LINKED DATA CONTRACTS TO SUPPORT DATA PROTECTION AND DATA ETHICS IN THE SHARING OF SCIENTIFIC DATA

In the light of new EU General Data Protection Regulation (GDPR)

Ensar Hadziselimovic
Kaniz Fatema
Harshvardhan Pandit
David Lewis

adaptcentre.ie

This Presentation

- Background and motivation
- Data Controllers and Data Subjects - legal perspective
- GDPR and the impact
- Data Subject Rights
- Ontologies
- Sharing scientific data
- Conclusion

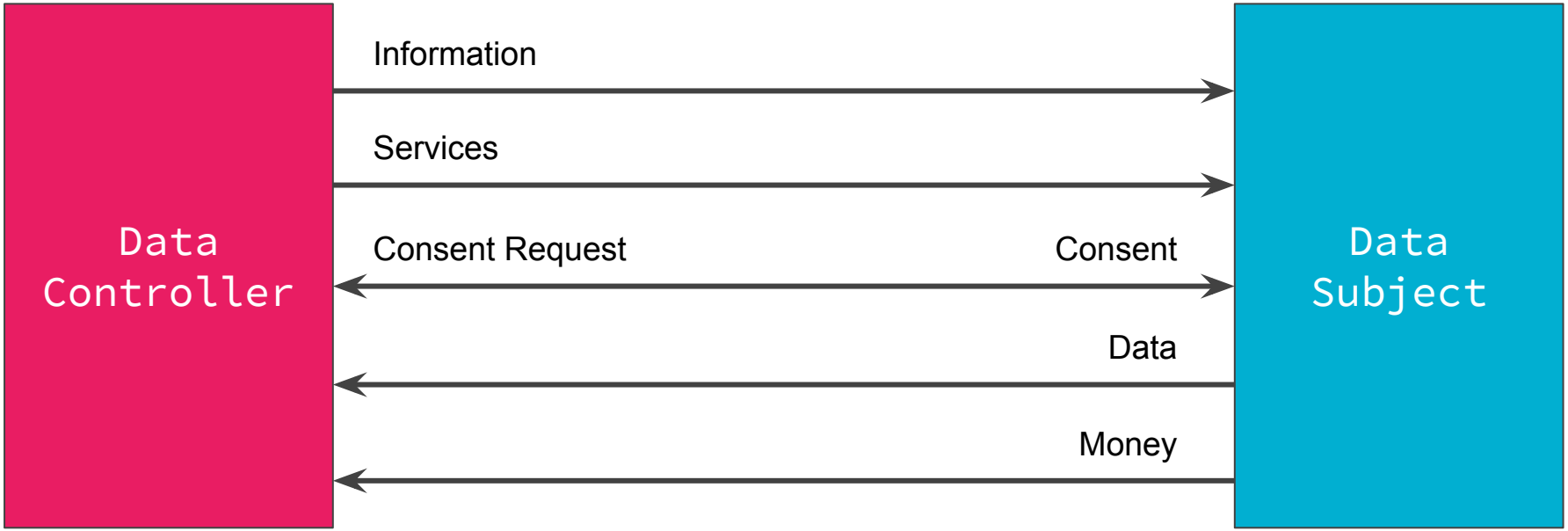
Background and Motivation

- Current rules and practices on academic research ethics ask for participant to provide written consent
- They vary from country to country
- Common ground is intention to protect the participants and researchers by stating the purpose and requesting a consent
- The EU’s adoption of GDPR imposes new requirements for tracking informed consent for the usage of any form of personal data

