

Fostering professionalism and integrity in research



**Final report of the Taskforce Scientific Integrity
Erasmus University Rotterdam
October 2013**

Editors final report:

Monique van Donzel, Geske Dijkstra and Finn Wynstra

Administrative support:

David van Ass, Esmee Arends and Eva Haaijer

Cover illustration:

Designapplause.com

Table of Contents:

Summary of recommendations	4	
Chapter 1	On the taskforce: background, mandate and projects	7
Chapter 2	Challenges in fostering professionalism and integrity in research	12
Chapter 3	Research Data Management	20
Chapter 4	Training	22
Chapter 5	Dilemma game	26
Chapter 6	Pledge-taking	28
Chapter 7	Seminars	32
Chapter 8	Process and output of PhD projects	34
Chapter 9	External relations	37
Chapter 10	Monitoring	40
Appendix 1	Report: Research Data Management at Erasmus University Rotterdam	
Appendix 2	PhD courses on scientific integrity	
Appendix 3	Procedure pledge-taking ceremony ERIM PhD students	

Summary of recommendations:

This chapter provides the summary of the recommendations as given in the chapters 3 to 10.

Each of the chapters reports on one of the sub-projects, which have been organized around key priorities following the table below (see for details Chapter 2).

Area/Project	Priorities	Policies for research data management	Awareness creation	Organized 'peer pressure' in the front end	Incentives and opportunities for peer review	Maintaining independence and impartiality	Improved transparency on procedures
Research data management		•					
Training	PhD and faculty training		•				
Dilemma game			•				
Pledge-taking			•				
Organised feedback	Seminars			•			
	Process and output of PhD projects			•			
External relations	Contract research					•	
	Media relations		•				
Monitoring	Integrity coordinators						•

Research data management:

- Appoint a **research data support officer at the central EUR level**. Dedicated support at a central level is crucial for continuity and for sharing good practices between institutes or disciplines.
- Draft a covenant to the extent that the deans of research or research directors and – at Erasmus Medical Centre – the heads of department **embed and evaluate research data management in their institute**. They should have room to describe how they will fulfil the data management minimum.
- Introduce the deans of research to the EUR data management minimum protocol.
- Provide **training and advice on research data management**: delegated to the university library (see below).
- Provide **services for storage and retention of research data**: delegated to SSC ICT (see below).
- Coordinate, supervise and stimulate **data management workflows and protocols**. In particular, refine the minimal EUR protocol such that it fits the institute's research discipline(s) and make it practical. Take care to explain that data management also includes documentation related to the research process.
- Socialise young researchers into responsible ways of working. Research data management is not something extra; the attitude should be that responsible data management is integral to professional research.
- Develop the library's virtual desk into a **front office** for researchers, as a central point of expertise in research data management. This includes training and collaboration with long-term archives (the back office). Close collaboration with the Research Support Office (currently under development) is foreseen.
- Develop and maintain an **activating data support website** and select/develop relevant courses/workshops.
- Create **safe storage and backup facilities** for individual researchers, as well as safe ways to share and collaborate on research data.
- Maintain and **offer expertise** to make research staff aware of advantages and risks of particular storage platforms and media, in collaboration with the front office.

Training:

- PhD training includes at least three **sessions on professional and ethical standards** and scientific integrity.
- Specific integrity issues related to particular research methods are also part of the PhD training program, either as integral part of these method courses or as a separate course.
- Scientific integrity becomes a **recurrent theme** during annual research days and should be integrated on a menu basis in all training programs for researchers.
- The **course on academic leadership** current offered to those attaining leadership positions should contain a module on scientific integrity.
- The Executive Board should encourage the VSNU to develop **on-line training resources**.
- A specific course should be offered to **university administrators** to make them more familiar with the formal processes involved in reporting and investigating misconduct.

Dilemma-game:

- The game is **part of the PhD training and faculty training sessions** on research integrity.
- All faculty members that are not already participating in the ‘standard’ training sessions (which mainly pertain to PhD students, tenure trackers and associate professors), regularly participate in a dilemma-game session (suggestion: every two years).
- **Senior faculty shows commitment** to openly discussing professionalism and integrity issues in such settings by participating in these dilemma game sessions.
- Directors of research should **actively encourage** the use of the dilemma-game; the exact way the game is used is up to institute/department/group leaders.

Pledge-taking (“integriteitsverklaring”):

- All EUR researchers, existing and new, take a **pledge related to professionalism and integrity** in research. Preferably, professors and senior researchers take up their leadership role by taking the pledge as the first group in their school/institute.
- New EUR employees involved in research are given “The Netherlands Code of Conduct for Scientific Practice” upon appointment.
- The Executive Board should **define the final procedure** and text of the pledge.
- Directors of research to implement the pledge-taking procedure in their local school/institute, so that every researcher has taken the pledge by December 2014.

Seminars:

- Senior researchers should be held responsible for a **frequent and active participation** in research seminars;
- Researchers should be obliged or strongly encouraged to present their work in an **early stage of research**, e.g. about the research design;
- Organization of an **annual graduate research** day at the school or university level.

PhD project process and output:

- The **supervision plan** should address transparency of authorship, and additional sources of advice/feedback.
- A discussion is facilitated on **guidelines for data collection**, and that these address the following integrity risks: replicability, verifiability, independence of data, fairness to subjects, professionalism in the use of data collection methods, fairness to researchers.
- **Plagiarism checks** are to be performed on at least two occasions during the doctoral education.
- The Executive Board revises the **doctoral regulations**, and considers a discussion in VSNU context about a further strengthening of the doctoral education as far as scientific integrity is concerned.

In particular, we suggest the following items to be addressed:

- transparency on the **contributions** of possible co-authors
- allowing the possibility for the inner doctoral committee to provide **formal feedback**
- ensuring that at least one member of the inner doctoral committee is from **outside EUR**, and at least one additional member of the plenary committee is from outside EUR
- the doctoral committee explicitly assesses the manuscript on the principles in the **VSNU code**

Contract research:

- In order to secure academic independence and a level playing field with other universities, the EUR should continue to push for **changes in the Dutch government's general terms for contract research** (ARVODI) that are currently in conflict with the right to publish results of contract research.
- Faculties and researchers should give **due attention to possible conflicts of interest** arising from contract research and from endowed chairs, and should avoid these conflicts as much as possible.
- It is impossible to define general rules on ownership of data collected in contract research, but in all contracts, due **attention should be given to defining data ownership**.
- Data collected during contract research should be stored appropriately and be available for peer review; the recommendations of chapter 6 apply for contract research as well.
- *Realized: the right to publish results of contract research is part of the approved general EUR Contract Terms for Commissioned Research*
- *Realized: EUR Contract Terms for Commissioned Research stipulate that commissioning parties can only publish outcomes of contract research after approval of the researchers.*

Media relations:

- All EUR researchers are made aware of **key professionalism and integrity considerations in media relations**.
- Central and local marketing & communication departments explicitly integrate these considerations in their policies, documentation and training.
- *Realized: The considerations identified have been included in a new brochure recently published by SMC.*

Monitoring:

- Appoint one **coordinator** per school for all integrity issues, but allow for the possibility to appoint one specific coordinator for scientific integrity.
- Following the suggested tasks and responsibilities as detailed (see Chapter 10).

Implementation and follow-up:

As Taskforce, we identify two possible modes of implementation and follow-up:

- *Decentral*: Deans/directors of research are responsible for implementation. The Executive Board identifies its key priorities among the Taskforce recommendations, and tangible, time-specific targets are integrated in the strategic covenants with the schools/institutes.
- *Coordinated*: As above, but a small university-wide advisory body remains which advises schools/institutes on implementation issues, and reports to the Executive Board on progress made.

We advise against the following:

- *Centralized implementation by a taskforce*: Professionalism and integrity are core issues, the implementation responsibility for which belongs to the existing management structures of the university/schools.
- *Implementation by the integrity coordinators*: The local integrity coordinators do not always have the position or time to supplement their counselling role with a more strategic, administrative role. Obviously, in individual cases, these roles may be combined.

Chapter 1: On the taskforce: background, mandate and projects

*“After all, we are being paid not to brush up our CV,
but to learn more about the truth.”
(Hilde Huizenga, employee UvA, interview in MindOpen, Summer 2012)*

1.1 Background

Following the ‘Stapel-affair’ and the growing discussion in academia in the Netherlands and beyond on fraud in science, several schools at Erasmus University were analyzing whether and how more attention should be given to professionalism and integrity in research. These initiatives gained even more urgency, when two internal investigations of misconduct in research were finalized in the course of 2012.

During the Summer of 2012, various schools and research institutes approached the university Executive Board with initiatives, most notably in the area of training and awareness. The Executive Board, in addition, had a keen interest in initiating plans for the management of research data. In order to reap synergies between the different initiatives where possible, a taskforce was established to identify, coordinate and develop key initiatives.

Starting from a focus on training and research data management, the taskforce then set out to identify a complete list of priorities. Based on several current reports mainly by the KNAW, the main categories for possible measures to foster professionalism and integrity in research were defined as follows¹:

- Awareness creation (training, pledge-taking) (1, 2, 3)
- Policies for research data management (1, 2, 3)
- Organized ‘peer pressure’ in the front end (1, 3)
- Incentives and opportunities for peer review (1, 2, 3)
- Maintaining independence and impartiality (objective reporting, data ownership) (3)
- Improved awareness/transparency of procedures (1, 2, 3)
- Consideration of integrity issues in appraisal talks (1, 3)
- Encouraging replication studies (2)

This list of categories provided an important element for the prioritization of possible projects by the Taskforce, as discussed in more detail below.

1.2 Assignment and mandate

The Taskforce Scientific Integrity has been established with the objective to raise awareness for and to develop proposals to help maintain scientific professionalism and integrity. The taskforce was established by the

¹ Numbers indicate in which specific reports these measures have been identified:

1: KNAW (2012), Zorgvuldig en integer omgaan met wetenschappelijke onderzoeksgegevens, Advies van de KNAW Commissie Onderzoeksgegevens (“Commissie Schuyt”), September

2: Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, www.commissielevelt.nl

3: KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013. www.knaw.nl

Executive Board of the Erasmus University Rotterdam on 28 June 2012, and would operate until September 2013.

The mandate of the taskforce has been to provide recommendations to the Executive Board regarding institutional measures to help foster professionalism and integrity in research. Such recommendations could range from quite specific, to relatively more generic – for instance, where heterogeneity in research practice and contexts across the EUR would ultimately require more tailored, local measures.

1.3 Process

Between August 2012 and October 2013, the Taskforce has held seven group meetings. During the first three meetings, the group made an inventory of problem areas and priority projects. This resulted in a list of nine main projects, as detailed below.

The Taskforce chair met regularly with the Rector Magnificus to discuss the scope and planning of the activities. At three moments, the Taskforce has reported in the meeting of research directors (“Onderzoeksdirecteuren overleg”). These presentations served to keep the research directors informed of the scope and progress of the activities of the Taskforce, and gather their input on the main projects.

Prof. dr Kees Schuyt, chairman of the “Landelijk Orgaan Wetenschappelijke Integriteit” (National Bureau for Scientific Integrity) and chairman of the KNAW Committee Schuyt (see Chapter 2) has generously provided advice on the activities of the taskforce and on the final report.

1.4 Projects

Based on a broad inventory of challenges faced at the different EUR schools and institutes, and a cross-check with the categories of measures of improvement identified by most notably the committees Schuyt and Levelt/Noort/Drenth (see 1.1), a list of priority projects was defined at the end of 2012. Nine projects were selected from this list, based on the shared urgency of the issue, and the possibility to develop general, university-wide recommendations.

As can be seen in Table 1.1, most of the categories identified were to be covered by one or several projects.

While two taskforce projects specifically address PhD students (training and the PhD project process), we should explicitly point out that there is no reason to suspect that PhD's encounter bigger or more frequent dilemma's in doing research, let alone are more susceptible to scientific fraud.

Still, every researcher's career and training starts with doing a PhD, and hence it is important to start there with paying more explicit attention to professionalism and integrity.

The project on training is also specifically targeting regular faculty.

There are three categories that have not resulted in specific projects by the Taskforce:

- Incentives and opportunities for peer review
- Consideration of integrity issues in appraisal talks
- Encouraging replication studies

In the current timeframe, these initiatives were thus defined as less urgent for the schools combined. This does not exclude the possibility that individual schools or institutes may want to address one or several of these. We provide some initial reflections in the final section of Chapter 2.

Table 1.1: Matching taskforce projects with key priorities

Area/Project	Priorities	Policies for research data management	Awareness creation	Organized 'peer pressure' in the front end	Incentives and opportunities for peer review	Maintaining independence and impartiality	Improved transparency on procedures	Consideration of integrity issues in appraisal talks	Encouraging replication studies
Research data management		●							
Training	PhD and faculty training		●						
Dilemma game			●						
Pledge-taking			●						
Organised feedback	Seminars			●					
	Process and output of PhD projects			●					
External relations	Contract research					●			
	Media relations		●						
Monitoring	Integrity coordinators						●		

The nine projects cover five main areas: research data management; feedback; training; external relations; and monitoring. The projects and their respective deliverables are detailed in Table 1.2. Working group members also included ‘specialists’ that were not member of the overall taskforce.

One project not discussed in this report concerns the exhibition Derivatives, featuring paintings by ISS colleague Peter van Bergeijk. This travelling exhibition investigates the issue of originality in the context of (self) plagiarism. The different views in the Arts and the scientific discourse are the point of departure for discovering how ideas that are identical can still be completely different and new, but also that ‘original’ works of art can be repetitive reproduction. By organising an exhibition, we want to raise awareness of (particular aspects of) research integrity.

Chapters 3 through 10 provide a detailed discussion and the task force recommendations for each of the different projects. Further general background for the taskforce initiatives is provided in the next chapter, by a brief review of recent studies and reports on fraud and questionable practices in research.

1.3 Taskforce members

The taskforce encompassed representatives from all EUR schools/institutes, representing a mix of active research faculty, research directors and policy advisors. Below are the names of the representatives, with functions specified where directly relevant to the work of the Taskforce.

1. Dr. Maarten van Dijck (ESHCC)
2. Prof. dr Geske Dijkstra (FSW)
3. Dr. Monique van Donzel, *Secretary* (ABD, senior policy advisor)
4. Drs. Gert Goris (UB, Head of Scientific Information Services)
5. Prof. dr Maria Grever (ESHCC, integrity coordinator)
6. Prof. dr Peter de Jaegere (EMC)
7. Dr. Rikard Juttmann (EMC, senior policy advisor research integrity)
8. Prof. dr Muel Kaptein (RSM, integrity coordinator)
9. Drs. Riëtte te Lindert (ABD, policy advisor/ Committee Scientific Integrity)
10. Dr. Sonja Meeuwsen (iBMG, integrity coordinator)
11. Prof. dr Ingrid Robeyns (FW, Director Dutch Research School of Philosophy)
12. Prof. dr Erik Schut (iBMG, director of research)
13. Prof. dr Irene van Staveren (ISS, chair PhD program)
14. Drs. Annet van der Veen (ESL)
15. Dr. Freddy van der Veen (FSW)
16. Prof. dr Bauke Visser (ESE, general director Tinbergen Institute)
17. Prof. dr Finn Wynstra, *Chair* (RSM, associate director ERIM)

Table 1.2: List of taskforce projects

Area	Project	Deliverables	Chapter	Working group members (groupleader in bold)
Research data management		<ul style="list-style-type: none"> - Inventory of user requirements for storing and archiving research data - Recommendations for data management protocols, workflows and platforms 	3	Maarten van Dijck (ESHCC), Monique van Donzel (ABD), Gert Goris (UB), Rikard Juttmann (EMC), Freddy van der Veen (FSW), Finn Wynstra (RSM/ERIM), Marjan Grootveld (DANS, projectleader) , Roel Otten (UB), Gerrit Jan de Bie (EBL)
Training	PhD and faculty training	<ul style="list-style-type: none"> - Research ethics course templates 	4	Bauke Visser (ESE/Tinbergen), Irene van Staveren (ISS), Annet van der Veen (ESL),
Dilemma game		<ul style="list-style-type: none"> - Ready-to-use group-based dilemma-game 	5	Monique van Donzel (ABD), Geske Dijkstra (FSW), Muel Kaptein (RSM), Finn Wynstra (RSM/ERIM)
Pledge-taking		<ul style="list-style-type: none"> - Recommendations scope and format pledge-taking ceremony 	6	Ingrid Robeyns (FW), Annet van der Veen (ESL)
Organised feedback	Seminars	<ul style="list-style-type: none"> - Overview of possible seminar forms, especially to support early feedback 	7	Sonja Meeuwsen (iBMG), Erik Schut (iBMG)
	Process and output of PhD projects	<ul style="list-style-type: none"> - Recommendations for additional checks and balances in PhD process, to increase early feedback 	8	Irene van Staveren (ISS), Maria Grever (ESHCC), Bauke Visser (ESE/Tinbergen)
External relations	Contract research	<ul style="list-style-type: none"> - Recommendations regarding regulation of data ownership and protection academic independence 	9.1	Geske Dijkstra (FSW), Sonja Meeuwsen (iBMG), Erik Schut (iBMG), Sadjie Theeuwes (ABD)
	Media relations	<ul style="list-style-type: none"> - Revised guidelines for researchers how to communicate with formal media 	9.2	Finn Wynstra (RSM/ERIM), Sandra van Beek (SMC)
Monitoring	Integrity coordinators	<ul style="list-style-type: none"> - Clarified task descriptions of local integrity coordinators 	10	Riette te Lindert (ABD), Finn Wynstra (RSM/ERIM)

Chapter 2: Challenges in fostering professionalism and integrity in research

*"It is the paradox of research that the reliance on truth
is both the source of modern science and engineering's enduring resilience and its intrinsic fragility."*

Walter E. Massey, Director National Science Foundation US, 1991

This chapter provides a brief review of the challenges that individual researchers, institutions and science at large are facing with respect to maintaining and fostering professionalism and integrity in research. This review is strongly based on recent reports pertaining to the general situation in academia in the Netherlands by institutes such as the Royal Netherlands Academy of Arts and Sciences (KNAW).

