

IMI2 821520 - ConcePTION

ConcePTION

**WP7 – Information and data
governance, ethics, technology
and data catalogue and quality
support**

D7.2 Description of the operational platform for data sharing and task management system

Lead contributor	Miriam Sturkenboom (UMCU)
	Email: m.c.j.sturkenboom@umcutrecht.nl
Other contributors	Rosa Gini (ARS)
	Marco Moscatelli (CNR-ITB)
	Petr Holub (BBMRI-ERIC)
	Caitlin Dodd (UMCU)
	Kirsten Lum (J&J)
	Marianne Cunnington (GSK)
	Morris Swertz (UMCG)
	Section 3: WP1 data strategy: Anja Geldhof (J&J), Helen Dolk (ULST), Maria Loane (ULST)
	Section 3: WP2 data strategy: Laura Yates, David Lewis, Eugene van Puijenbroek, Bernke te Winkel
	Section 3: WP3 data strategy: Isabelle Huys (KUL), Michelle
	Section 3: WP4 data strategy: Theresa Rosario-Jansen, Mats Hansson
	Section 3: WP7: data strategy: Miriam Sturkenboom (UMCU), Marianne Cunnington (GSK)

Due date	30-09-2019
Delivery date	29-6-2020
Deliverable type	R
Dissemination level	PU

Description of Work	Version	Date
First draft section 3 (TMS) (Rosa Gini, ARS)	V0.2	16 September 2019
Incorporated comments by Marco Moscatelli (CNR-ITB) and Peter Holub /BBMRI-ERIC) on section 3	V0.3	2 October 2019
Miriam Sturkenboom, First draft Sections 1,2,4, edits section 3	V0.4	31-10-2019
Caitlin Dodd, edits, additions section 5	V0.5	21-01-2020
Caitlin Dodd, edits, additions section 5	V0.6	26-03-2020
Miriam Sturkenboom, summary, review, lay-out and finalization section 5	V0.7	April 15, 2020

Document History

Version	Date	Description	Non-contributor reviewers (if applicable)
V0.7	April 24 2020	Draft	Helen Dolk (ULST) Maria Loane (ULST), Morris Swertz (UMCG), Anna Pierini (CNR-IFC) Helle Wallach Kildemoes (UOSL), Chuntao Wu (Sanofi), Hildrun Sundseth (EIWH), Christine Damase-Michel (CHUT) Helle Kieler (KI), Sue Jordan (USWAN), Anja Geldhof (JAN)
V1.0	June 29, 2020	Final Version	Miriam Sturkenboom (UMCU)

Table of contents

Document History	2
Table of contents	3
Abbreviations	5
Partners	6
Summary	7
1. Introduction and background	9
2. Technical infrastructure	11
2.1 Catalogue	11
2.2 Task management system	12
2.3 Digital Research Environment	12
3. Task Management System	13
3.1 Background	13
3.2 Requirements	13
3.3 Review of available solutions	14
3.3.1 TASKA	14
3.3.2 PopMedNet ^{TN}	14
3.3.3 Request Tracker for BBMRI-ERIC	14
3.3.4 Comparison of different task management systems	15
3.4 Design, development and test of selected task management systems	15
4. Work Package 1, 2, 3, 4, 5, 7 needs	18
4.1 Work package 1	18
4.2 Work package 2	20
4.3 Work package 3	21
4.4 Work package 4	23
4.5 Work package 5	25
4.6 Workpackage 7	25
5. Remote Research Environment (platform) requirements & testing	28
5.1 Requirements	28
5.1.1 Access to the remote research platform	28
5.1.2 Private workspaces within the remote data platform	28
Figure 9: Outline of proposed folder structure on the platform	29
5.1.3 Analytical tools available on the remote data platform	29
5.1.4 Upload	29
5.1.5 Download	29
5.2 Review of Potential Solutions	30
5.2.1 Virtual Research Workspace (University Medical Center Groningen-UMCG)	30
5.2.2 anDREa Azure DRE (UMC Utrecht, Radboud University and Erasmus Medical Center)	31
5.2.3 UKSeRP UK Secure eResearch Platform (Swansea University Medical School)	33
5.3 Testing of the platforms	34
5.3.1 Features to test	34
5.3.2 Testing status	35
5.3.3 Testing results	36
5.4 Choice of RRE/platform	40
6. References	41
Instructions to manage a task	42

in the ConcePTION Task Management System.....42

Abbreviations

CDM:	Common data model
DAP:	Data access provider
ELSI:	Ethical Legal and Social Implications
ETL:	Extraction transform load
ENTIS:	European Network of Teratology Information Services
FAERS:	FDA Adverse Event Reporting System (FAERS)
FAIR:	Findable, Accessible, Interoperable and Re-usable
GDPR:	General Data Protection Regulation
ISO:	International Organization for Standardization
MAH:	Marketing authorization holder
MS:	Multiple Sclerosis
OMOP	Observational Medical Outcomes Partnership
RA:	Rheumatoid Arthritis
RRE:	Remote research environment
RT:	Request tracker
SAP	Statistical Analysis Plan
SFTP:	secure file transfer protocol
SLE:	Systemic lupus erythematosus
TMS:	Task Management System
VAERS:	Vaccine Adverse Event Reporting System
VRW:	Virtual Research Workspace
WP:	Work-package

Partners

1. Universitair Medisch Centrum Utrecht (UMCU; Academic; NL)	27. Stichting ENTIS - European Network Teratology Information Services Foundation (ENTIS; Association; NL)
2. University of Ulster (ULST; Academic; UK)	28. University of Manchester (UMAN; Academic; UK)
3. Uppsala Universitet (UPPS; Academic; SE)	29. Stockholm County Council Health and Medical Care Administration (SCC; Public Health; SE)
4. BIOBANKS AND BIOMOLECULAR RESOURCES RESEARCH INFRASTRUCTURE CONSORTIUM (BBMRI - ERIC; AU)	30. Elevate Health (Elevate; SME; NL)
5. Alma Mater Studiorum - Università di Bologna (UNIBO; Academic; IT)	31. National Institute for Health and Welfare (NIHW; Academic; FI)
6. The European Institute for Innovation through Health Data (i~HD; Research association; BE)	32. Consiglio Nazionale delle Ricerche – Institute of Clinical Physiology (CNR-IFC; IT)
7. Katholieke Universiteit Leuven (KUL; Academic; BE)	33. Fundacion Para El Fomento De La Investigacion Sanitaria Y Biomedica De La Comunitat Valenciana (FISABIO; Academic; ES)
8. Medicines and Healthcare products Regulatory Agency (MHRA; Regulator; UK)	34. ttopstart BV (TTOP; SME; NL)
9. Stichting Lareb (LAREB; Public Health; NL)	35. RIJKSINSTITUUT VOOR VOLKSGEZONDHEID EN MILIEU National Institute for Public Health and the Environment (RIVM; Public Health; NL)
10. Agenzia Regionale di Sanita - Tuscany (ARS; Public Health; NL)	36. The European Medicines Agency (EMA; Regulator; EU)
11. Karolinska Institutet (KI; Academic; SE)	37. Janssen Pharmaceutica NV (JAN; EFPIA, NL)
12. The Newcastle upon Tyne Hospitals NHS Foundation Trust (NUTH; Academic; UK)	38. Novartis Pharma AG (NVS; EFPIA, CH)
13. University of KwaZulu-Natal (UKZN; Academic; ZA)	39. Sanofi Aventis Recherche et Développement (SANOFI; EFPIA; FR)
14. UNIVERSITETET I OSLO (UOSL; Academic; NU)	40. Takeda Development Centre Europe Ltd (Takeda; EFPIA; UK)
15. St. George's, University of London (SGUL; Academic; UK)	41. GLAXOSMITHKLINE RESEARCH AND DEVELOPMENT LTD (GSK; EFPIA; UK)
16. European Institute for Women's Health Limited (EIWH; NGO; IE)	42. Eli Lilly and Company Limited (Lilly; EFPIA; UK)
17. Academisch Ziekenhuis Groningen (UMCG; Academic; NL)	43. Pfizer Ltd. (Pfizer; EFPIA; UK)
18. Orcion B.V. (ORC; SME; NL)	44. Bristol-Myers Squibb Company Corp (BMS; EFPIA; USA)
19. Institut National de la Santé Et de la Recherche Médicale (INSERM; Academic; FR)	45. MERCK KOMMANDITGESELLSCHAFT AUF AKTIEN (Merck; EFPIA; DE)
20. BioNotus (Bionotus; SME; BE)	46. Novo Nordisk A/S (Novo; EFPIA; DK)
21. Centre Hospitalier Universitaire de Toulouse (CHUT; Academic; FR)	47. UCB BIOPHARMA SPRL (UCB; EFPIA; BE)
22. Università degli Studi di Ferrara – University of Ferrara (FERR; Academic; IT)	48. AbbVie (AbbVie; EFPIA; USA)
23. Swansea University (USWAN; Academic; UK)	49. Ellegaard Göttingen Minipigs A/S (Ellegaard; EFPIA; DK)
24. UNIVERSITE DE GENEVE (University of Geneva) (UNIGE; Academic; CH)	50. Covance (Covance; EFPIA; UK)
25. European Forum for Good Clinical Practice (EFGCP; NGO; BE)	51. Teva Pharmaceutical Industries Ltd. (Teva; EFPIA; IL)
26. The Synergist ASBL (Synergist; NGO; BE)	

Summary

This deliverable is describing the architecture and tools that we plan to use in the ConcePTION ecosystem to efficiently, robustly and transparently generate evidence from a wide variety of data sources. We describe 5 different steps that can be taken to conduct a study using the ConcePTION platform (remote research environment) and tools

Figure 1 (below) describes how the different architectural pieces that WP7 has created fit together to allow different types of data to be transformed into evidence in the ConcePTION ecosystem. Each of the individual pieces and requirements are described.

Step 1

Based on a study outline a study team starts by searching for potential data access providers (DAP) that have adequate data using the catalogue. The catalogue will contain meta-data, and some quality indicators. Through the data catalogue negotiator, conditions for access can be agreed between the study team and the DAP in a private manner.

Step 2

The study team can use the following documentation from the ConcePTION toolbox to develop a study protocol:

- Interviews with DAP
- Common data model for health care data and surveillance data
- ETL process by DAP
- Data dictionaries from DAP
- Event forms for specific events
- Experience to extract events from DAPs
- Statistical analyses plan
- Coding conventions

Step 3

The study team can use the ConcePTION task management system to send documents, ask and monitor the execution of certain tasks that are required from DAPs or other study team members for the study.

Step 4

DAPs are responsible for

- Review of protocol
- Access to data
- Providing access to data for quality indicators
- Obtaining of governance approval for study protocol
- Review of statistical analysis plan
- Transforming the data into the common data model
- Updating ETL scripts
- Execution of R scripts
- Review of output

Step 5

DAPs upload the results to the Remote Research Environment (RRE), which can be accessed by the study team to further analyse and pool the results. DAPs and study team members will review

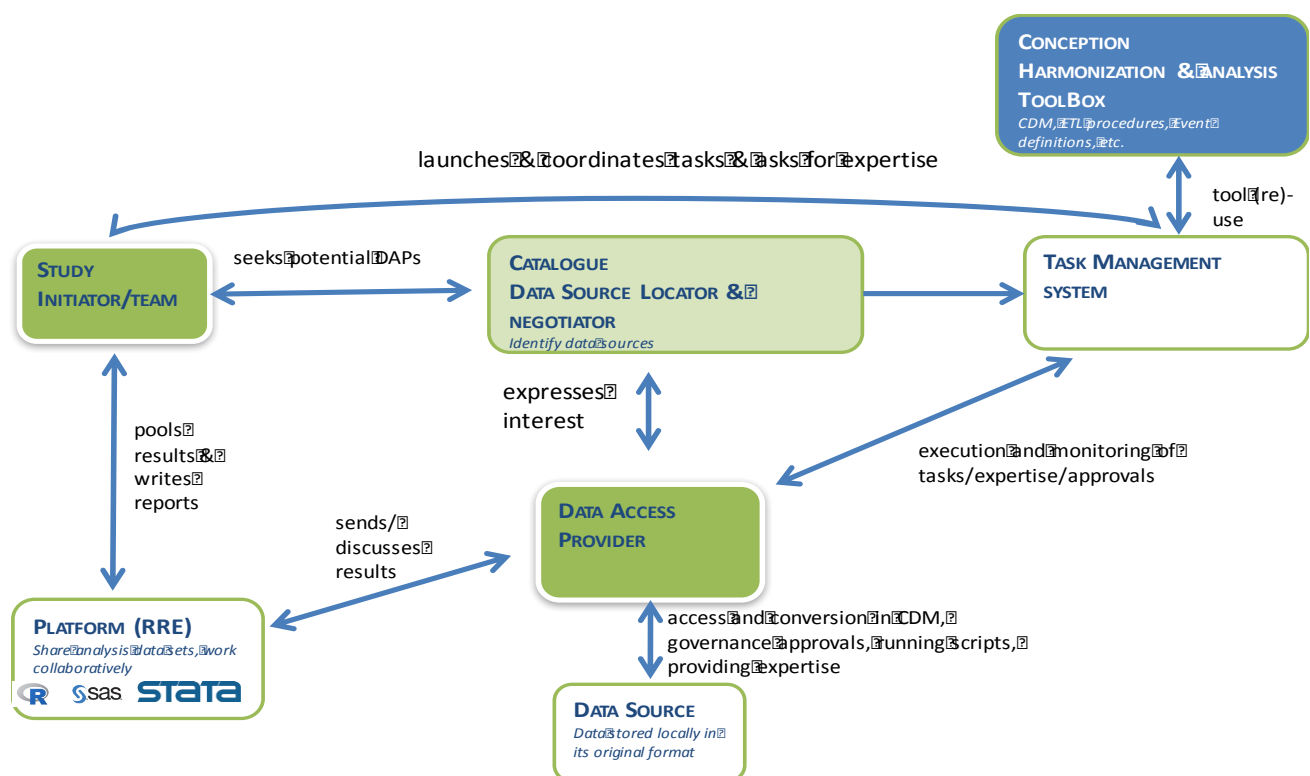


Figure 1: Interaction diagram to use ConcePTION architecture for a study

These steps can be used for each protocol-based pregnancy related study in ConcePTION (WP1, 2, 7) and for any additional studies (e.g. through sustainability task force) that may be done using the ecosystem. This deliverable describes the different pieces of architecture: their requirements and testing.

1. Introduction and background

ConcePTION aims to create an ecosystem for the rapid and robust generation of evidence on the use and safety of medications in pregnancy and during lactation, using both existing and newly generated real-world data.

The real-world data that will be transformed into evidence is coming from various sources

1) Existing data

- a. Pharmacovigilance data e.g. reports of adverse events in the mother or child following drug exposure prior to or during pregnancy, or during lactation. This data may also capture events following paternal exposure.
- b. Research data e.g. Drug exposure pregnancy registries aimed to recruit women who are exposed to specific drug(s) and are followed up prospectively
- c. Existing health care and health surveillance data
 - *Healthcare claims databases* – created for operational health care purposes and billing of costs on defined population that is followed over time (for example drug dispensing claims)
 - *General practice databases* – electronic medical records provided by General Practitioners (GPs) on defined population that is followed-up prospectively
 - *Demographic/population databases* – includes the population registry, resident registry, date of birth/death, migration
 - *Linkable Registries*: relevant outcome/exposure data collected for a specific purpose when they can be linked to an underlying population file that defines follow-up time
 - Medical Birth Registries
 - Specific Disease or outcomes surveillance registries e.g. congenital anomaly registries such as EUROCAT, cancer registries, infectious disease surveillance, death
 - Child surveillance databases –growth and developmental records as measured by community child health teams/public health nurses
 - Educational databases - created for operational education administration purposed for example school results, special educational needs and attendance
 - Registry of disability - created for insurance purposes and service delivery
 - Immunization registries

2) Newly collected data

- a. Human: human reported data on pregnancy (prospective)
- b. Milk: human milk samples
- c. Blood: human blood samples
- d. Animal: animal milk and blood data on milk transfer
- e. Human cells

The trajectory and steps of conversion of data into evidence is depicted graphically in the following graphic (figure 2).

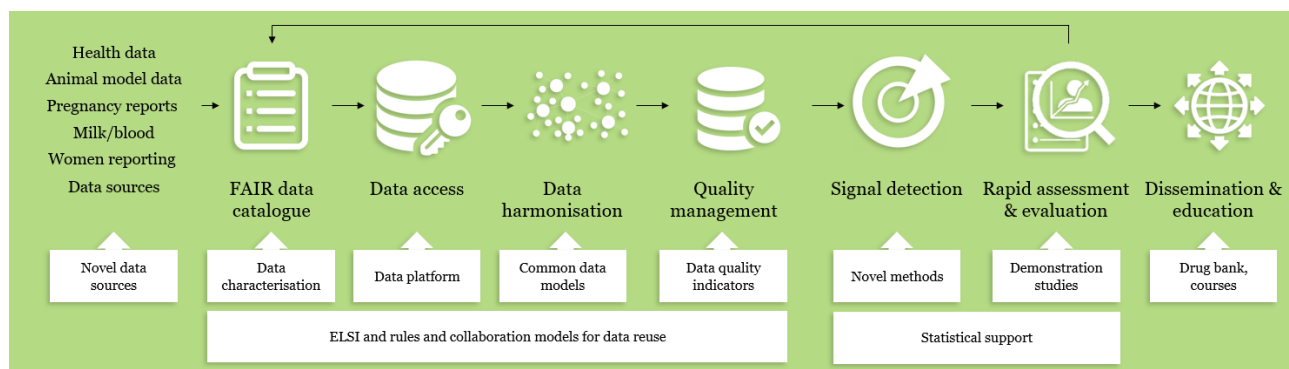


Figure 2: Central to the ConcePTION ecosystem is a dedicated workflow that allows for identification, verification and transformation of data into evidence

The ConcePTION ecosystem relies on 6 different dimensions:

1. The societal and ethical dimension
2. The technical infrastructure
3. The data dimension
4. The evidence generation (analytical) dimension
5. The communication & evidence dissemination dimension
6. The sustainability dimensions

In this deliverable we discuss the technical infrastructure that will support the management and generation of evidence from different sources of data. Before we do this, we will describe the data dimension and approach.

