AARC Authentication and Authorisation for Research and Collaboration

Policies for the processing of personal data

Authentication and Authorisation for Research and Collaboration

On behalf of the NA3/AARC – Uros Stevanovic
KIT

FIM4R, Vienna
20 Feb 2017

Data protection in research and education
Processing of personal data

- Collaboration between scientists inevitably crosses administrative domains and borders
  - Europe
  - Outside of Europe
- Access to services:
  - Computing instances
  - Storage
  - Processing applications
- Providing services involves processing of personal data
  - Accounting
  - Monitoring
  - Collaboration
  - Security

Disclaimer

- We’re NOT lawyers
- This is not a legal advice
- However, we did and are communicating with lawyers

Disclaimer

- We’re NOT lawyers
- This is not a legal advice
- However, we did and are communicating with lawyers
- Again, this is NOT a legal advice

Scope

- Collection of usage data in research infrastructures (RI) and e-infrastructures
- Correlating resource usage to people and groups
- Accumulation of usage data across countries (and continents)
- Collection and processing of personal data for incident response

NOT in scope

- Policies regarding personal data in research data sets
  - medical data
- Attribute release:
  - E.g. attributes provided by the IdP to an SP → DPCoCo
- However, if attributes contain Personally Identifiable Information (PII), data protection policies still apply

Assumptions:

- All activities undertaken on the infrastructures are assumed as “professional” work, i.e. everything researcher employed by a “legal entity” (university, institute, etc.) does, binds the organisation in it (maybe even implicitly)

Background

- Initial overview, described in "Requirements on data to protect from AAI, community, resource providers and e-infrastructure":
  - Current data protection policies
  - Types of personal data
  - Current national legal framework
  - https://aarc-project.eu/wp-content/uploads/2015/11/MNA3.2-AccountingDataProt-20151130.pdf
- General Data Protection Regulation (GDPR)
  - Adopted 14th April 2016, goes into force 25th May 2018
  - Legally binding for all Member States, without the need for parliament ratification
  - Data Protection Directive 94/46/EC – still valid

GDPR changes

- Apply to the processing of personal data by controllers and processors in the EU, regardless where it takes place
- Penalties – up to 4% of annual global turnover or 20M€ (whichever is greater)
- Consent – conditions are strengthened (clear and plain language, explicitly related to the processing, easy to withdraw)
- Breach notification
- Privacy by design
- Right to be forgotten
- Data Protection Officers
- Right to access

Info: www.eugdpr.org

Legal terms

- Personal data
  - Article 4(1) - Any information relating to an identified or identifiable natural person (‘data subject’), i.e name, IP address, email, geolocation, identifier to physical, gender, mental, economic etc. identity
- Processing
  - Article 4(2) – any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, i.e. collection, recording, organisation, structuring, storage, etc.
- Data Controllers and Data Processing
  - Article 4(7) – Data controller is any person (natural or legal) or entity that decides for which purpose personal data is processed
  - Article 4(8) – Data processor is any person (natural or legal) who process the data on behalf of the controller
  - Data controller is the responsible party that must ensure that all processing of personal data complies with the GDPR
- In the case of research communities, data controllers are most common
- Joint controllers – Article 26(1): “two or more controllers jointly determine the purposes and means of processing”

Rules for data processing, ensured by data controller

- Personal data must be processed legally and fairly
- It must be collected for explicit and legitimate purposes and used accordingly
- It must be adequate, relevant and not excessive, updated when necessary
- Data controllers must ensure that data subjects can rectify, remove or block incorrect data about themselves
- Data that identifies individuals (personal data) must not be kept any longer than strictly necessary
- Data controllers must protect personal data against accidental or unlawful destruction, loss, alteration and disclosure, particularly when processing involves data transmission over networks. They shall implement the appropriate security measures
- These protection measures must ensure a level of protection appropriate to the data

Examples

- VO manager (organisation to which VO belongs is a data controller)
  - If VO designates more people that decides which and how data are processed, they’re joint controllers
- Identity Providers (IdPs) providing identity information data to Service Providers (SPs)
  - Organisation to which IdP belongs is a data controller
  - However, SPs are also data controllers (in most cases)
- Hub-and-spoke and mesh federations:
  - Mesh federations – no entity between IdP and SP
  - Hub-and-spoke – Proxy that acts both as an SP and an IdP, so it is both a data controller and a data processor

Cloud Computing

- Wide deployment of the cloud computing model for commercial services
- Examples for research sector – Helix Nebula and INDIGO DataCloud
- With cloud implementations there is personal data exchanged
- Responsibility, and liability, for fulfilling legal obligations remains with the client, for “private” and “public” cloud
- However, there may be difficulty in ensuring compliance
  - Example, multiple layers of dynamic service provision, with several providers
- Actions for creating code of conducts under way, still work in progress
- Existing policies may be sufficient, for the time being

Purpose of processing

- Controller decides the purpose (Article 5(b))
  - Information security
  - Monitoring
  - Capacity planning
  - Security, incident response
  - Invoicing

- All previously mentioned rules apply
  - Informing the user
  - Data minimisation
  - Data integrity

Legal ground from processing of personal data

- Six distinct grounds for processing (Article 6), with two most relevant for research use cases
  - Article 6.1(a) - User consent: “the data subject has given consent to the processing of his or her personal data for on or more specific purposes”
  - Article 6.1(f) – Legitimate interest: “processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.”
- GDPR has strict requirements for user consent
  - Freely and clearly given, revocation at any time, given for a limited purpose
  - However, if a researcher needs access to services to perform their job, is the consent really “freely given, specific, informed and unambiguous indication of the data subject’s wishes” (from Article 4(11))
  - Data protection by design and by default (Article 25), data minimisation
- Relying only on the user consent is not good enough

