



Contents

| | |
|---|-----------|
| 1 Introduction | 3 |
| 1.1 Motivation | 3 |
| 1.2 Current situation | 4 |
| 1.3 Why is a national approach desirable? | 4 |
| 2 Deliverables plan | 5 |
| 2.1 Overall aim of the Data4lifesciences program – True North | 5 |
| 2.2 Goal | 5 |
| 2.3 Deliverables and approach | 5 |
| 2.4 Fundamental concept | 8 |
| 2.5 Organization | 8 |
| 3 Time schedule | 9 |
| 3.1 Long term (2020+) | 9 |
| 3.2 Medium term (2018-2020) | 9 |
| 3.3 Short Term (2017) | 9 |
| 4 Budget plan | 10 |
| 4.1 Resources/deliverables expected from or created in collaboration with other WPs | 10 |
| 5 Organization plan | 11 |
| 5.1 Project team | 11 |
| 5.2 Stakeholders | 11 |
| 6 Communication plan | 12 |
| 7 Risk plan | 13 |

1 Introduction

1.1 Motivation

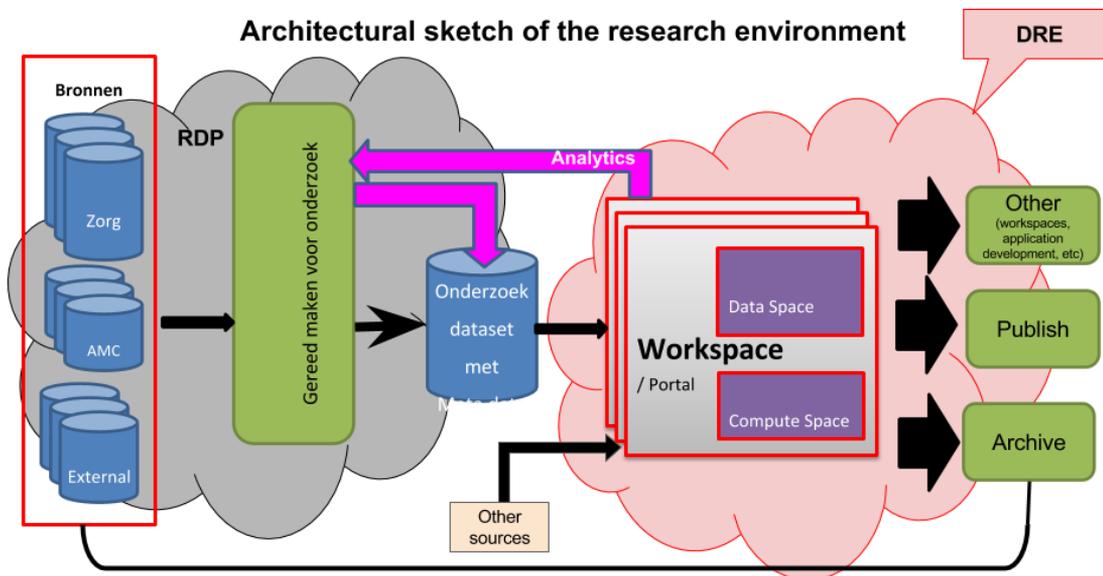
In 2015, the first version of the WP4 Data4lifesciences description of work was released. WP4 can be summarized as:

Empowering the biomedical research community to share, use, analyze and archive data in a sustainable way

-- CONNECTING RESEARCH

Looking from the architectural sketch of the research environment developed by Data4lifesciences WP2, the scope of WP4 is depicted in the red area and deals with three questions:

- How to make data available for research to the workspace of the researcher?
- What should a workspace look like?
- How to move data from a workspace to an archive and/or making it available for other (future) studies?



Source: Hans van de Berg (WP2)

Key drivers for Data4lifesciences WP4 as well as the Research ICT initiatives in the UMCs are:

- Ambition / vision to be an international research top institute requiring the ability to reuse and share data with collaboration partners.
- Risk and impact of law and regulations; “Wet op Bescherming Persoonsgegevens”
- The pressure from research funders and the general public to provide complete traceability of data, and full reproducibility of the results of studies (in particular in view of some recent years scientific fraud cases).