The key principles in ConcePTION are that we work as a distributed data system, and that analytics go to the data rather than the data to the analytics.

2. Technical infrastructure

The information technology architecture that is provided by WP7 to the ConcePTION ecosystem has three key components (see figure 3)

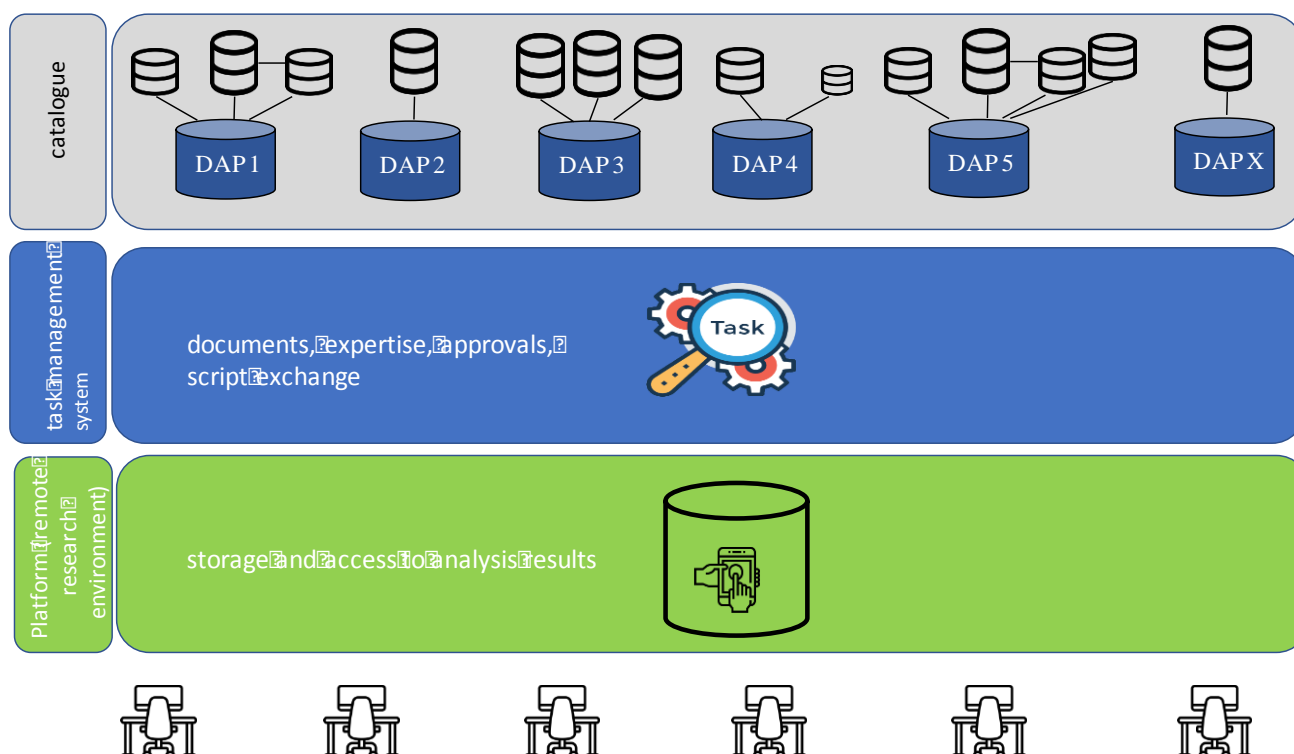


Figure 3: Overview of the three IT components of the ConcePTION architecture (from top to bottom: catalogue, task management system and platform, and below the users of the results)

2.1 Catalogue

A FAIR data catalogue is a meta-data management tool designed to contain searchable meta-data describing organizations that can provide access to specific data sources.

The catalogue will comprise:

- Essential meta-data descriptions of each organization that can provide access and the data sources they can access.
- Variable and data harmonization information including a searchable interface for extract, transform, and load (ETL) procedures designed by DAPs and the data dictionaries underlying these ETL procedures for each data source
- Dashboard for querying of data characterization output
- A private negotiation service to arrange data access

The catalogue requirements and details are described in more detail in D7.1 and the first prototype is described in D7.6 and will not be further described in this deliverable.

2.2 Task management system

Task management is the process of managing a task related to specific projects through its life cycle. The ConcePTION evidence generation work flows will require extensive task management to work with the many different data access providers for data characterization and demonstration studies. A detailed description of the task management system that we use in ConcePTION is provided in section 3.

The task management system supports the following activities that are taken by WP7 with each of the DAPs (both those participating in WP7, 1, or 2) in the process of transparent and systematic evidence generation:

Provision of information to set up the site to be a potential ConcePTION DAP

- 1) Request to provide information on the characteristics of the data sources that the organization can access for the conduct of studies on drug use and effects during pregnancy/lactation (so called “meta-data”) to be included in the ConcePTION catalogue
- 2) Retrieval of the data dictionaries (table structures and variables) of the data sources that the organization can access for the conduct of studies on drug use and effects during pregnancy/lactation (so called “meta-data”) to be included in the ConcePTION catalogue
- 3) Request to participate in an interview on data source structure
- 4) Request to provide documentation on the governance and privacy requirements/assessments (of the data sources that the organization can access for the conduct of studies on drug use and effects during pregnancy/lactation

Actions related to studies

- 5) Requests to review study protocol and statistical analysis plan per study
- 6) Request input on DAP prior experience in extractions of specific events of interest
- 7) Request to submit protocols for governance approvals required to participate in specific studies
- 8) Request to review the output of the data characterization scripts and fit for purpose for specific use protocol.
- 9) Request to design the ETL to transform local data into the ConcePTION CDM, using the ConcePTION ETL templates
- 10) Request DAPs to implement the ETL required for the protocol into the ConcePTION CDM
- 11) Request to share and provide the ETL script for the study
- 12) Request to run R-scripts (which will be provided) to execute the protocol related data transformation as specified in the statistical analysis plan
- 13) Request to submit the output of the R-scripts to the ConcePTION platform and review
- 14) Request to approve the outputs for release in reports and communication

2.3 Digital Research Environment

In order to pool aggregated results from the R-scripts that are being sent to the DAPs to generate results, and to allow multiple partners and stakeholders to work together in a collaborative manner we will require the use of a remote research environment (RRE) with ISO certified and GDPR proof services for remote collaborations.

A priori requirements for this platform are as follows:

- Access to the application server only using two-factor authentication.

- Hosting of private workspaces for each study in the ConcePTION project accessible only by named individuals in the study team. Each will have its own secured area to share data and results and provide access through remote desk tops clients. Within each of these project folders, each DAP should have a folder into which they can upload results. Output of analysis scripts should be able to be uploaded through a secure file transfer protocol (SFTP)¹.
- The RRE should offer several analytical tools (R, SAS, and Stata), word processing software, and utilities and have adequate back-up and security processes.
- Downloading of files should be blocked for security and only final results can be exported after approval by an administrator.

To better select the platform prior to contracting we have requested the needs and the data management strategies of the various work packages (1-5, 7). These are described in section 4. We have identified different platforms that are available with partners and tested them, this is described in chapter 5.

3.Task Management System

3.1 Background

Many activities connected to evidence generation in ConcePTION will require interaction among a large number of actors belonging to different organizations. For this reason, it was deemed necessary to incorporate in the ConcePTION infrastructure a web interface that supports such interactions, namely, a task management system.

3.2 Requirements

During Month 3, the following list of requirements was created based on previous experience with the task management system developed in the IMI-EMIF project (Almeida, 2019).

Each user has a profile, identified by name, email address, and organization

- Notification system via email
- Profiles of task requestor and task owner dependant on task or project (group of tasks)
- Each task is defined and has a status (to be started, ongoing, partially completed, completed), tasks can be reopened if the result is to be updated
- For each user, a unique space is available
- Functionality of accepting or discussing a task with the requestor is available
- Communication referring to each task/project is available
- Links to tasks are available, where they are described in detail with instructions and deadlines
- Functionalities to upload documents associated with tasks (such as replies) and/or to mark the task as completed
- For task requestors functionalities are available:
 - to create tasks, with instructions and deadlines
 - to assign tasks to task-customized groups of task owners, and discuss with them individually

¹ In SFTP both the commands and data are encrypted in order to prevent passwords and other sensitive information from being transferred over the network. SFTP requires that the client user must be authenticated by the server and the data transfer must take place over a secure channel (Secure Shell-SSH)

- to visualize the status of each task/project

3.3 Review of available solutions

A recent paper contained an extensive review of available task management systems (Almeida, 2019), in the Description of Action (DoA) of ConcePTION the TASKA and PopMedNet solutions were described as options for task management. This was complemented during Month 3 with a discussion with partners in UMCU, GSK, BBMRI-ERIC, UMCg and ARS. Based on this, three solutions were analysed for possible adoption in ConcePTION:

- The TASKA system developed in IMI-EMIF
- The PopMedNetTM solution of the US Sentinel System
- The Request Tracker (RT) ticketing service for BBMRI-ERIC

3.3.1 TASKA

The TASKA system had been created during the IMI-EMIF project, an IMI project that was completed in 2017 and aimed to provide an infrastructure for data re-use for epidemiological studies in Europe. TASKA was developed by Aveiro University (Portugal). Correspondence between the requirements and the functionality of the last release of TASKA was assessed. Results showed that the TASKA prototype had been successfully used to support a study in EMIF, and it came with a useful graphical representation of the state of tasks and of studies based on several dependent tasks. However, the functionalities of communication between task requestors and task owners were still to be refined. Contact with the Aveiro University showed that after the end of the IMI-EMIF Project, development had not progressed and support had been discontinued. The system was released with a General Public License (GPL) and it was therefore potentially possible for a WP7 partner to revive it and further develop. Time to create a team to restart development and obtain a fully functional deployment was estimated to be long.

3.3.2 PopMedNetTM

PopMedNetTM (PMN) is an open source software platform developed by the Harvard Pilgrim Health Care Institute (HPHCI) and used by Sentinel, National Patient-Centred Clinical Research Network (PCORnet), the Biologics and Biosimilars Collective Intelligence Consortium (BBCIC), and other health data networks initiatives. Access to the website www.popmednet.org clarified that currently the requirements of ConcePTION are met by the system. Moreover, it was clear that an active community is maintaining a rich wiki with detailed documentation, and that there are at least two initiatives supported by PMN that are not coordinated by HPHCI. Due to the open source availability of the platform, direct costs for adoption are in principle null, however functionalities seem to come with a dependence on Microsoft technologies and it is unclear whether all of them are free of charge. Moreover, time to adaptation of a complex system to the needs of ConcePTION was estimated to be in the order of several months.

3.3.3 Request Tracker for BBMRI-ERIC

BBMRI-ERIC suggested that their Request Tracker (RT) system maintained by CNR-ITB (Italy) could be easily repurposed to meet the requirements of ConcePTION, potentially in a matter of days. The system is used for

operating the BBMRI-ERIC Helpdesk for various services. The system manages different queues to identify the responsible of the various services and allow a rapid and adequate support.

3.3.4 Comparison of different task management systems

Table 1 shows the results of the feasibility assessment that was conducted. As a result, at Month 4 repurposing of RT was chosen as the optimal solution for ConcePTION.

Table 1: Results of comparisons of feasibility between different task management systems

System	Native correspondence to requirement	Currently supported	Developers in ConcePTION	Direct cost associated with adoption	Time to adaptation
TASKA	Partial	No	No	None	Months
PopMedNet ^{TN}	Complete	Yes	No	Unknown	Months
BBMRI ticket service	Almost complete	Yes	Yes	None	Days

3.4 Design, development and test of selected task management systems

Once feasibility was assessed, a quick round of design, development and test of the repurposed RT was conducted. Different solutions were implemented to adapt the ticket system to a task system for ConcePTION. The dashboard was configured to highlight the received tasks and their due date for the responders. In the requestor dashboard the global request was also present.

An automatic pipeline was developed to generate the ‘children (sub)’ task and assign them to specific responders. A basic course of action was envisioned which was deemed sufficient to support the WP7 activities of Year 1:

Requestor of a task

- creates a new group of responders (and, as an option, a queue: queue= several tasks grouped together)
- creates a new task ('parent' task) in the 'management' queue and sets the custom fields referring to the new group (and queue, if any) and the deadline

then the course of action goes as follows:

- the deadline is inherited by the ‘children’ tasks
- all the responders receive a notification and can access the 'parent' task all the responders can comment
- all the responders can reply
- requestor can change deadline of a single responder
- all the responders can set their task to 'closed'
- all the responders can re-open the task
- requestor can view the status of all the ‘children’ tasks
- requestor can comment and reply to all the ‘children’ tasks

Testing

Three rounds of development and test were conducted during Month 5. The final system was deployed during Month 5, and presented to the DAPs during a round of conference calls.

In Figure 4 the login page is reproduced. Access can be obtained through WP7leads.

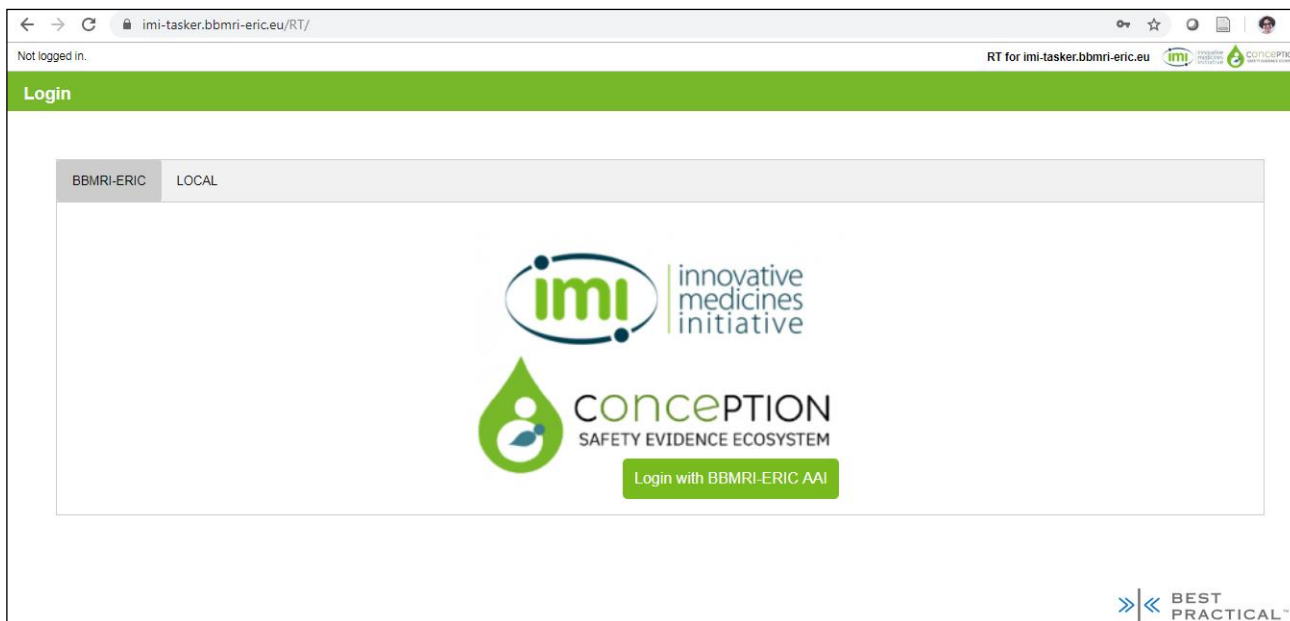


Figure 4. Login page for task management system

Once the user is in the system, the personal home page shows, in three different sections, the active tasks where the user is responder, the active tasks where the user is requestor (if a task has many responders, the ‘children’ tasks are grouped), and the closed tasks. Each task is associated with some information and can be clicked to open the task page. See Figure 5.