We should explicitly point out that the brief for the taskforce was not to identify and analyse prevalence and causes of research misconduct and questionable research practices within Erasmus University. Within the limited timeframe, the taskforce therefore started out from recent analyses for Dutch academia at large (as referenced below). Then, through extensive debates in the first few months within the taskforce, but also with the communities within our respective schools and institutes, we agreed on a shared understanding of the main challenges in fostering (nurturing, maintaining) professionalism and integrity in research.

2.1 Research integrity

Research integrity can be defined as “the quality of possessing and steadfastly adhering to high moral principles and professional standards, as outlined by professional organizations, research institutions and, when relevant, the government and public”². In the Netherlands, research integrity is most explicitly defined in “The Netherlands Code of Conduct for Scientific Practice”³, which is applicable to every university scientist in the Netherlands. This ‘VSNU Code’ specifies five principles⁴:

- **Scrupulousness**
 - Scientific activities are performed scrupulously, unaffected by mounting pressure to achieve.
- **Reliability**
 - A scientific practitioner is reliable in the performance of his/her research and in the reporting, and equally the transfer of knowledge through teaching and publication.
- **Verifiability**
 - Whenever research results are publicized, it is made clear what the data and the conclusions are based on, where they were derived from and how they can be verified.
- **Impartiality**
 - The scientific practitioner heeds no other interest than the scientific interest.
- **Independence**
 - Scientific practitioners operate in a context of academic liberty and independence. Insofar as restrictions of that liberty are inevitable, these are clearly stated.

The VSNU code is relatively concise, compared to most other codes of research conduct. Reviewing an international set of such codes, the Royal Netherlands Academy of Arts and Sciences recently summarized these codes in terms of seven principles⁵:

² Steneck, N.H. (2006). Fostering integrity in research: definitions, current knowledge, and future directions. *Science and Engineering Ethics* 12: 53–74; quote from p. 55

³ Association of Universities in the Netherlands (VSNU), 2012. www.vsnu.nl

⁴ The introduction to the code specifies “transparency” as an overarching principle

⁵ KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013. www.knaw.nl

- **Honesty** – complete reporting;
- **Fairness** – respectful treatment of colleagues;
- **Objectivity** – subjective perceptions should not influence research;
- **Reliability** – using accepted methods of inquiry and analysis;
- **Scepticism** – professional doubt in exercising control and monitoring;
- **Accountability** – answering to other researchers and society at large;
- **Openness** – providing access where possible to methods, data and results.

Based on a comparison of these principles and the current VSNU code, the KNAW then recommends to supplement the code with the explicit principles of ‘honesty’ and ‘accountability’.

It is not our task here to evaluate this proposal, but even though explicating these principles may be useful, one could argue that ‘honesty’ is very closely related to ‘scrupulousness’ and ‘verifiability’. ‘Accountability’ may indeed be a complementary principle, although it is mainly the specific meaning that the KNAW report adheres to ‘accountability’, including accountability for the choice of research topics, which makes this a specific principle. This shifts the discussion somewhat from research integrity to research ethics; from doing things right, to doing the right things.

As a taskforce, we have chosen to focus on research integrity ‘proper’, and also for that reason to stick to the principles outlined in the VSNU code.

2.2 Fraud versus questionable research practices

Research misconduct or fraud is typically defined as consisting of three forms: fabrication (of data, results), falsification, and (self-) plagiarism (‘FFP’). Questionable research practices (QRP) form the more or less ‘grey zone’ of less-than-ideal scientific practices, and include among others the following⁶:

- Dropping observations or data points from analyses based on a gut feeling that they were inaccurate;
- Failing to present relevant data that contradict one’s own research;
- Publishing the same data or results in two or more publications without full disclosure;
- Inadequate record keeping related to research projects.

While a clear definition of fraud is important from a legal perspective, the definition is less important from the perspective of “truth-finding”. Plagiarism is less problematic in terms of truth-finding than various questionable research practices, such as an opaque procedure for dropping outliers⁷.

Another argument for addressing both research misconduct and QRP lies in the observation that researchers who commit research misconduct are typically starting by committing QRP (flawed or sloppy science), as has been demonstrated in the Stapel-affair and the recent misconduct case at Erasmus Medical Center⁸.

Addressing QRP could therefore, ultimately, also help to prevent misconduct.

For these reasons, the Taskforce has explicitly defined as its objective scope raising awareness for and developing proposals to help foster scientific *professionalism and integrity*, addressing both opportunities for research misconduct and questionable research practices (QRP).

⁶ See for specific examples related to experimental research: Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, pp. 49-52. www.commissielevelt.nl

⁷ Steneck, N.H. (2006). Fostering integrity in research: definitions, current knowledge, and future directions. *Science and Engineering Ethics* 12: 53–74

⁸ Erasmus MC Commissie Vervolgonderzoek (2012), Rapport vervolgonderzoek naar mogelijke schending van de wetenschappelijke integriteit, September 2012, Erasmus Medisch Centrum, <http://www.erasmusmc.nl/1172194/2090115/rapp.vervolgondz.wi>

The importance of focusing on questionable research practices is echoed by a key study published in Nature some years ago: “Our findings suggest that (...) scientists engage in a range of behaviors extending far beyond FFP that can damage the integrity of science. *Attempts to foster integrity that focus only on FFP therefore miss a great deal.*”⁹

Currently, the term “responsible conduct of research” is becoming increasingly used to refer to conduct that both avoids FFP and QRP. We have adopted the term scientific professionalism and integrity as a synonym for responsible conduct of research.

2.3 Prevalence

Despite the recent attention to highly-publicized cases of research misconduct, such as the Stapel case that was brought to light in the Fall of 2011, various reports claim that, without proper scientific research, it cannot be firmly established that fraud and questionable research practices are becoming more prevalent in Dutch academia¹⁰. While further research may be necessary, there are several indications that misconduct and questionable research practices are quite widespread:

- A large study commissioned by the US Office of Research Integrity estimates that about 1.5 % of all US biomedical research published is fraudulent¹¹;
- In a meta-analysis of 18 survey studies that directly ask scientists on the prevalence of research misconduct, average admission rates were 3 % (self-rated), respectively 17 % (non-self); regarding the prevalence of questionable research practices (6 studies), average admission rates were 10 % (self-rated) and 29 % (non-self)¹²;
- Based on a survey among US psychology researchers, estimates are that more than 50 % of researchers fail to report all dependent measures; collect more data after seeing whether results are significant; selectively report positive results; and claim to have predicted an unexpected finding¹³;
- A study of 2,047 article retractions in biomedicine between 1975-2012 revealed that 67 % of these were due to misconduct, and that these retractions have increased 10-fold since 1975 as percentage of published articles)¹⁴; one of the reasons for these increasing retraction rates is also that the ‘barriers to retraction’ have decreased – i.e. there are now more reasons for retraction than before (e.g. duplicate publication)¹⁵.

Other studies provide corroborating evidence (of the increase of QRP), such as the gradual disappearance of negative research findings (especially in the social sciences), suggesting that “research is becoming less pioneering and/or that the objectivity with which results are produced and published is decreasing”¹⁶.

Regarding the situation in the Netherlands, Dutch newspaper NRC studied investigations of scientific misconduct between January 2005 (when the VSNU code was implemented) and January 2012 at 12 universities and 8 academic medical centers. (Of the 102 investigations reported, 8 occurred at EUR/EMC.)

⁹ Martinson, B.C., Anderson, M.S. and De Vries, R. (2005). Scientists behaving badly. *Nature*, 435: 737-738. *Emphasis added*

¹⁰ KNAW (2012), Zorgvuldig en integer omgaan met wetenschappelijke onderzoeksgegevens, Advies van de KNAW Commissie Onderzoeksgegevens (“Commissie Schuyt”), September, p. 17

¹¹ Gallup Organization. (2008). Observing and reporting suspected misconduct in biomedical research. Retrieved from http://ori.hhs.gov/sites/default/files/gallup_finalreport.pdf

¹² Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS One*, 4 (5): 1-11

¹³ John, L.K., Loewenstein, G., Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, 23 (5): 524-532

¹⁴ Fang, F.C., Steen, R.G., Casadevall, A. (2012) Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109: 17028–17033

¹⁵ Steen R.G., Casadevall, A., Fang, F.C. (2013) Why Has the Number of Scientific Retractions Increased? *PLoS ONE*, 8(7): e68397

¹⁶ Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90: 891-904; quote from p. 891.

More than half of these investigations related to (suspected) cases of plagiarism and authorship disputes. While these statistics alone provide little evidence of a growing problem of research misconduct in the Netherlands, it is interesting that (as far as can be documented) only 27 investigations provided support for misconduct claim, 16 of which led to sanctions. Apparently, there is also quite some disagreement what constitutes research misconduct and/or it is very difficult to prove misconduct.

Considering our earlier argument that questionable research practices (QRP) can be equally problematic in terms of ‘truth-finding’, these numbers do suggest that the issue of research misconduct and ‘sloppy science’ is to be taken seriously – not only in light of the scientific “mission” but also to uphold public confidence in science. Even though there is no evidence that public trust (confidence) in science, at least in the Netherlands, is substantially declining, the Royal Netherlands Academy of Arts and Sciences also concludes that constantly paying attention to the possible risk of such a decline is pertinent¹⁷.

2.4 Contributory factors

Contextual factors

There is some research to suggest that misconduct is more likely to occur in highly competitive environments¹⁸. The committees investigating the Stapel-affair have not in detail addressed the often-mentioned “publish or perish” factor as a contributory factor. The Committee Schuyt concludes that, in the absence of proper studies, it is difficult to establish whether this pressure has led to increased fraud.

As a Taskforce, we do believe that the *signals* of publication pressure and its possible relation to (increased incidence of) fraud should be taken seriously. It is very unlikely that the increased publication pressure does not lead to more sloppy science and QRP. For instance, research in the US has found that positive findings are more likely in a more competitive environment¹⁹. Following our earlier argument, publication pressure could thus ultimately lead to misconduct, as (some) researchers start to engage in ever increasing ‘darker shades of grey’.

Another indirect effect of publication pressure on misconduct and its delayed discovery, is time pressure that researchers experience in conducting reviews of others’ work, both internally and externally. As pointed out, the Stapel-investigators were clearly astonished that many of the blatant errors in Stapel’s work had not been uncovered in the review process. A recent analysis of fraud causes also established that the review process rarely led to the exposure of such cases²⁰. This latter study explains this by pointing towards the intrinsic “trust” that reviewers have in the work of their peers (as opposed to, for instance, members of integrity investigation committees), but time pressure is an alternative explanation.

In a recent study conducted at Rotterdam School of Management, for instance, junior researchers admit feeling under pressure and vulnerable in relation to the Tenure Track system and the associated publication pressure. This is not to say that publication pressure is necessarily a bad thing, but a “linear assessment” of research quality based on the number of publications does carry the risk of an attenuation of the assessment of that quality.

¹⁷ KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013. www.knaw.nl

¹⁸ Swazey, J.P., Anderson, M.S., & Seashore Lewis, K. (1993). Ethical problems in academic research. *American Scientist*, 81, 542-553

¹⁹ Fanelli, D. (2010). Do pressures to publish increase scientists’ bias? An empirical support from US States data. *Plos One*, 5, e10271

²⁰ Stroebe, W., Postmes, T., Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, 7 (6): 670-688

More importantly in the current context, if the (increased) pressure on publications is effectively to be maintained, it should be supplemented with more “soft controls”, such as increased attention to continuous enactment of core values, coaching/mentoring, and an open debate on principles of a professional and integer research process.

The investigation of the Stapel-affair puts great emphasis on “the failure of scientific criticism” as a factor contributing to sloppy and fraudulent science. While the authors of the report explicitly point out that this analysis cannot be automatically extrapolated to the whole field of social psychology (let alone: science at large), in conjunction with other analyses they may provide useful pointers. Specifically, they point out the following factors²¹:

1. Sheer statistical incompetence;
2. Neglect by internal and peer reviewers of fundamental scientific standards and methodological requirements;
3. Reviewers being strongly in favour of telling an elegant and compelling story, frequently at the expense of the necessary scientific diligence;
4. Reluctance to replicate prior studies;

Finally, the Committee Schuyt also cites four ‘lessons learned’ on preventing misconduct from the widely praised book of Jonathan Marks²²:

- Fraud will be relatively easy, if there are no clear procedures on assessing and dealing with fraud, and if authorities have an interest in covering up fraud cases;
- There is reluctance to raise suspicions, as whistleblowers (rather than the perpetrators) usually are harmed in their subsequent scientific career²³;
- There is an all too easy acceptance of ‘mistakes’ and ‘sloppiness’, particularly raised as apology by co-authors involved (cf. the findings of the Stapel investigation);
- The strong hierarchy in academia makes it particularly difficult to prove and deal with cases of fraud committed by senior faculty.

Personal factors

Besides contextual factors, personal factors may also contribute to flaws and misconduct in research, although there is no substantial research evidence to speak of yet. Adopting the ‘rational choice’ perspective, the decision to engage in sloppy research or research misconduct can be seen as a trade-off between expected benefits and costs. Within this perspective, some authors suggest that the propensity to engage in fraud and flawed science is strongly related to the *degree of self-control*. Important components therein are²⁴:

- Propensity for risk-seeking
- Immunity to reactions of others
- Disregard for long-term consequences

Also in the analysis of the Stapel-case, the investigative committees place strong emphasis on the personality of the key actor: “The first and most important group of factors (explaining the failure of the regular scientific criticism) must be sought in Mr Stapel’s working method and research environment ... He

²¹ Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, PP. 53-4. www.commissielevelt.nl

²² Marks, J. (2009). Why I am not a Scientist. Berkeley: University of California Press

²³ Here, we want to emphasise the excellent example of EMC, where the dean took over the role of claimant from the initial whistleblower in the recent case of misconduct

²⁴ Komter, A. (2012). Laakkbare wetenschap. Over alledaagse verleidingen en normoverschrijding in de wetenschap. *Mens & Maatschappij*, 4, 415-436

used his position of great prestige and power to commit fraud and to stifle any possible doubt about his methods.²⁵

Other sources also indicate that two factors nearly always co-exist in case of scientific fraud: a researcher who perceives career pressures and who has a clear idea of how research results should look like²⁶.

As discussed in the introductory chapter, the mandate of the Taskforce was to identify institutional measures to help foster research professionalism and integrity. As such, we have decided not to focus on these personal factors.

Obviously, there could be institutional measures to address personal factors, such as personal assessments when recruiting PhD students or new faculty. At this point, however, we deem the scientific evidence too limited to be able to provide any detailed recommendations on this point. However, it may be useful to monitor emergent research on this issue, and to fine-tune recruitment and other HRM policies accordingly.

2.5 Areas for improvement

Recent reports have indicated several areas of improvement; ways in which the core principles of professionalism and integrity research can be protected and fostered.

The Levelt/Noort/Drenth Committee provides the following recommendations²⁷:

- All new staff and PhD students informed of **code of conduct**;
- **Training** in research integrity should be provided to (master and) PhD students;
- There should be **tighter review** of PhD thesis projects, a.o. by appointing two supervisors;
- **Replication** should be a more common research practice;
- **Data archiving and access** should be an intrinsic element in the publication process;
- Scientific integrity committees should deal with cases in secrecy, but ultimately provide **public information** on numbers and the nature of cases.

