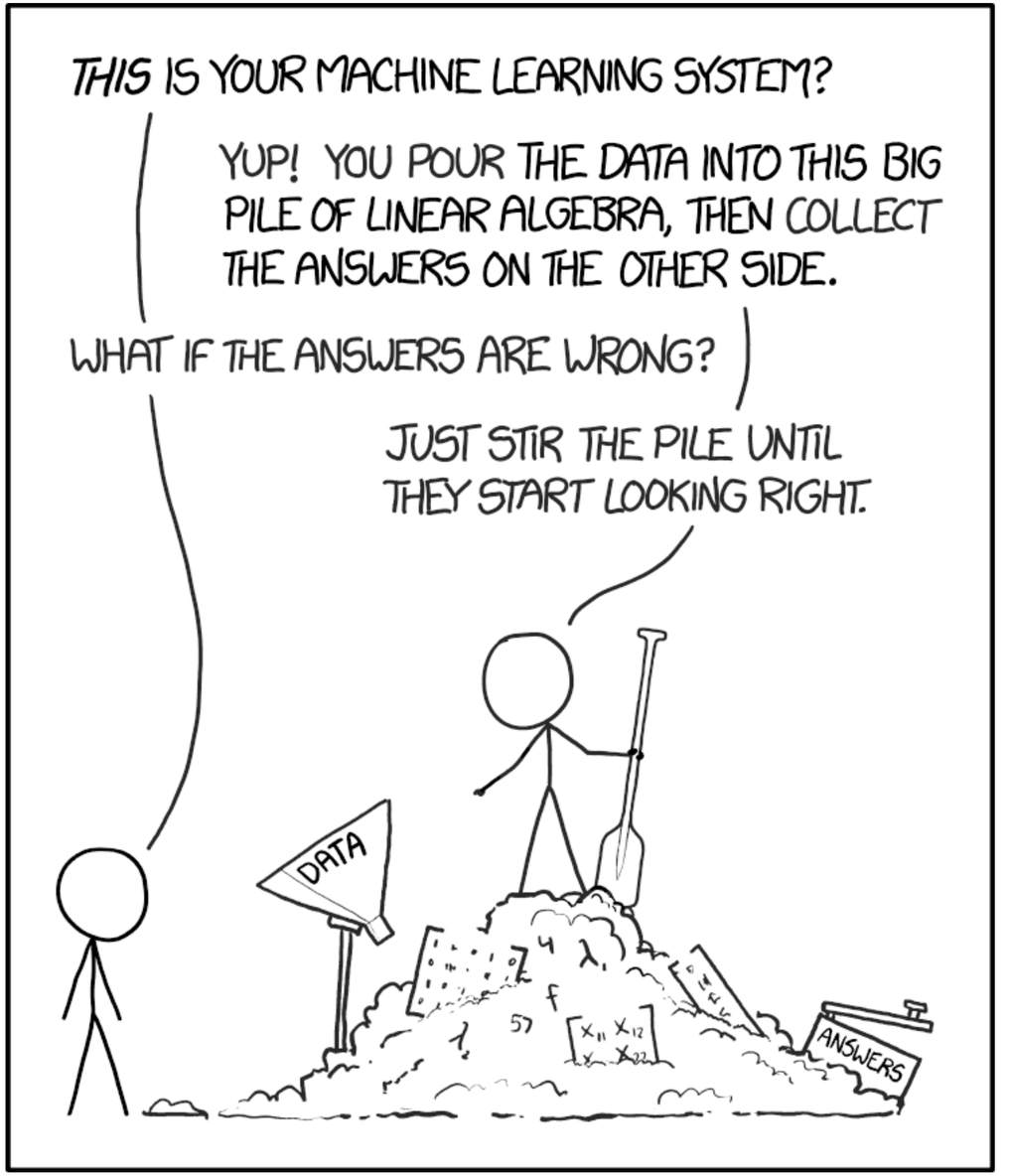
assist with other GDPR compliance

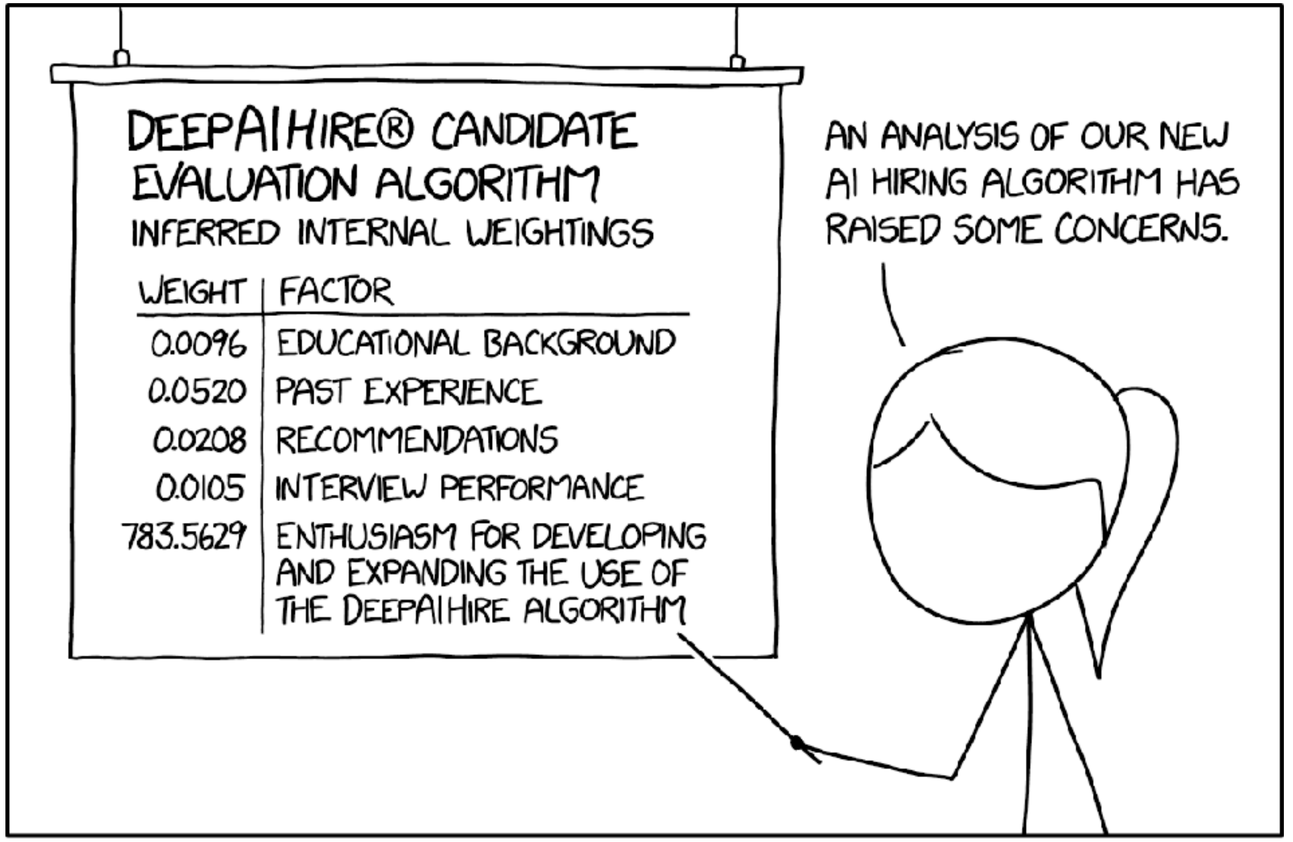
practical application

SO MUCH OF "AI" IS JUST FIGURING OUT WAYS TO OFFLOAD WORK ONTO RANDOM STRANGERS.