1.2 Current situation

This information has been obtained through work visits to all 8 UMCs and Maastricht Clinic that took place between April 7th and May 12th 2016. The key questions for these visits were:

- What activities are being or will be run regarding Research ICT?
- What were and are the learning experiences / challenges?
- How can the UMCs help one another and learn from each other?

Right now there are five initiatives that can be considered as (the development of) workspaces:

- The workspace from LifeLines in Groningen
- The virtual machines of Vumc
- The workspaces of RSRCH BV (deployed in Radboudumc)
- The workspaces of Surfsara
- The collaborative efforts of BBMRI-NL and CTMM TraIT

How to move data to and from a workspace is generally solved locally. Many studies rely on risky data carriers. Progress is being made to change this for the better.

Right now most researchers get little to no support to ensure sufficient security, privacy, traceability and reproducibility. Most of the researchers have not sufficient knowledge and experience with security and privacy challenges, nor the resources or tooling to address these challenges.

Even among WP4 members, there is a confusion of tongues. Definitions must be aligned in order to make progress.

The preferred and most successful way of implementation is through the showcase approach. The overall key challenge seems to be how to scale up. Pilots tend to be successful but seem to have a low adoption rate, or run into legal/organizational obstacles.

1.3 Why is a national approach desirable?

The scope of WP4 deals with three questions:

- How to make data available for research to the workspace of the researcher?
- What should a workspace look like?
- How to move data from a workspace to an archive and/or making it available for other (future) research?

From a UMC point of view it makes sense to address the WP4 scope on a national level, because:

- researchers have to work together and must get their data in an easy, cheap and relatively fast way = key challenge for all UMCs
- from a UMC perspective, the scope of WP4 is precompetitive
- (re)inventing the wheel 8+ times not only is more expensive, it also increases the odds of poor implementations
- one national standard simplifies making focused make/buy decisions

2 Deliverables plan

2.1 Overall aim of the Data4lifesciences program – True North

André Dekker(WP5 lead) has defined the overall aim, or True North, for the Data4lifesciences program. The wording has been adapted somewhat for WP4 to the following statement:

What all biomedical data and analytic methods/tools should be available
Who by every institute in the world
Why for every valid research or related question
When now and forevermore
Where in a scalable, secure, distributed environment
How taking into account FAIR principles
with a flexible SAAS infrastructure
with a user-friendly app store and user interface
with full protection of privacy of participants
with full control of the local owner (e.g. UMC)
and, if possible, made available open access

2.2 Goal

The overall aim translates into the goal of WP4:

Empowering the researcher community to share, use, analyze and archive data in a sustainable way

-- CONNECTING RESEARCH

2.3 Deliverables and approach

The activities, deliverables and approach have been derived from the work visits and in particular the [overall report of the work visits](#).

| # | Activity / Deliverable | Approach |
|---|------------------------|--|
| 1 | Communication plan | <ol style="list-style-type: none"> 1. Identify the audiences and their needs 2. Plan structurally the communication 3. Ad-hoc inform when an opportunity / need arises |
| 2 | Common WP4 glossary | <p>Collect concepts and definitions. Go through a peer-review with representatives of each UMC and other stakeholders (e.g. SURFsara, RSRCH BV). Four or more representatives in favor of a definition results in adaptation of the definition</p> <p>Changes will go through a change procedure.</p> <p>Merge with other WP initiatives like WP2.</p> |