The report by the KNAW Committee Schuyt was originally focused on research data management practices, but (due to disclosure of the Stapel-affair in fall 2011) also included an analysis of possible ‘threats’ to scientific integrity. The committee recommendations can be summarized as follows²⁸:

- The principles in the VSNU code need **better communication and embedding**;
- Individual level: professionalism and integrity as self-evident but also **consciously and continuously maintained**;
- Institution (research groups>directors of research>schools>universities): responsible for the **conditions** to promote professionalism and integrity (through measures relating to recruitment, stimulating debate, training/coaching, performance appraisal, pledge-taking);

The Schuyt report made the following recommendations related to management of research data:

- Given the **great variety** between disciplines, general conclusions about management of research data are not possible;
- At the institutional level: increased **peer pressure** in the front on data management practices;

²⁵ Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, p. 37. www.commissielevelt.nl

²⁶ Goodstein, D. (2010). *On fact and fraud. Cautionary tales from the front lines of science*. Princeton: Princeton University Press. Cited in: Abma, R. (2013). *De publicatie fabriek. Over de betekenis van de affaire-Stapel*. Nijmegen: Vantilt

²⁷ Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, p. 37. www.commissielevelt.nl

²⁸ KNAW (2012), Zorgvuldig en integer omgaan met wetenschappelijke onderzoeksgegevens, Advies van de KNAW Commissie Onderzoeksgegevens (“Commissie Schuyt”), September

- Earlier phases in the research cycle are relatively free from external control, which may lead to particular **risks** ;
- The **need for additional controls varies** by discipline and the type of research (e.g. very exploratory vs. large scale testing);
- Random **surveys** of ongoing and completed projects by groups/institutes for **learning about best practices**;
- Learned societies and disciplines: upholding and where necessary **tightening norms** for scrupulousness and integrity and rules on data access.

We strongly agree with one of the main conclusions of the Stapel-investigation, which are strongly in line with the recommendations of the Committee Schuyt published two months earlier : “shortcomings identified must not result in organized distrust or overblown bureaucracy that unnecessarily impedes scientific work, but rather in the creation of a research environment in which researchers are encouraged, through **coaching, training and effective controls**, to take account of the rules for careful and honest academic research.²⁹” (emphasis added).

Finally, in May 2013, the KNAW published the report “Vertrouwen in Wetenschap” (“Trust in Science”). This report reflected on the earlier recommendations of the committees Levelt/Noort/Drenth and Schuyt, and provided a set of recommendations for fostering professionalism and integrity in research³⁰:

1. Create more opportunity for **peer pressure**, an institutional responsibility;
2. Develop **protocols** for collecting, managing and sharing **research data**;
3. Create **constant attention** to integrity through a.o. training, particularly in the graduate program;
4. Create sufficient opportunity for peer review; review efforts should be acknowledged in capacity planning and evaluation protocols;
5. Research integrity should be an integral element in **job appraisal talks**;
6. Experiment with **pledge-taking** ceremonies;
7. Create awareness for (international) **differences** in codes and traditions regarding integrity;
8. Create regulation regarding possible **conflicts of interest**.

These recommendations to universities are complemented with recommendations that are primarily directed towards the KNAW, VNSU and NWO:

9. Monitor the awareness regarding the **VSNU code**, and efforts to enhance this awareness;
10. **Accreditations** should spend more attention on management of research data and integrity.

Finally, one comment that has been raised at various moments throughout the debate the past two years is the notion that research misconduct is best prevented by working in teams that openly and extensively discuss their research process and findings. While it is hard to provide an exhaustive set of specific recommendations on this point (the various reports are also relatively silent on this point), it is one principle (cf. the Erasmus University core principle of “teamwork”) to keep in mind throughout. For instance, it is one principle underlying the Levelt/Noort/Drenth recommendation to have two supervisors for each PhD. Teamwork may also an element in considering the introduction and awarding of recognition and prizes for good research³¹.

2.6 Responsibilities different actors

One of the key elements in the discussion in the Netherlands at the moment, is the call for more explicitly defined institutional responsibilities for fostering professionalism and integrity in research. As the KNAW

²⁹ Levelt Committee, Noort Committee, Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. November, p. 55. www.commissielevelt.nl

³⁰ KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013, pp. 50-4, www.knaw.nl

³¹ Professor De Dreu also recommended this, e.g. to NWO, at the presentation of the Schuyt report

comments: "In its current form, the (VNSU) code creates the impression that integrity is solely the responsibility of the individual researcher." The KNAW acknowledges that various organisations are now taking more responsibility, but mainly in a reactive fashion³².

As Taskforce, we have also been tasked with developing recommendations on *institutional measures*. As will become clear in the subsequent chapters, our recommendations are primarily targeted to the Executive Board and the deans and directors of research of the individual schools and institutes. However, it will be clear that in the implementation of many recommendations, the commitment and effort is required from research leaders (project leaders, group leaders, chairs etc) and ultimately, individual researchers themselves.

We do want to point out that senior faculty, in their formal and informal capacity as research leaders, have a special role to fulfil. Senior faculty are not only conducting research themselves, but should lead the way in practicing professional and integer research, and openly discussing the associated dilemmas. The Stapel investigation has also questioned the fact that the whistleblowers were PhD researchers – and not senior faculty, who should have had the expertise and the courage to identify and raise the suspected fraud. Just treating professionalism and integrity in research as an issue pertaining to their own research activities neglects the important leadership role of senior faculty.

2.7 Areas for future measures

We think there is a number of additional measures that schools and faculties should consider, but which we did not elaborate yet in any of the following chapters.

- First, we think it is important that integrity issues become a standard agenda item in appraisal talks. For example, the appraiser could ask for integrity dilemmas the appraised has faced during the research process (or in teaching). These dilemmas and the chosen solutions can then be discussed.
- Second, there should be more attention and recognition for peer review. Peer reviewing of articles (and in some areas, books) is very important work from an integrity point of view. However, the current incentive structure does not stimulate high quality reviewing. Researchers receiving many review requests are usually the ones performing well on this task, but by doing this, they can dedicate less time to their own research. As a start, all reviewing activities and including the level of the journal involved could be registered, as for example is already current practice in the University of Antwerp. The next step is to include reviewing in the research performance evaluation system.
- Third, it is important to encourage replication studies. Also in this area, the current incentive structure is not favourable. Schools, faculties and research leaders should consider ways to change this.

³² KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013, pp. 49-50, www.knaw.nl

Chapter 3: Research Data Management

This chapter contains the summary of the report of the working group, "Research Data Management at Erasmus University Rotterdam". This report attached in its entirety, as Appendix 1. This subproject was addressing the (organisationally and technically) most complex issue of all the issues addressed by the Taskforce, and has required substantial time investments. At the same time, these investments represent only the "groundwork" for necessary further refinement and implementation. To facilitate these further initiatives, the appendix presents a quite comprehensive report of the work done so far.

A solid research data infrastructure at the Erasmus University Rotterdam (EUR) has several stakeholders who play a role in responsible and reproducible research. This section relates the main recommendations made in the current report to these stakeholders and their responsibilities. The *Rector Magnificus* and the *Executive Board* are responsible for a coherent and effective EUR research data management approach. The following recommendations pertain to the Rector Magnificus and the Executive Board (mainly based on Chapter 5, Appendix 1):

- appoint a research data support officer at the central EUR level. Dedicated support at a central level is crucial for continuity and for sharing good practices between institutes or disciplines.
- draft a covenant to the extent that the deans of research or research directors and – at Erasmus Medical Center – the heads of department embed and evaluate research data management in their institute. They should have room to describe how they will fulfil the data management minimum.
- introduce the deans of research to the EUR data management minimum protocol.
- provide training and advice on research data management: delegated to the university library (see below).
- provide services for storage and retention of research data: delegated to SSC ICT (see below).

The *deans/directors of research* and *heads of department* at the various faculties, graduate schools and institutes of the EUR are responsible for the organisational embedding of data management. The following recommendations pertain to them (more in Chapter 5, Appendix 1):

- coordinate, supervise and stimulate data management workflows and protocols. In particular, refine the minimal EUR protocol such that it fits the institute's research discipline(s) and make it practical. Take care to explain that data management also includes documentation related to the research process.
- socialise young researchers into responsible ways of working. Research data management is not something extra. Rather, the attitude should be that professional research goes hand in hand with responsible data management.

Researchers are responsible for storing data and documentation at various moments during a study. The minimum that must be stored consists of both the raw data and the data underlying any submitted or published publication, the project plan, documentation that describes and explains major changes to the earlier plan(s), as well as the submitted version of the publication (see Chapter 5). It is essential that *PhD supervisors* and *research group leaders* are also role models for young researchers. Similarly, commitment and the willingness to share good practices are more important than protocols and covenants.

The *university library* has been given responsibility for raising awareness for data management and for providing training and advice for research data management. The following recommendations pertain to the university library (more in the more in Sections 3.4, 4.3 and Chapter 7, Appendix 1):

- develop the library's virtual desk into a front office for researchers, as a central point of expertise in research data management. This includes training and collaboration with long-term archives (the back office). Close collaboration with the Research Support Office (currently under development) is foreseen.
- develop and maintain an activating data support web site and select or develop relevant courses and workshops.

SSC ICT has been given responsibility for providing services for safe storage and retention of research data. The following recommendations pertain to *SSC ICT* (more in Chapters 3 and 4, Appendix 1):

- create safe storage and backup facilities for individual researchers, as well as safe ways to share and collaborate on research data.
- maintain and offer expertise to make research staff aware of advantages and risks of particular storage platforms and media, in collaboration with the front office.

These recommendations and the division of responsibilities are based on an online EUR-wide survey, interviews and meetings with various researchers and experts from within and outside of the EUR, as well as on data management documentation provided by other universities. When these recommendations are implemented and research data management continues to be taken seriously, the EUR will at least be on a par with other data-aware universities.

Chapter 4: Training

Taskforce recommendations training:

- The PhD training program includes at least three training sessions on professional and ethical standards and scientific integrity;
- Specific integrity issues related to particular research methods are also part of the PhD training program, either as integral part of these method courses or as a separate course;
- Scientific integrity becomes a recurrent theme during annual research days and should be integrated on a menu basis in all training programs for researchers;
- The course on academic leadership current offered to those attaining leadership positions should contain a module on scientific integrity.
- The Executive Board should encourage the VSNU to develop on-line training resources.
- A specific course should be offered to university administrators to make them more familiar with the formal processes involved in reporting and investigating misconduct.

It has become clear that following standards of scientific professionalism and integrity is not automatic. This means that continuous attention for this topic is necessary. It should be an essential element of training, not only for junior researchers but also for faculty members in general and for research leaders in particular.

Over the past year, the Taskforce has also become increasingly aware that standards of professionalism and integrity are not universal, and that even for one and the same method different research groups may adopt different standards. This should be explicitly acknowledged and addressed, particularly in the respective methodology courses. At the same time, graduate schools would do well by addressing this heterogeneity by organising a more explicit discussion between different methodology teachers.

4.1 Junior researchers

It is important that doctoral students are aware of the professional standards for scientific research in their fields. These standards can be passed on in formal course work and in the professional relationship with their supervisors, other faculty, and members of the academic community in general. Training professional standards and scientific integrity cannot be uniform across different research fields as the pitfalls and dilemmas are different for each field and even for different research methods. However, scientific integrity should be an identifiable part of the training program of all doctoral students. We recommend graduate schools and/or faculties to take the following two steps.

1. Every graduate school or faculty should offer at least three sessions in which students get an introduction to professional behaviour and scientific integrity. These sessions should include:
 - Providing information on relevant documents and the general principles presented in them;
 - Reflection by PhD student on her or his own research practice;
 - Discussion of some cases of clear and less clear lack of integrity;
 - Playing the dilemma game.

Professionalism and integrity dilemmas can be discussed concerning the following topics: the relationship supervisor – doctoral student, sharing ideas, co-authorship, commenting other people's research, departmental citizenship, and refereeing. These sessions can be led by, e.g., the "integrity coordinator" of a faculty or of the university, by faculty specializing in ethical issues or by faculty teaching methods courses. Appendix 2 provides some information on research integrity courses currently offered in the graduate schools EMC and ERIM.

2. Graduate schools or faculties should make sure that integrity aspects of the most common methods used in a field of inquiry are discussed. A discussion of such matters can either be part of existing methods courses or can take place in two or three separate sessions on professional use of research methods. In the latter case, the sessions are best coordinated and preferably co-developed with methods course teachers. Topics to be discussed include the professional attitude towards and the integrity issues surrounding collection, storage, cleaning, and analysis of data, and the archiving of data and do-files (or protocol files); the presentation of findings in academic journals, and the presentation of existing literature.

The right balance between general and specific sessions depends heavily on the organization of the graduate program and possibly on the school. In particular, graduate schools that offer many specialized methods courses are probably best off by focusing the discussion of scientific integrity and professionalism on what is specific for the material taught in each course. Schools that offer very few methods courses are advised to organize separate sessions on professional standards and integrity issues concerning the main (quantitative, qualitative) research methods.

Part of these sessions may be conducted in an on-line format. In an early phase, the Taskforce decided not to embark on developing an EUR-specific on-line training module, primarily because more priority was given to the dilemma-game and because developing such a module could be quite costly. However, on-line instructions and webinars are becoming increasingly widely available and could provide a suitable environment for participants particularly to learn more about the basic issues regarding professionalism and integrity, before engaging in more interactive debate³³. One of the problems, though, is that many of these on-line resources are geared towards particular (North-American) contexts. (This is for instance the case when referring to the role of research integrity officers.) The Taskforce recommends to the Executive Board that in the context of the VSNU, efforts are made to develop alternative on-line training resources for the Dutch context.

After having received these courses and sessions, the PhD students are ready for the pledge (see Chapter 6).

4.2 Senior Faculty

It is important that faculty adhere to the professional standards for scientific research in their fields. For many reasons, such adherence may at times be below par. To protect the reputation of academic research, colleagues, and oneself / themselves, we recommend the following:

1. Faculties make scientific integrity a recurring theme of annual research days. Faculty should check and discuss whether the organization and culture in a faculty stimulates the highest professional standards in all stages of academic research. To start the discussion, we recommend the discussion of recent cases of scandals and playing the dilemma game.
2. In addition to the discussion on annual research days, faculties should offer a menu approach for training on Scientific Integrity, in which different training programs for senior staff³⁴, PhDs (as discussed in 4.1) and research master students are offered. Depending on the target group professionalism and integrity dilemmas can be discussed concerning the following topics: the relationship supervisor – doctoral student, sharing ideas, co-authorship, references, commenting other people's research, departmental citizenship, refereeing, presentation of findings in academic journals and issues surrounding collection, storage,

³³ For instance, Nick Steneck (who was involved in establishing the Office of Research Integrity in the US, offers such webinars through Epigeeum. See also: <http://www.youtube.com/watch?v=rBLwC62LwhU>.

³⁴ For this category a distinction can be made between current senior and new senior staff, and see 4.3

cleaning, analysis and archiving of data. In discussions among seniors, the focus should be primarily on their responsibility and their role in the large grey area in conducting science, where people may be violating aspects of the Integrity Code without even being aware. Faculties could also involve the dilemma game in their training programs.

4.3 Research leaders

Guaranteeing scientific integrity is pre-eminently the responsibility of senior staff in universities, especially those with leadership positions in research. These include directors of research schools, leaders of research programmes and thesis supervisors but also those with leadership roles in the research process, such as journal or book editors and journal referees.

It is good practice in Erasmus University that Faculty members who attain leadership positions take the course on Academic Leadership. We recommend that part of that course is dedicated to academic integrity. Aim of the module on scientific integrity should be to reflect on one's skills to conduct, lead and supervise research in an integer way, and to transfer these capabilities to other researchers.

The reflection on skills to be an integer research leader includes getting to know the relevant behavioural codes, recognition of one's power position, and the equal treatment of researchers no matter their formal status or age, sex, and cultural background. The transfer of capabilities to practice scientific integrity in research can take place in two ways: through being a role model and through coaching.

1. Being a role model requires continuous reflection on one's own practice as researcher, research leader and supervisor. Playing the dilemma game can be a useful tool, in particular if cases concerning academic leadership are selected.
2. The transfer of scientific integrity skills through coaching is done on two parallel levels, relational and task focused. Coaching in relationships between researchers requires being able to listen well, to respect and support researchers in their development towards independent researchers, to recognize dilemma's, to signal risks, to clarify norms and to give and receive feedback. In addition, the research leader as coach must dare to engage in difficult (awkward?) conversations on possible violations of scientific integrity by researchers under her direction, but also as referee and editor of papers by other researchers. The task focused coaching role concerns concrete steps in the research process. It involves taking responsibility for the full process, including training financing, publication and the division of tasks between researchers and supervisors; being accountable transparent on one's own behaviour; and signalling, discussing and acting on possible violations of scientific integrity during the process.