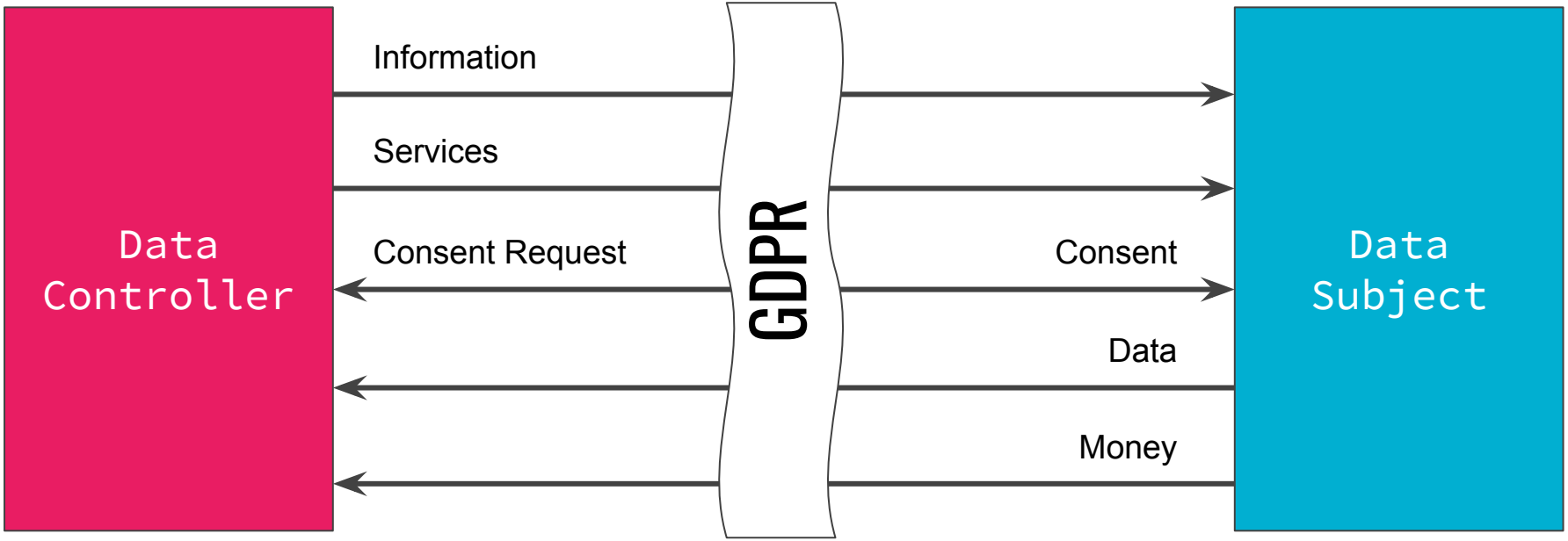
GDPR - Quick Facts

- Replaces 1995 EU Directive which regulates the processing of personal data within the European Union
- Adds to European Legislation from 2011 “Cookie Law”
- Will come into force on the 25th May 2018, replacing the existing data protection framework under the EU Data Protection Directive.
- Unlike a directive, it does not require national governments to pass any enabling legislation, and is thus directly binding and applicable

Data Controllers vs Data Subjects



Data Controllers vs Data Subjects



GDPR - The Good, the Bad and the Ugly

- GDPR is here mainly to protect the Data Subject rights (“identified or identifiable natural person”)
- There is a strong requirement on data controller, who is responsible for handling of personal data, to be able to demonstrate to regulators that any data subject’s personal data has been correctly processed
- Tracking data usage and keeping records
- Organizations in breach of GDPR can be fined up to 4% of annual global turnover or €20 Million (whichever is greater).

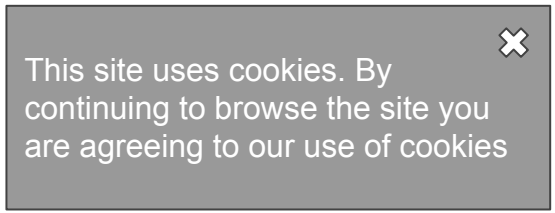
GDPR - The Good, the Bad and the Ugly

- When applied to research data, GDPR potentially imposes more rigorous and legally enforceable requirements on the collection and processing of data from individuals than current research ethics practices
- Ambiguity in certain clauses: data ‘portability’, ‘common machine-readable format’
- Incompatibility of organisational systems between data controllers / providers