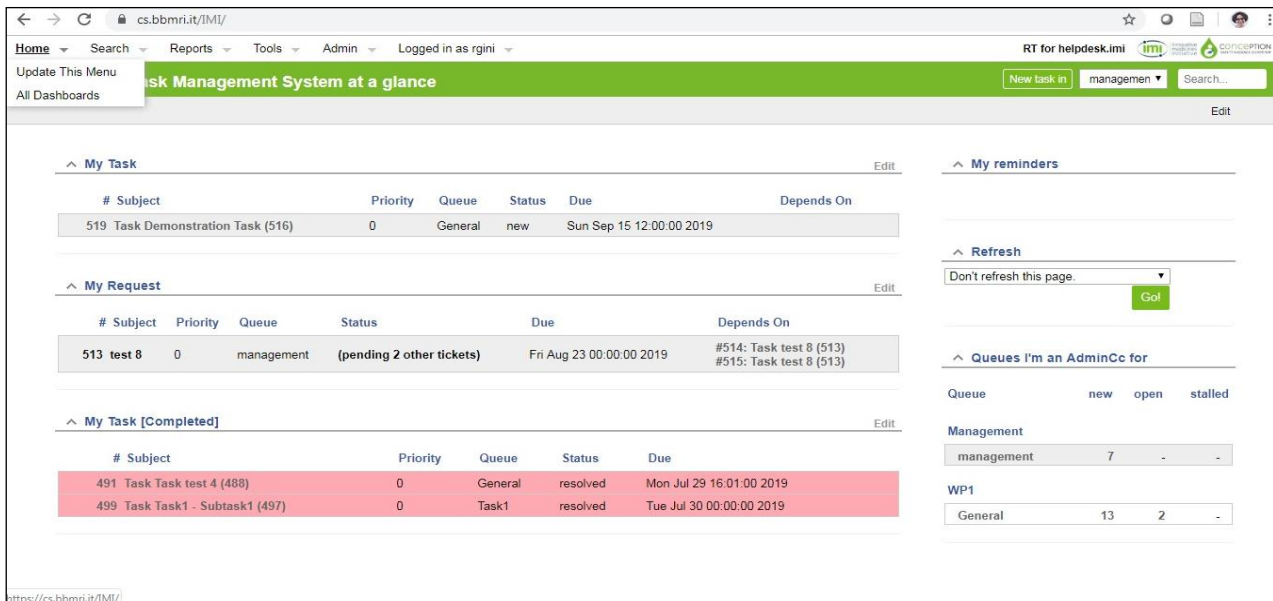


Figure 5. Personal home page

The task page contains several tabs. In the main tab the summary information is contained: identifier of requestor, deadline, attachment, history of communication. See Figure 6.

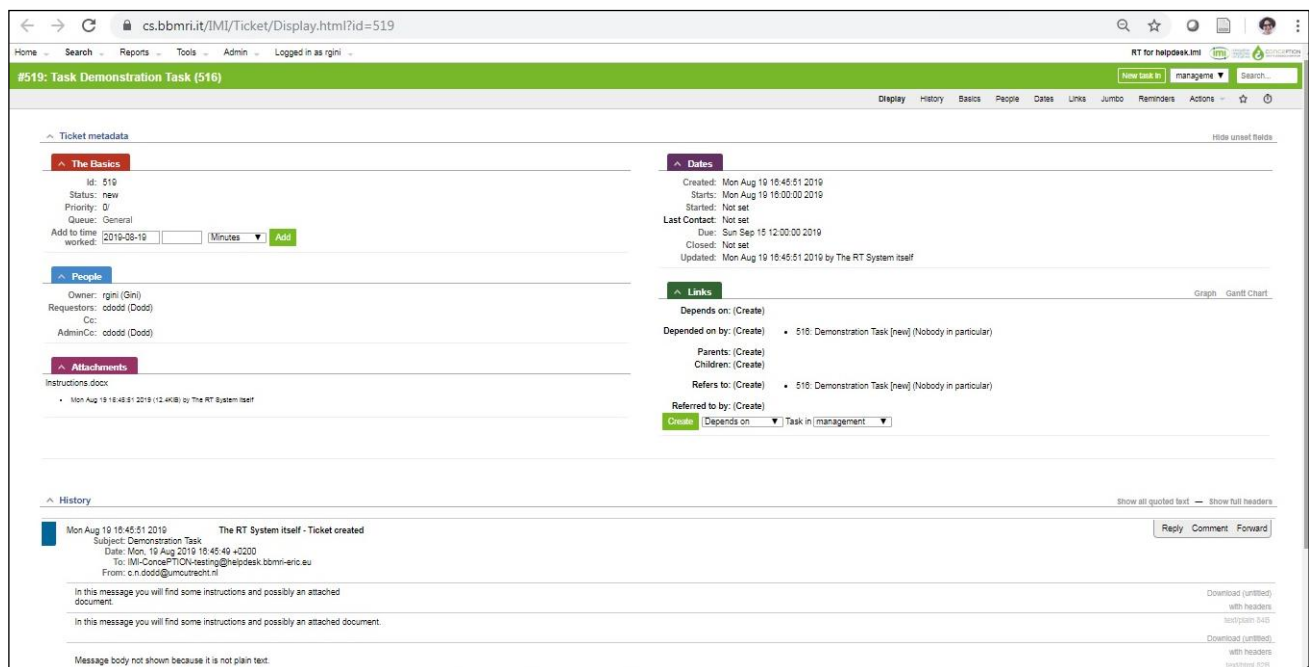


Figure 6. Main page of a task.

The other tabs of the task page contain more detail about the task. The task page is where all the interaction between requestor and responder is recorded, in particular messages and documents are stored. In the tab 'Action' it is possible to 'resolve' the task, that is, declare it is completed. Once a task is resolved it can be re-opened and interaction can be renewed.

4. Work Package 1, 2, 3, 4, 5, 7 needs

The ConcePTION workplan has been divided in 8 work packages (see figure 7). Work packages 1-4 will deal with generation of evidence from different sources of data, WP5 will summarize and disseminate the evidence, WP6 will discuss and get feedback on acceptability of the tools/solutions from stakeholders, and WP7 is providing the ethical, IT system, and quality dimensions. WP8 deals with project management and coordination. Whereas all WP will deal with some type of data, we focus in this section on the WPs that will deal with transformation of original health/animal/ data into evidence.

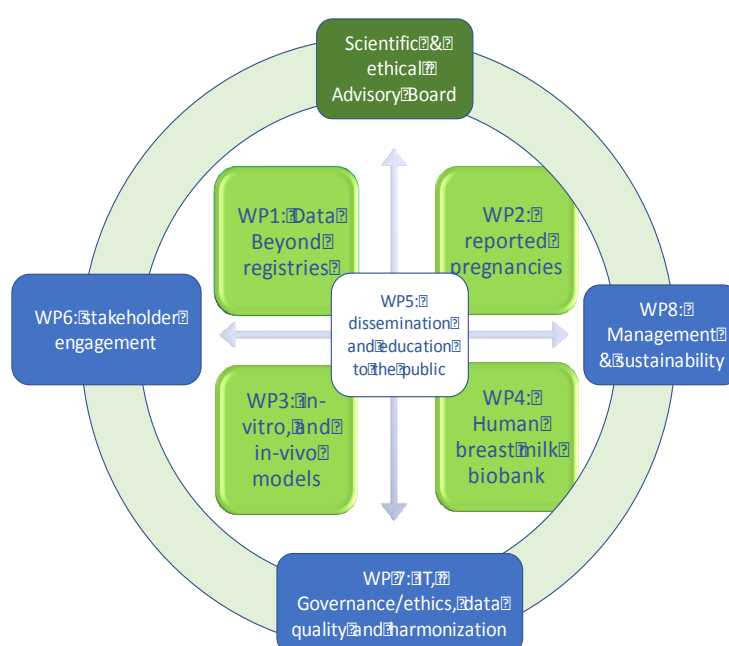


Figure 7: Workplan according to work packages in the ConcePTION project

To collect the needs and approaches in different WPs we created and submitted a data strategy survey template to each work package. Below we give an extraction of the key data strategies in each work package, the full templates are available in the data management plan.

4.1 Work package 1

Table 2: Summary of the data that may be utilized in WP1 demonstration studies

What is the purpose of the data collection/generation and its relation to the objectives of the project? To develop methods and access to data for better harmonized and standardized generation of evidence on drug safety in pregnancy and perform demonstration studies
What types of data will the project generate/collect? Four types of data source are foreseen 1. EUROMediCAT registries (population-based registries of congenital anomaly with information on medication exposure in the first trimester of pregnancy). Sometimes these registries can be linked to other electronic databases 2) Electronic health record data, education data and civil registration data 3) Prospective cohort data collected for research purposes 4) Population-based registries other than EUROMCAT, e.g. Cerebral Palsy registries cancer registries)

<p>Note: Prospective cohorts and pregnancy registries which are collected for the sole purpose of medication safety monitoring will be captured by WP2</p> <p>The data types to be collected include:</p> <ul style="list-style-type: none"> - Metadata to describe databases (in a data source catalog) - Aggregate data to be held on a data platform, resulting from data analysis <p>N.B. Individual data will be held within the contributing data source institutions only.</p>
<p>What is the origin of the data?</p> <p>1) Data included in EUROMediCAT registries:</p> <ul style="list-style-type: none"> - Population-based registries for surveillance of congenital anomalies <p>2) Electronic health record data</p> <ul style="list-style-type: none"> - Data captured for medical recording or billing purposes in the process of routine care <p>3) Prospective cohort data set up for research purposes for example</p> <ul style="list-style-type: none"> - Rotterdam Study - Danish Cohort - Norwegian Cohort <p>.....</p> <p>4) Registries other than EUROCAT – for surveillance and research purposes</p>
<p>Will you re-use any existing data and how?</p> <p>Yes, for items 1-4 all data is about re-use of data collected for routine care or specific research.</p> <p>Re-use will comprise</p> <ul style="list-style-type: none"> - Use of individual level anonymized data on a local secure platform - Distributed analytics of data transformed locally in a Common Data Model (CDM)
<p>What is the expected size of the data?</p> <p>Not clear yet there are many data on births locally, but only aggregated results will be shared and should not be big in terms of Bytes</p>
<p>To whom might it be useful ('data utility') in the project and afterwards?</p> <p>1) Methods development for analysis and determining optimal methods and study design approaches in different scenarios within ConcePTION: will have long term value for many stakeholders – researchers, Health Care Providers, patients, regulators, industry.</p> <p>2) Enables accessible harmonized data repository to study medication safety in pregnancy and lactation for all stakeholders wanting to give advice about drug use in women</p> <p>3) The results of the Demonstration projects will provide medication safety evidence to be published in peer reviewed journals and will have immediate value to stakeholders.</p>

Table 3: process of cleaning, transforming and analysis of the data in WP1

Type of data to be collected	What type of data cleaning is needed?	Who is responsible for data cleaning/quality?	What type of data transformation/analysis do you anticipate	What software will be used for cleaning, transformation and analysis?	Where/by whom will the analysis be conducted?	What standards will you use for code development / access and re-use?
EUROMediCAT Registries	Validation of EUROMediCAT registry data will be conducted at the Coordinating Centre in collaboration with the registries	EUROMediCAT/ DAPs	Structural transformation will be conducted locally and semantic harmonization will be done by WP7. Statistical analysis plan (SAP) will be generated by WP1	STATA	Semantic harmonization by WP7	
Electronic health record data		DAPs			Semantic harmonization by WP7	
Prospective Cohort data		DAPs			Semantic harmonization by WP7	
Registries outside of EUROMediCAT	Conversion to the lower level CDM	DAPs			Semantic harmonization by WP7	

Please explain if you wish:

- Structural transformation will be conducted locally and semantic harmonization (harmonization of content) will be done by WP7 using different algorithms. WP1 will describe in the Statistical analysis plan (SAP) for each demonstration project the variables for the level 2 CDM
- WP1 has a task to define algorithms for exposures and outcomes in the demonstration projects together with task 7.7
- People locally will convert data to the low-level CDM, WP7 will work with databases to help accomplish this.
- Standardization and characterization of databases is done by the DAPs under the instruction of WP7, while study-specific variables will be defined in collaboration with task 1.3, task 7.7 and the definitions task force
- Individual level data will be stored locally. Aggregate data will be made available to specific partners on the ConcePTION secure Platform.

4.2 Work package 2

Table 4: Summary of the type of data that will be utilized by WP2

What is the purpose of the data collection/generation and the relation to the objectives of the ConcePTION project? To develop methods and access to reported pregnancy data for better harmonized and standardized generation of evidence on drug safety in pregnancy
What types and formats of data will WP2 generate/collect? Four types of data are foreseen <ol style="list-style-type: none"> 1. Available data from publicly available international spontaneous reporting systems: EUDRAVIGILANCE, FAERS, VAERS, Vigibase 2. Available spontaneous reporting data from pharmacovigilance centers locally 3. Available Pregnancy registry data & prospective cohort datasets 4. Newly collected data in ConcePTION on neurodevelopmental outcomes
What is the origin of the data? Ad 1) Available data from publicly available international spontaneous reporting systems: EUDRAVIGILANCE, FAERS, VAERS, Vigibase EUDRAVIGILANCE: spontaneous reports sent to EMA (small molecules and biologics), from marketing authorization holders, patients, Health care professionals, pharmacovigilance centers, lawyers. Regarding all products licensed in the European Union FAERS: spontaneous reports sent to FDA on drugs from marketing authorization holders, patients, Health care professionals, pharmacovigilance centers. Regarding all products licensed in the USA VAERS: spontaneous reports sent to US-CDC on vaccines from marketing authorization holders, patients, Health care professionals, pharmacovigilance centers. Regarding all products licensed in the USA VIGIBASE: spontaneous reports of suspected adverse drug reactions, collected by national drug authorities in over 110 countries Ad 2) Available spontaneous reporting data from pharmacovigilance centers locally LAREB: spontaneous reports including narratives, from NL potentially other pharmacovigilance sites Ad 3) Pregnancy registry data & cohorts ENTIS reported pregnancy exposure & follow-up data, at different ENTIS sites in Europe Data collection from clinical research groups (i.e. Rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis) at different institutions_ Regulatory mandated pregnancy exposure registries at companies, based on clinical research organization collected data Pregnancy & birth cohorts (independent of drug exposure), with patient reported data and follow-up (www.birthcohorts.net) Ad 4) Cohort of women and children followed-up after pregnancy exposure specifically for ConcePTION to test new neurodevelopmental outcomes
Will you re-use any existing data and how? Yes, for items 1-3 all data is about re-use of data collected for routine care, surveillance or specific research. Re-use will comprise <ul style="list-style-type: none"> - Use of individual level anonymized data on a secure platform - Distributed analytics of data transformed locally in a CDM

What is the expected size of the data? Spontaneous reporting data publicly available (not restricted to pregnancy) - EUDRAVIGILANCE: 15 million reports - VIGIBASE: 20 million reports - FAERS: 17 million reports - VAERS: 350,000 reports Updates will be made during the course of ConcePTION - LAREB: 3000 reports - Pregnancy registries & cohorts: unclear - New data: unclear
To whom might it be useful ('data utility') in the project and afterwards? Methods development for analysis: useful for ENTIS, ConcePTION, pharmacovigilance centers, EMA, FDA, MAH, CDC, head of medicines agencies etc. research groups. Accessible harmonized data repository on pregnancy reports will be useful for all stakeholders wanting to give advice about drug use in women

Table 5: process of cleansing, transforming and analysis of the data in WP2

Type of data to be collected	What type of data cleaning is needed?	Who is responsible for data cleaning/quality	What type of data transformation/analysis do you anticipate	What software will be used for cleaning, transformation and analysis?	Where/by whom will the analysis be conducted?	What standards will you use for code development / access and re-use?
FAERS/EUDRAVIGILANCE	Extraction Deduplication , completion mapping	Luc, HJ, Caitlin Dodd	Transformation into CDM	R	LAREB/U MCU	Public SAP Publicly available R code in repository (WP7)
Prospective cohort data	unknown		Transformation into CDM	R	?	Public SAP Publicly available R code in repository (WP7)
LAREB data	None additional	LAREB	None	None	LAREB	
Newly collected data neurodevelopmental	unknown	Unknown	unknown	unknown	unknown	

4.3 Work package 3

Table 6: Summary of the type of data that will be utilized by WP3

What is the purpose of the data collection/generation and its relation to the objectives of the project? To develop, characterise, validate and apply a non-clinical testing platform for reliable prediction of drug concentrations in human breast milk (and plasma; milk/plasma ratio) along with systemic drug exposure in breastfed infants.
What types and formats of data will the project generate/collect? <i>In silico</i> data, <i>in vitro</i> cell culture data and <i>in vivo</i> animal data Clinical data to be used will come from WP4 DOC, XLS, CSV, PDF, TXT, PPT, JPG, ZIP, PNG, EPS, AVI, MPG
What is the origin of the data? Experimental <i>in silico/in vitro/in vivo</i> animal data
Will you re-use any existing data and how? Yes, literature data, data obtained in related external projects (e.g. opportunistic sampling in infants), and clinical data generated in WP4
What is the expected size of the data? 1Tb
To whom might it be useful ('data utility') in the project and afterwards?

All that need to predict drug exposure in human milk to inform risk assessment on medication use during breastfeeding (e.g. Health Care Industry: pharmaceuticals, physicians; Health Authorities, Patients, Academic Researchers, etc...).