We recommend the University Board to instruct the Human Resources Department responsible for the course on academic leadership to have this course as soon as possible include a module on Integrity in Academic Leadership.³⁵

4.4 University administrators

The Taskforce recommends that a specific training session should be offered to university administrators to make them more familiar with the formal processes involved in reporting and investigating misconduct. This would include (selected members from) the Executive Board, Deans, the confidential advisor, (possible) members of the Committee Scientific Integrity (CWI), and the

³⁵ By mid-2013, The Bureau Eva Wiltingh BV that for some years now has offered this course for the EUR, did not yet include such module in its offer

local integrity coordinators. Experiences of Taskforce members and the LOWI demonstrate that university administrators are not always familiar with the VSNU code of conduct and the processes related to the LOWI. The session should focus on what administrators in their respective roles should do, and should *not* do.

Chapter 5: Dilemma game

Recommendations dilemma-game:

- The game is part of the PhD training and faculty training sessions on research integrity;
- All faculty members that are not already participating in the ‘standard’ training sessions (which mainly pertain to PhD students, tenure trackers and associate professors), regularly participate in a dilemma-game session (suggestion: every two years).
- Senior faculty shows commitment to openly discussing professionalism and integrity issues in such settings by participating in these dilemma game sessions.
- Directors of research should actively encourage the use of the dilemma-game; the exact way the game is used is up to institute/department/group leaders.

Central in the whole debate on fostering professionalism and integrity in research are dilemmas; in many situations, there is no clear ‘right’ or ‘wrong’. As argued before; the most important improvements in terms of researchers’ ultimate goal of ‘truth-finding’ are to be found in addressing questionable research practices. Therefore, one of the guiding principles underlying the work of the taskforce has been to help stimulate an open debate on what is questionable and what not.

There are various dilemmas that may arise in such a debate: Can I exclude particular observations from my research? Can I use exactly the same data set for multiple papers? Should I agree on a colleague being a co-author on a paper to which she has not made a significant contribution?

The taskforce has developed a game that aims to support researchers in further developing and honing your own “moral compass”, by exposing individual researchers to such dilemmas in a group discussion.

The game lets researchers consider, choose and defend (and possibly reconsider) alternative courses of action regarding a realistic dilemma regarding professionalism and integrity in research.

Participants will also come to appreciate the dilemmas that others are faced with, how they resolve them and the reasoning behind these solutions. The game encourages participants to discuss issues relating to professionalism and integrity, and to help one another to find solutions for their own dilemmas.

The game can be used in a variety of settings. It can be used in a course setting, for instance with a group of PhD students. Or it can be used in a research strategy meeting of a department or institute. Depending on the objectives, it may be used primarily as an exercise to let people exchange opinions and experiences, or also as a step towards defining more formally defined principles, on for instance co-authorship. Often, it may be effective to let participants come up with their own dilemmas, after playing a number of dilemmas from the game.

Whichever setting or objective, the game may be helpful in bringing attention to “The Netherlands Code of Conduct for Scientific Practice” (Association of Universities in the Netherlands, 2012), which is applicable to every university scientist in the Netherlands.

The 75 dilemmas included in the game have been collected through sessions at different EUR schools, and among researchers who use different research strategies and who are in different stages of their careers. In that way, we have aimed to develop a set of dilemmas that are relevant to a diverse population of researchers. While the dilemmas are based on actual cases, they should be recognizable and relevant to many researchers. Should you wish, you can preselect a particular set of dilemmas to ‘play’, based on for instance a particular phase of the research process you want the discussions to focus on.

The dilemma game is available (in English) in a ‘tangible’ box-format, one box for each group of 4 participants. It is also available for download on www.eur.nl/integrity, both in English and Dutch.

Chapter 6: Pledge-taking

Taskforce recommendations pledge-taking:

- All EUR researchers, existing and new, take a pledge related to professionalism and integrity in research. Preferably, professors and senior researchers take up their leadership role by taking the pledge as the first group in their school/institute.
- New EUR employees involved in research are given “The Netherlands Code of Conduct for Scientific Practice” upon appointment.
- The Executive Board should define the text of the pledge.
- Directors of research should be responsible to implement the pledge-taking procedure in their local school/institute, so that every researcher has taken the pledge by December 2014.

One of the strategies that have been suggested to combat scientific misconduct in academic research is the introduction of an academic pledge.³⁶ In this report we address the effectiveness and desirability of such a pledge, and which form it could take. In particular, we aim to answer the following questions:

1. What evidence do we have that pledge-taking can contribute to scientific integrity? If there is evidence, or if there are other good reasons to believe that a pledge could contribute to scientific integrity, then which form should it take in order to be effective and to be able to command support among the relevant groups?
2. What examples exist of existing practices of pledge-taking, and what can be learnt from these practices and experiences?
3. If the answer to the first question is affirmative, then how could an academic pledge for the Erasmus University look like? Who should take the pledge, and at which point in their scientific career should they take the pledge?

6.1 Can pledge-taking strengthen scientific integrity?

The rationale behind taking a pledge is that it would create more awareness regarding the ethical norms which apply to scientific conduct in the Netherlands. All scholars who are employed by Dutch universities have to adhere to the “Integrity Code” which has been developed by the VSNU, the Association of Dutch Universities.³⁷ Despite that all new employees may be given the VSNU Integrity Code upon appointment, anecdotal evidence suggests that many scholars have never read this code, and some have even never heard of it. They may have received a copy when they became employed at the university (most likely during their appointment with the Personnel Officer when they signed their contract, together with a large number of other forms and leaflets of information). Moreover, not all scholars have an employment contract with the university: an increasing number of scholars pursuing a PhD-degree do not have an employment contract, and some scholars too have only a hospitality agreement and no employment contract.

The rationale behind a pledge is that it would affirm and thereby strengthen the social norm that scientists ought to respect the values that are discussed in the VSNU code. In general terms, it would help scholars to raise questions in case they see someone else engage in unethical behaviour, since the scientific community has collectively and publically declared that it will not engage in such behaviour. It could help to back up a scholar who is asked by another scholar to engage in behaviour that violates the code. For example, it could make the difference in case a junior scholar (e.g. PhD

³⁶ See. E.g. *Rapport Schuyt*, sectie 3.6.4 & p. 79; J. van der Meer (2013) ‘Pleidooi voor een onderzoekerseed’, *Nederlands Tijdschrift voor Geneeskunde*, 157: A6064

³⁷ http://www.vsnu.nl/wetenschappelijke_integriteit.html

student) is put under pressure by a more senior scholar to engage in scientifically unethical behaviour, especially if that senior scholar stands in a position of authority to the junior scholar. In addition, there is a large grey area in conducting science, where people may be violating some aspect of the Code without even being aware. In those circumstances, having taken a pledge, and knowing that your colleagues have also taken a pledge, may help to create more awareness and make create a professional ethos and atmosphere in which one can openly talk about doubts and dilemma's one encounters regarding proper scientific conduct.

Some scholars are sceptical about the effectiveness of such a pledge since it is regarded as merely symbolic and hence there are no real incentives to change one's behaviour. This critique can take two forms.

The first critique is that if there are no material consequences for violating the ethical code, or if there is no support in helping well-intended scholars to help them behave ethically in cases of doubt, then what use is the pledge? Is not merely symbolic, so that the universities can claim that they did what they could to prevent scientific misconduct? This critique is valid, but it can be countered by stressing that a pledge should never be the only part of a strategy to increase scientific integrity. It must be one element of a larger package, as is the case in the proposals worked out by the Taskforce.

The second critique entails that whether or not as part of a larger package of measures, if the material rewards of unethical scientific behaviour are strong enough (in terms of jobs, promotion, status, etc.), then these material incentives will be much stronger than the symbolism of a pledge. Yet this materialist view underestimates the scientifically proven behavioural effects that promises have.³⁸ If, in a community where people know each other, they publicly and collectively make a promise, then there exists a strong social norm that one ought to keep that promise. A pledge regarding scientific conduct will have that effect too. Pledges have been shown to be effective in changing quite dramatic social norms (such as female genital cutting in Africa), but only if the pledge was public, and if the pledge was part of a larger dialogue about the reasons to change the social norm.³⁹

6.2 Examples of existing pledges

Pledge taking has existed for a long time in many other professions, such as the Hippocratic Oath among medical doctors, or the oath of judges and lawyers. There are important similarities between these professions and researchers. First, these professions have a set of ethical norms and values that they want to be shared by all and to guide the behaviour of all professionals in that particular profession. Second, all of these professions consist of professionals who have to pursue the truth, and who should not cut corners under external pressures (whether financial or otherwise). Third, in all of these professions we want the professionals to act independently and trustworthy. These are some of the reasons why these professions have had the practice of a pledge for a long time, and this would also explain why we should take the idea of a pledge for researchers seriously.

In academia, pledge taking is not standard, yet it is not entirely new either. While we are not aware of all instances of pledge taking in academia, the following four examples (two from the Netherlands, and two from abroad), are relevant.

Within the Netherlands, the Erasmus Research Institute of Management (ERIM) at the Erasmus University Rotterdam, introduced the pledge for the first time in 2013. This pledge-taking was the

³⁸ For an overview, see Ch. Biccieri (2006), *The Grammar of Society*, Cambridge University Press, chapter 5.

³⁹ Bicchieri, Descartes Lectures, Tilburg November 2012.

concluding part of a course on academic skills and values that PhD students at ERIM had taken. The text of that pledge reads as follows:

I hereby declare to uphold the ethos of good scientific research and to apply throughout my scientific activities the principles described in The Netherlands Code of Conduct for Scientific Practice: Scrupulousness, Reliability, Verifiability, Impartiality, and Independence. I will apply these principles in my own work, and will endeavour to promote these principles among other scientists and in particular my direct colleagues.

The second Dutch example is the introduction, since January 2013, of a reminder of the values of the VSNU code, to any new professor who takes up his or her appointment at the University of Leiden. When in Leiden a professor holds his or her inaugural lecture, the public lecture itself is preceded by a *laudatio* read by the Rector Magnificus for the new professor and the other professors attending the ceremony, which now ends with a reminder to keep up to the values and norms that are entailed in the VSNU code.

The two foreign examples are also interesting. First, at the University of Gent (Belgium), all members of the academic staff who are given tenure, have to take an oath whereby they promise they will obey the Belgian Law (this applies to all Belgian civil servants). This is not a pledge that is particularly aimed at preventing unethical behaviour, yet in a certain sense this can also been seen as superfluous since every inhabitant of Belgium needs to comply with the law.

Second, and more relevant to the case of academic pledges, in Austria all Master-students (*Magister*) and PhD-students are to take a pledge when they graduate. In this pledge they promise that they will continue to serve the sciences, that they will pursue truth and will not suppress or commit fraud with scientific knowledge.⁴⁰

One thing that can be learnt from these existing pledges is that if taking a pledge has become a standard practice in a profession, *which applies to all*, then the initial resistance that one may expect will vanish. In addition, the two recent Dutch examples show that various forms are possible, and that indeed the recent pledge introduced at the ERIM in Rotterdam could be combined with the reminder read by the Rector Magnificus for full professors that Leiden University introduced.

6.3 How could an academic pledge for the Erasmus University look like?

We suggest two options: option A introduces a pledge for all researchers, while Option B introduces the pledge for PhD-students at the end of their first year in combination with a reminder of the VSNU code for full professors at their inaugural lecture.

Option A

Option A would consist of the pledge that was introduced at ERIM early 2013⁴¹, which would be taken by all PhD students at the Erasmus University during their first year of PhD studies. The procedure adopted at ERIM is briefly described in Appendix 3.⁴²

We propose that taking a course on scientific integrity should be part of the first year graduate training of PhD students (whether they have an employment status with the university or not), and that they pledge is the closing part of this course.

⁴⁰ <http://www.uni-graz.at/zv1www/mi001115a.html> - section 2: 'Sponsionsfeier' (last accessed 4 September 2013).

⁴¹ <http://www.erim.eur.nl/news/detail/2987-erasmus-researchers-take-a-pledge/>

⁴² If a different formulation for the pledge is chosen, we recommend that it is at a minimum consistent with the VSNU code, since that code is already a binding ethical norm for all scientific employees at the Dutch Universities

One can debate whether ‘external PhD students’ (PhD students who don’t have an employment contract but rather have a hospitality agreement) should also take the pledge, since one may be worried whether having to take a course would put them off. Yet the question is whether this fear is justified. After all, they are conducting research, and hence the code should apply to them to. Moreover, their situation demands extra attention, since to the best of our information, they are now not even given the VSNU code.

However, it will strengthen the credibility and support for the pledge if *all* members at the EUR will take the pledge. New scholars, such as PhD student or postdoctoral students, may rightly argue that most high-profile cases of fraud were not conducted by their peers, but rather by full professors. Moreover, there is no good reason to exempt any currently already employed scholar from taking the pledge, since the reasons for taking the pledge apply equally to all scholars.

In order to facilitate that eventually all academic scholars will have taken the pledge, the Directors of Research at all faculties should organize once or twice a year a day on scientific integrity, which could include a lecture on this topic and the dilemma game. That day, too, should be closed with taking a pledge. Scholars should be given a certificate that they have taken the pledge, and if they change universities, not be required to take the pledge again.

Option B

Still, one major objection to Option A is that it will require a major undertaking to get all researchers through a course or workshop on scientific integrity and take the pledge. There are good reasons to believe that many scholars don’t need to be reminded of the pledge, and forcing all of them to take the pledge may be counterproductive.

The alternative proposal is therefore to only ask PhD-students to take the pledge at the end of their first year, combined with the reminder of the VSNU code for professors who give their inaugural lecture. Option B would very likely be able to command more support from the scientific community and it would not *only* target new members of the scientific community (PhD students) but also those who have been given the most power (professors).

6.4 Conclusion

Having considered the evidence for the effects of public pledge taking, and the arguments for the different implementation options, we finally advise the Executive Board to demand that all schools implement a procedure where all researchers take a pledge.

The Taskforce finds it of key importance that by engaging in an open debate, and by signing a pledge, also senior scholars publicly demonstrate their commitment to scientific professionalism and integrity.

Senior scholars have to realize, also in this respect, that through their behaviour they act as ‘role model’ for junior scholars. Therefore we recommend that professors and senior researchers take up their leadership role by taking the pledge as the first group in their school/institute.

Not extending the pledge beyond PhD students would send the wrong signal that this is an issue that senior faculty cannot be ‘burdened’ with.

As is argued elsewhere, also for (senior) faculty, the public pledge should be part of a wider dialogue, such as a dilemma game or training session (see Chapters 3 and 4).

Chapter 7: Seminars

The organization of effective feedback and peer pressure during all stages of research is of critical importance to safeguard scientific integrity. This is one of the main conclusions by both the Schuyt and Levert Committee.

The taskforce has dedicated specific attention to seminars, as a particularly strong instrument to organize feedback and interaction on research projects, and between researchers at all levels, both within the faculty and outside. A seminar culture enhances interaction, and can also add to creating an atmosphere of discussion and of confidence.

To investigate how feedback is organized by the different faculties of Erasmus University, representatives of each faculty were asked to respond to a short survey. Based on the responses, the taskforce formulated a number of recommendations in order to implement seminars into the research culture of a faculty.

The questions of the survey were the following:

1. How is feedback at the different stages of a research project organized within your faculty? Which activities take place, how frequent and systematic and at what level (faculty, department, research group), and with or without external peers?
2. To your opinion, is the current organization of feedback: (i) effective; (ii) sufficient?
3. Which ideas do you / does your faculty have to increase the effectiveness of feedback in safeguarding scientific integrity?

The response to these questions can be summarized as follows:

1. There are substantial differences across and within faculties in the way feedback is organized, but all faculties stress the important role of internal research seminars next to external feedback through presentations at international workshops and conferences. Most attention is being paid to feedback to PhD students. Feedback on research by post-docs and senior researchers typically is much less structured and often without any obligation. This poses the risk that post-docs and senior researchers are allowed to work isolated from both internal and external peer pressure (see the findings by both the Schuyt and Levert Committee).
2. Several faculties report that the internal feedback activities (e.g. research seminars) suffer from low participation rates, especially among senior researchers. This reduces the effectiveness and quality of the feedback and reduces the potential for senior researchers to act as a role model for junior researchers. Another frequently reported deficiency is that feedback at the first stages of a research project is often limited, while during this stage of research important choices and decisions must be made. This lack of feedback in an early research stage is in line with the findings by both the Schuyt and Levert Committee.