| # | Activity / Deliverable | Approach |
|---|--|---|
| 3 | Requirements checklist and guidelines for workspaces | <ol style="list-style-type: none"> 1. Create a list with possible features based on grant and audit requirements for studies 2. Create a grouped and easy to use/understand checklist 3. Create guidelines for: <ol style="list-style-type: none"> a. workspace users (e.g. self-assessment) b. suppliers how to implement it and create specific reports based on chosen features of a study |
| 4 | Data governance for workspaces | <p>Based on #2, write according to #1 process a data governance for workspaces with respect to:</p> <ul style="list-style-type: none"> • Moving data to a workspace • Data usage in a workspace • Moving data from a workspace |
| 5 | DRE showcases | <p>Follow-up on the promise of the Dean O&O meeting to have one or two DRE showcases per UMC</p> <ul style="list-style-type: none"> • Align/facilitate/coordinate where opportune • Collect and distribute experiences & lessons-learned • Provide a stage for the showcases |
| 6 | Focus meetings/ workshop | <p>Organize a few focus DTL meetings and/or a workshop to share current best practices in the area of data integration and data analysis platforms. The output will be a whitepaper describing currently available solutions, and outlining the roadmap to a more integrated approach across the UMCs.</p> |
| 7 | Subgroup Workspaces | <p>Organizing meetings for those actively preparing the implementation and utilizing/improving workspaces to exchange ideas, learning experiences, and to align effort.</p> <p>Actively approach UMCs to get their active involvement</p> |
| 8 | Scale-out pilot | <p>Pilot to</p> <ul style="list-style-type: none"> • test if RSRCH DRE users can scale out to SURFsara HPC Cloud when more compute (8+ cores) is required; also addresses an important finding in the due-diligence • find out what it takes to have access to national e-Infrastructures |

| # | Activity / Deliverable | Approach |
|----|---------------------------|---|
| 9 | Elixir-NL | Contributing in writing up the grant proposal where Workspaces / the DRE plays an important role. |
| 10 | Other showcases | See for the deliverables and requirements the text below this table. #5 have a greater freedom to deviate from these deliverables and requirements. |

A (DRE) showcase¹ is a project that clearly demonstrates the functionality in a for the target group meaningful way.

- Supporting the science community using a federated approach
- Integral approach of research and care/application
- Integral data management in a single cohesive, compliant digital research environment
- 'Research Environment as a Service'
- Supporting standards and best practices
- Self-service and unburdening (compliance on the background)
- Flexible and scalable (also small studies with small budgets)

To prevent vendor locking, showcases are encouraged to choose from a list of different environments/solutions; provided more than one solution is available for their needs.

The deliverables of a showcase are:

- Set of requirements / user stories and why they are of value to whom
- Working product / service
- Wrap up document containing
 - best practice on how to implement
 - lessons learned
 - product/service backlog
 - validation of scalability & usability
 - validation of requirements like security, compliance & privacy

All the above documentation will be made publicly available unless NDA is applicable. When an NDA is applicable, a high level summary document will be made available. This will include a demonstration movie of the working product / service.

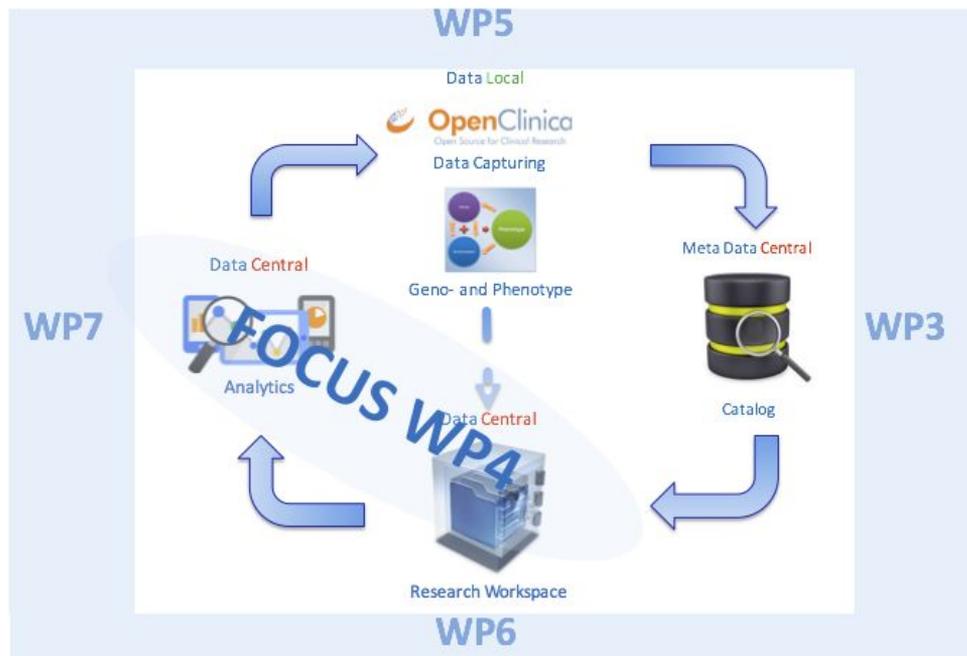
A showcase must comply with:

- Reference Architectures like [Reference Architectures for Workspaces](#)
- [NFU Terms & Definitions](#)

¹ Source: [NFU Terms & Definitions](#)

2.4 Fundamental concept

A central data solution is one of the possibilities, but the federated solution appears to be the emerging trend in the long run. In practice most likely a combination of central, federal and a hybrid will be needed to address the specific needs of studies now and in the coming years. Sharing data will take place in a secure environment after being harmonized and, if necessary, anonymized through a TTP.



2.5 Organization

- Every UMC has provided a contact person for WP4
- The project leaders of WP4 keep close contact with the other WPs
- When opportune to collaborate with other Work Packages, people of those Work Packages will be invited to participate and provide a liaison role
- National coordination exists between the local UMC programs to implement the required interfaces to the national infrastructure. Collaboration with the following entities is in place:
 - o CTMM/TraIT for integration of clinical, imaging and experimental data for the support of translational research.
 - o BMMRI-NL and PSI for national coordination and collection of biobank data.
 - o DTL Data for data stewardship and synergies with other life sciences domains.

3 Time schedule

3.1 Long term (2020+)

By 2020 the program's True North will be realized: a large scale facility for sharing and analysis of biomedical research data in the Netherlands suited for international collaboration.

3.2 Medium term (2018-2020)

- Link up with other research infrastructure related projects creating synergy
 - national level BBMRI-NL2.0
 - LifeLines, developed in close collaboration with the UMCG
 - Vancis Research Environment as a Service, a national service currently developed by Vancis/Aridhia
 - OpenEarth, developed outside the field of healthcare, an initiative of the 3 TU's
- Obtain structural and innovation budget(s) for (parts of) the WP4 research infrastructure.

3.3 Short Term (2017)

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---------------------------|---------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Update Project plan 2017 | | | | | | | | | | | | |
| 1) Communication plan | | | | | | | | | | | | |
| Project plan 2018 | | | | | | | | | | | | |
| 2) Common NFU glossary | | | | | | | | | | | | |
| 3) Requirements checklist | | | | | | | | | | | | |
| 4) Data governance WS | | | | | | | | | | | | |
| 5) Dean O&O showcases | To be planned | | | | | | | | | | | |
| 6) Focus meetings | To be planned | | | | | | | | | | | |
| 8) Scale-out pilot | | | | | | | | | | | | |
| 9) ELIXIR-NL | | | | | | | | | | | | |
| 10) Other showcases | To be planned | | | | | | | | | | | |
| WP4 team meetings | | | | | | | | | | | | |
| 7) WP4 subteam meetings | | | | | | | | | | | | |
| Mentor meeting | | | | | | | | | | | | |

