«Protection of Personal Data in High Performance Computing Platform for Scientific Research Purposes»

LUDOVICA PASERI, SÉBASTIEN VARRETTE and PASCAL BOUVRY

June 17th, 2021
OUTLINE

1. Research Aim and Methodology
2. Concepts Overview
   - High Performance Computing (HPC)
   - Open Science paradigm
3. Data Protection and Scientific Research
4. Case Study: ULHPC @ University of Luxembourg
5. Case Study Outcomes
6. Conclusions
1. RESEARCH AIM and METHODOLOGY

How can an HPC platform used for scientific research purposes be compliant with the data protection framework, ensuring the tenets of the Open Science paradigm?

2. CONCEPTS OVERVIEW

High Performance Computing (HPC)

- **High Performance Computing (HPC)**
  - practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation
  - **Objective**: solve large problems in science, engineering, or business.

Uni.lu HPC Facility Usage (2020)

Ex. in DIGITAL HISTORY:
- historical data and governmental reports
- personal records of Luxembourg’s citizen activity WW1/2

Ex. in COMPUTER SCIENCE:
- Digital Twins for Industry 4.0
- Digital Surveillance of Luxembourg’s Tier-IV datacentres by autonomous drone swarms

**THE OPEN SCIENCE PARADIGM**

- Openness of every phase of research cycle;
- More open, global & collaborative research;
- Social engagement.

**European level:**

Brussels, 19.2.2020
COM(2020) 66 final

**International level:**

On 27 October 2020:
**Joint appeal for Open Science**

"As open as possible, as closed as necessary".

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066


3. DATA PROTECTION and SCIENTIFIC RESEARCH

DEFINITION OF SCIENTIFIC RESEARCH (Recital 159):

“For the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research. In addition, it should take into account the Union’s objective under Article 179(1) TFEU of achieving a European Research Area. Scientific research purposes should also include studies conducted in the public interest in the area of public health”.

a space of free movement for scientific research across Europe (circulation of knowledge and researchers)

“5th European freedom”

the freedom of free movement of knowledge and researchers across Europe established in 2008, in the Council of Ljubljana

### SPECIFICITIES OF PROCESSING FOR SCIENTIFIC RESEARCH PURPOSES

<table>
<thead>
<tr>
<th>Art.</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>89</td>
<td>Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes</td>
</tr>
<tr>
<td>5(1)b</td>
<td>Purpose limitation</td>
</tr>
<tr>
<td>5(1)e</td>
<td>Storage limitation</td>
</tr>
<tr>
<td>9(2)j</td>
<td>Processing of special categories of personal data</td>
</tr>
<tr>
<td>14(5)b</td>
<td>Duty to provide information about processing</td>
</tr>
<tr>
<td>17(3)d</td>
<td>Right to erasure</td>
</tr>
<tr>
<td>21(6)</td>
<td>Right to object</td>
</tr>
</tbody>
</table>

- **Art. 63**: possible derogation from certain rights, if they make it impossible or damage the purpose of processing;
- **Art. 64**: possibility of processing PARTICULAR CATEGORIES OF DATA (subject to compliance with “appropriate additional measures”);
- **Art. 65**: a list of “appropriate additional measures” that the controller must necessarily implement in the processing of data for scientific research purposes.

---

4. CASE STUDY: ULHPC @ UNIVERSITY OF LUXEMBOURG

L. PASERI, S. VARRETTE AND P. BOUVRY. "Protection of Personal Data in HPC Platform for Scientific Research Purposes" (APF 2021)
5. CASE STUDY OUTCOMES

CHALLENGES IDENTIFIED:

- The principle of transparency and related information duties
- The meta-principle of accountability
- The right of access by the data subjects
- The principle of data minimisation
- The security of processing (emphasis on data movement)

5. CASE STUDY OUTCOMES

SUGGESTIONS PROPOSED

**Organisational approach**
- Standardisation of procedures:
- **ACCOUNTABILITY**: internal act or guidelines, demonstrating (1) continuity; (2) awareness of choices and demonstrability as a **META-PRINCIPLE**
- Procedure for the **communication** of information pursuant to art. 13 and 14 GDPR;
- Procedure for handling any requests from data subjects for the exercise of rights.

**Technical approach**
- **DATA MINIMIZATION**: two-steps approach:
  1) I do not use personal data (!?);
  2) pseudonymization or anonymization of data; consider always the given case!
- **SECURITY and DATA MOVEMENT**: for authentication and authorizations: compartmentalization; GDPR security level < Information Security (ISO, etc).

**Cooperative approach**
- Similar issues in similar environments;
- approach able to take into account the specificities of the different realities: as a moment of **convergence**.

5. CONCLUSIONS

 Procedures need to be standardised
 Technical adaptations, accountability for choices
 Cooperative approach to similar issues

PERSPECTIVES / OPEN ISSUES:
• Anonymisation: is it enough? does it limit research? can synthetic data be a solution?
• Lack of technical solutions for HTC/HPC (ex. data movement auditing / control);
• on security, additional issues if HPC is connected with cloud computing.

Thank you very much for the attention!

LUDOVICA PASERI, ludovica.paseri2@unibo.it
SÉBASTIEN VARRETTE, sebastien.varrette@uni.lu
PASCAL BOUVRY, pascal.bouvry@uni.lu

https://hpc.uni.lu