The following recommendations are made to increase the effectiveness of feedback:

- Department leaders should be made responsible for organizing periodic research seminars and for looking after that all junior researchers regularly present their work;
- Senior researchers should be held responsible for a frequent and active participation in research seminars;
- Research seminars should be chaired by a tenured/senior researcher, comments should be noted and the presentation should be evaluated afterwards with the presenting researcher;
- Researchers should be obliged or strongly encouraged to present their work in an early stage of research, e.g. about the research design;
- Discussants should be invited to provide comments during a research seminar on the research design or draft paper that is submitted to the discussant;

- To improve participation research seminars should be ‘horizontally scheduled’ (i.e. same day, time, frequency) and be added to the electronic agendas of all researchers;
- Organization of an annual graduate research day at the faculty or university level (e.g. similar that organized by the faculty of social sciences / psychology).

Chapter 8: Process and output of PhD projects

Recommendations PhD project process and output:

- We recommend that the supervision plan addresses three points: transparency of authorship, additional sources of advice; additional feedback.
- We recommend that a discussion is facilitated on guidelines for data collection, and that these address the following integrity risks: replicability, verifiability, independence of data, fairness to subjects, professionalism in the use of data collection methods, fairness to researchers;
- We also recommend that plagiarism checks are performed on at least two occasions during the doctoral education;
- The taskforce recommends that the Executive Board revises the doctoral regulations, and considers a discussion in VSNU context about a further strengthening of the doctoral education as far as scientific integrity is concerned. In particular, we suggest the following items to be addressed:
 - transparency on the contributions of possible co-authors,
 - allowing the possibility for the inner doctoral committee to provide formal feedback
 - at least one member of the inner doctoral committee is from outside EUR, and on top of that, at least one additional member of the plenary committee is from outside EUR
 - the doctoral committee explicitly assesses the manuscript on the five principles as laid down in the VSNU code

This chapter elaborates on a number of issues that contribute to an atmosphere of providing and receiving constructive feedback on research, in particular for PhD candidates. It also gives recommendations on inclusion of checks and balances as part of the supervision process (supervision plan). A separate discussion on academic leadership and the role and responsibility of senior staff members and supervisors was already provided in Chapter 2.

8.1 Supervision of PhDs

PhD supervision involves a hierarchical relationship with risks of dependence and power abuse. This relationship may therefore potentially violate various scientific integrity values. Professional supervision requires the promotor⁴³ to be a role model, guiding the PhD student to become an independent, good researcher. Teamwork and the responsibility of role model require supervisors to be open to different viewpoints and to have open discussions.

Impartiality and fair play in the supervision relationship are best served when the PhD student has access to alternative sources of advice and feedback besides the promotor, or at least a person to go to in the case of problems. Alternative sources of advice and feedback will also strengthen the value of reliability in the supervision process.

We therefore propose that the supervision plan addresses the following three points:

1. Transparency of authorship: The supervision plan should include in the planning for every paper and chapter of the dissertation a clarification of the authorship of these. Supervisors and PhD student should agree on the substance of the contributions provided by co-authors of each paper and chapter.
2. Additional sources of advice: This can be organized in various ways. First, the programme may appoint an ombudsman, as an independent staff member available for confidential consultation. Second, the supervision by the promotor may be complemented by advice from an informal advisor, who is not a co-promotor but could give valuable input and serve as a sounding board.

⁴³ Supervisor or promotor: in the Dutch system generally a full professor who chairs the doctoral examination.

- Third, the supervision plan may require the PhD student to present his or her work outside the university. This may be done at seminars, workshops and conferences where feedback can be expected (see also section 7.2 below specifically on seminars). Finally, a more formal way to organize plurality in advice is to appoint a second promotor or co-promotor. The supervision plan, or additional agreement, should make explicit what the tasks and responsibilities of the supervision team members are, as well as the potential benefits of their contributions.
3. Additional feedback: Next to research presentations outside the university, the supervision plan could also agree on internal public presentations of the work in progress. One of these could be a pre-defence, well ahead of the actual defence, during which feedback will be collected and integrated in the manuscript.

8.2 Data collection process

The process of data collection entails integrity risks. These risk range from data fabrication to insufficient anonymity of subjects. Erasmus Medical Center has advanced procedures, which regulate data collection on humans and test animals. The other research and graduate schools do not have such detailed procedures and protocols yet. Funding agencies increasingly require applicants to fill out a form with questions on research ethics in data collection.

We recommend that all graduate/research schools but also faculties facilitate a discussion on guidelines for data collection. We recommend that these discussions address the following integrity risks: replicability, verifiability, independence of data, fairness to subjects, professionalism in the use of data collection methods, fairness to researchers.

Recommendations on research data management, including setting up and implementing guidelines and protocols for data collection at faculty and / or department level are included in chapter 6 of this report.

8.3 Plagiarism checks

Furthermore, we recommend that all graduate and research schools perform plagiarism checks on at least two occasions during the doctoral education. Using one of the many available tools to check plagiarism, it is relatively easy to establish whether the material is indeed produced by the PhD candidate. There are two key reasons for such checks. First, it enables beginning researchers to learn from instances of sloppy referencing and non-deliberate copying of texts, with the help of plagiarism test results. Second, a PhD degree is awarded by Erasmus University, which has a strong academic reputation, and any suspicion of plagiarism, including of non-deliberate forms, should be prevented. As such, this measure relates directly to the prevention of fraud.

For PhD candidates who submit their own research proposal: The taskforce recommends having the proposal checked for plagiarism, to assess whether the work is indeed by the same candidate, and whether there is any overlap with work of others. In case plagiarism is clearly established, this should be a reason to categorically reject the research proposal as well as the candidate.

For all PhD candidates: This concerns a plagiarism check on the draft version of the final version of the manuscript. The results of the check should be shared with the supervisor and the candidate. In case of doubtful results of the plagiarism check, the candidate will be asked to reflect and comment upon the results.

In case of serious results, and if the candidate does not provide a satisfactory reply for the plagiarism, sanctions should be possible. The ultimate sanction is removal from the doctoral programme and denial of the possibility to obtain a PhD degree from the Erasmus University. In addition, such cases

will be recorded with the Executive Board (and the VSNU, for PhD candidates funded from 1st stream funds).

8.4 Doctoral regulations ('promotiereglement')

The current doctoral regulations do not pay particular attention to scientific integrity. There are several moments in the doctoral curriculum where this can be an issue. The taskforce recommends that the Executive Board revises the doctoral regulations, and considers a discussion in VSNU context about a further strengthening of the doctoral education as far as scientific integrity is concerned. The following issues play a role in this discussion:

- 1. Transparency of authorship of the thesis (article 4.5.2.c)**

Every discipline has their own culture about co-authorship for parts of the thesis. While in some disciplines it is unusual to have the promotor as the co-author, in other disciplines the empirical chapters have more than one author (such as promoter, co-promotor, fellow PhD candidates in the same research programme). The taskforce recommends as much transparency as possible on the contributions of each (co-)author, not just for the inner doctoral committee, but also for the plenary committee and the future readers of the thesis.

- 2. Possibility of providing feedback by the inner doctoral committee (article 6.3.3)**

- 3. According to the current doctoral regulations, the inner doctoral committee is not allowed to provide content feedback.** This means that, formally, any shortcomings identified by the inner doctoral committee cannot be taken into account in the final version of the manuscript. This is a missed opportunity, both in terms of scientific and scholarly quality, but also regarding scientific integrity, in case improvement on that level is desirable. The taskforce recommends revision of the relevant article, allowing the possibility for the inner doctoral committee to provide feedback. This does not include the possibility to give a *conditional* approval: the incorporation of the feedback remains the responsibility of the PhD candidate and promotor(s). The feedback provided by the committee member is administered by the secretary of the committee and may be used in the defence deliberations.

- 4. Minimum number of non-EUR members of the committee (articles 6.1.6 and 7.1.4)**

The current doctoral regulations state in article 6.1.6 that at least half of the members of the inner doctoral committee should be members of staff of Erasmus University. Article 7.1.4 states that at least half of the members of the committee should be staff members of EUR. This means in practice that all members can be affiliated to EUR, and above that, from the same faculty and the same research programme. This does not seem favorable to impartiality. The taskforce recommends that at least one member of the inner doctoral committee is from outside EUR, and on top of that, at least one additional member of the plenary committee is from outside EUR.

- 5. Intensifying the criteria for admission to the graduation ceremony on the basis of the assessment of the inner doctoral committee (articles 6.3.3, 6.3.4 and 6.3.6)**

The current doctoral regulations stipulate that the assessment by the inner doctoral committee is only of a qualitative nature. Scientific integrity should be part of this assessment. The taskforce therefore recommends that the inner doctoral committee explicitly assesses the manuscript on the five principles as laid down in the VSNU code. Should there not be unanimity among the committee members on approval of the thesis, and the reason for this is a possible violation of one of the five values of scientific integrity, then the thesis should be rejected at this stage. The current regulations only state that a majority of votes is necessary for rejection of the thesis.

Chapter 9: External relations

Taskforce recommendations contract research:

- In order to secure academic independence and a level playing field with other universities, the EUR should continue to push for changes in the Dutch government's general terms for contract research (ARVODI) that are currently in conflict with the right to publish results of contract research.
- Faculties and researchers should give due attention to possible conflicts of interest arising from contract research and from endowed chairs, and should avoid these conflicts as much as possible.
- It is impossible to define general rules on ownership of data collected in contract research, but in all contracts, due attention should be given to defining data ownership.
- Data collected during contract research should be stored appropriately and be available for peer review; the recommendations of chapter 6 apply for contract research as well.
- *Realized: the right to publish results of contract research is part of the approved general EUR Contract Terms for Commissioned Research*
- *Realized: EUR Contract Terms for Commissioned Research stipulate that commissioning parties can only publish outcomes of contract research after approval of the researchers.*

Taskforce recommendations media relations:

- All EUR researchers are made aware of key professionalism and integrity considerations in media relations.
- Central and local marketing & communication departments explicitly integrate these considerations in their policies, documentation and training.
- *Realized: The considerations identified have been included in a new brochure recently published by SMC.*

9.1 Contract research

There are several potential integrity issues involved in doing contract research – research commissioned and paid for by a commercial or non-commercial third party. These issues can be subsumed under two labels, scientific independence and data management in contract research. This leads to the following list:

- a) The right to publish the outcomes of commissioned research
- b) Avoidance of conflicts of interests between commissioning and executing research
- c) Preventing abuse of research outcomes by commissioner
- d) Establishing rules on ownership of data used and collected in contract research
- e) Guaranteeing appropriate storage and availability of data for peer reviewing

In 2005, the KNAW published a report on contract research⁴⁴, containing an inventory of practices as well as some important recommendations, in particular in the area of securing academic independence. Most universities, including EUR, proved to have guidelines regarding contract research. But by then, the EUR was one of the four (out of fourteen) universities not having model contracts or general conditions for contract research. This has now changed. During academic year 2012/2013 another EUR working group was operating with the assignment to elaborate general contract terms for commissioned research. Members of the Taskforce on Scientific Integrity were

⁴⁴ KNAW, Wetenschap op bestelling: Over de omgang tussen wetenschappelijk onderzoekers en hun opdrachtgevers. KNAW-Werkgroep opdrachtonderzoek, 2005

able to connect with that other group and to provide inputs for the Terms and for the Explanatory Notes (for internal use only). These EUR Contract Terms for Commissioned Research and Explanatory Notes have been approved by the Board in March 2013.⁴⁵

a) The right to publish the outcomes of commissioned research

The new Terms address issue a) by including the right to publish the outcomes of commissioned research (article 8, 9.2 and pages 6 and 7 of Explanatory notes). In itself this is good, but it is perhaps not in all cases sufficient. As the KNAW already observed in 2005, this right to publish may be in conflict with general terms that are held by commissioners, and in particular the Dutch government. These general terms, called ARVODI (Algemene RijksVOorwaarden Dlensten), hold that intellectual property rights of the research outcomes is with the commissioner (art. 23.1), and that researchers must ask approval from the commissioner before publishing results (art. 23.6). According to the Explanatory Notes of the EUR contract terms, it is up to each individual researcher or research group to find a solution for this. However, not accepting ARVODI implies loss of competitiveness vis-à-vis other universities. For this reason the KNAW (2005, p. 9) recommended that the government accept the KNAW “Declaration of scientific independence” (including the right to publish results) for all government-commissioned research, thus achieving the de facto elimination of currently used conditions that are in conflict with this independence.

By-mid 2013 this has not happened yet. However, the Taskforce contacted the EUR representative in the University Lawyers Network of the VSNU. On his initiative, representatives of VSNU and the Ministry of Education and Culture are going to meet with representative of the Ministry of Internal Affairs, in order to start the discussion on changing article 23 of ARVODI.

In order to guarantee academic independence and a level playing field with other universities, we recommend that the EUR continues to push for changes in ARVODI that are currently in conflict with the right to publish results of contract research.

b) Avoidance of conflicts of interests between commissioning and executing research

A conflict of interest may arise in two occasions. First, EUR researchers may at the same time have an affiliation with agencies commissioning research, as employee or as executive or non-executive board member. In this case, EUR researchers should preferably not be involved in implementing the contract research. Secondly, researchers appointed on an endowed chair financed by an agency or company may be expected to carry out research on the area of activity of the agency or company.

Currently, the EUR does not have rules to avoid these possible conflicts of interest. The working group on defining general contract terms has considered the matter, but discovered that it was impossible to define such rules in a way acceptable to all different faculties. The Taskforce recognizes these difficulties as well. As a result, we recommend that faculties give due attention to these two possibilities of conflict of interests and attempt to avoid them as much as possible. The same holds for researchers involved.

c) Preventing abuse of research outcomes by commissioner

The contract terms include rules for publishing results by the commissioner (article 9.3 and page 7 of the Explanatory notes). Commissioner can only publish data analyses and research outputs after approval from researcher. This issue has been dealt with sufficiently.

⁴⁵ The Terms only apply for the Woudestein faculties. EMC has its own terms.

d) Establishing rules on ownership of data used and collected in contract research

It was quickly realized that it is very difficult to establish general rules regarding data ownership in contract research. Yet it is important that ownership is defined at the start. We therefore recommend that in all contracts, due attention is given to data ownership.

e) Guaranteeing appropriate storage and availability of data for peer reviewing

For contract research it is important that collected data are stored appropriately and are available for peer review. However, there is no difference with other research on this matter, so we recommend that the rules and guidelines as stipulated in chapter 6 of this report are applied.

9.2 Media relations

In recent times, researchers and universities are increasingly (encouraged to be) active in communicating with the media and the general public with regards to research.

Scientific professionalism and integrity also implies that research and research outcomes are communicated in a responsible way. This implies: scrupulously, reliably, verifiably, impartially and independently.

Recently, the KNAW has emphasized that besides these five core principles, research should also act *honestly* (e.g. in terms of acknowledging the limitations of their research) and responsibly (e.g. be able and willing to defend the choice for a specific research design)⁴⁶. These two principles also have a bearing on research communication.

In communication on research, these principles play a role in the following main considerations:

- Choosing the appropriate **moment**. In general, one should be reluctant to approach the media on the basis of early research outcomes, for instance from a pilot study. Be restrictive in approaching the media with research that has not yet undergone peer review or has not yet been published. Some media will even be unwilling to publish research findings in that stage.
- Determining who has the **initiative**. As a general, the (institution of the) first author of a publication or the research leader should have the lead in contacting the media. In any case, make explicit **arrangements** in case of multiple researchers being involved, also internationally.
- Indicating the **stakeholders** by whom the research is possibly financed, facilitated and/or commissioned. Where relevant, indicate whether and to what extent these stakeholders have had influence on the design, execution and reporting of the study.
- Be honest with respect to the **limitations** of a study, for instance when research is conducted in a certain period or population.
- Report results **reliably and verifiably**; for instance, do not leave out important findings, and do not speculate extensively.

⁴⁶ KNAW (2013), *Vertrouwen in Wetenschap*. Advisory report, May 2013. www.knaw.nl

Chapter 10: Monitoring

Taskforce recommendations monitoring:

- Appoint one coordinator per school for all integrity issues, but allow for the possibility to appoint one specific coordinator for scientific integrity.
- Following the suggested tasks and responsibilities as detailed below.

10.1 Background

The goal of this project was to (re-)define the monitoring role of the integrity coordinators at EUR.

Integrity coordinators were first appointed at the introduction of the Erasmus University Code in 2002. The code formulates the core values and core responsibilities of both employees and students, but it does not contain detailed rules of conduct. It serves as a reference with regards to everyone's responsibility within the work environment. **Professionalism, Teamwork** and **Fair Play** are the integrity values of the EUR.

In principle, every faculty / institute counts one integrity coordinator, who serves as the contact person for matter relating to the integrity code. As part of the package of measures to foster professionalism and integrity at Erasmus University, the taskforce has reviewed the tasks, responsibilities and expectations of the integrity coordinators, as main contact point for integrity issues within each faculty / institute. The taskforce (re-)initiated a meeting of the EUR integrity coordinators for input and feedback.