Table 7 description of the process of cleansing, transforming and analysis of the data in WP3

Task	Type of data to be collected	What type of data cleaning is needed?	Who is responsible for data cleaning?	What type of data transformation/analysis do you anticipate	What software will be used for cleaning, transformation and analysis?	Where/by whom will the analysis be conducted?	What standards will you use for code development / access and re-use?
3.1	Literature lactation data	As defined by the process of systematic literature review: removal of irrelevant data, duplicates, standardization	3.1 participants	NA	NA	NA	NA
	Literature data from existing in vitro/in silico models	Removal of irrelevant data, duplicates, standardization	3.1 participants	NA	NA	NA	NA
3.2	In vitro animal and human cell culture data	Exclusion of data from experiments not passing quality indicators	KUL, UNIBO, BioNotus	Calculation of rates, clearances; dealing with BLQ values; statistical analysis, modelling	Excel, R, Graphpad Prism	KUL, UNIBO, BioNotus	NA
3.3	In vivo experimental animal data	Exclusion of data from experiments not passing quality indicators	UNIBO	Calculation of rates, clearances; dealing with BLQ values; statistical analysis, modelling	Excel, R, Graphpad Prism	KUL	NA
3.4	In silico data	Data transformation as part of data processing	KUL, BioNotus	Simulation-based outputs	Modelling and simulation software (R, Simcyp, PK-SIM, NONMEM)	KUL, BioNotus	NA
	Clinical data from literature, related external studies and/or WP4	removal of irrelevant data, duplicates, standardization	KUL	NA	NA	NA	NA
3.5	Literature data on predicted infant drug exposure	removal of irrelevant data, duplicates, standardization	KUL	NA	NA	NA	NA
	No data collection (data from WP4)	Data transformation	KUL	NA	NA	NA	NA
3.6	No data collection	NA	NA	NA	NA	NA	NA
3.7	No data collection (generation of protocol)	NA	NA	NA	NA	NA	NA

Task	Type of data to be collected	What type of data cleaning is needed?	Who is responsible for data cleaning?	What type of data transformation/analysis do you anticipate	What software will be used for cleaning, transformation and analysis?	Where/by whom will the analysis be conducted?	What standards will you use for code development / access and re-use?
3.8-3.9	No data collection	NA	NA	NA	NA	NA	NA

4.4 Work package 4

Workpackage 4 leads compiled the data strategy questionnaires and provided a written response as well as a graphic

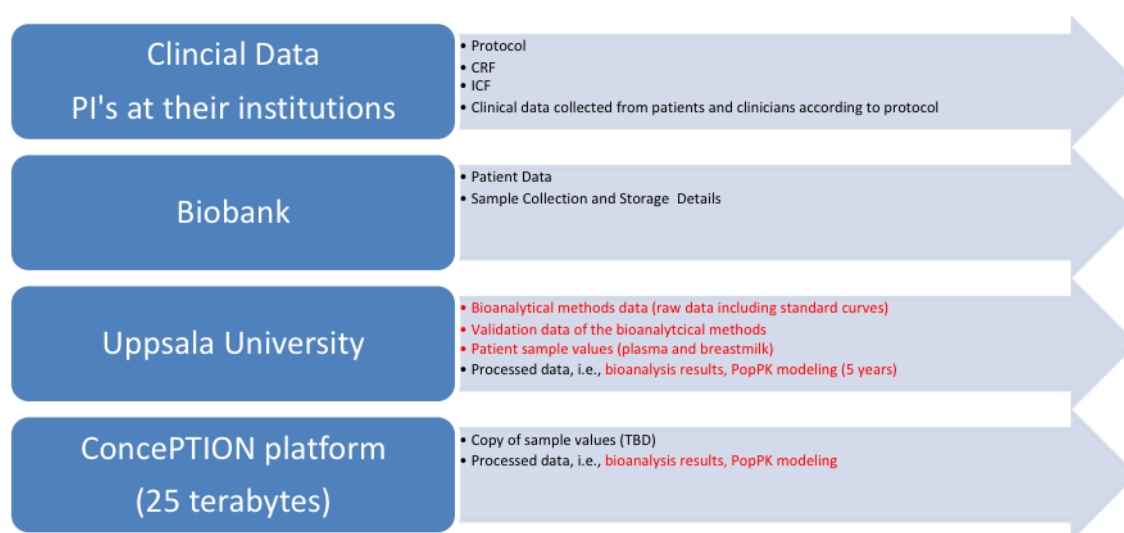


Figure 8: Overview of data strategy in WP4

Table 8: Summary of the type of data that will be utilized by WP 4

<p>What is the purpose of the data collection/generation and its relation to the objectives of the project?</p> <p>Assessment of drug (and active metabolites) concentration values in milk of breastfeeding women as part of the overall objective of ConcePTION to demonstrate feasibility of performing research with samples collected at the BBMRI breast milk biobank and bioanalytical center.</p>
<p>What types and formats of data will the project generate/collect?</p> <p>Bioanalysis data. Rawdata are reported in form of text file (.txt) and as excel file (.xls). The processing data will be in format of MassLynx TargetLynx or any other software used for processing of mass spectrometry Rawdata (alternative in Thermo instrument).</p> <p>The popPK modelling data and models will be processed in specific software (NONMEM, MonoLix or any other). The reports will be in word format (.docx), and final versions is Acrobat Reader format (.pdf file).</p> <p>Please specify if the data will be in SAS datasets</p> <p>Please specify if the preferred software used for PK analyses and modeling</p> <p>Please indicate number of users for the ConcePTION database UPPS – 8.</p> <p>Master protocol</p> <p>Informed Consent Form Master Template</p> <p>Clinical trial data from collection and analysis of blood and breastmilk from women – data from each clinical sites conducting breast milk sampling (stored at ConcePTION platform, UPPS at Biobank)</p>

- Study sites information
- Study personnel information
- Protocol
- Informed Consent
- Case report form
- Ethics Board Approvals
- Screening information
- Inclusion/exclusion data
- Woman's disease or condition
- Medication dose, frequency and exposure
- Concomitant medications
- Adverse events
- Subject disposition
- Protocol deviations
- Demographic data
- Date of birth of the child
- Gestational age of the child
- Weight and height of mother
- Weight of baby
- Time of last dose of studied medication
- Date and time of sampling for milk and blood
- Volume of breast milk from which sample was taken
- Collection Method

Samples stored (samples shipped to and stored at UPPS Biobank)

- Name of site
- Type of sample (plasma/breast milk)
- Date of sample collection
- Date of shipment to biobank
- Date of receipt to biobank
- Noticeable problems at reception
- Specifications met for shipment (yes/no according to IATA recommendations)
- Notification that samples are frozen when received
- Storage temperature of samples
- Processing time
- Deviations from protocol
- Storage temperature at site
- Freeze-Thaw-Cycles
- Storage time before analysis/withdrawal
- Processing times
- Records of measuring instruments
- Records of methods validation
- Records of the maintenance of the equipment

Data on storage at UPPS (original data stored at UPPS Biobank)

- Date and time of reception
- Freezer location
- Storage temperature (also with temperature alarms on freezers)
- Freeze-Thaw-Cycles
- Date and time of retrieval
- Withdrawals for analysis
- LIMS data – all biobank data will be in LIMS and traceability of samples and shipments will be managed in the LIMS

Bioanalytical data (original data stored at UPPS)

- Drug concentration value in breast milk
- Drug concentration value in plasma from mother
- Drug concentration value in plasma/blood from infant
- Analytical methods validation results
- Quality data

PK data (Data stored on ConcePTION platform)

- AUC_τ: Area under the curve over a dosing interval
- Cav: Average concentration over a dosing interval, equal to $Cav = AUC_{\tau} / \tau$
- C_{max}: Maximum observed drug concentration

<ul style="list-style-type: none"> tmax: Time of the maximum observed concentration λ_z: First-order terminal elimination rate constant, calculated from a semi-log plot of the milk (plasma) concentration vs time curve t1/2: First-order terminal elimination half-life, calculated as $0.693/\lambda_z$
Population PK modeling data (Data stored in ConcePTION database)
Calculated infant dose and relative infant dose (Data stored in ConcePTION database)
What is the origin of the data? Pregnant and breastfeeding women from 5 studies of different drugs
Will you re-use any existing data and how? Data from medical records will be collected <i>de novo</i> at clinical sites.
What is the expected size of the data? Local: 25 terabytes On platform only results
To whom might it be useful ('data utility') in the project and after wards? For monitoring purposes regarding pharmacovigilance, for academic research and for drug companies. For physicians and patients to understand medication exposures through breast milk.

The process for data cleaning and data transformation in WP4 was not yet specified

4.5 Work package 5

Table 10: Summary of the type of data that will be utilized by WP5

What is the purpose of the data collection/generation and its relation to the objectives of the project? Get end users input and insights on needs, experience and preferences for information about drug use during pregnancy and lactation
What types and formats of data will the project generate/collect? Ad hoc surveys and focus groups, literature reviews Knowledge bank
What is the origin of the data? 1. Newly collected data through surveys or focus groups 2. Literature references and synthesis of evidence in knowledge databases developed by local TIS
Will you re-use any existing data and how? TBD (information from existing knowledge databases)
What is the expected size of the data? 1. Unclear but expected to be limited (Surveys in thousands of end users, focus groups of limited size) 2. unclear (note: check with existing KB e.g. Lareb)
To whom might it be useful ('data utility') in the project and after wards? Accessible harmonized repository on evidence regarding drug use in pregnancy and breastfeeding will be useful for all stakeholders wanting to give advice about drug use in women

The process for data cleansing and data transformation in WP5 was not yet specified

4.6 Workpackage 7

Table 11: Summary of the type of data that will be utilized by WP7

What is the purpose of the data collection/generation and its relation to the objectives of the project? To collect information on ethical requirements to share data from pregnant women and data access providers, to assess governance issues and information, to characterize data sources and to assess whether they are fit for purpose to provide evidence on drug safety in pregnancy
What types and formats of data will WP7 generate/collect? Task 7.1: Reward models: ethico legal documents from data access providers, and interview data from DAPs participating in the ConcePTION project Format: video mp3, .docx Task 7.2: Governance: codes of conduct for analysis of different types of data that are used in ConcePTION Format: .docx Task 7.3: Ethics: interview data from pregnant women who consent to participate Format: mp3, .docx

<p>Task 7.4: Catalogue: non-curated publicly available information on EUROMediSAFE data sources, curated data on organizations, responsible persons and data sources Format: .Xls; Catalogue (Molgenis)</p> <p>Task 7.5: Common data models: common data models for use by WP1 and 2 Format: .docx, .xls</p> <p>Task 7.6: Data characterization: approaches all the DAPs for WP1 and 2 that agree to participate in data characterization, data collected will comprise indicators of the data (aggregate). Event definition templates to find codes to extract events /data/ Simulated data to develop programs Format: .csv .docx</p> <p>Task 7.7: Algorithm quantification: Will assess impact of the use of different algorithms to extract events and evaluate novel validation manners Format: .csv .docx</p> <p>Task 7.8: Double coding of scripts: in this task codes for distributed analyses will be created and hosted. Simulated data for programming Format: .csv ; .txt; R, GitHub repository</p> <p>Task 7.9: Reprotox data: collection of repro-toxicology data from rats and rabbits extracted from regulatory submission Format: .csv; .docx</p> <p>Task 7.10: Ethics issues: collection/repository of documents Format: .docx; .pdf</p>
<p>What is the origin of the data?</p> <p>Task 7.1: DAPs in the consortium as partners or third parties</p> <p>Task 7.2: Existing codes of conduct that are publicly available or owned by consortium members</p> <p>Task 7.3: Pregnant women</p> <p>Task 7.4: EUROMediSAFE (data from internet) (non-curated), survey data with DAPs partners in consortium (curated)</p> <p>Task 7.5: Common data model development by WP7, 1 and 2, based on prior CDM and those in Sentinel, OMOP, EUROCART, EUROLINKCAT</p> <p>Task 7.6: Data access providers that participate in the data characterization & algorithm validation/</p> <p>Task 7.7: Data access providers that participate in the data characterization & algorithm validation</p> <p>Task 7.8: Tools written by the WP7 participants</p> <p>Task 7.9: Data collected by CBG and RIVM</p> <p>Task 7.10: Documents generated by different WP and organizations</p>
<p>Will you re-use any existing data and how?</p> <p>Yes, for the data characterization we will access data that is available at the sites of the data access providers. This will be done in a distributed manner. Data will stay local and data characterization scripts will be sent to the DAPs, these scripts will generate results that will be sent to the secure platform</p>
<p>What is the expected size of the data?</p> <ol style="list-style-type: none"> 1) WP7 anticipates that the data that is shared as results will be small in terms of bytes 2) Simulated data will be hosted on the platform to develop scripts
<p>To whom might it be useful ('data utility') in the project and after wards?</p> <ol style="list-style-type: none"> 6) WP7 will support methods development and conduct of demonstration studies by WP1 and 2 7) WP7 will also systematically characterize datasources and assess whether they are fit for purpose, this will increase the transparency and potentially take away criticisms about quality of RWE

Table 12: description of the process of cleaning, transforming and analysis of the data in WP7

Type of data to be collected	What type of data cleaning is needed?	Who is responsible for data cleaning?	What type of data transformation/ analysis do you anticipate	What software will be used for cleaning, transformation and analysis?	Where/by whom will the analysis be conducted?	What standards will you use for code development / access and re-use?
7.1 Interviews	Recording and transcribing	UMCU	Summarize and reflect based on ethical framework	Word	UMCU other 7.1 partners	None
7.2 Code of conducts/guidance	None	NA	NA	NA	iHD, GSK	Existing standards/guidances
7.3 Ethics interviews	Recording and transcribing	UMCU	Summarize and reflect based on ethical framework	Word	UMCU	None

7.4 Catalogue details	Review and uploading of information and documents	DAPs, WP1 partners, WP7	Uploading of data collected in questionnaire format into the catalogue	Excel	UMCG	Open source /Molgenis
7.5 Results of analyses on health data in DP	Quality indicators of data to support Demonstration projects	DAP	Transformation of original data into CDM, transformation of data in CDM into analysis tables And pooling & final analysis of results	Locally into CDM: SAS, SQL, Stata, R From CDM into analysis table: R From analysis table into final results: WP1/2 partners	Locally ETL from original to CDM: DAP From CDM into analysis table: UMCU, ARS, J&J On platform: Analysis: WP1/2	ETL will be available on catalogue Data transformation script: open source Final analysis script: double coded
7.6 Data characterization	Quality indicators of data	DAP	Transformation of original data into CDM, transformation of data in CDM into analysis tables And pooling & final analysis of results	Locally into CDM: SAS, SQL, Stata, R From CDM into analysis table: R From analysis table into final results	Locally ETL from original to CDM: DAP From CDM into analysis table: UMCU, ARS, J&J On platform: Analysis: UMCU, ARS, DAP	ETL will be available on catalogue Data transformation script: open source
7.7 Algorithm validation	Algorithm development, verification of impact	DAP	Analysis of different combinations of components to define event	From CDM into analysis table: R From analysis table into final results	From CDM To platform: Analysis: UMCU, ARS, DAP	Data transformation script: open source
7.8 Double coding	None	NA	Double coding of scripts	R	UMCU/ARS, J&J	ENCePP
7.9 Repro-toxicology	Extraction from files	CBG	Extracting and transforming into Excel	Excel	CBG, RIVM	GLP
7.10 Ethical follow-up	None	None	None	NA	NA	NA

5. Remote Research Environment (platform) requirements & testing

5.1 Requirements

WP7 had put the following requirements in the Description of Action for a remote data platform:

“A state-of-the-art digital research environment with ISO certified and GDPR proof services for remote collaborations will be subcontracted and operated. Access to the application server will be only allowed using two-factor authentication. The environment will be able to host multiple research projects, each with its own secured area to share data and results and provide access through remote desk tops clients. The infrastructure will offer several analytical tools (e.g. R, SQL database, Shiny, Stata) word processing software, and utilities”

5.1.1 Access to the remote research platform

ConcePTION partners and third parties should be able to access the remote data platform using two-factor authentication. This access should be independent of the IP address of the user and the operating system in use on the user’s computer. Access should be automated and should not require the intervention of an administrator nor the use of any additional hardware.

Within the remote data platform or remote research environment (RRE), folders should be structured based upon the principles of protection of data integrity and independence of protocols within ConcePTION.

5.1.2 Private workspaces within the remote data platform

The environment should be able to host multiple workspaces: WP1/7, WP2/7, WP3/4, WP7, and sustainability projects, each with its own area to share data. Within each of these workspaces, folders will be made available for each protocol. For example, within the WP1/7 folder, each demonstration project will have its own dedicated area. In other words, the workspace will have a per-protocol structure. Within each protocol folder, participating data access providers will have spaces into which data can be uploaded. Because the data access providers participating in each demonstration project may be different, as will the content and structure of the data uploaded, copying of data from one protocol space to another will be prohibited. See figure 9 for the proposed folder structure.