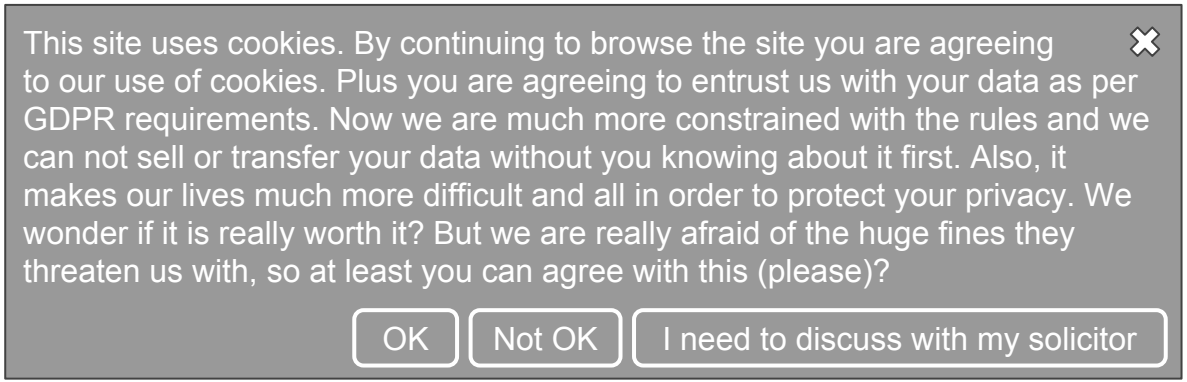
GDPR - The Good, the Bad and the Ugly

- Potential organisational issues for smaller businesses and individuals in roles of data controllers
- Expressing the rules and regulations on per-use cases