The taskforce also discussed whether or not dedicated officers should be appointed for scientific integrity, harassment, etc. As it became clear in the discussion that the coordinator is the first point of contact, and that he / she essentially has a referring role, it was recommended that a coordinator for integrity in the broad sense, including scientific integrity, is appointed.

10.2 References to rules and regulations

- EUR Integrity Code: www.eur.nl/english/eur/publications/integrity (May 2013)
- Information on integrity issues and examples of dilemmas: [Integrity Code](#)
- Doctoral Regulations (in particular article 2.5 on responsibility for the thesis; May 2013): http://www.eur.nl/english/ab/registrars_office/phd_defence_ceremonies/doctoralregulations/
- EUR scientific integrity complaints procedure: can be found [here](#)
- E-mail address scientific research confidential advisor (prof. P.J.F. (Patrick) Groenen): vertrouwenspersoon.onderzoek@eur.nl⁴⁷
- The Netherlands Code of Conduct for Scientific Practice: '[The Netherlands Code of Conduct for Scientific Practice](#)'
- National Board for Research Integrity: <http://www.knaw.nl/Pages/DEF/28/889.bGFuZz1FTkc.html>

10.3 Taskforce recommendations on integrity coordinators

- One contact person (integrity coordinator) per faculty / institute for integrity matters in general, including scientific integrity, as originally defined in 2002;
- Adoption of the tasks and responsibilities for integrity coordinators by the Executive Board in September 2013;

⁴⁷ http://www.eur.nl/english/eur/publications/integrity/integrity_coordinators/ (May 2, 2013)

- Implementation of the recommendations of the taskforce, as well as of the decisions of the Executive Board in relation to scientific integrity is the responsibility of each faculty / institute;
- The general website on integrity (www.eur.nl/integrity; www.eur.nl/integriteit) will be adapted and expanded with information about integrity coordinators, scientific integrity and the various projects of the taskforce.

Recommended tasks and responsibilities integrity coordinators:

General

- The faculty / institute board is responsible for integrity;
- An integrity coordinator will be appointed within the following faculties / institutes / sections: each of the faculties, iBMG and ISS, ABD / SSC's, University Library;
- One contact person at central level is appointed (ABD) for the integrity coordinators, also chair to the meeting of integrity coordinators;
- Appointment of an integrity coordinator is for a period of two years, with a possibility for extension for another two years by the administration of the relevant faculty / institute;
- The integrity coordinator produces a yearly report of the activities within the department; the contact person at central level produces an overarching report for EUR as a whole;
- Action plans are discussed at the meeting of integrity coordinators;
- The integrity coordinators meet at least once a year, chaired by the contact person at central level, or on request of one of the coordinators;
- An annual workload of 40 hours is available for the integrity coordinator.

Tasks

- The integrity coordinator stimulates and enhances a culture of integrity within his / her department / organisation, which induces integrity, and where matters concerning integrity can be addressed.
- The integrity coordinator ensures proper and regular attention for matters relating to integrity within his / her department / organization.
- The integrity coordinator stimulates awareness of the integrity code among staff members, students and departments, and ensures that they are up to date regarding rules and regulations on integrity.
- The integrity coordinator is responsible for the coordination of activities on integrity within his / her department / faculty.
- The integrity coordinator functions as information point to students and staff members within his / her department on the broad domain of integrity matters.
- The integrity coordinator refers people to the relevant officer / contact person / counsellor in case of a question about integrity (to a counsellor in case of harassment, or to the integrity counsellor in case an integrity breach is suspected).
- The integrity coordinator contributes to the overall EUR policy on integrity and sits on the working group of EUR integrity coordinators.

Rights

- The integrity coordinator will have sufficient means and trust to be able to perform his / her tasks;
- The integrity coordinator shall in no way be disadvantaged as a result of exercising his / her role as integrity coordinator;
- The dean / vice-dean / rector ISS / secretary Executive Board / librarian shall inform the integrity coordinator of the relevant developments and integrity issues in his / her department.

Profile of the integrity coordinator

The integrity coordinator:

- Has a permanent appointment with Erasmus University;
- Has a position with sufficient weight and seniority within the department, and is a serious interlocutor with the board of the faculty / institute;
- Has an approachable personality;
- Can function as a spider in the web;
- Is aware of the rules, regulations, procedures and activities, EUR-wide and at faculty level, regarding integrity, or is willing and able to acquire the relevant knowledge.

Appendix 1: Report: Research Data Management at Erasmus University Rotterdam

(See separate report)

Appendix 2: PhD courses on scientific integrity

This appendix contains:

1) The announcement of the 2012-2013 integrity course for PhD students from Erasmus MC. More information can be obtained from Dr. Suzanne van de Vathorst;

2) The Course Manual for the 2012-2013 ERIM course on scientific integrity
The Tinbergen Institute – Graduate school and research institute (economics, econometrics and finance) also offers scientific integrity courses. Information can be obtained from Prof. Dr. Bauke Visser.

Announcement Integrity course for Erasmus MC PhD students:

Integrity in Research, does it exist?

Science is a competitive field. There is a lot at stake, reputations, careers, money. The downside is that there is a lot of pressure and this may lead to quarrels on authorships, data massage, incomplete informed consent, and etcetera. When is a researcher a person of integrity? And is it possible to combine integrity with a successful scientific career?

The department of Medical Ethics and Philosophy organises two courses on research integrity in 2012. The first, English-spoken, course will be held on three consecutive Tuesday evenings, on April 10th , April 17th and April 24th from 17.30 – 20.30 hr. A second, Dutch-spoken course will be given in September, on the 4th, 11th and 18th.

Sandwiches and drinks will be provided.

During this course we will exchange experiences, analyse cases and discuss ethical dilemmas. We will examine the ethical questions that are raised by working in a scientific and institutional setting. Examples of such issues are the dependency of junior researchers on their seniors, cooperation between research groups, authorships, handling data, presenting results. Problems that are raised by this are often seen as being of the type: “good people, unsupportive environments?”, and “strong temptations, weakness of will, stringent regulations?” Should we put our trust on personal integrity or rely on robust supportive systems?

The course is offered free of charge, on the condition that the student attends all meetings and is able to conclude with a certificate. Cancellation after April 1st, failure to attend all meetings or to write a satisfactory concluding essay, will be followed by an invoice for 200 euros send to the head of the Department.

We will invite experienced guest speakers. A reader with background articles will be distributed. The student will have to submit an individual essay, discussing the newly acquired knowledge, at the end of the course, as part of the course. If the student has been present at all three meetings, and completed the essay, he/she will receive two ECTS.

The number of participants is limited to 20. It is advisable to subscribe ASAP, subscription for the April-course closes on March 12th 2012. Subscription for the September course (in Dutch) will close the 3rd of August. Cancellation not possible after August 31st.

Questions re the course, and subscription to the course can be addressed to dr Suzanne van de Vathorst (dept Medical Ethics and Philosophy) s.vandevathorst@erasmusmc.nl (Hoboken, building AE , room 340. On subscribing please name your department, the head of your department, and give a short description of your research project .

Scientific Integrity: On scrupulousness and integrity

ERIM PhD/MPhil program 2012-2013

Detailed course manual

Prof dr Finn Wynstra (fwynstra@rsm.nl); T10-54, Tel. 81990; coordinator)

Prof dr Patrick Groenen

Course assistant: David van Ass (vanass@rsm.nl)

Like in any profession, scientists are frequently faced with integrity dilemmas:

Can I exclude particular observations from my research? Can I leave out certain statistics from the analysis I report? Can I use the same data set, or “idea”, in multiple papers? Should I agree on a colleague being a co-author on a paper to which (s)he has not made a significant contribution?

This course aims to expose you to such integrity dilemmas, and to support you in developing and honing your own “moral compass”. We do this by discussing the context of the principles, values and rules such as they apply to the field of management research in general, and to our university and research institute in particular. But we also do this by discussing specific dilemmas with you, and in particular by letting you discuss among yourselves.

As such, the course serves as a foundation for other courses, in particular the various research methodology courses and the course on Publishing Strategy, during which you will discuss in more detailed terms the various dilemmas you will be facing in your PhD studies and possible further academic career.

The course consists of four sessions, built around three key elements:

- Providing information
- Self-reflection
- Debate

In some of the sessions, we will conduct a dilemma “game”, in which you will discuss in groups actual dilemmas and possible courses of action to deal with such a dilemma. The objective of the game is to let you explore how different people think about various dilemmas, and let you develop, explicate and hone your own moral compass in these matters. The purpose of this exercise is to learn how to form and formulate an opinion on a dilemma in scientific conduct.

Assignments:

All assignments need to be done prior to the session for which they are listed. For the assignments 2 and 3, you need to form teams of four, based on a shared interest in one of the following research strategies:

- Survey
- Existing data study
- Experiment
- Qualitative study
- Quantitative/modeling study

Please inform the course assistant by email of the composition of your team by 2 November; your group will then also receive a number.

Course workload:

The course load is set at 1 ects, equaling 28 hours of study. Roughly, this course load consists of the following elements:

- Attending the sessions: 4 * 3 hrs = 12 hrs

- Preparing the readings, videos: $4 * 2 \text{ hrs} = 8 \text{ hrs}$ (note: readings not equally distributed over sessions)
- Assignments (one individual, two group-based): $1 + 2 + 5 = 8 \text{ hrs}$

Session I:	Scientific Integrity and Research Misconduct: an Introduction
When:	31 October, 9-12
Where:	T3-25

Description:

In this first of four sessions, we discuss the notion of integrity and scientific integrity. We identify and discuss the main forms of research misconduct: fabrication, falsification and plagiarism.

In the middle part of this session, we do a debrief of The Lab.

Finally, we pay attention to the particular challenges and responsibilities that you as MPhil or PhD students are facing.

Topics:

- Notions of Integrity
- Scientific Integrity
- Research misconduct: Fabrication, falsification and plagiarism
- Promoting ethical behavior: principles versus rules
- Being a PhD student at ERIM: roles, responsibilities and values

Preparation:

Carefully study the assigned readings.

Assignment 1 (individual):

The Lab: Avoiding Research Misconduct is a Virtual Experience Interactive Learning Simulation (VEILS) program. Participants will assume one of four playable roles: a graduate student, a postdoctoral student, a principal investigator, or a research integrity officer. In each segment, the character has to make decisions about how to handle possible research misconduct. The story spins off in different directions, depending upon the choices participants make as the character. The decisions that each character makes have consequences that not only affect that character's future, but also the future of others in the lab. Each choice or combination of choices brings results that must be dealt with.

Your assignment is to play the role of Kim Park, graduate (PhD) student in the "The Lab":

<http://ori.hhs.gov/TheLab/TheLab.shtml>.

Note the critical decision points, and be prepared to recall and argue for your choices during class.

Literature:

- Steneck, N.H. (2006). Fostering integrity in research: definitions, current knowledge, and future directions. *Science and Engineering Ethics* 12: 53–74.
- National Academy of Sciences (2009). Introduction to the Responsible Conduct of Research. In: *On Being a Scientist: A Guide to Responsible Conduct in Research* (3rd Ed.), pp. 1-3.
- National Academy of Sciences (2009). Advising and Mentoring. In: *On Being a Scientist: A Guide to Responsible Conduct in Research* (3rd Ed.), pp. 4-7.
- National Academy of Sciences (2009). Research Misconduct. In: *On Being a Scientist: A Guide to Responsible Conduct in Research* (3rd Ed.), pp. 15-18.
- EUR Integrity Code.
- Academy of Management, ethics video series:

- Plagiarism:
http://www.youtube.com/watch?v=qDaKowkuuy&list=PL65B059BC12E75502&index=2&feature=plpp_video

Background reading:

- Kaptein, M. and Wempe, J. (2011). Three General Theories of Ethics and the Integrative Role of Integrity Theory. Working paper, Social Science Research Network, SSRN-id1940393. (52 pp).

Session 2:	Beyond Falsification, Fabrication and Plagiarism
When:	28 November, 9-12
Where:	T3-25

Description:

In the first part of today's session, we extend the discussion of research misconduct to more nuanced forms of misconduct. We introduce and discuss in groups a number of dilemmas related to questionable research practices. We review such questionable research practices along the different phases of (empirical) research: collection, processing, analysis and archiving of data.

The session also discusses the code of conduct applied by all Dutch universities, and some of the practical issues around reporting misconduct.

In the middle part of this session, we discuss your assignments.

At the end of this session, senior PhD students will share their own experiences in facing integrity dilemmas in conducting and reporting research.

Topics:

- The Grey Zone: Questionable Research Practices
- Personal characteristics and integrity risks

Preparation:

Carefully study the assigned readings.

Assignment 2 (team):

Questionnaires are widely used to measure traits of the respondents. For example, a company may be interested to measure job satisfaction of its employees by asking several questions related to job satisfaction. Summing the answers to the job satisfaction items yields a job satisfaction score for each individual. In this way, a job satisfaction scale can be constructed. Of course, additional validation of these items is necessary before they could be used.

The goal of this assignment is to **develop scales** that could be used in a survey to measure the susceptibility of respondents to several aspects of (scientific) misconduct. Each scale consists of 7-10 statements and respondents use a 7-point Likert scale (1 = not like me at all, 7 = very much like me). An example statement for an item could be: "I enjoy presenting my work at conferences." It is considered good practice to have the statements not all formulated in the same direction to avoid respondents answering all items positively (or negatively). It is also common to formulate the items not too directly because they may lead to socially desirable answer instead of the opinion of the individual.

We distinguish five scales around five topics related to (scientific) misconduct. This assignment is done in groups of 4 students. Each group produces a set of 7-10 statements such that should cover the topic/scale assigned to your group.

- Groups 1 and 6: *Sensitivity to peer pressure*.
Here, we would like to know to what extent the respondent is sensitive to social pressure. One would also like to know whether or not the respondent prefers working on his/her own.
- Groups 2 and 7: *Selling your scientific work*.
Is the respondent willing to bend reality so that the story becomes better?
- Groups 3 and 8: *Accuracy*.
How accurate is the respondent when it comes to organizing his/her scientific work?
- Groups 4 and 9: *Power issues*.
In case the respondent has power, in what situations would he/she exert the power? Would this also be done to his/her own benefit?
- Groups 5 and 10: *Sensitivity to own gain*.
This sensitivity should be considered in the context of paying taxes, financial declarations, or other forms of personal gain or recognition, etc.

For each item of your scale, provide a brief explanation of what you intend to measure with this item. Upload your proposal for the items of your scale to BlackBoard, by Friday 23 November, 17.00.

Before coming to the session, download and review the proposal of the other group assigned to the same topic/scale and prepare comments.

Literature:

- Martinson, B.C., Anderson, M.S. and De Vries, R. (2005). Scientists behaving badly. *Nature*, 435: 737-738.
- Association of Universities in the Netherlands (VSNU). The Netherlands Code of Conduct for Scientific Practice. (10 pp)

Session 3: Integrity in Management Research and Publication Ethics

When: 16 January, 14-17

Where: T3-24

Description:

In this session, we first discuss specific integrity risks and dilemmas in management research.

For illustration, we also review some specific ethics code in the management research field.

In the middle part of this session, we again introduce and discuss in groups a number of dilemmas.

In the final part, we deal specifically with publication ethics.

Topics:

- Integrity in Management Research
- Examples of ethics codes in management research
- Publication ethics

Preparation:

Carefully study the assigned readings and the on-line videos.

Assignment:

No assignment for this session – note assignment 3 for the final session!

Literature:

- Academy of Management. Code of Ethics. (6 pp)

- Albert, T. and Wager, E. (2003) How to handle authorship disputes: a guide for new researchers. Committee on Publication Ethics (COPE), Report 2003. (3 pp)
(<http://publicationethics.org/files/u2/2003pdf12.pdf>)
- Academy of Management, ethics video series:
 - Reporting research
http://www.youtube.com/watch?v=fbFqLvuosbM&list=PL65B059BC12E75502&index=6&feature=plpp_video
 - Authorship
http://www.youtube.com/watch?v=l3wEmi1rMeQ&list=PL65B059BC12E75502&index=1&feature=plpp_video
 - Conference papers and presentations
http://www.youtube.com/watch?v=azwpp6kFcq0&list=PL65B059BC12E75502&index=5&feature=plpp_video
 - Slicing the data in publications
http://www.youtube.com/watch?v=gGo9oK8v0bc&list=PL65B059BC12E75502&index=3&feature=plpp_video
 - Publishing in journals
http://www.youtube.com/watch?v=NmsiDzgvQyM&list=PL65B059BC12E75502&index=4&feature=plpp_video
 - Global ethics in publishing
http://www.youtube.com/watch?v=2tyAXzrHW98&list=PL65B059BC12E75502&index=8&feature=plpp_video

Session 4: Integrity and scrupulousness in different research strategies

When: 30 January, 14-17

Where: T3-17

Description:

In the first part, we ask you to report on the interview assignments. (A selection of) video interviews will be shown in class.