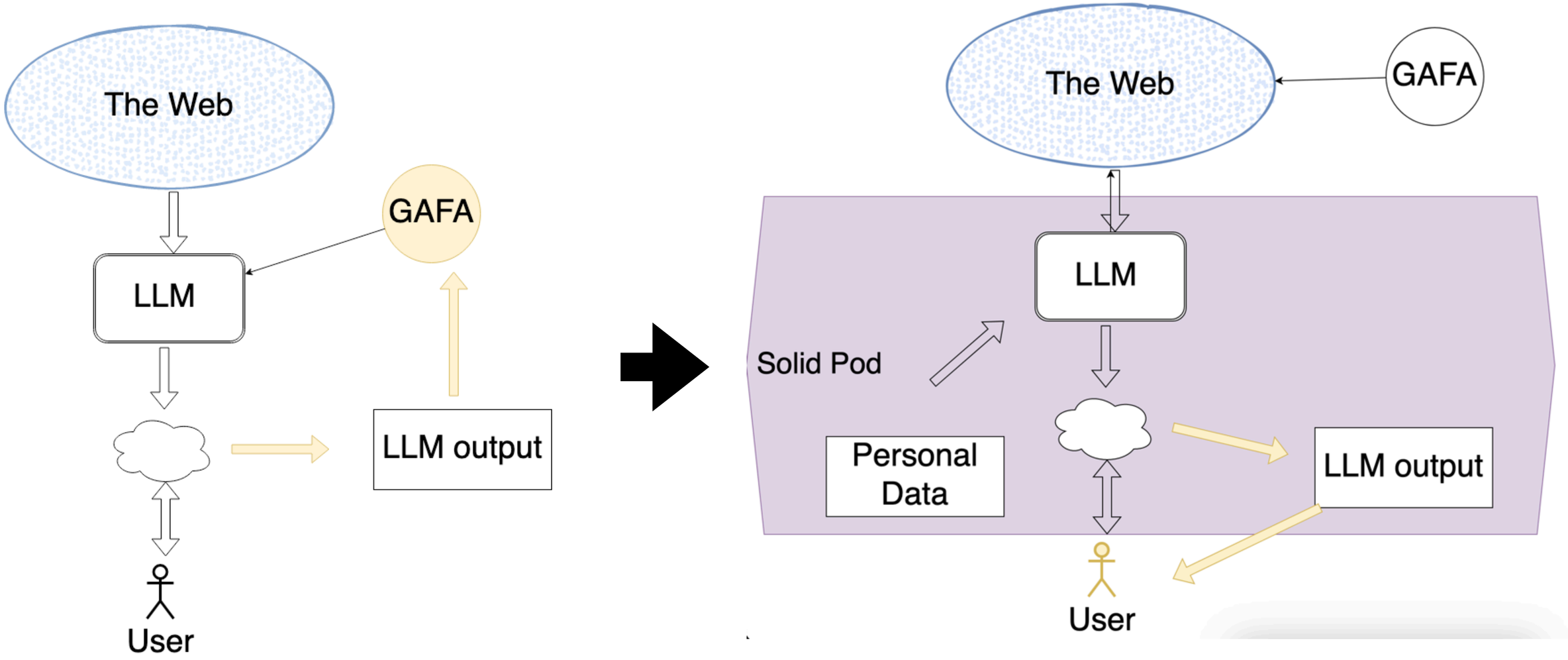


ENGINEERING TIP:
WHEN YOU DO A TASK BY HAND, YOU CAN TECHNICALLY SAY YOU TRAINED A NEURAL NET TO DO IT.