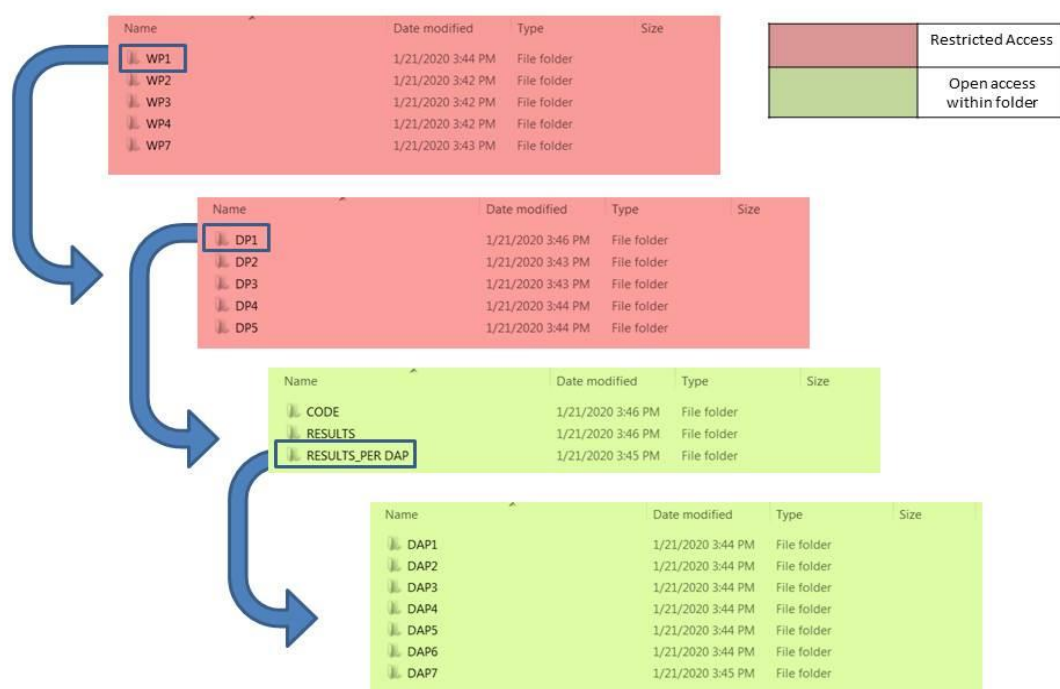


Figure 9: Outline of proposed folder structure on the platform

5.1.3 Analytical tools available on the remote data platform

The RRE should offer analytical tools, specifically R, Stata, SAS, Excel, and word processing utilities. Updating of packages must be permitted and documented.

5.1.4 Upload

Upload of data or documents will be possible only into those folders to which a user has write access. Uploads should be date and time stamped and the user who made the upload should be recorded. Ideally, this data should be automatically generated in a readme file to accompany each upload. The user uploading the data should amend this readme file to explain the purpose of the upload. Output of scripts which generate aggregated analysis files against local data, or produce analysis results, should be able to be uploaded through remote desktop clients. The infrastructure will use a secure file transfer protocol (SFTP)²

5.1.5 Download

Download of results or data is possible only by the owner of the data (the DAP who originally uploaded). All other downloads require administrator approval and are limited to analysis results.

² In SFTP both the commands and data are encrypted in order to prevent passwords and other sensitive information from being transferred over the network. SFTP requires that the client user must be authenticated by the server and the data transfer must take place over a secure channel (SSH)

5.2 Review of Potential Solutions

Various organizations offer remote cloud-based services, within WP7 we have first explored with ConcePTION partners since this would facilitate the process and partners would provide services at actual rather than commercial costs.

5.2.1 Virtual Research Workspace (University Medical Center Groningen-UMCG)

Through BBMRI-ERIC we have accessed and described the remote workspace at UMCG. The Virtual Research Workspace (VRW) of the University Medical Center Groningen has been developed by, and is technically maintained by the University of Groningen (RUG, CIT department).

The Virtual Research Workspace (VRW) is a virtual research environment for collaboration amongst different researchers (working within and outside the UMCG). In the VRW, researchers can work safely together on a research project. The original data stays within the VRW, and researchers can access the data in the VRW from different locations (even outside the UMCG). In the VRW you will find storage facilities, as well as software and applications for analysis such as Matlab, R, Stata, and SPSS. It is possible to request that additional software be added to the VRW. The VRW offers two services: A basic service (Service Based Computing) and a premium service (Virtual Desktop Infrastructure). The premium service offers dedicated resources (memory, vCPU, VGPU).

Legal issues: Using the Virtual Research Workspace is safe and secure according to the GDPR. The RWP is ISO 27001 certified (information security). Original data cannot be removed from the RWP.

A general animation video on the Virtual workspace can be found in this YouTube link:

<https://www.youtube.com/watch?v=YZL1X-1AzHg>

Research Workspace provides an integrated and secured research environment allowing researchers to access their requested dataset. In addition, the Research Workspace contains a number of tools to analyse the data and to store the results.

Using the Research Workspace is like working on a normal computer using Windows 10. However, some actions are restricted because of dataset security measures, and can only be done by the Data Manager (administrator) on the researcher's request.

- Researchers may not export or import data. This can only be done by the data manager.
- Researchers may not install applications.

The Research Workspace requires the free Citrix Receiver plugin to be installed. The plugin is available for different desktop platforms (Windows, Linux, Mac). The installation procedure has to be performed only once for each Desktop PC on which the UMCG Research Workspace will be used (<https://www.citrix.com/downloads/citrix-receiver/>)

Access to the workspace requires a OneGini account.

There are multiple research workspaces within the platform. Following successful login, users are asked to choose the relevant Research Workspace by selecting their study.

Following this selection, users are prompted to enter an SMS token. This token will be received on the mobile phone number linked to the user's Research Workspace account.

5.2.2 anDREa Azure DRE (UMC Utrecht, Radboud University and Erasmus Medical Center)

The anDREa Research Environment is available through the anDREa consortium, a collaboration between the Dutch University hospitals Radboudumc Nijmegen, Erasmus MC Rotterdam, and UMC Utrecht. The consortium owns, develops, distributes and promotes the product that is enclosed in its name: the Digital Research Environment (Azure DRE for short).

The Digital Research Environment (DRE) is a cloud based, globally available research environment where data is stored and organized securely and where researchers can quickly generate workspaces to collaborate in. Within these workspaces, researchers have preinstalled applications at their disposal, as well as the ability to bring own tooling. Globally available and accessible 24/7.

The DRE facilitates users to collaborate on research projects in a safe, yet flexible compute and storage environment. The architecture of the DRE allows researchers to use a solution within the boundaries of data management rules and regulations. Although General Data Protection Regulation (GDPR) and Good (Clinical) Research Practice still rely on researchers, the DRE offers tools to more easily control and monitor which activities take place within projects.

Within the DRE platform each of the projects you are a member of consists of a separate, secure folder, called a 'workspace'. Each workspace is completely secure, so researchers are in full control of their data. Each workspace has its own list of users, which can be managed by its administrators.

Each workspace is fully scalable with regard to data quantity and computing power, thereby supporting anything from small to complex multi-center, multisource studies.

Thanks to the flexibility of the Azure cloud platform, on which the DRE has been built, it offers a variety of virtualized hardware, including, but not limited, to:

- Standalone virtual Windows / Linux machines (e.g. 72 cores, 144 Gb RAM);
- Compute Clusters (HPC);
- Web-Servers;
- Several storage options (tables/files/SQL/etc.).

Hereby researchers pay for computing capacity by the second, with no long-term commitment or upfront payments, and can increase or decrease compute capacity via self-service³.

The research environment is GDPR compliant, requires two factor authentications for access, and has been penetration tested. It allows for storing of all data types, formats, and volumes with archiving of entire workspaces. All work is conducted within workspaces on virtual machines to which users have administrator rights. Virtual machines can be scaled up and down.

Management of users is conducted by workspace owners who may be principal investigators of a project or another researcher chosen to perform administration. The research environment allows international colleagues to collaborate with self-service independent IP-independent access at any time. The environment also allows

³ <https://www.radboudumc.nl/en/research/radboud-technology-centers/data-stewardship/digital-research-environment/what-is-the-digital-research-environment>

for construction of a meta-data catalogue and workspace to workspace transfer. Finally, the research environment provides a tool to track resource consumption as payment is charged according to usage⁴.

Upload to the file system within the workspace is possible for any user directly from his or her own computer. Owners of workspaces may download files while workspace users must request download permission from the workspace owner.

Azure DRE is suitable for many types of research, data analysis and collaboration. The product has been developed within the healthcare domain, but will realize general applicability through continuous development.

- Clinical data, images, -omics data, small and large volume data, SPSS, Matlab, R and advanced/self-made tools, batch processing.
- Clinical trials, pre-clinical studies, non-human studies, biobanks, registries.
- Quality of care studies, apps for care givers

Software available on the research environment by default includes:

- Windows Server 2016 Datacentre
- 7-Zip version 18.05
- Adobe Acrobat Reader DC version 19.01
- ATLAS.ti version 8.3
- Google Chrome version 71
- SPSS Statistics version 25
- Microsoft Office Enterprise version 2007
- MPC-HC version 1.7.1
- Notepad++ version 3.7
- Python Launcher version 3.7
- R version 3.5.1
- R Studio version 1.1.46

DRE applications available for installation include:

- Docker Toolbox version 18.03
- Endnote version 8
- F4 version 3.1
- Fiji ImagJ version 1.52
- Git version 2.21
- Knime version 3.7
- Matlab version 2014
- Notepad++ version 7.6.1
- Power BI desktop version 2.7
- Putty version 0.70
- Python Anaconda 3 version 2018.12
- SQL Server version 2017
- SQL Server Management Studio version 17
- R version 3.6.1
- Stata version 16
- Storage Explorer version 1.9⁵

⁴ <https://www.andrea-consortium.org/media/1013/andrea-digital-research-environment-info-pack.pdf>

⁵ https://wiki.researchenvironment.org/index.php/Main_Page

An informational video is available here: <https://www.andrea-consortium.org/about-andrea/>

5.2.3 UKSeRP UK Secure eResearch Platform (Swansea University Medical School)

The UKSeRP is available through the Swansea University Medical School. The research platform has been designed to be very user friendly. Each organization's particular security considerations that require specialist solutions to meet specific needs can benefit from the bespoke design and implementation service to create a tailored product specifically to meet those requirements. UKSeRP works closely with users to gain a full understanding of their requirements in order to set up a personalized UKSeRP.

Once established, data owners can begin to upload and prepare their data, create projects, set up user accounts and make the necessary preparations to start sharing the data. Researchers or analysts can then request an online account and, if the owner approves, can easily create them. The research environment provides a secure desktop with a range of statistical research software.

Data Owners Retain Full Control. With UKSeRP the data custodians retain complete ownership and remain in full control of their data at all times with the ability to define roles and to set preferred access controls and project permissions to ensure that particular data sets are accessed only by pre-approved individuals as defined by data owners.

UKSeRP is already pre-approved as an accredited platform for HSCIC (Health and Social Care Information Centre) data sharing framework agreements and has been granted the ISO27001 'kite mark'. Full audit trails of use are available upon request at any time.

UKSeRPs' solution to the problem of data access with data security is its secure remote access facility. Approved users can access data from any desktop from anywhere in the world at any time, or from an established 'safe zone' if preferred, so that research activity isn't restricted.

UKSeRP gives users a straightforward way to migrate large amounts of data into a highly regulated secure environment. Once data has been uploaded into UKSeRP, it can be automatically analyzed in order to highlight any anomalies in the data such as missing fields or whether any fields are corrupted etc. giving researchers the opportunity to improve the quality of their data before use.

A semi-automated metadata cataloguing system that has been built with the user in mind to eliminate the need for training and associated costs. No-one knows more about the content of the data being uploaded than the owners who are best positioned to augment the data sets with descriptors of what each dataset contains. Approved users can then quickly identify data sources that can be linked for specific research questions with ease and increase productivity.

The metadata catalogue allows researchers to attach supporting documentation to datasets giving researchers additional intelligence to quickly select the most appropriate data sets for their projects

UKSeRP comes with a wide-ranging suite of analysis tools. For data storage, UKSeRP offers the flexibility to choose the most appropriate option from DB2, SQL Server, PostgreSQL, Hadoop or NoSQL. UKSeRP is a robust, large-scale high-performance data storage and analysis platform for large complex data sets. There is scalable infinite capacity for data to be stored in a safe neutral shared environment that has the capability to integrate diverse forms of data from multiple record systems for analysis.

5.3 Testing of the platforms

5.3.1 Features to test

In general, we will test the following features of a candidate platform:

1. There must be a concept of ‘study folder’. A study folder can have subfolders, but everyone who has access to the folder has access to all the subfolders.
2. It must be impossible for a person in two separate study teams to copy from one study folder to another, that is, study folders should be isolated.
3. Upload and download of results must be governed by an administrative actor, and there must be procedures to govern this;
4. in the platform there should be interfaces that support upload and download.

Tests are conducted from the perspective of each user:

DAP perspective

- Wiki (easy to read and use)
- Filesystem
 - Data uploading interface
 - Access: ease of installation and use
 - Load: ease of use
 - Request uploading on the platform: test the procedure in terms of ease of use,
 - Have the data uploaded
 - Data downloading interface
 - Security
 - Downloading should be impossible
 - logging of downloads should be inspectable (date, who downloaded, filenames, folder from which they were downloaded, file dimension)
- Virtual machine
 - Access to the machine (velocity, easiness)
 - Access to filesystem
 - Organization (easy to use and understand)
 - Check integrity of uploaded data
 - Ability to restrict access and downloading
 - Ability to verify access & logging
 - Check that data access is as agreed (there must be an easy way to inspect list of users that are allowed to access, to copy, and to modify)

Study team perspective

- Wiki (easy to read and use)
- Virtual machine
 - Access to the machine (velocity, easiness)
 - Access to filesystem
 - Organization (easiness of use and documentation)
 - Check multiple people can access to relevant folders at the same time
 - Check that all the folders requested have the correct access rights, in terms of access, copy, and modify

- Analyse uploaded data
 - Speed
 - Memory use (in case of very heavy computations going on in parallel, such as the test of the developers' group)
 - Existing software (R, SAS, Stata, ...)
 - Installation of additional software
- Filesystem
 - Data downloading
 - Security
 - Downloading should be impossible
 - logging of downloads should be inspectable (date, who downloaded, filenames, folder from which they were downloaded, file dimension)

Data administrator perspective

- Wiki (easy to read and use)
- Filesystem
 - Data uploading interface
 - Access: ease of installation and use
 - Load: ease of use
 - Request uploading on the platform: test the procedure in terms of ease of use,
 - Have the data uploaded
 - Data downloading interface
 - Security
 - Downloading is possible and is automatically logged
 - Can inspect file before downloading
- Easiness to maintain access and create reports

5.3.2 Testing status

5.3.2.1 Virtual Research Workspace (University Medical Center Groningen)

Testing of the VRW began in October 2019 and proceeded through interactions with UCMG informatics, creation of test user accounts, and discussions of research environment functionality. UCMG informatics staff showed willingness to share resources and knowledge of the VRW with WP7. WP7 researchers were able to create accounts and log in to the research environment. However, due to unavailability of dedicated staff for administration of the VRW at UMCG, access to the research environment, especially for rapid feedback on technical issues, proved difficult. For this reason, testing on the VRW was halted in January 2020.

5.3.2.2 anDREa Azure DRE (Radboud University, UMC Utrecht, and Erasmus Medical Center)

Formal testing of the anDREa DRE began in January 2020 through interactions with anDREa developers and staff. Testing proceeded in February 2020 with creation of accounts for test users. In March 2020, formal testing according to the protocol described in section 4.3.1

5.3.2.3 UKSeRP UK Secure eResearch Platform (Swansea University Medical School)

Discussion of the UKSeRP platform between WP7 and Swansea began in February 2020. Research platform staff at UKSeRP showed willingness to share resources and to open discussion with WP7 regarding use of the platform. To date, testing of this platform has not proceeded beyond these initial discussions due to uncertainty surrounding the impact of Brexit and on the use of a UK-based research environment.

5.3.3 Testing results

5.3.3.1 anDREa Azure DRE

In March 2020, WP7 tested the anDREa research platform in the roles of data access provider (DAP), research environment administrator (ADMIN), and study team member (STM). Results of the test are provided below. Following testing, WP7 has proceeded in discussion with developers of the platform to address outstanding questions.

Test 1. Data Access Provider (DAP) and Administrator (ADMIN)

Course of action

1. DAP tests access to the DRE
2. DAP tests filesystem and downloading
3. DAP tests access to the Virtual Machine
4. Administrator grants/does not grant permission and does/ does not download

Results from the DAP perspective

1. Access

DAP will receive the email to create an account. The English version of the greeting should be the only one for this project. It should be made more clear which email address to use for the Microsoft interface through Radboud. It should also be clear that users should change their passwords on first login to something they can remember. DAP had to reset her password. Double-authentication using mobile phone went smoothly. The wiki is very useful. DAP and ADMIN both used it throughout the test.