In the second part of this session, we have a panel of experts from the ERIM community who will discuss how different research strategies pose different risks in the different phases of (empirical) research (collection, processing, analysis and archiving of data).

At the end of this session, each of you will personally sign a pledge.

Preparation:

Carefully study the assigned readings.

Assignment 3 (team):

In your teams, you have the task to **interview a faculty member** (assistant/associate/full professor) who is actively engaged in the research strategy of your choice. Preferably, this should be an ERIM member, but also other faculty members from the EUR can be interviewed.

The interview should focus on the main risks and dilemmas regarding scientific (specifically: research) integrity, and the best ways to promote integrity. Sample questions could be:

- *What "perils" are present for your line of research / what is the "gold standard" of good behavior?*
- *How can we best stimulate attention to scientific scrupulousness and integrity? What is the role of individuals; research groups; schools/universities; associations; reviewers/journal editors..?*

The interview may be reported in writing, but we particularly encourage you to video the interview so that we can show and discuss the interview during this class session.

Literature:

- ERIM report “Management of Research Data” (38 pp.)

Note: This is a working document; the formal status of the report is still to be determined.

Appendix 3: Procedure pledge-taking ceremony ERIM PhD students (January 2013)

The ceremony formed the conclusion of the PhD course on research integrity (see Appendix 2). In total, it lasted about 20 minutes (with some 40 people taking the pledge), and took place in a regular lecture theatre.

The course leader explained the procedure and then showed the text of the pledge on a projector:

The exact text of the pledge is:

I hereby declare to uphold the ethos of good scientific research and to apply throughout my scientific activities the principles described in The Netherlands Code of Conduct for Scientific Practice: Scrupulousness, Reliability, Verifiability, Impartiality, and Independence. I will apply these principles in my own work, and will endeavour to promote these principles among other scientists and in particular my direct colleagues.

Subsequently, each of the course participants was called to the front of the room.

Each person was then asked:

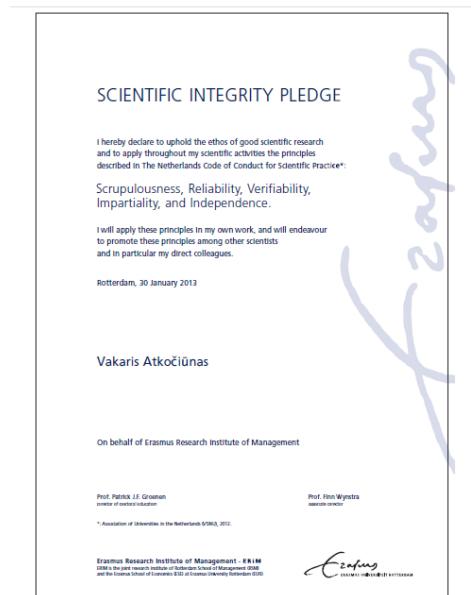
Do you, XX, promise to uphold the principles on ethical research behaviour as outlined in the pledge?

People were answering with: Yes, I do. (no hand-raising)

Each participant then received two pledges to sign, one to keep and one for the records (of ERIM).

The course leaders had signed the pledges already earlier.

Participants also received a booklet with the VSNU code of conduct.



Research Data Management at Erasmus University Rotterdam

Table of Contents

1	Summary: responsibilities and recommendations.....	2
2	Background and structure of the project	3
3	Inventory of user requirements and data storage	4
3.1	Set up.....	4
3.2	Survey results	5
3.3	Assessment and discussion	9
3.4	Recommendations based on the survey results	10
4	Safe storage and collaboration.....	10
4.1	Criteria and candidate platforms	11
4.2	@wEURk storage cluster	11
4.3	Dutch Dataverse Network	12
4.4	Recommendations	12
5	Data management workflows and protocols	13
5.1	Approach taken	13
5.2	What should be stored.....	14
5.3	Introduction and implementation of a data management protocol	17
5.4	Stakeholders, responsibility and supervision.....	19
5.5	Discussion	19
5.6	Recommendations	20
6	Erasmus Behavioural Lab pilot.....	21
6.1	Goal	21
6.2	Setup.....	21
6.3	Current workflow	22
6.4	Analysis phase	22
6.5	The proposed new EBL system.....	22
6.6	Proof-of-Concept.....	23
6.7	Recommendations	23
7	Awareness raising: website, front office and workshop	24
7.1	The data support web site.....	24
7.2	The virtual information desk	25
7.3	Workshop and courses	26
8	Cost estimations	26
8.1	Recommendation	28
9	Appendix 1 – Respondents and project team.....	29
10	Appendix 2 – Survey questionnaire.....	30
11	Appendix 3 – Findings from the Dutch Dataverse Network pilot	39
12	Appendix 4 – General recommendations for storing research data.....	45
13	Appendix 5 – Digitalisering en opslag van onderzoekdata	46
14	Appendix 6 – Concept data checklist for various disciplines	47
15	Appendix 7 – Raming structurele kosten na de EBL-pilot.....	49
16	Appendix 8 – Front office budget	50
17	Appendix 9 – Storage criteria matrix.....	52

1 Summary: responsibilities and recommendations

A solid research data infrastructure at the Erasmus University Rotterdam (EUR) has several stakeholders who play a role in responsible and reproducible research. This section relates the main recommendations made in the current report to these stakeholders and their responsibilities.

The *Rector Magnificus* and the *Executive Board* are responsible for a coherent and effective EUR research data management approach. The following recommendations pertain to the Rector Magnificus and the Executive Board (mainly based on Chapter 5):

- appoint a research data support officer at the central EUR level. Dedicated support at a central level is crucial for continuity and for sharing good practices between institutes or disciplines.
- draft a covenant to the extent that the deans of research or research directors and – at Erasmus Medical Center – the heads of department embed and evaluate research data management in their institute. They should have room to describe how they will fulfil the data management minimum.
- introduce the deans of research to the EUR data management minimum protocol.
- provide training and advice on research data management: delegated to the university library (see below).
- provide services for storage and retention of research data: delegated to SSC ICT (see below).

The *deans of research/research deans* and *heads of department* at the various faculties, graduate schools and institutes of the EUR are responsible for the organisational embedding of data management. The following recommendations pertain to them (more in Chapter 5):

- coordinate, supervise and stimulate data management workflows and protocols. In particular, refine the minimal EUR protocol such that it fits the institute's research discipline(s) and make it practical. Take care to explain that data management also includes documentation related to the research process.
- socialise young researchers into responsible ways of working. Research data management is not something extra. Rather, the attitude should be that professional research goes hand in hand with responsible data management.

Researchers are responsible for storing data and documentation at various moments during a study. The minimum that must be stored consists of both the raw data and the data underlying any submitted or published publication, the project plan, documentation that describes and explains major changes to the earlier plan(s), as well as the submitted version of the publication (see Chapter 5). It is essential that *PhD supervisors* and *research group leaders* are also role models for young researchers. Similarly, commitment and the willingness to share good practices are more important than protocols and covenants.

The *university library* has been given responsibility for raising awareness for data management and for providing training and advice for research data management. The following recommendations pertain to the university library (more in Sections 3.4, 4.4 and Chapter 7):

- develop the library's virtual desk into a front office for researchers, as a central point of expertise in research data management. This includes training and collaboration with long-term archives (the back office). Close collaboration with the Research Support Office (currently under development) is foreseen.
- develop and maintain an activating data support web site and select or develop relevant courses and workshops.

SSC ICT has been given responsibility for providing services for safe storage and retention of research data. The following recommendations pertain to SSC ICT (more in Chapters 3 and 4):

- create safe storage and backup facilities for individual researchers, as well as safe ways to share and collaborate on research data.
- maintain and offer expertise to make research staff aware of advantages and risks of particular storage platforms and media, in collaboration with the front office.

These recommendations and the division of responsibilities are based on an online EUR-wide survey, interviews and meetings with various researchers and experts from within and outside of the EUR, as

well as on data management documentation provided by other universities. When these recommendations are implemented and research data management continues to be taken seriously, the EUR will at least be on a par with other data-aware universities.

2 Background and structure of the project

Erasmus University Rotterdam (EUR) intends to offer its researchers a solid infrastructure for storing and archiving research data. Research data that is stored safely and well-documented can be made available to reviewers, fellow-researchers, or even the public at large – depending on a proper authorisation system. This ensures that researchers can be credited as well as be accountable for their results and that replication of research and studies is possible. Furthermore, research data that is available and reusable may eventually contribute to the advancement of science and increased citation rates for researchers. The project reported on here, however, has focused on storing data for the sake of replicability and accountability.

The envisioned data infrastructure provides both mid-term and long-term data storage; the former is most relevant during a study or a research project, while the latter is appropriate in case of (nearly) finished or published research. Data management in these stages is different. We will refer to these data storage situations as storing and archiving, respectively. Neither is a matter of mere technology; instead, data awareness is of the essence and the use of relevant technology must be embedded in the EUR organisation and in academic workflows. The project described in the current document will affect many people and multiple parts of the organisation. This in turn calls for communication as well as for tailoring to various kinds of research strategies with their respective data types.

This project has been carried out as a subproject of the EUR Taskforce Scientific Integrity (TSI). The objective of the taskforce has been to raise awareness for and to develop proposals to help maintain scientific professionalism and integrity. The current project has started on 24 April 2013 with a meeting of the members of the steering committee and the members of the project team (see Appendix 1). It has run for three months.

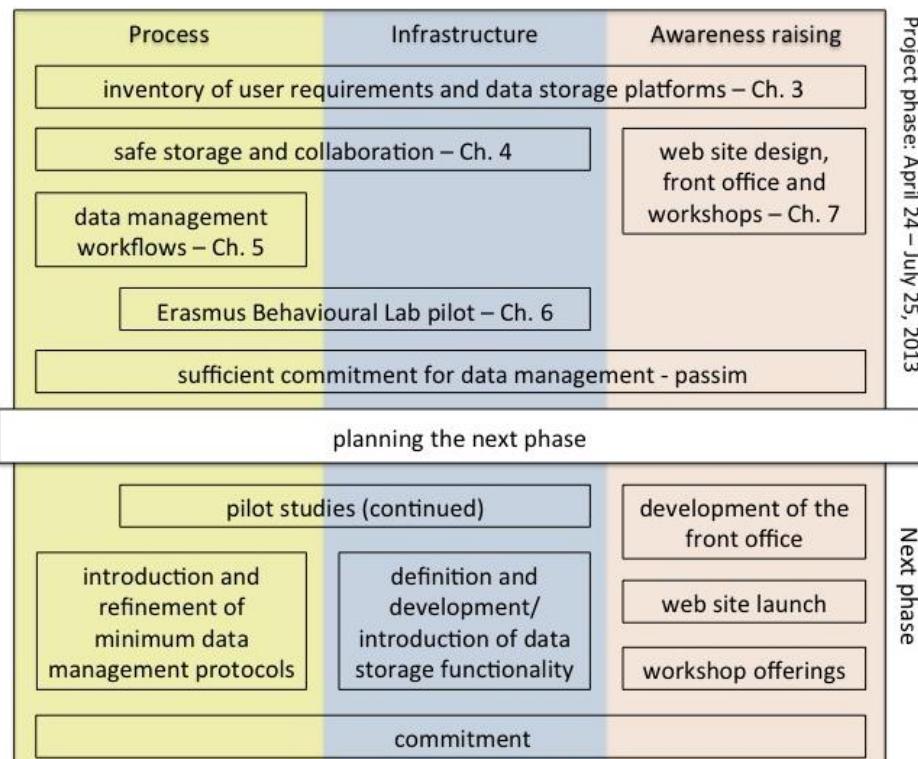


Figure 1 Work breakdown structure. The number of a building block refers to a chapter in this report.

Figure 1 illustrates the project's work breakdown structure, with activities in three work packages: Awareness Raising, Infrastructure, and Process. As was already motivated in the project plan, the EUR's ambition exceeds the time available for this project. Therefore the bottom half of Figure 1, from "Planning the next phase" onward, is a projection beyond the scope of the project reported on.

3 Inventory of user requirements and data storage

3.1 Set up

In the project plan the goal of this component was expressed as follows (summarised): "There is an overview of what EUR researchers actually use and need for storing their data during research in progress. An initial inventory in the form of a campus wide survey and interviews should be carried out to distil researchers' requirements. The survey and the interviews also serve as awareness raising instruments."

Between 23 May and 5 June 2013 an online survey was conducted amongst all EUR research staff concerning

- Data types, storage, backup
- Metadata and data management protocols
- Requirements and training needs

The survey built on a research data management survey done in the ADMire project¹ carried out at the University of Nottingham. One question was taken from a recent iBMG survey on data storage. Another question was inspired by the structured data management interviews developed in the CARDS project² carried out by the KNAW, several universities and SURF. The online survey was tested and revised before being distributed amongst all EUR researchers, including research staff at the Erasmus MC. The overall response was 25% (n=763). The endorsement email by the Rector Magnificus to research directors and deans of research in schools and institutes, and in turn their email to individual researchers contributed to this overall response. More than half of the respondents have entered their names, which the project team interprets as confirming the reliability of the answers. Response rates in schools and institutes differ (Figure 2).

Survey responses on storage have been discussed with a representative of SSC ICT (a) to assess possible safety risks of storage and backup behavior of researchers and (b) to explore possible alternative storage and backup facilities provided and serviced by SSC ICT (see Section 3.3).

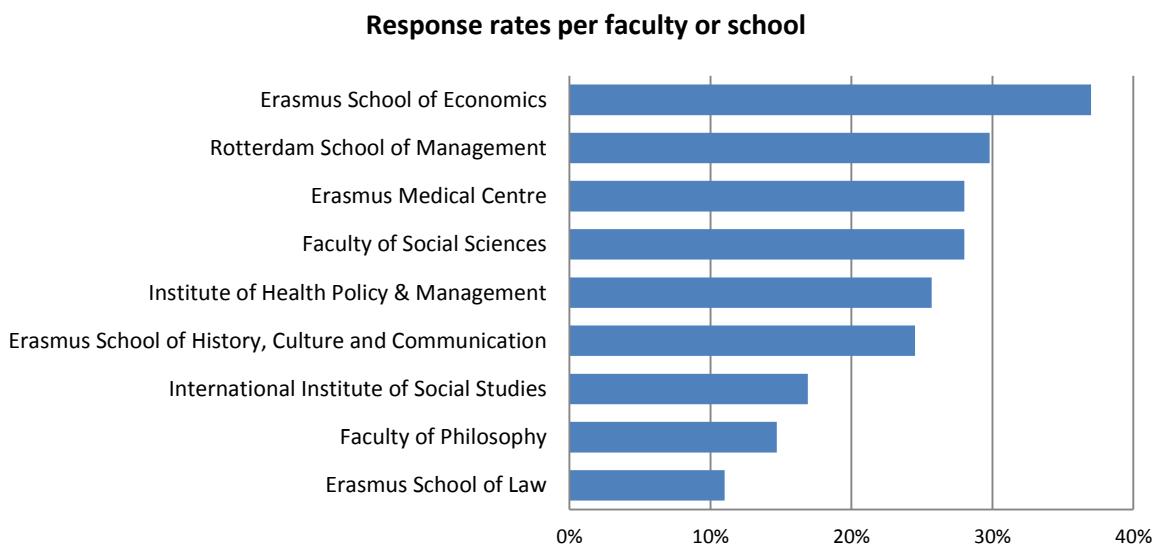


Figure 2 Survey reponse rates in faculties and schools

¹ <http://admire.jiscinvolve.org/wp/project-outputs/>

² <http://www.surf.nl/nl/projecten/pages/cards.aspx>

3.2 Survey results

The questionnaire can be found in Appendix 2. In this section we report on answers to the main multi- or single-choice survey questions. Where relevant, answers from the open text categories are used as well.

3.2.1 Data types, storage and backup

There is a rich variety of produced data types. Survey data and experimental data are the most widely - but not dominantly - produced data types at the EUR (Figure 3).