Before



After



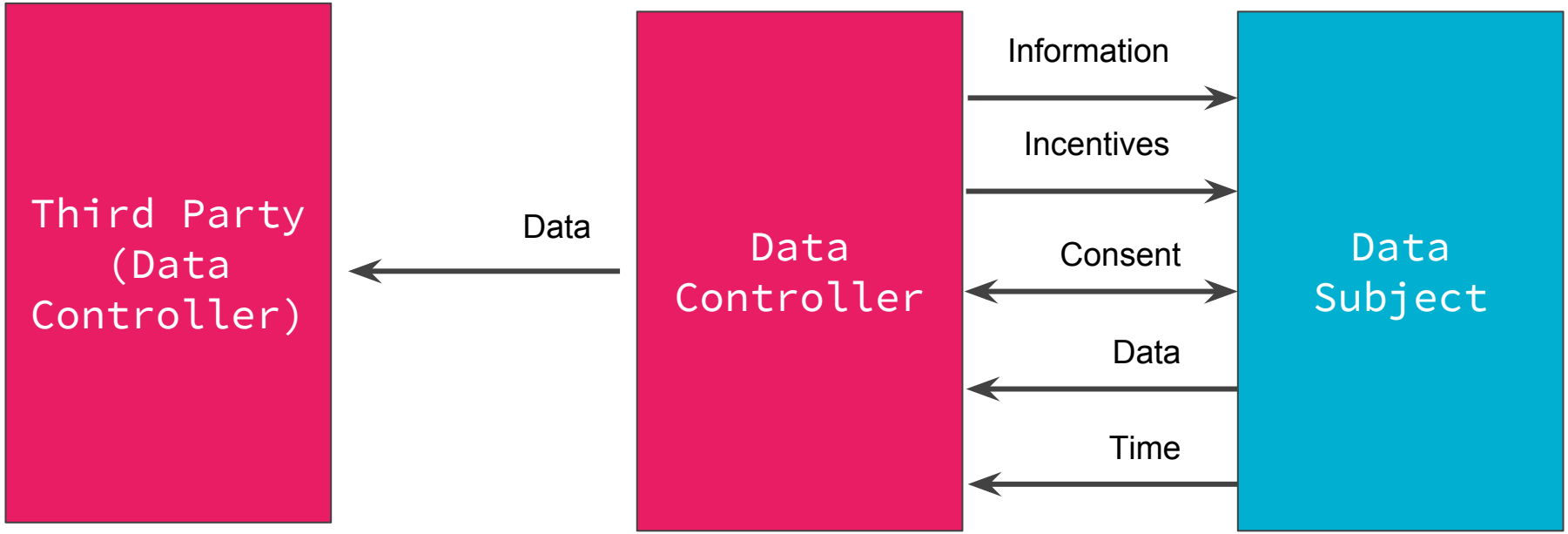
Data Subject Rights in GDPR

- Right to information “concise, transparent, intelligible and easily accessible”
- Right to file a subject access request (SAR)
- Right to rectification
- Right to erasure (the “right to be forgotten”)
- Right to data portability
- Right to restrict access
- Right to object
- Right to know about data being shared with third parties

Data Subject Rights in GDPR

- Right to **information** “concise, transparent, intelligible and easily accessible”
- Right to file a subject **access** request (SAR)
- Right to **rectification**
- Right to **erasure** (the “right to be forgotten”)
- Right to data **portability**
- Right to **restrict** access
- Right to **object**
- Right to know about data being **shared** with third parties

Data Controllers and Third Parties in Scientific Data Sharing



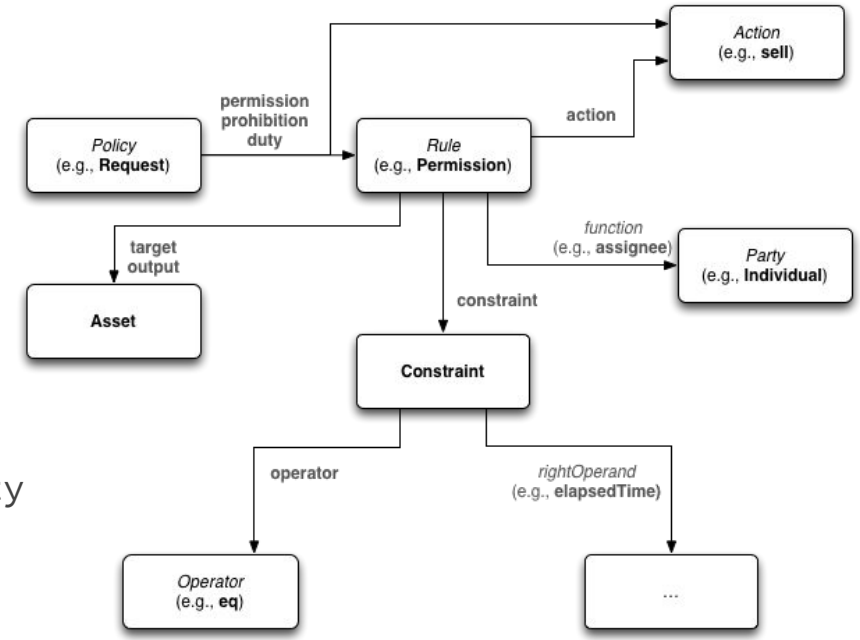
Ontologies

- ODRL (Open Digital Rights Language)
  - Expressing digital rights management
  - Used for access and usage conditions
- DataID
  - Describing simple and complex datasets in an interoperable way
  - Extension to the Data Catalog Vocabulary (DCAT)
  - Adds features such as dataset hierarchies, permissions, distribution and machine-readable licensing information
  - Integrates provenance information from the PROV Ontology
- DUV
  - Tracking, sharing, and persisting dataset usage

ODRL

https://www.w3.org/community/odrl/
https://www.w3.org/2016/poe/

- Classes
  - Permission, Prohibition
  - Asset, Party
  - Policy, Privacy
- Properties
  - inheritAllowed, inheritFrom
  - constraint, relation
  - attributedParty, consentingParty
- Concepts
  - grantUse, annotate
  - anonymize, attribute
  - derive, distribute