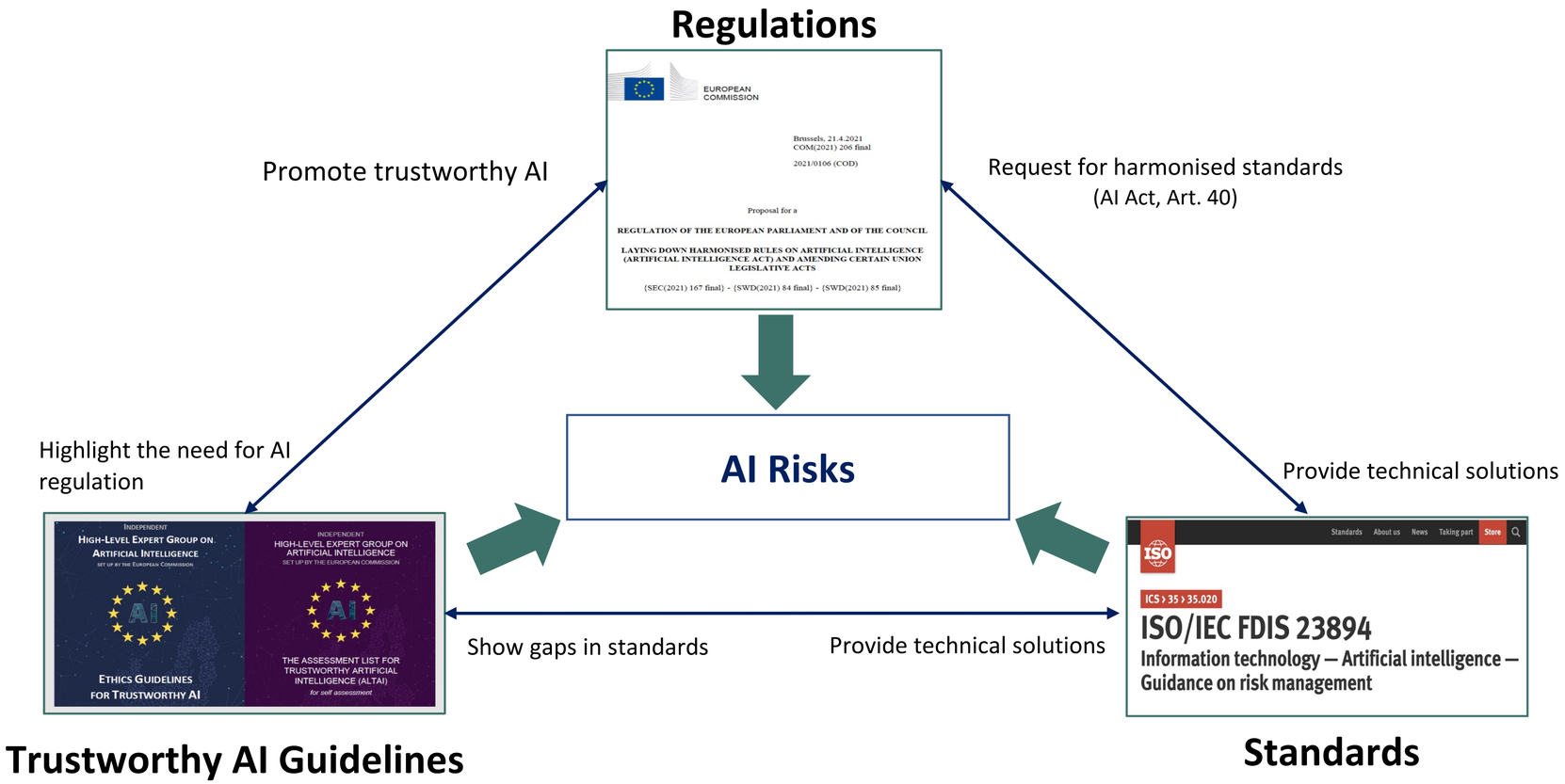

THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS


DEEPAIHIRE® CANDIDATE EVALUATION ALGORITHM
INFERRED INTERNAL WEIGHTINGS
WEIGHT | FACTOR
0.0096 | EDUCATIONAL BACKGROUND
0.0520 | PAST EXPERIENCE
0.0208 | RECOMMENDATIONS
0.0105 | INTERVIEW PERFORMANCE
783.5629 | ENTHUSIASM FOR DEVELOPING AND EXPANDING THE USE OF THE DEEPAIHIRE ALGORITHM
AN ANALYSIS OF OUR NEW AI HIRING ALGORITHM HAS RAISED SOME CONCERNS.

Vision - Tim Berners-Lee



Efforts Addressing AI Risks



AIRO: Ontology for representing AI Risks | Delaram Golpayegani et al. | SEMANTiCS 2022 | contact:sgolpays@tcd.ie | https://w3id.org/AIRO

AIRO Requirements

Describing High-Risk AI Systems

Questions to identify whether an AI system is high-risk according to Annex III

Question	concept	Relation with AISystem
What techniques are utilised in the system?	**AI Technique**	usesAITechnique
What domain is the system intended to be used in?	**Domain**	isAppliedWithinDomain
What is the intended purpose of the system?	**Purpose**	hasPurpose
What is the application of the system?	**AI Application**	hasApplication
Who is the intended user of the system?	**AI User**	hasAIUser
Who is the subject of the system?	**AI Subject**	hasAISubject
In which environment is the system used?	**Environment Of Use**	isUsedInEnvironment

ANNEX I
ARTIFICIAL INTELLIGENCE TECHNIQUES AND APPROACHES referred to in Article 3, point 1

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods.