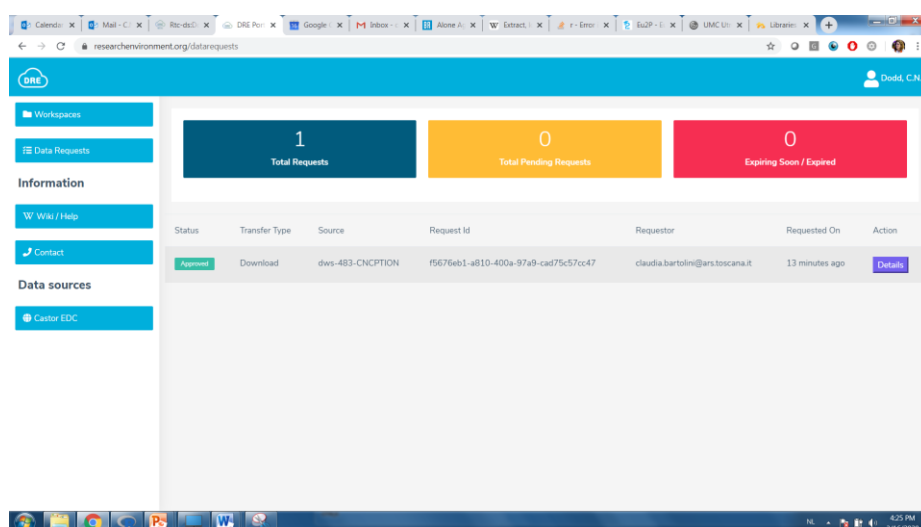
2. Filesystem

- Data uploading interface
 - Access: ease of installation and use
 - DAP can access the file system in the DRE
 - Load: ease of use
 - Using the cloud ‘upload’ icon DAP was able to upload a file. She was not able to using the ‘drag and drop’ method described in the wiki. When a file is uploaded, it is automatically uploaded into a new ‘inbox’ folder. Inside the ‘inbox’ folder, is a ‘transfer’ folder with date and time containing the uploaded data. With the drag-and-drop option, a window appears asking DAP if she wants to open or download the file. After uploading a second file, DAP is asked if she would like to upload the file to the already existing ‘inbox’. When she does so, there is a new ‘transfer’ folder within the ‘inbox’ folder with the date and time.
 - Request uploading on the platform: test the procedure in terms of ease of use,
 - No request required. She can upload the file and move it to another folder within the workspace but she can’t drag and drop the file to upload it, as Wiki said.
 - Have the data uploaded

Notes: DAP can transfer a file to another workspace. (it is only possible to copy and then paste the file, not “move”). Does this mean that DAPs can copy data from workspace to workspace if they are members of more than one? This will require clarification

○ Data downloading interface

- DAP selected the file she wanted to download and chose ‘submit download request’. ADMIN (admin) can then see the download request in her ‘Data Requests’ tab. There are options to message the requester via message (this prompts Skype which may be blocked if not installed or if there are popup blockers). It is also possible to send an email. ADMIN sent a message using outlook through the workspace but DAP did not receive it.
 - Security
 - ADMIN approved the request and DAP was then able to download by returning to her data requests tab. ADMIN and DAP are both able to view their data requests and the status of each request.
 - logging of downloads should be inspectable (date, who downloaded, filenames, folder from which they were downloaded, file dimension)
- ADMIN and DAP can see the status, date of request, and size of the file. But cannot see if it has been downloaded. DAP can see the size of the file when the request is pending.



3. Virtual machine

- DAP
 - Virtual machine takes a long time to start up (2-5 minutes)
 - Running the virtual machine requires download of a .rdp file. DAP cannot enter the virtual machine. DAP receives a notification that her password is incorrect 3 times. “Authentication error – remote computer cannot be accessed. Maybe the password is expired. Update if expired. Contact administrator or technical support.”. Also does not work using the virtual machine in DAP’s institution.
 - DAP is using firefox and switches to Chrome for further testing, but is still not able to use the virtual machine. Is the password for the virtual machine the same as that for the workspace as a whole? Does it matter that she is not the administrator of her PC?
 - Finding a help email address to address this problem is not straightforward through the wiki. Instead, when a person becomes a new user, the help email address is included in the welcome message. DAP

wrote to the help email address and received a response within 20 minutes that her account had been reset.

- Access to the machine (velocity, easiness)
 - Access to filesystem
 - ADMIN was able to login to the virtual machine and to see the uploaded files in the ‘data’ drive under ‘This PC’.
 - Organization (easy to use and understand)
 - Check integrity of uploaded data
 - Ability to restrict access and downloading
 - Ability to verify access & logging
 - Check that data access is as agreed (there must be an easy way to inspect list of users that are allowed to access, to copy, and to modify)
- List of users is visible in the workspace

Test 2. Study Team Member (STM) and Administrator (ADMIN)

Course of action

1. STM tests the virtual machine (e.g. using scripts to generate simulated data and process them)
2. STM tests filesystem in downloading
3. Administrator grants/does not grant permission and does/ does not download

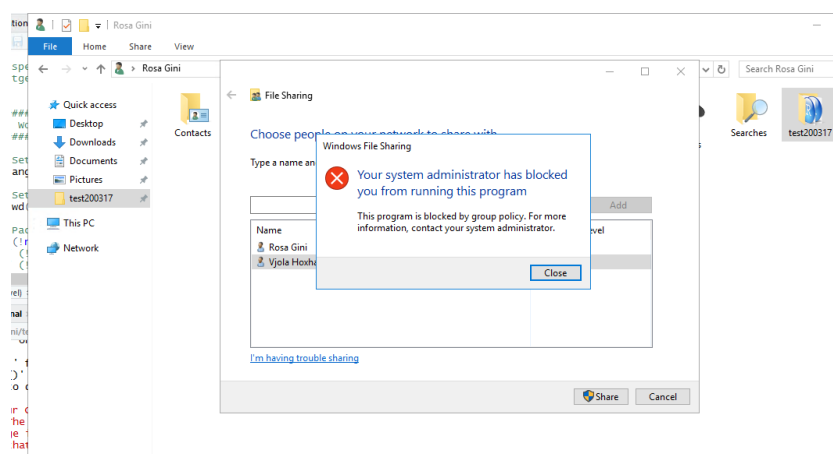
From the study team perspective

1. Wiki (easy to read and use)

2. Virtual machine

- Access to the machine (velocity, easiness)
Can be done. When exiting, you need to re-download a new instance of the virtual machine. Private files are still there (virtual machine has not been ‘wiped clean’).
- Access to filesystem
STM can see the files in a common folder, STM cannot see who generated them. I can delete them with no problem. STM can create a folder in this space, and save files here.

ADMIN is not able to access STM private folder. STM created a subfolder there and tried to share it with ADMIN. The answer is supplied in screenshot below



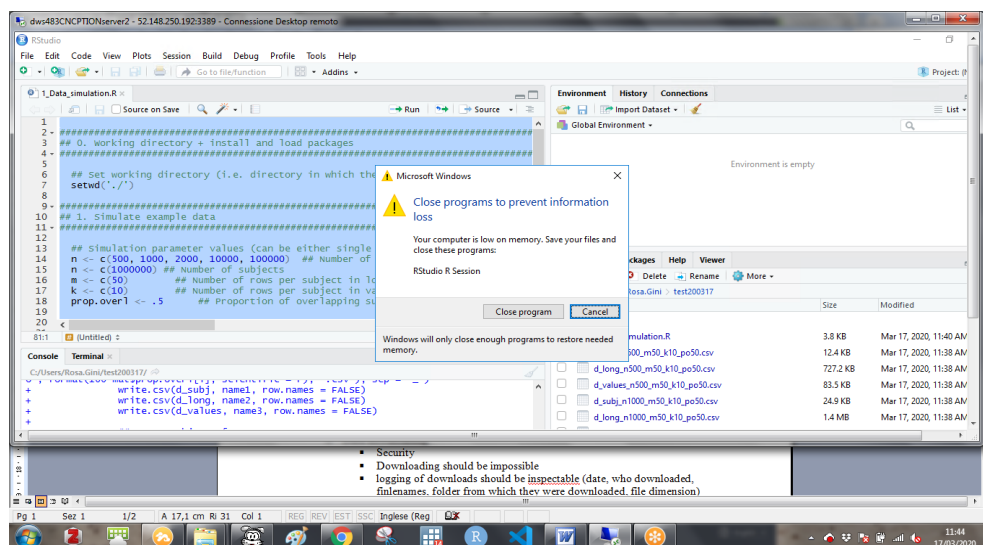
Would it be possible to allow a researcher to share private files with other users (researchers or owners)?

ADMIN was not able to access STMs' private files but then ADMIN went to C:/Users/STM.Gini ADMIN and tried to open it. ADMIN saw a window pop up which says "Do you want to make permanent changes to this folder?" ADMIN pressed Complete and then ADMIN could access the folder. All this process was done without notifying STM of my actions (as an owner ADMIN could access her private files without asking for her permission). She tried to access my private files but she couldn't do that (she is specified as researcher in the workspace). Both using same virtual machine.

- Organization (easiness of use and documentation)
 - Easy
- Check multiple people can access to relevant folders at the same time
 - True
- Check that all the folders requested have the correct access rights, in terms of access, copy, and modify
- Analyse uploaded data
 - Speed
 - Memory use (in case of very heavy computations going on in parallel, such are the test of the developers' group)

When running the generation of simulated data for 1,000,000 subjects, STM received this message

When STM closed the session, STM realised R has given this message



- Existing software (R, SAS, Stata, ...)

R, SPSS and Matlab are there, STM could not see SAS nor Stata

- Filesystem
 - Data downloading
 - Security
 - Downloading should be impossible

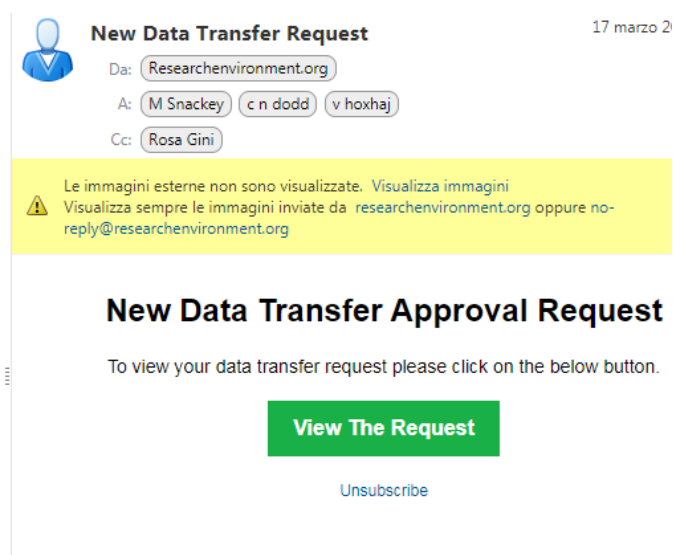
From within the virtual machines, STM moved 3 large data files to the 'data' folder, where STM was able to create a subfolder

Then from the workspace STM could see those files, and select them and request for download. STM was not requested to specify who STM was asking this, nor why

Later STM received an email from ADMIN, but there was no automatic reference to the fact that she was the owner of my request nor any reference to the request itself (even though STM knows that ADMIN generated the email from an interface that was generated from the request itself).

When STM send the request for download ADMIN could access that request in the Data request tab. There ADMIN could see her name, email, the request ID, time of request, the names of the files she wanted to download, their size. There was no explanation about the reason behind downloading but there were two buttons (message: which opened a skype and email: this what we tried) which we can use to ask why this download is being requested before approving.

Later STM received an email with a confirmation that the transfer was requested



After ADMIN approved, STM received a confirmation

5.4 Choice of RRE/platform

Based on the available platforms among ConcePTION partners, the requirements and the testing results WP7 has decided that the anDREa platform may be easiest accessible most suitable for the needs of the consortium and cheapest.

The anDREa platform complies with the requirements, but would need some additional documents and improvement around documentation of downloading of data. A request for technical support around this has been communicated to the developers of the platform.

6. References

Almeida JR, Gini R, Roberto G, Rijnbeek P, Oliveira JL. TASKA: A modular task management system to support health research studies. BMC Med Inform Decis Mak. 2019 Jul 2;19. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6604289/>

Appendix 1: Instructions for managing a task in the TMS



Instructions to manage a task
in the ConcePTION Task Management System
version 1.0
3 June 2020

Authors: Rosa Gini (ARS), Marco Moscatelli (CNR_ITB), Caitlin Dodd (UMCU)

Introduction

This document is meant to support you in managing a task you have created in the Task Management System (TMS) of ConcePTION. To understand how to create a task, follow the instructions in the document “Instructions to create a task in the ConcePTION Task Management System”.

This document recommends that all the interaction happens via the TMS interface. It is intended that the interaction is also supported by notifications via email, but this is not covered in this manual.

The document describes how to

1. perform surveillance of status of parent and children tasks
2. manage children subtasks
 - 2.1. send a message to one responder organization
 - 2.2. reassign a child subtask to a new responder from the same organization
 - 2.3. extend the deadline for one responder
 - 2.4. resolve the child subtask for one responder
 - 2.5. send a bulk update to many children subtasks
3. manage the parent task
 - 3.1. reaccess a child subtask that was resolved
 - 3.2. extend the deadline or resolve the parent task
 - 3.3. add a new child subtask to an existing parent

1. Perform surveillance of status of parent and children tasks

To check the status of the task, access your home. You will see the task and its list of subtasks, as follows.

My Request Edit					
#	Subject	Priority	Queue	Status	Due
881	finalize interview answer sheet	0	management	(pending 10 other tickets)	Fri May 29 00:00:00 2020
#882: Task finalize interview answer sheet (881) #883: Task finalize interview answer sheet (881) #884: Task finalize interview answer sheet (881) #885: Task finalize interview answer sheet (881) #886: Task finalize interview answer sheet (881) #887: Task finalize interview answer sheet (881) #888: Task finalize interview answer sheet (881) #889: Task finalize interview answer sheet (881) #890: Task finalize interview answer sheet (881) #891: Task finalize interview answer sheet (881) #892: Task finalize interview answer sheet (881) #893: Task finalize interview answer sheet (881) #894: Task finalize interview answer sheet (881) #895: Task finalize interview answer sheet (881) #896: Task finalize interview answer sheet (881) #897: Task finalize interview answer sheet (881)					

The number after 'pending' indicates how many responder organisations are still to complete the task.

You can interact with the task at two levels: either the overall parent task, or the single children subtasks.

To manage the children subtasks, click on the link 'pending other XX tickets'

My Request Edit					
#	Subject	Priority	Queue	Status	Due
881	finalize interview answer sheet	0	management	(pending 10 other tickets)	Fri May 29 00:00:00 2020
#882: Task finalize interview answer sheet (881) #883: Task finalize interview answer sheet (881) #884: Task finalize interview answer sheet (881) #885: Task finalize interview answer sheet (881) #886: Task finalize interview answer sheet (881) #887: Task finalize interview answer sheet (881) #888: Task finalize interview answer sheet (881) #889: Task finalize interview answer sheet (881) #890: Task finalize interview answer sheet (881) #891: Task finalize interview answer sheet (881) #892: Task finalize interview answer sheet (881) #893: Task finalize interview answer sheet (881) #894: Task finalize interview answer sheet (881) #895: Task finalize interview answer sheet (881) #896: Task finalize interview answer sheet (881)					

You will access the screen below.

Home Search Reports Tools Admin Logged in as rosa.gini@ars.toscana.it RT for imi-tasker.bbmri-eric.eu

Found 10 tickets [New task in](#) [managemen](#) Search...

Edit Search Advanced Show Results Bulk Update Chart Feeds Activity Reports Gantt Chart

#	Subject	Owner	Queue	Status	Due	New messages
885	Task finalize interview answer sheet (881)	cecile.droz@u-bordeaux.fr (Droz-Perroteau)	General	open	Fri May 29 00:00:00 2020	No
887	Task finalize interview answer sheet (881)	gdepaoli@dundee.ac.uk (De Paoli)	General	open	Fri May 29 00:00:00 2020	No
889	Task finalize interview answer sheet (881)	jaysultana@gmail.com (Sultana)	General	open	Fri May 29 00:00:00 2020	New
893	Task finalize interview answer sheet (881)	ma.loane@ulster.ac.uk (LOANE)	General	open	Fri May 29 00:00:00 2020	No
894	Task finalize interview answer sheet (881)	miriam.gatt@gov.mt (Gatt)	General	open	Fri May 08 00:00:00 2020	No
895	Task finalize interview answer sheet (881)	nvm@unife.it (Julie Neville)	General	open	Tue May 26 00:00:00 2020	No
897	Task finalize interview answer sheet (881)	schink@leibniz-bips.de (Schink)	General	open	Fri May 29 00:00:00 2020	No
898	Task finalize interview answer sheet (881)	s.e.jordan@swansea.ac.uk (Jordan)	General	open	Fri May 22 00:00:00 2020	No
899	Task finalize interview answer sheet (881)	tduarte@idiapigol.org (Duarte Salles)	General	open	Fri May 29 00:00:00 2020	No
900	Task finalize interview answer sheet (881)	ve@clin.au.dk (Ehrenstein)	General	open	Fri May 29 00:00:00 2020	No

Don't refresh this page. [Change](#)

Press ? to view keyboard shortcuts.

This screen shows the status of each responder organization.

The pink color indicates an organisation whose deadline has expired. The word 'New' in the last column indicates that there is a new message. In order to enter the page of a specific subtask and interact with the corresponding organization, click on the corresponding subtask number or subject.

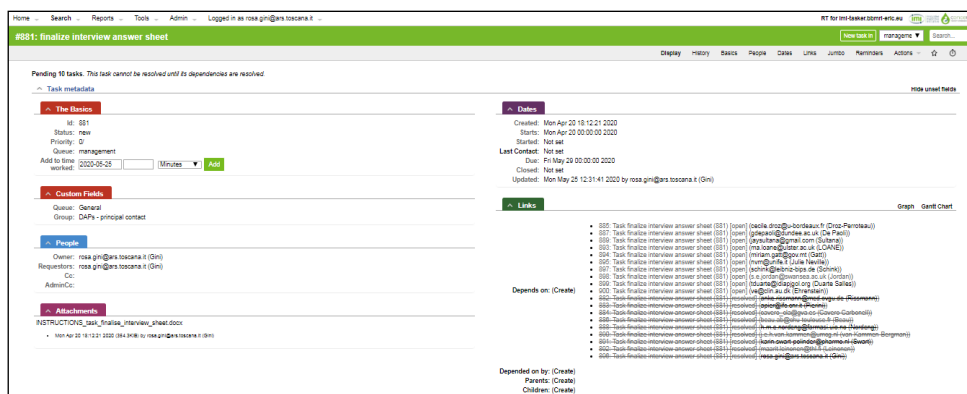
#	Subject	Owner
885	Task finalize interview answer sheet (881)	cecile.droz@u-bordeaux.fr (Droz-Perroteau)
887	Task finalize interview answer sheet (881)	gdepaoli@dundee.ac.uk (De Paoli)
889	Task finalize interview answer sheet (881)	jaysultana@gmail.com (Sultana)
893	Task finalize interview answer sheet (881)	ma.loane@ulster.ac.uk (LOANE)
894	Task finalize interview answer sheet (881)	miriam.gatt@gov.mt (Gatt)
895	Task finalize interview answer sheet (881)	nvm@unife.it (Julie Neville)
897	Task finalize interview answer sheet (881)	schink@leibniz-bips.de (Schink)

For details on how to manage the children subtasks, go to section 2.