DPRL: Data Protection Rights Language

- Gentle extension to ODRL, by the means of templating
- Overcoming technological obstacles in following the criteria set by GDPR
- Licensing-oriented action concepts on ODRL are insufficient

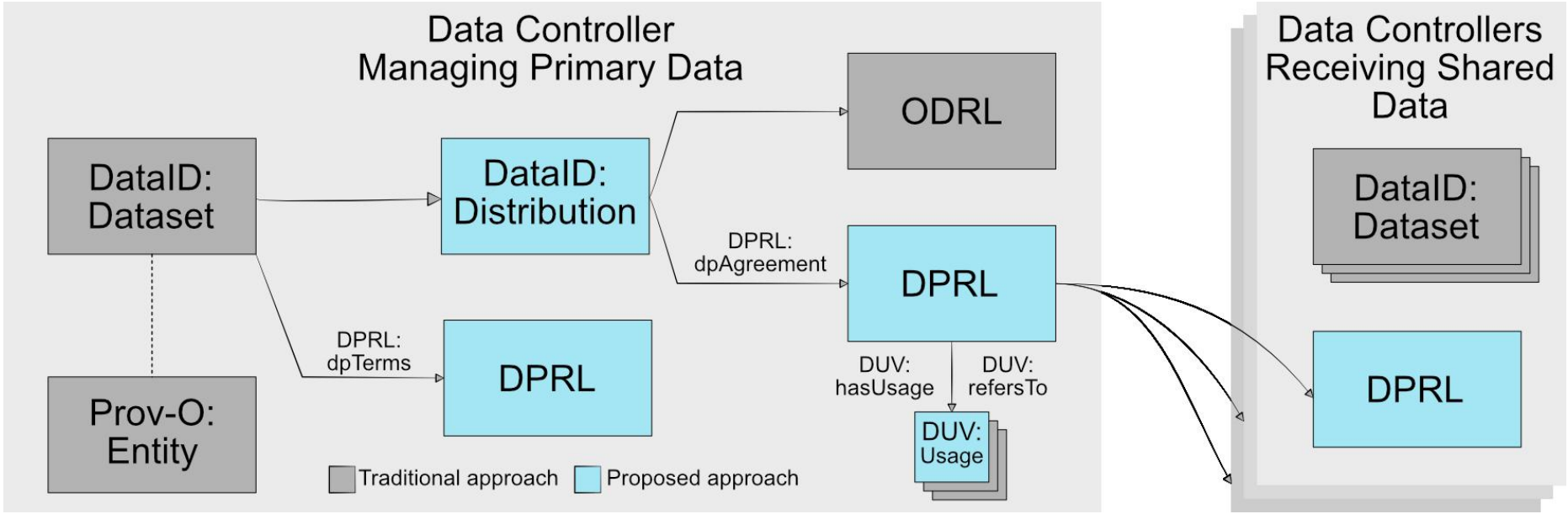
DPRL concepts:	dpAccess	dpRectify	dpErase	dpPort	dpRestrict	dpObject
ODRL concepts:	read/use	update	delete	export/transfer	-	-

DPRL properties:	dpAgreement	dpTerms		
ODRL properties:	permission	constraint	**DPRL classes:**	dpRight dpObligation

Sharing Scientific Data

- Currently there is no established method commonly used between the academic institutions when sharing scientific data
- GDPR data portability clause requires that the data subject can access their data in common machine-readable format
- GDPR requires not only the provenance tracking of data processing to support compliance
- Tracking extends to the data controller of any other organisation with which the data is shared

Open data architecture for sharing the scientific data between academic institutions



Future Work

- Further develop DPRL through fuller integration with data set processing provenance tracking using PROV-O and the use of SPARQL over DPRL and other DataID components
- GDPR compliance checks
- Sharing linked data itself through Linked Data Rights ontology
- Design of future open scientific data API and platforms, such as those previously developed for publication metadata in the OpenAire project

that's all for today

questions?

Slides:
https://goo.gl/8A8ghU