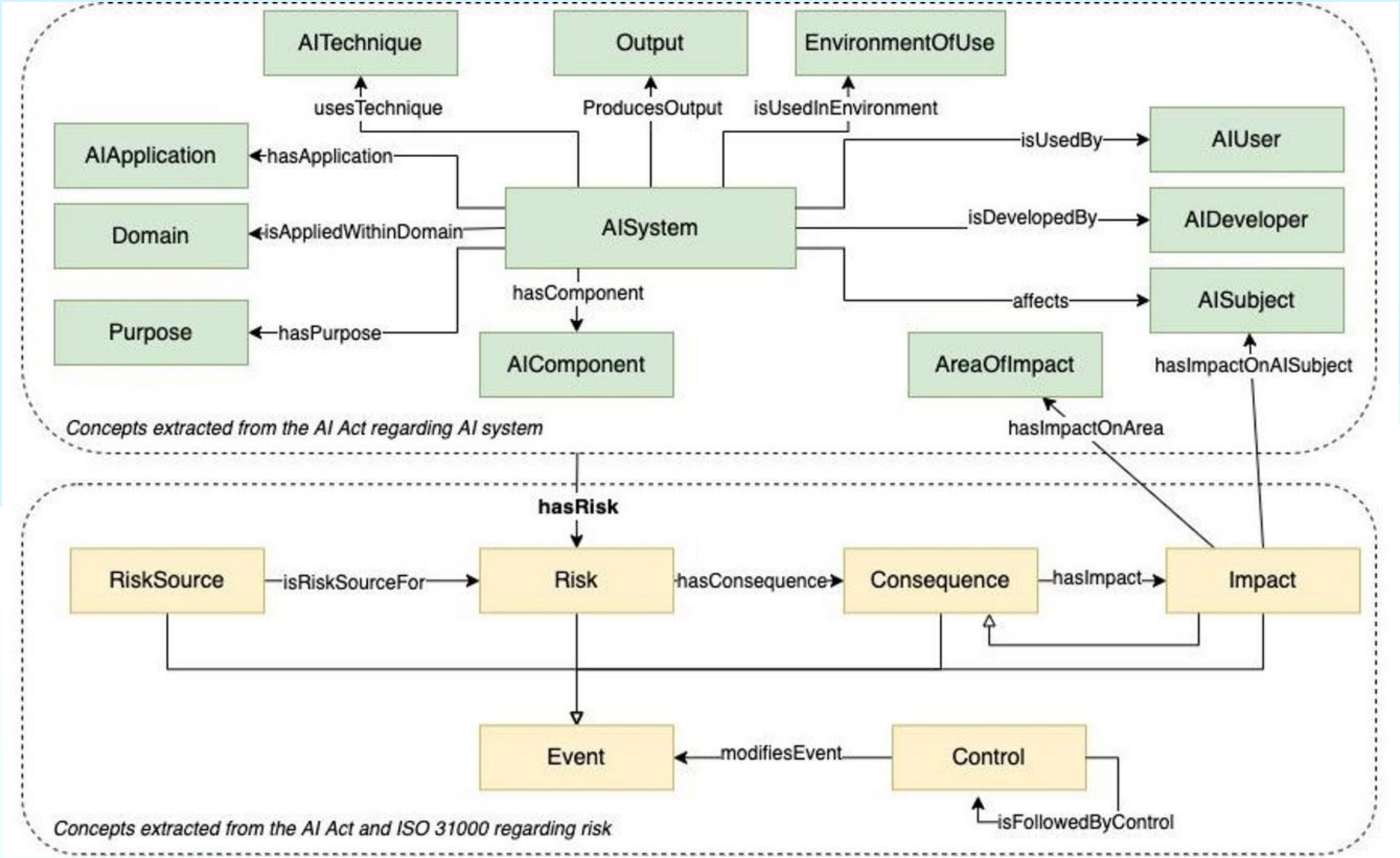
ANNEX III
HIGH-RISK AI SYSTEMS REFERRED TO IN ARTICLE 6(2)

High-risk AI systems pursuant to Article 6(2) are the AI systems listed in any of the following areas:

1. Biometric identification and categorisation of natural persons:
   (a) AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons;
2. Management and operation of critical infrastructure:
   (a) AI systems intended to be used as safety components in the management and operation of road traffic and the supply of water, gas, heating and electricity.
3. Education and vocational training:
   (a) AI systems intended to be used for the purpose of determining access or assigning natural persons to educational and vocational training institutions;
   (b) AI systems intended to be used for the purpose of assessing students in educational and vocational training institutions and for assessing participants in tests commonly required for admission to educational institutions.
4. Employment, workers management and access to self-employment:
   (a) AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests;
   (b) AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships.
5. Access to and enjoyment of essential private services and public services and benefits:
   (a) AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for public assistance benefits and services, as well as to grant, reduce, revoke, or reclaim such

10

AIRO: Ontology for representing AI Risks | Delaram Golpayegani et al. | SEMANTiCS 2022 | contact:sgolpays@tcd.ie | https://w3id.org/AIRO


AITechnique
Output
EnvironmentOfUse
usesTechnique
ProducesOutput
isUsedInEnvironment
AIApplication
hasApplication
isUsedBy
AIUser
Domain
isAppliedWithinDomain
AISystem
isDevelopedBy
AIDeveloper
Purpose
hasPurpose
hasComponent
affects
AISubject
AIComponent
AreaOfImpact
hasImpactOnAISubject
Concepts extracted from the AI Act regarding AI system
hasImpactOnArea
hasRisk
RiskSource
isRiskSourceFor
Risk
hasConsequence
Consequence
hasImpact
Impact
Event
modifiesEvent
Control
isFollowedByControl
Concepts extracted from the AI Act and ISO 31000 regarding risk

Identification of High-Risk AI Systems

```
PREFIX airo: <https://w3id.org/AIRO#>
SELECT  ?system ?technique ?domain ?purpose
        ?application ?user ?subject ?environment
WHERE {
        ?system a airo:AISystem ;
                airo:usesTechnique ?technique ;
                airo:isUsedWithinDomain ?domain ;
                airo:hasPurpose ?purpose ;
                airo:hasApplication ?application ;
                airo:isUsedBy ?user ;
                airo:affects ?subject ;
                airo:isUsedInEnvironment ?environment . }
```

AIRO concept	
`AISystem`	uber's real time id check
`AITechnique`	machine learning techniques
`Domain`	employment
`Purpose`	biometric identification of drivers to decide on contract termination
`AIApplication`	facial recognition
`AIUser`	uber driver
`AISubject`	uber driver of bame background
`EnvironmentOfUse`	work related relations

1. Biometric identification and categorisation of natural persons:

   (a) AI systems intended to be used for the 'real-time' and 'post' re identification of natural persons;

4. Employment, workers management and access to self-employment:

   (a) AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests;

   (b) AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships.

- Manual analysis

High Risk

SHACL Shapes for Automatic Identification of High-Risk AI

- ”Rules” to determine whether AI satisfies conditions for being “high-risk”
- Choose your favourite flavour of rule languages & mechanisms

- We chose **SHACL**
- Why:
  - Flexible, Standardised
  - Extensible with plugins/features
  - Built-in documentation of outputs
  - Integrate to instead check outputs e.g. another rule engine

- We implement SHACL shapes for clauses defined in Annex III that determine high-risk
- Validation is to NOT satisfy the expressed criteria

```
1  @prefix dash: <http://datashapes.org/dash#> .
2  @prefix sh: <http://www.w3.org/ns/shacl#> .
3  @prefix airo: <https://w3id.org/AIRO#> .
4  @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
5  :AnnexIII-1
6      a sh:NodeShape ;
7      sh:targetClass airo:AISystem ;
8      sh:message "High-Risk AI System as per AI Act Annex III-1"@en ;
9      sh:description "Biometric Identification of Natural Persons"@en ;
10     sh:not [
11         a sh:PropertyShape ;
12         sh:path airo:hasPurpose ;
13         sh:class airo:BiometricIdentification; ] .
```

Your Data, Your AI

Towards a Decentralised Future WITH SEMANTIC WEB

Harshvardhan Pandit | Knowledge and Engineering | TCD 2023 | https://harshp.com/presentations