Legal ground from processing of personal data

- Six distinct grounds for processing (Article 6), with two most relevant for research use cases
  - Article 6.1(a) - User consent: “the data subject has given consent to the processing of his or her personal data for on or more specific purposes”
  - Article 6.1(f) – Legitimate interest: “processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.”
- GDPR has strict requirements for user consent
  - Freely and clearly given, revocation at any time, given for a limited purpose
  - However, if a researcher needs access to services to perform their job, is the consent really “freely given, specific, informed and unambiguous indication of the data subject’s wishes” (from Article 4(11))
  - Data protection by design and by default (Article 25), data minimisation
- Relying only on the user consent is not good enough

→ Legitimate interest

Release of personal data to third parties

- Article 6.1(f) “in effect requires a balancing of the legitimate interests of the controller, or any third parties to whom the data are disclosed, against the interests or fundamental rights of the data subject” (Opinion 06/2014, Article 29 Data Protection Working Party)
- Any sharing of personal data within the infrastructure must be considered “release” of personal data to a third party (this makes almost all entities data controllers)
  - Log files, accounting records, community membership information
- Release of personal data is permitted, under certain safeguards:
  - Informing the user
  - Performing a balancing test, i.e. the stronger the legitimate interest and the less harm the processing does to the interest of the data subject, the greater the likelihood the activity will be lawful
  - Examples: attribute release has a positive impact on a user (accessing the service); security incidents are legitimate interests of a service provider
  - Legitimate interest should NOT be treated as a last resort, nor be automatically applied

Release of personal data outside the EU

- Infrastructure operation almost invariably involves transfer of personal data outside EU
- Recognized countries (Andorra, Isle of Man, Argentina, etc.) – not many, US is not recognized
- Conditions for transfer:
  - User needs to be informed of the safeguards
  - Custom exchange model needs an explicit approval from a data protection authority
  - “Binding Corporate Rules” (BCR)
  - “Standard Data Protection Clauses”
  - “Approved Codes of Conduct”
- Explicit user consent – an option, requirements may not make it suitable
- “Safe Harbor” – invalidated by ECJ in 2015
- “EU-US Privacy Shield” – not applicable to research environment

Codes of conduct

- Potentially viable, especially with the GDPR
- Need to be very specific, need to be approved → ongoing effort in REFEDS and GÉANT
- Potential difficulties in distributed infrastructures:
  - Presenting unified view to users, communities, resource managements….
  - lack of signed contracts
  - no ability to sign legally binding documents
  - lack of central hub

- However, if structured in way that the new Code of Conduct model is applicable, “safer” to use than other models → further work ongoing
- Good potential example – Cloud Select Industry Group (C-SIG)

Standard data protection clauses (Model Contracts - MC)

- Controller-to-Controller or Controller-to-Processor
- Needs a signed agreement, legally binding
- Hard to implement for infrastructures
  - Many independent entities
  - Hard to scale
  - Contracts “by proxy” not legally acceptable
- It is suitable in certain contexts
  - Procurement of commercial cloud services
- If the structure permits, perfectly adequate solution
- Use and application is standard

Binding Corporate Rules

- Drafted by the organisation, unlike MC
- Legally binding
- Approved by the appropriate data protection authority (DPA)
- Must contain the following elements:
  - Privacy principles including transparency, data quality, security of the data, etc.
  - Tools to measure the effectiveness, for example, audits, training, complaint handling, etc.
  - Clear statement that the BCR are binding
- Adopting procedure:
  - Organisation communicates with a lead data protection authority only
  - Organisation drafts the BCR addressing WP29 criteria
  - Lead authority consults relevant DPAs, i.e. DPAs not covered by EU law
  - EU cooperation closes
  - Once BCR is considered final, organisation may start using it
- BCR is drafted by and applicable to a legal organisation → most infrastructures are not

BCR-like model

- Infrastructures are not legal organisations → “real” BCR not applicable
- Following the model and guidance of the BCR may be still applicable
- Drafting the Use Policy, following the eight principles and WP29 guidance
- Relatively lightweight framework
- Of course, assignment of risk between participants not legally enforceable, nor governed by legal documents
- Risk assessment
  - Potential harm to the user is very low
  - Processed personal data are not sensitive (usually mail, name, affiliation)
  - Such personal data are already public (publications, grants, etc)
  - Accounting data may already be open to consultation by persons with (legitimate) interest
  - Using “balancing test”, cost vs risk vs feasibility
  - “Say what we do, and do what we say”
- Providing a complete set of rules, limiting exchange of data to “inside” entities

Conclusion

- Exchange of personal data within EU – framework is already present
- Model contracts, BCR, code of conduct – release of data to non EU countries
- No clear winner, probably all should be used with modifications/caveats
- NOT BCR, but BCR-like approach
- One possible solution – Code of Conduct, Article 46.2(e) "an approved code of conduct … with binding and enforceable commitments of the controller … in the third country to apply the appropriate safeguards …" without requiring any specific authorisation from a supervisory authority
  - Discussion still ongoing
  - REFEDS effort for attribute release outside EU
- More info:
  - https://aarc-project.eu/
  - https://aarc-project.eu/wp-content/uploads/2016/12/AARC-DNA3.5_Recommendations-for-Processing-Personal-Data_2016_11_07_v4_DG.pdf

Thank you

Any Questions?

uros.stevanovic@kit.edu

https://aarc-project.eu

The work leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 653965 (AARC).