Eindrapportage pilot – DCC voor Praktijkgericht onderzoek

Pilotgegevens	
Titel van pilot	Data in Privacygevoelig Praktijkgericht Onderzoek (DPPO)
Hogeschool	Avans University of Applied Sciences
Contactpersoon	Dr. Miaomiao Zhou
Emailadres	m.zhou1@avans.nl
Datum	03/16/2023

1 – Beschrijving van pilot

- *Geef een korte beschrijving van de pilot*
- *Wat was het doel van de pilot?*
- *Welk vraagstuk of probleem heb je aangepakt en hoe (doorlopen stappen/proces)?*

Every day, in different applied research domains, GDPR-sensitive data are being produced. To properly manage this data, not only special tooling is needed, but a GDPR-compliant data management infrastructure is highly necessary. Recently several Data Research Environments (DRE) have been made available via different service providers. It becomes a struggle for applied research data stewards to compare, select and introduce the services to researchers in different domains.

This project aims to compare the most commonly used DREs from selected providers on the aspects of GDPR-compliance. This is an explorative study to find a way to understand DRE services concerning GDPR sensitive data. The main stakeholders are data stewards from different applied research domains.

In this project we collected the main questions from the data stewards as:
1, what are the requirements from the research concerning the GDPR-sensitivity ?
2, what types of policies do the DRE services follow and how are they linked to GDPR sensitive data?
3, what details of services from these providers are needed to fulfill the needs of the research?

To answer the questions we designed 4 work packages to carry out the project in 3 steps.
WP1, Collection of GDPR-sensitive data examples/requirements from different work fields
WP2, Design of DRE service evaluation tools concerning GDPR sensitive data
WP3, Analysis of data management infrastructures from different resources
WP4, Project management and dissemination

2 – Resultaat & impact

- *Wat zijn de resultaten? Welk resultaat heb je bereikt? Wat is er klaar?*
- *Wat heb je geleerd?*
- *Hoe ga je hiermee verder?*
- *Wat is de impact op (het faciliteren) van praktijkgericht onderzoek?*

From WP1, Three sets of example data was collected together with their consent form templates. These sets were from 3 different study fields: 1, molecular study on inherited disease; 2, human behavior; 3, medical treatment practice. Besides, a simulated research participation consent was made by using the Avans general consent template with a user case of behavior disorder. From these data we extracted the needs for GDPR sensitive data concerning data management.

This resulted in the two tools we made in WP2. The tool "DPPO_Gap_Analysis_policy_Tool_Avans" focuses on assessing the details of service organization policy on GDPR sensitive data. The tool "DPPO_service_requirement_tool" helps to analyze the service contents concerning the requirements to take care of DGPR sensitive data. Both tools were constructed by conducting a life circle analysis of the GDPR sensitive data in WP1 by using existing/commercial and/or in-house tools. These two resulted tools are qualitative studies which need to be conducted via interviews. The Gap

Analysis tool has a simple scoring panel which helps to quantify the results, thus the analysis across platforms could be compared to each other to a certain extent.

From the third work package, two most well-known DRE service providers, SURF research cloud + drive, and Andrea were invited to take the interviews by using the tools designed in WP2. The results were analyzed to see if data stewards could gain enough insights into these two services while selecting DREs for their own research data. The conclusion was affirmative:

- by using the results of these analysis, the data stewards could understand the service details in a structured way
- by conducting the interviews, especially by using the DPPO_service_requirement_tool, the data stewards gain more insights of what detailed services they can expect. This concerns both the standard service and the tailored service.
- By using the DPPO_Gap_Analysis_policy_Tool the data stewards can construct a helicopter view on what's happening with the service, the results are transparent and stimulates discussion.

Thus we conclude that the two tools generated from this project are sufficient to provide preliminary aid to the data stewards and researchers looking for GDPR-compliant ICT infrastructures.

The tools are made by Excel. This is not very convenient for long term usage. Therefore, in the future, the tools should be made available as an online service. The same counts for analysis updates, because the policies and services of providers may update rapidly, it would be more user-friendly if the tools are made online, renewable and with proper backup plans.

This study brings a general impact to the data steward community. It has been a long-lasting wish from the community to have more help in ICT service evaluation. The tools and analytical routes we took in this project could be generalized for other purposes concerning data stewardship. The results of the project brings insight to the service providers too. This project helps them to improve services concerning the needs of the specific group of users.

A few stakeholder meetings with SURF and UMCU have been conducted in the past years to review the subject. New developments and pilots are being proposed during the meetings. We are satisfied with the continuation of the subject, this shows a positive impact to the work fields.

3 – Diensten

- *Welke SURF-diensten waren onderdeel van de pilot?*
- *Hoe zijn deze diensten ingezet of gebruikt binnen de pilot?*

In this project, SURF research cloud and research drive were used.

We received timely support from the service. Especially, because the type of analysis requests information from higher organizations, more support were provided by the management team in answering the inquiry forms. We sincerely appreciate this cooperation.

4 – Betrokkenen en inzet

- *Wie waren betrokken bij de pilot (rollen/functies)?*
- *Welk lectoraat(en) of kenniscentra waren betrokken bij de pilot?*
- *Welke inzet en expertise is vanuit SURF geleverd?*
- *Urenverantwoording graag aanleveren in de verstrekte Excelfile*

Involvement table Avans:

Name	Role	Position	Affiliation	Expertise
Miaomiao Zhou	Project leader	Associate professor	Center of expertise biobased economy	Bioinformatics, big data, AI and data stewardship
Leon van Ekeren	Researcher	Data steward	Learning Innovation Centre	Data stewardship
Antal Derene	Researcher	Data steward	Learning Innovation Centre	Digital infrastructure and Data stewardship
Sjef Ollers	Researcher	Data steward	Learning Innovation Centre	Data stewardship
Bas van der Heijden	Adviser	Privacy officer	Service, ICT and Facilitation	Legalisation of digital data privacy

The excel file has been submitted in Q1 2023. During this project, SURF delivered expertise in digital infrastructure setting and consultancy.

5 - Lessons learned

- *Zijn er lessons learned?*
- *Zijn er aandachtspunten of verbeterpunten?*

As a multidisciplinary project we planned the tasks to be carried out by a relatively diverse team within Avans. However, during the first phase of the project, we experienced unexpected member withdraw. Some was caused by diseases, some caused by career rearrangements. From the formation of the project till the conclusion, only M. Zhou and L. van Ekeren stayed as the original members. The consortium had to wait for long before the new members were hired. This led to management difficulties and delivery delay.

Hereby we would like to express our appreciation to the fund organization, with your understanding, support and extension, we could eventually complete the tasks as planned.

In the future, with similar projects, we will plan more parallel work packages to ensure efficiency. Furthermore, we will give a better project framing to avoid subjects outside of the scope. It would be helpful to avoid monotonic consortium too. By including members from other Universities of Applied Sciences, in case of Avans member withdrawing we could have more options to decrease the impact. Lastly, we hope to keep up this effective communication with SURF. Without the communication and support this project would never have been possible.

6 - Kennisdeling en disseminatie

- *Hoe zijn de resultaten van het pilot beschikbaar gesteld binnen de eigen hogeschool?*
- *In welke vorm is de opgedane kennis beschikbaar gesteld aan andere hogescholen?*

The results of the project are submitted to DATAVERSENL, access will be granted upon requests. The two tools created in this project are available for the community as long as they are only used for research purposes.

A few presentations have been given during the data steward (DCC-PO) gatherings. We hope to bring more impact by sharing the results, experiences and tools to the community by attending more congresses via WP4.