4 Budget plan

4.1 Resources/deliverables expected from or created in collaboration with other WPs

The following resources, deliverables can be drawn from, or will be (co-)created in close collaboration with other work packages within the program

| Personnel |
|---|
| Project support for 2017 |
| Resources for running the showcases |
| Creating one reference architecture for the whole life cycle of data / study to be used by all WPs |
| Aligning and merging all glossaries into one maintained for everybody accessible glossary, and ensuring everybody knows and has easy access to this glossary (especially outside the WP community) |
| Data stewardship guidelines (WP1) Guidelines that match and can be used/applied within the standardization and validation element |
| Discovery and access to data (WP3) The proposed model catalog has an important role in the workspaces, collaboration between WP3 and WP4 is required. |
| Using clinical data for research (WP5) Extraction of data from clinical sources, metadating and harmonization. |
| Good research practice (WP6) Provides the necessary conditions for multicenter research, collaboration is needed to ensure the deployment of tools & procedures; examples: <ul style="list-style-type: none"> - Toolbox for documents that can be approached via the workspace - National TTP service to be implemented through the workspace |
| Facilities for high-throughput data processing (WP7) Aligned with SAAS solution and workspace solutions |
| Coordinate access to experts and support (WP9) Regularly meeting between WP9 & WP4 to assess where it is opportune to work together as opportunities arise from Workspace progress or needs uncovered in UMCs by WP9. |
| Materials & Services |
| None foreseen |

5 Organization plan

5.1 Project team

| Name | Home Institute/Project | Project Role |
|---------------------|------------------------|-------------------------------------|
| Arnoud van der Maas | RadboudMC | Lead |
| Bert van Ooijen | Erasmus MC | UMC representative |
| Erik Flikkenschild | LUMC | UMC representative |
| Ethan den Boer | Erasmus MC | UMC representative |
| Hans Brouwers | RUG | Representative |
| Hans van den Berg | AMC | UMC representative |
| Harry Pijl | UMCU | Lead |
| Igor Schoonbrood | MUMC | UMC representative |
| Irene Nooren | SURFsara | Representative |
| Morris Swertz | UMCG | Linking pin WP3 |
| Patrick Lubbers | NKI | Representative |
| Paula Jansen | UMC Utrecht | Linking pin WP1 |
| Petra Overveld | LUMC | UMC representative |
| Rob Bieringa | CIT | Representative |
| Rob Cornelisse | LUMC | Representative Medical Intelligence |
| Ronald Schijndel | VUMC | UMC representative |
| Sanne Soer | Maastru | UMC representative |
| Stefan van Aalst | flow4U | Assistant-to Leads |

5.2 Stakeholders

| Contact | Stake |
|-------------------|----------------------|
| André Dekker | MUMC |
| Martin Kalshoven | LUMC |
| Joost van Kempen | Radboudumc |
| Hans van den Berg | AMC |
| Jan-Willem Boiten | Program manager D4LS |
| Jeroen Beliën | WP2 |
| | |
| Gerrit Meijer | NKI Mentor |
| | |
| | |
| | |

6 Communication plan

Information on the progress and results of Data4lifesciences WP4 will be actively and openly communicated within the Data4lifesciences operational board (and program committee when relevant), related stakeholders/partners/projects like BBMRI-NL, TraIT as well as the wider community. This includes providing access to work and finalized documents.

Given the number of stakeholders and information needs, the communication plan is now a separate deliverable of this plan with a target date for March 2017.

7 Risk plan

Below a risk matrix is given which identifies the major risks as can be seen at this moment in the project. The risk plan is to formally review the risk list every 6 months so that new risks can be added, hazard can be re-estimated and actions be taken.

| Risk description | Probability | Impact | Hazard (P*I) | Action |
|---|-------------|--------|--------------|--|
| 1. The in kind contribution is usually placed on top of the normal workload | 9 | 8 | 72 | Ask for specific timeslot(s) per month dedicated for WP4 work other than WP4 meetings |
| 2. Finding studies that can act as a showcase ** | 9 | 9 | 81 | 1) Create criteria that a study has to meet 2) Create an offer 3) Casting a wide net by project leaders and project members to actively search out |
| 3. Insufficient funding to support the showcases | 9 | 9 | 81 | Get a support budget for each showcase |

** The study that wants to be a showcase, should not see the contribution of WP4 as a solution to funding nor as the vehicle to get a go on the study.