To manage the parent task instead, from the home screen click on the link at the left.

My Request						Edit
#	Subject	Priority	Queue	Status	Due	Depends On
881	finalize interview answer sheet	0	management	(pending 10 other tickets)	Fri May 29 00:00:00 2020	#882: Task finalize interview answer sheet (881) #883: Task finalize interview answer sheet (881) #884: Task finalize interview answer sheet (881) #885: Task finalize interview answer sheet (881) #886: Task finalize interview answer sheet (881) #887: Task finalize interview answer sheet (881) #888: Task finalize interview answer sheet (881) #889: Task finalize interview answer sheet (881) #890: Task finalize interview answer sheet (881) #891: Task finalize interview answer sheet (881) #892: Task finalize interview answer sheet (881) #893: Task finalize interview answer sheet (881) #894: Task finalize interview answer sheet (881) #895: Task finalize interview answer sheet (881) #896: Task finalize interview answer sheet (881)

You will access the screen below

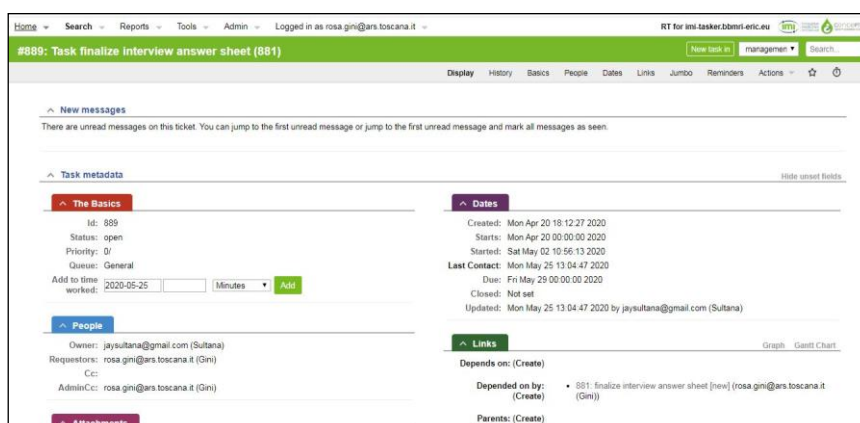


For details on how to manage the parent task, go to section 3.

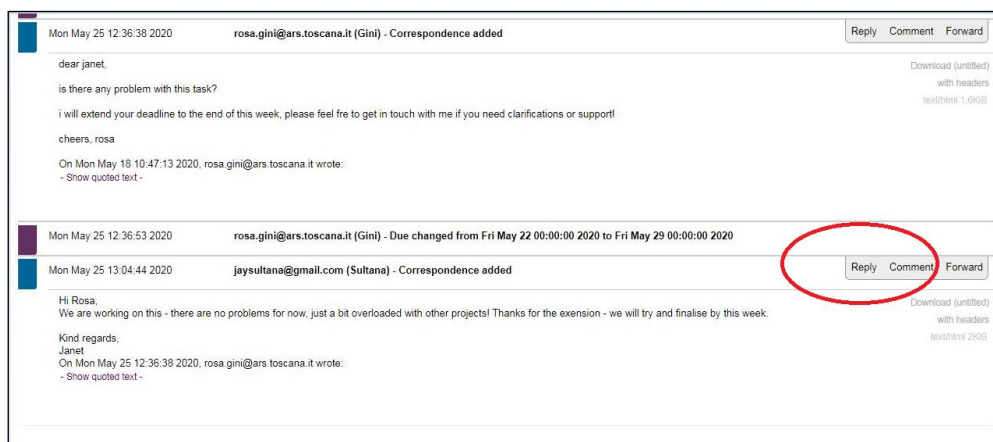
2. Manage children subtasks

2.1. Send a message to one responder organization

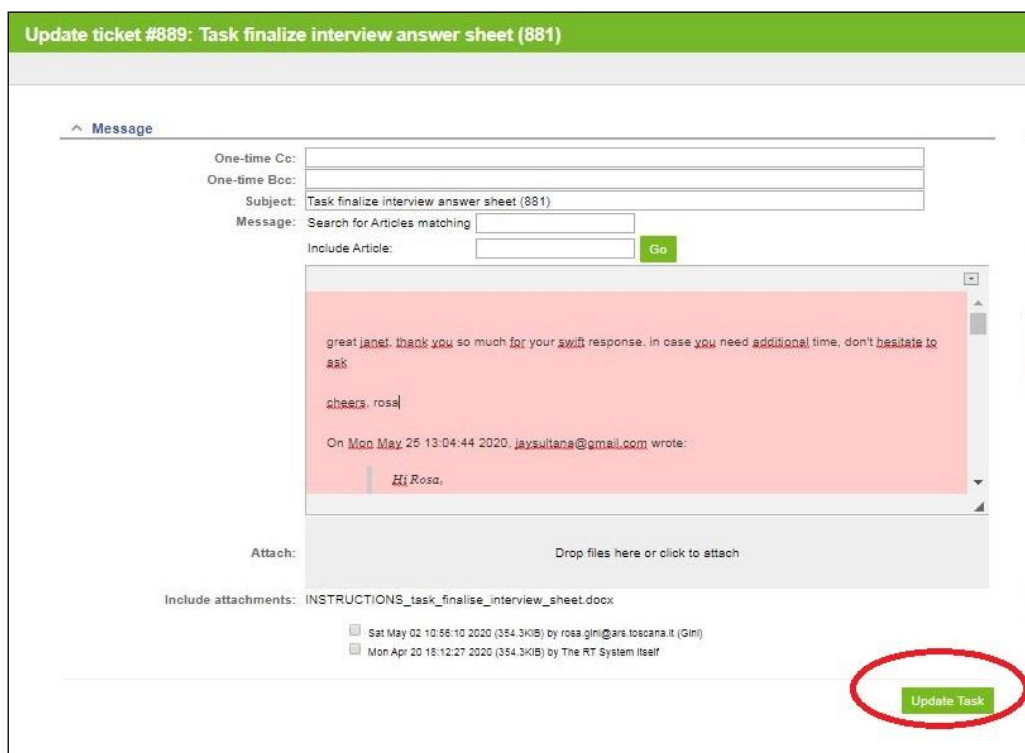
This screen shows the home of the subtask.



At the bottom of this screen you can find the history of the correspondence with the specific subtask. To reply to the last message, click on 'Reply' as in the following

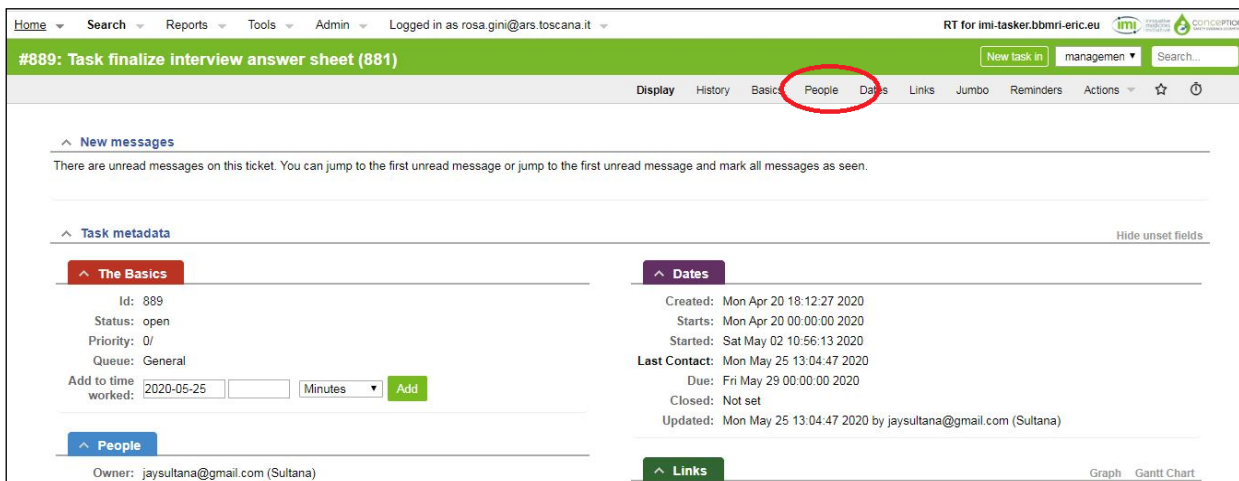


This will open the following interface where you can write your answer and click on 'Update task' to send your message. In the 'One-time Cc' box you can add email addresses of people you would like to be cc'ed (they don't need to be registered in the TMS)-



2.2. Reassign a child subtask to a new responder from the same organization

If the responder requests to reassign the subtask to a new responder, click on the 'People' tab at the top of the subtask home, as follows



Home Search Reports Tools Admin Logged in as rosa.gini@ars.toscana.it RT for imi-tasker.bbmri-eric.eu

#889: Task finalize interview answer sheet (881) New task in management Search...

Display History Basics **People** Dates Links Jumbo Reminders Actions ☆ ⌵

^ New messages
There are unread messages on this ticket. You can jump to the first unread message or jump to the first unread message and mark all messages as seen.

^ Task metadata Hide unset fields

The Basics

Id: 889
Status: open
Priority: 0/
Queue: General
Add to time worked: 2020-05-25 Minutes Add

Dates

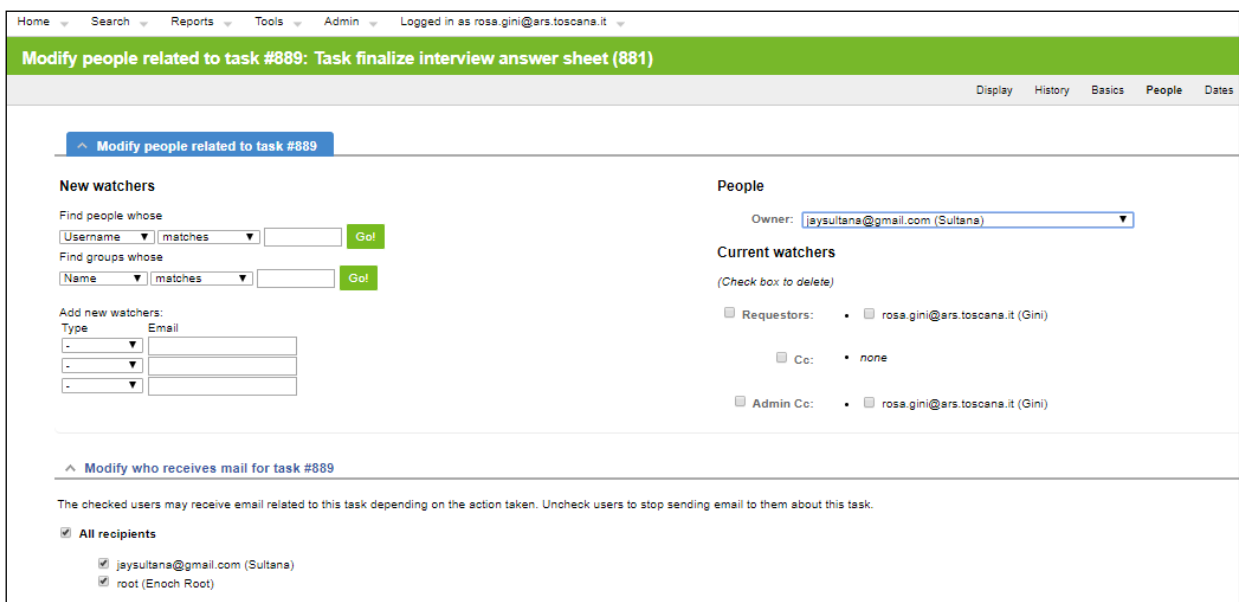
Created: Mon Apr 20 18:12:27 2020
Starts: Mon Apr 20 00:00:00 2020
Started: Sat May 02 10:56:13 2020
Last Contact: Mon May 25 13:04:47 2020
Due: Fri May 29 00:00:00 2020
Closed: Not set
Updated: Mon May 25 13:04:47 2020 by jaysultana@gmail.com (Sultana)

People

Owner: jaysultana@gmail.com (Sultana)

Links Graph Gantt Chart

You will enter this screen



Home Search Reports Tools Admin Logged in as rosa.gini@ars.toscana.it

Modify people related to task #889: Task finalize interview answer sheet (881) Display History Basics **People** Dates

^ Modify people related to task #889

New watchers

Find people whose
Username matches Go!

Find groups whose
Name matches Go!

Add new watchers:
Type Email
- -
- -
- -

People

Owner: jaysultana@gmail.com (Sultana)

Current watchers
(Check box to delete)

☐ Requestors: rosa.gini@ars.toscana.it (Gini)

☐ Cc: none

☐ Admin Cc: rosa.gini@ars.toscana.it (Gini)

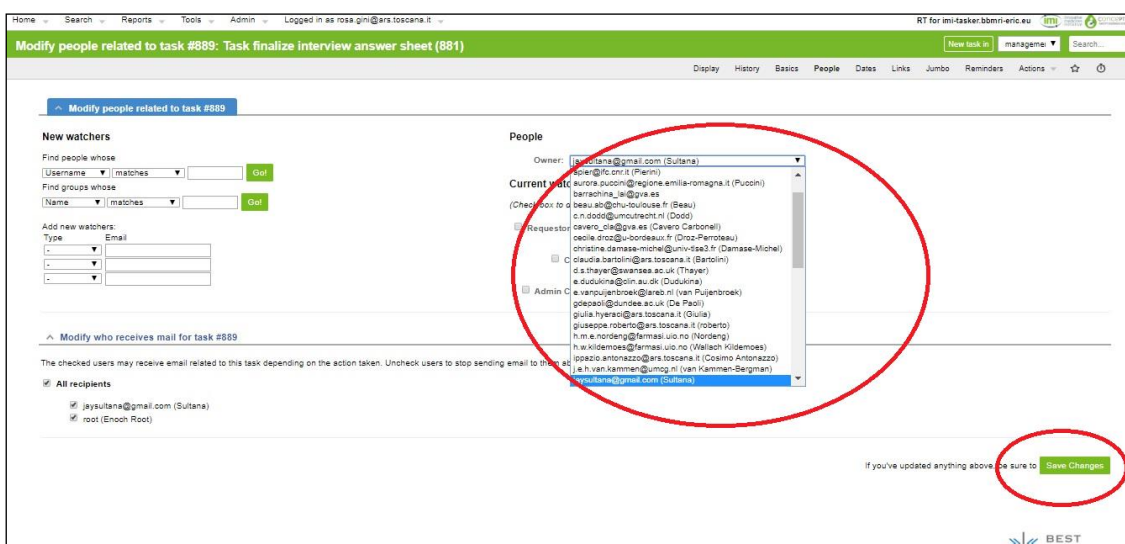
^ Modify who receives mail for task #889

The checked users may receive email related to this task depending on the action taken. Uncheck users to stop sending email to them about this task.

☒ All recipients

☒ jaysultana@gmail.com (Sultana)
☒ root (Enoch Root)

Click on 'Owner' to assign a new respondent from the list, and click on 'Save changes' to save the change, as shown below.



Home Search Reports Tools Admin Logged in as rosa.gini@ars.toscana.it RT for imi-tasker.bbmri-eric.eu

Modify people related to task #889: Task finalize interview answer sheet (881) New task in management Search...

Display History Basics **People** Dates Links Jumbo Reminders Actions ☆ ⌵

^ Modify people related to task #889

New watchers

Find people whose
Username matches Go!

Find groups whose
Name matches Go!

Add new watchers:
Type Email
- -
- -
- -

People

Owner: jaysultana@gmail.com (Sultana)

Current watchers
(Check box to delete)

☐ Requestors: rosa.gini@ars.toscana.it (Gini)

☐ Cc: none

☐ Admin Cc: rosa.gini@ars.toscana.it (Gini)

^ Modify who receives mail for task #889

The checked users may receive email related to this task depending on the action taken. Uncheck users to stop sending email to them about this task.

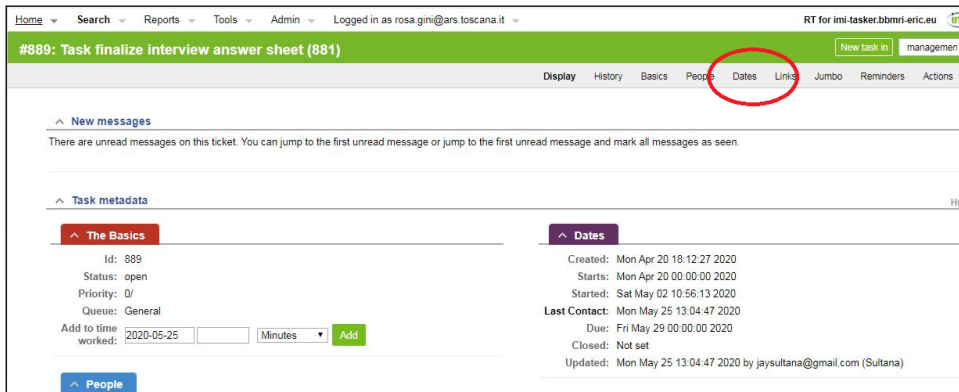
☒ All recipients

☒ jaysultana@gmail.com (Sultana)
☒ root (Enoch Root)

If you've updated anything above be sure to **Save Changes**

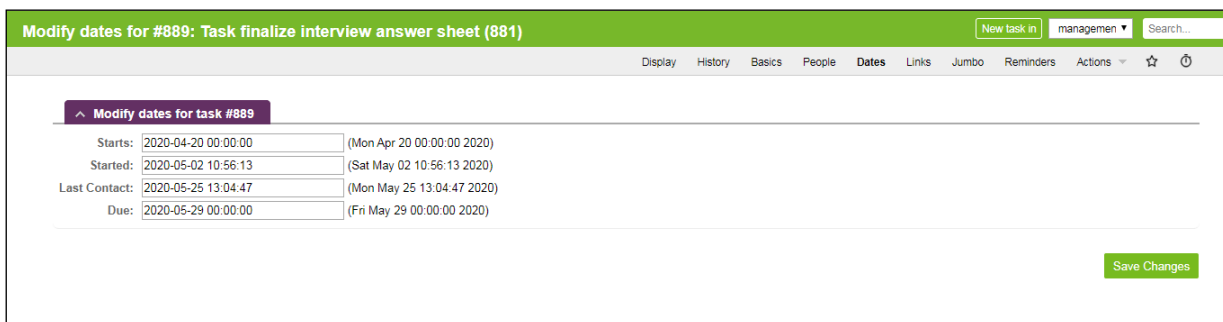
2.3. Extend the deadline for one responder

To extend the deadline of a respondent, click on the tab 'Dates' at the top of the subtask home, as follows



The screenshot shows the RT for imi-tasker.bbmri-eric.eu interface. The top navigation bar includes Home, Search, Reports, Tools, Admin, and a login status for rosa.gini@ars.toscana.it. The main header displays the task title '#889: Task finalize interview answer sheet (881)' and a 'New task in' button. Below the header, a tabbed interface is visible with tabs for Display, History, Basics, People, Dates, Links, Jumbo, Reminders, and Actions. The 'Dates' tab is currently selected and highlighted with a red circle. The main content area is divided into two sections: 'Task metadata' and 'Task details'. The 'Task metadata' section includes fields for Id (889), Status (open), Priority (0), Queue (General), and an 'Add to time worked' field. The 'Task details' section shows a 'Dates' tab with a list of dates: Created (Mon Apr 20 18:12:27 2020), Starts (Mon Apr 20 00:00:00 2020), Started (Sat May 02 10:56:13 2020), Last Contact (Mon May 25 13:04:47 2020), Due (Fri May 29 00:00:00 2020), Closed (Not set), and Updated (Mon May 25 13:04:47 2020 by jaysultana@gmail.com (Sultana)).

You will enter this screen



The screenshot shows the 'Modify dates for #889: Task finalize interview answer sheet (881)' screen. The top navigation bar includes Home, Search, Reports, Tools, Admin, and a login status for rosa.gini@ars.toscana.it. The main header displays the task title '#889: Task finalize interview answer sheet (881)' and a 'New task in' button. Below the header, a tabbed interface is visible with tabs for Display, History, Basics, People, Dates, Links, Jumbo, Reminders, and Actions. The 'Dates' tab is currently selected and highlighted with a red circle. The main content area is divided into two sections: 'Task metadata' and 'Task details'. The 'Task metadata' section includes fields for Id (889), Status (open), Priority (0), Queue (General), and an 'Add to time worked' field. The 'Task details' section shows a 'Dates' tab with a list of dates: Created (Mon Apr 20 18:12:27 2020), Starts (Mon Apr 20 00:00:00 2020), Started (Sat May 02 10:56:13 2020), Last Contact (Mon May 25 13:04:47 2020), Due (Fri May 29 00:00:00 2020), Closed (Not set), and Updated (Mon May 25 13:04:47 2020 by jaysultana@gmail.com (Sultana)).

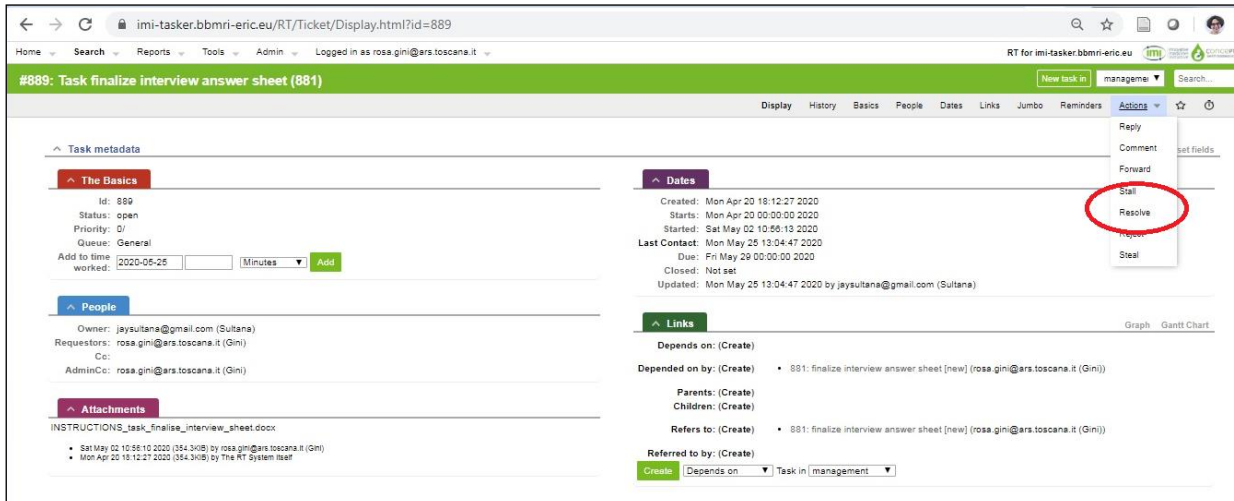
Click the field 'Due' to modify the deadline, and on 'Save changes' to make you changes permanent.



The screenshot shows the 'Modify dates for #889: Task finalize interview answer sheet (881)' screen. The top navigation bar includes Home, Search, Reports, Tools, Admin, and a login status for rosa.gini@ars.toscana.it. The main header displays the task title '#889: Task finalize interview answer sheet (881)' and a 'New task in' button. Below the header, a tabbed interface is visible with tabs for Display, History, Basics, People, Dates, Links, Jumbo, Reminders, and Actions. The 'Dates' tab is currently selected and highlighted with a red circle. The main content area is divided into two sections: 'Task metadata' and 'Task details'. The 'Task metadata' section includes fields for Id (889), Status (open), Priority (0), Queue (General), and an 'Add to time worked' field. The 'Task details' section shows a 'Dates' tab with a list of dates: Created (Mon Apr 20 18:12:27 2020), Starts (Mon Apr 20 00:00:00 2020), Started (Sat May 02 10:56:13 2020), Last Contact (Mon May 25 13:04:47 2020), Due (Fri May 29 00:00:00 2020), Closed (Not set), and Updated (Mon May 25 13:04:47 2020 by jaysultana@gmail.com (Sultana)). The 'Due' field is highlighted with a red circle, and a date picker calendar is open, showing the month of May 2020. The 'Save Changes' button is also highlighted with a red circle.

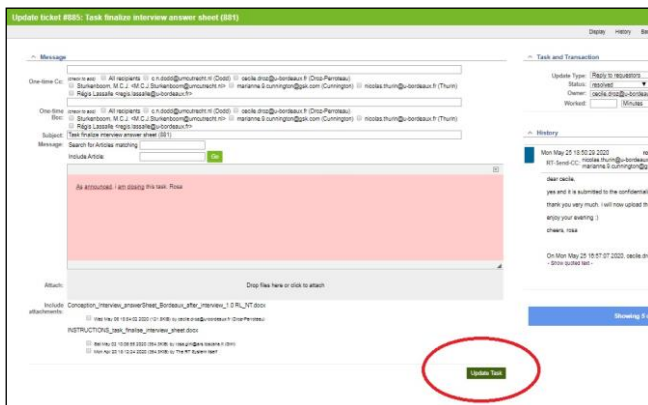
2.4. Resolve the child subtask for one responder

To close a subtask, click on 'Actions>Resolve', as follows.



The screenshot shows the RT ticket display page for task #889: "Task finalize interview answer sheet (881)". The "Actions" menu is open, and the "Resolve" option is highlighted with a red circle. The page includes sections for Task metadata (The Basics, Dates, People, Attachments), Links (Depends on, Depended on by, Parents, Children, Refers to, Referred to by), and a "Create" button.

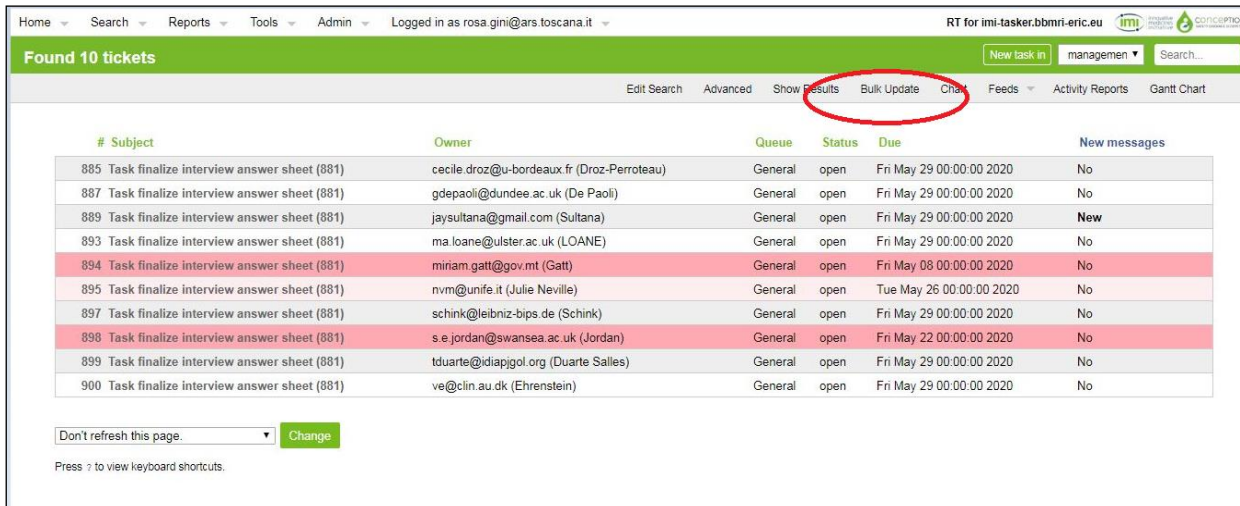
You will enter a screen similar to the correspondence screen. Announce closure and click on 'Update task'.



The screenshot shows the "Update ticket #889: Task finalize interview answer sheet (881)" screen. The "Update Task" button is highlighted with a red circle. The page includes a message input area, a "Send" button, and a "History" section.

2.6 Send a bulk update to many children subtasks

To send the same message to all or a subsets of the responders, access this screen (see section 1) and press 'Bulk update', as follows



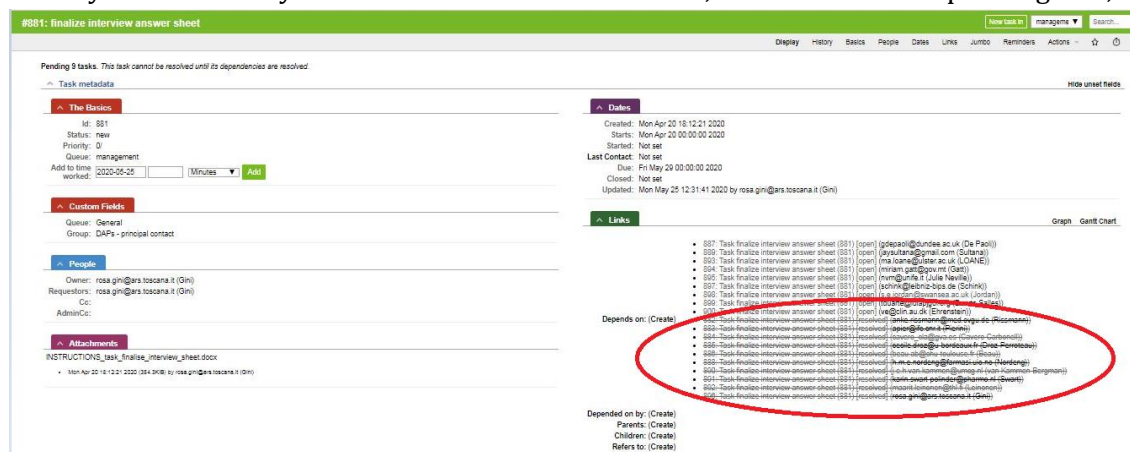
The screenshot shows the RT ticket list page. The "Bulk Update" button is highlighted with a red circle. The page displays a table of tickets with columns for #, Subject, Owner, Queue, Status, Due, and New messages.

#	Subject	Owner	Queue	Status	Due	New messages
885	Task finalize interview answer sheet (881)	cecile.droz@u-bordeaux.fr (Droz-Perroteau)	General	open	Fri May 29 00:00:00 2020	No
887	Task finalize interview answer sheet (881)	gdepaoli@dundee.ac.uk (De Paoli)	General	open	Fri May 29 00:00:00 2020	No
889	Task finalize interview answer sheet (881)	jaysultana@gmail.com (Sultana)	General	open	Fri May 29 00:00:00 2020	New
893	Task finalize interview answer sheet (881)	ma.loane@ulster.ac.uk (LOANE)	General	open	Fri May 29 00:00:00 2020	No
894	Task finalize interview answer sheet (881)	miriam.gatt@gov.mt (Gatt)	General	open	Fri May 08 00:00:00 2020	No
895	Task finalize interview answer sheet (881)	nvm@unife.it (Julie Neville)	General	open	Tue May 26 00:00:00 2020	No
897	Task finalize interview answer sheet (881)	schink@leibniz-bips.de (Schink)	General	open	Fri May 29 00:00:00 2020	No
898	Task finalize interview answer sheet (881)	s.e.jordan@swansea.ac.uk (Jordan)	General	open	Fri May 22 00:00:00 2020	No
899	Task finalize interview answer sheet (881)	tduarte@diapigol.org (Duarte Salles)	General	open	Fri May 29 00:00:00 2020	No
900	Task finalize interview answer sheet (881)	ve@clin.au.dk (Ehrenstein)	General	open	Fri May 29 00:00:00 2020	No

3. Manage the parent task

3.1. Reaccess a child subtask that was resolved

The home of the parent task shows one row the child subtask, and those which have been resolved already are barred. If you want to access the closed task, click on the corresponding row, as follows



The screenshot shows the 'finalize interview answer sheet' task page. The 'Links' tab is active, displaying a list of child tasks. A red circle highlights a row for task ID 881, which is the parent task being accessed. The list includes tasks such as 'Task finalize interview answer sheet (881)' and 'Task finalize interview answer sheet (882)'. The 'Dates' tab is also visible, showing the task's timeline.

3.2. Extend the deadline or resolve the parent task

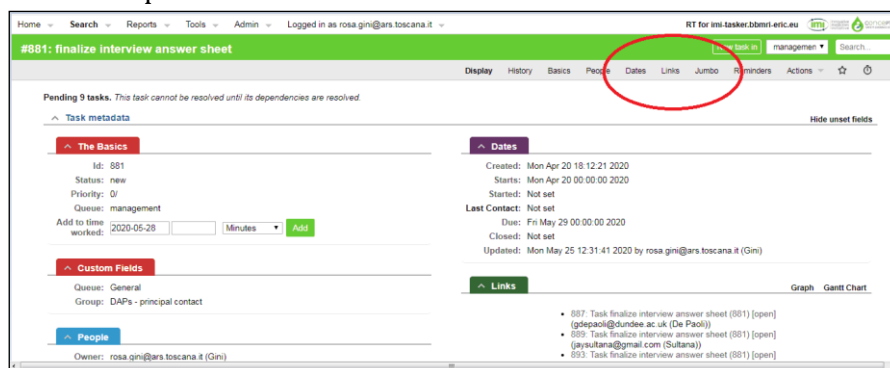
The parent task can have the deadline extended, as in the case of a subtask (see Section 2).

Only when all the children subtasks have been resolved can the father task be closed. To close the task do as indicated in Section 2.

3.3 Add a new child subtask to an existing parent

If a new partner joins after a parent task has already be launched, proceed as follows to create a new child subtask for the existing parent

1. Create a new task with the same name for the new partner, and take note of its number
2. Access the parent task and click on the tab 'Links'



The screenshot shows the 'finalize interview answer sheet' task page. The 'Links' tab is active, displaying a list of child tasks. A red circle highlights the 'Links' tab in the top navigation bar. The list includes tasks such as 'Task finalize interview answer sheet (881)' and 'Task finalize interview answer sheet (882)'. The 'Dates' tab is also visible, showing the task's timeline.

3. You are in the following screen; enter the number of the new child task in the window 'Referred to by' (you will be prompted with a list of matching numbers)

4. Similarly for the window 'Depended on by'

- At the bottom of the page click on 'Save changes'

52