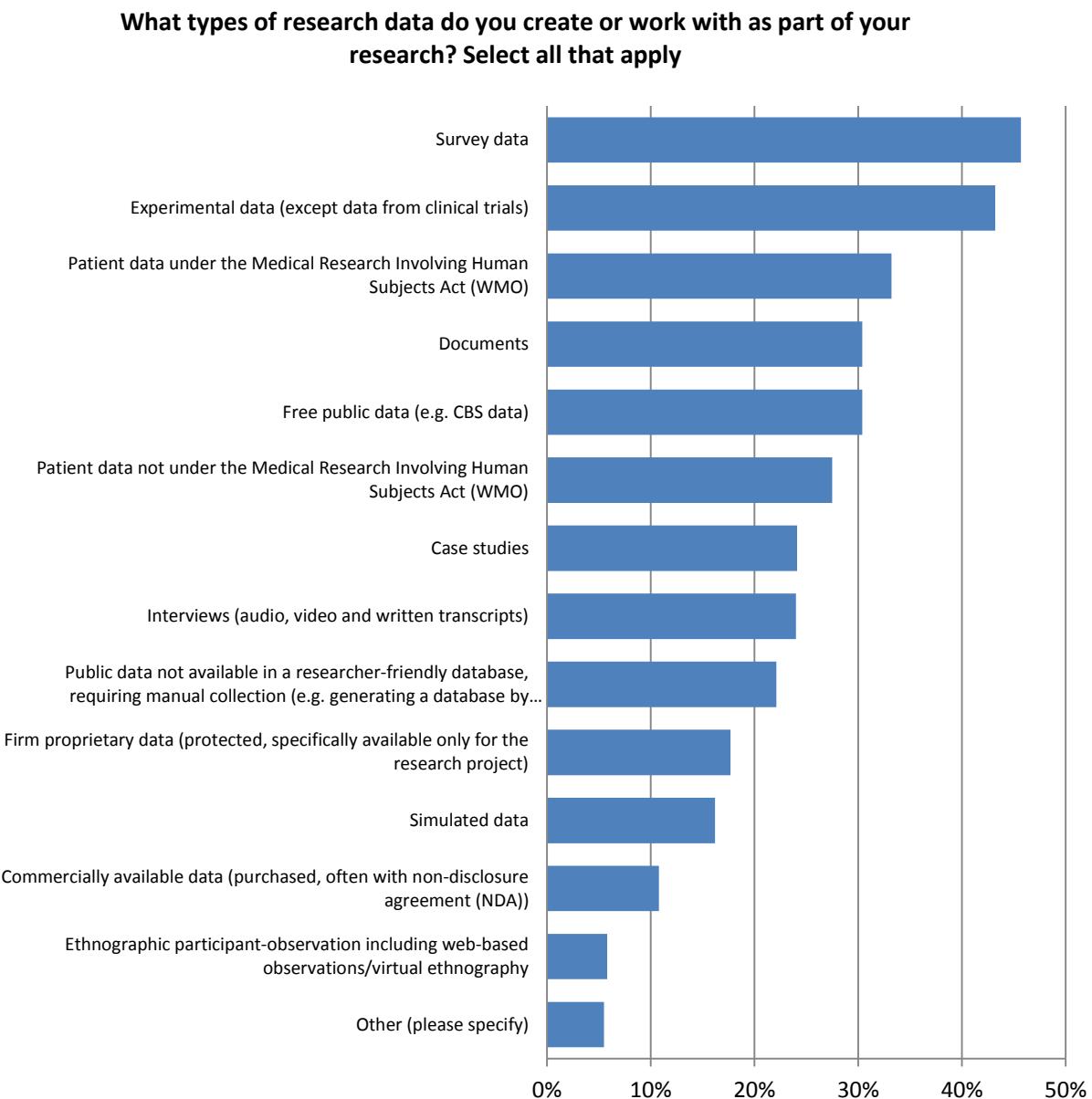


Figure 3 Types of research data

Researchers store or back up their data mainly on hard disks, either on their campus computer, computer at home, their notebook or on an external drive (Figure 4). Web-based services and USB/Flash drives are commonly used media as well. The shared drives/university servers (network drives) are mentioned by 30-40% of the researchers as storage or backup media.

Asked about the frequency with which respondents back up their data, almost a third answers to back up on an ad-hoc basis (Figure 5). However, broken down to school/institute level, we see that data at the Erasmus Medical Centre is backed up more systematically (Figure 6). On average 15% of all responding researchers do not know if their data is being backed up.

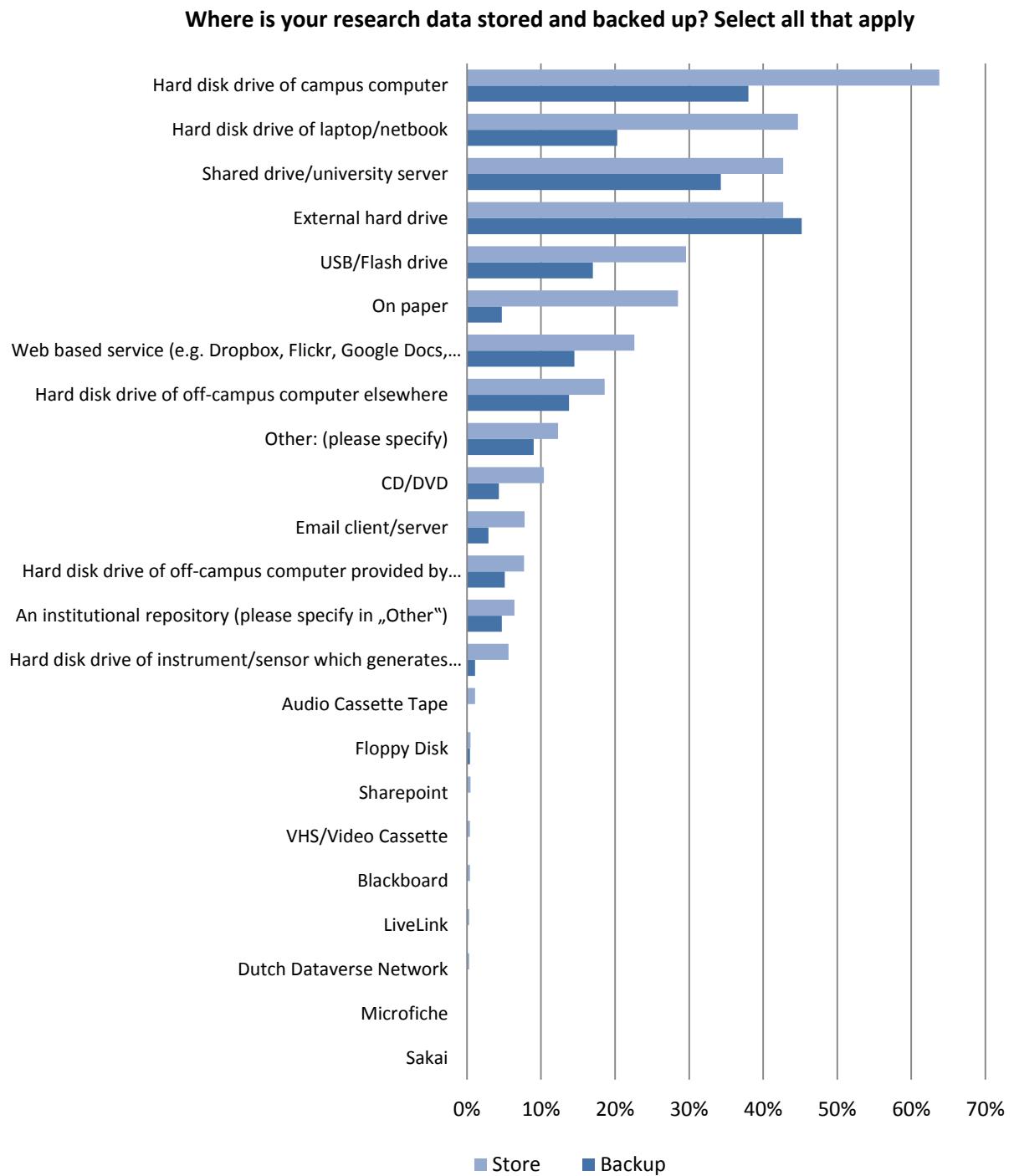


Figure 4 Data storage and backup locations

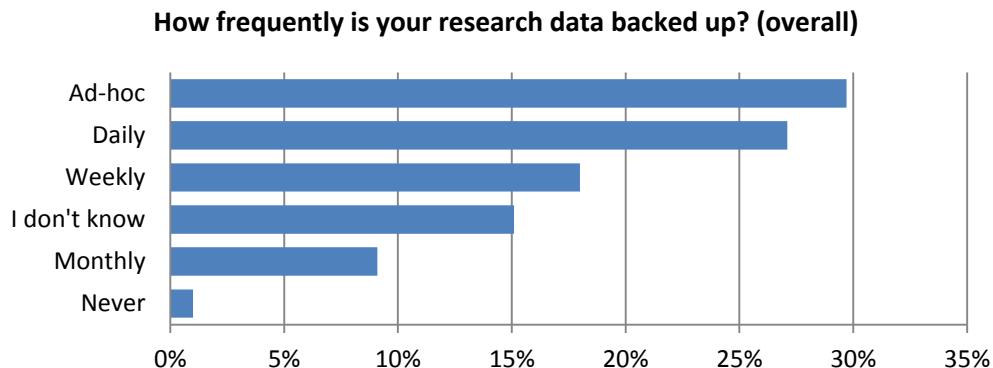


Figure 5 Backup frequency - overall

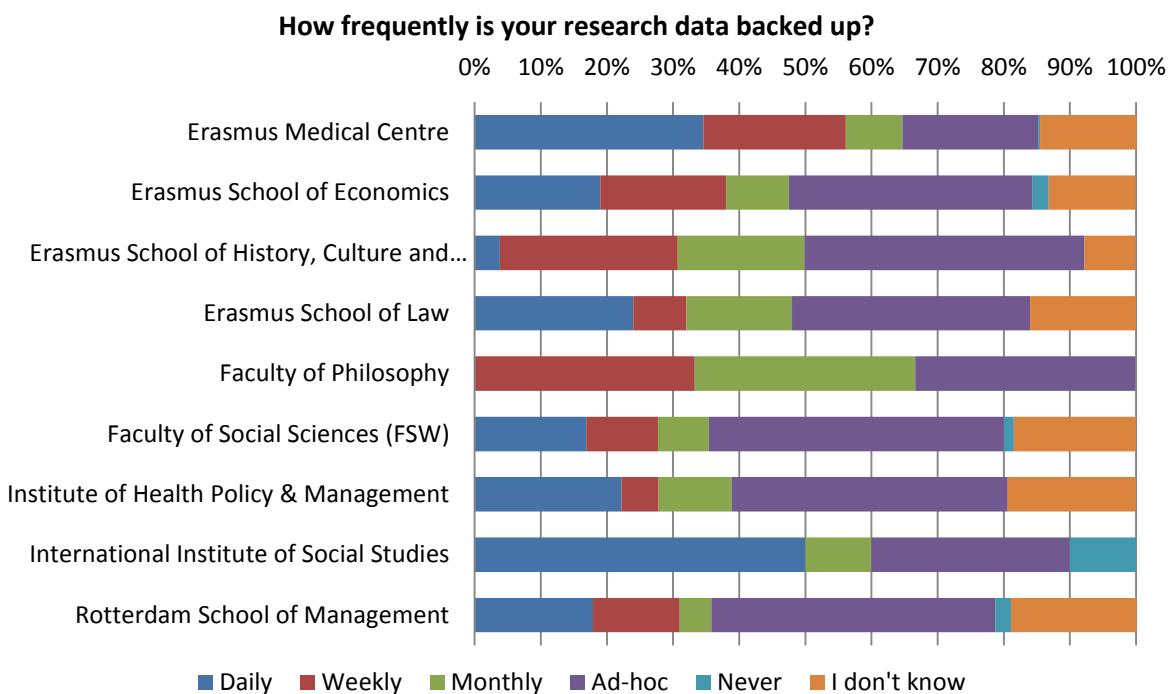


Figure 6 Backup frequency per school or faculty

The self-reported estimated average volume of produced research data circles around 50 GB per researcher (Figure 7). Of course answers vary per scholar and faculty or school. On average 1 in 7 researchers does not know the volume of his or her data.

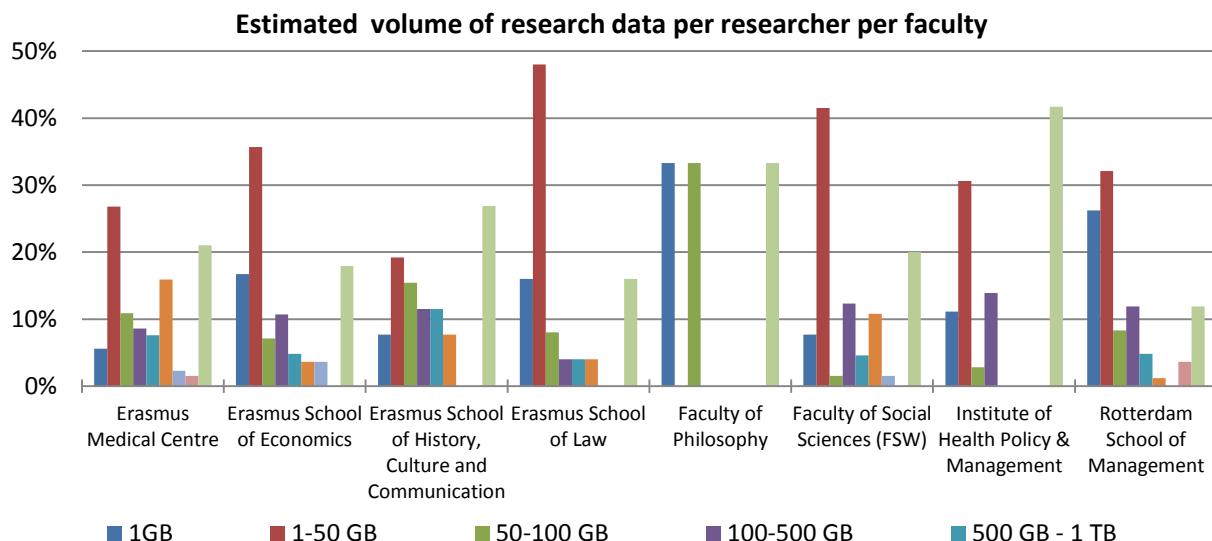


Figure 7 Estimated volume of research data per researcher

3.2.2 Research data management

Metadata is important for finding data and for linking data to other data (interoperability). “Metadata” means literally “data about data”. Metadata is traditionally found in the card catalogues of libraries. As information has become increasingly digital, metadata is also used to describe digital data using metadata standards specific to a particular discipline. By describing the contents and context of data, the quality of the original data is greatly increased. 45% of the researchers assign metadata to their data. 60% of these researchers do not use standards or guidelines for metadata. So about 30% of all respondents use standards or guidelines to describe their data.

About 50% of all respondents answer they are not aware of any requirements from funders, school/institute et cetera concerning a data management plan (Figure 8).

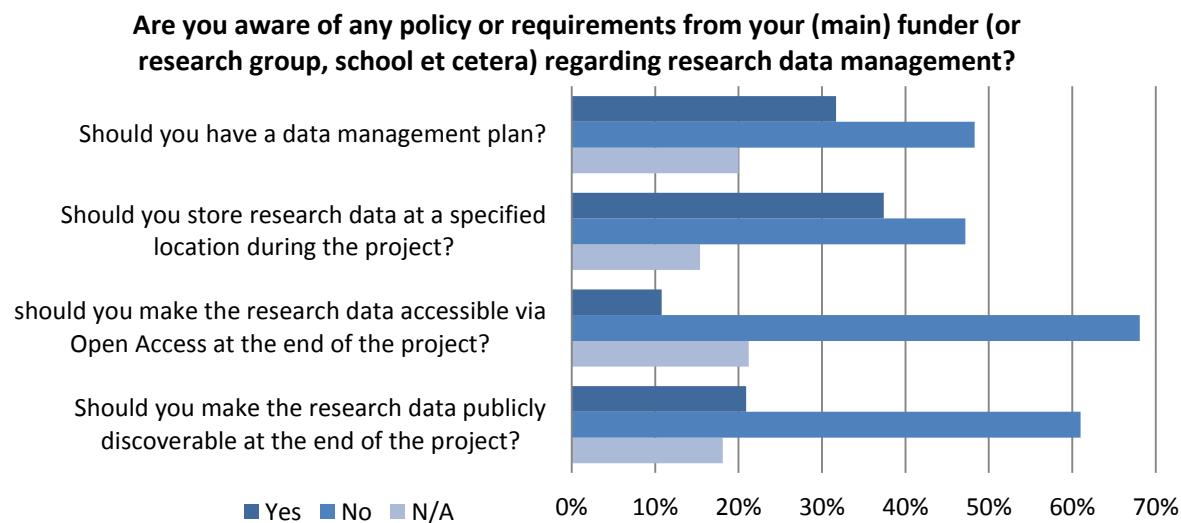


Figure 8 Awareness of requirements regarding data management

3.2.3 Researchers' requirements and training

The survey answers paint a clear picture on researchers' needs concerning data management (Figure 9). Firstly, researchers indicate the need for more storage capacity and better tools to share data during research projects. Secondly, they would like information and advice about research data management in the form of a website. Thirdly, they are interested in data management training on a variety of

topics, especially about developing a data management plan, documenting data/creating metadata and intellectual property right (Figure 10).

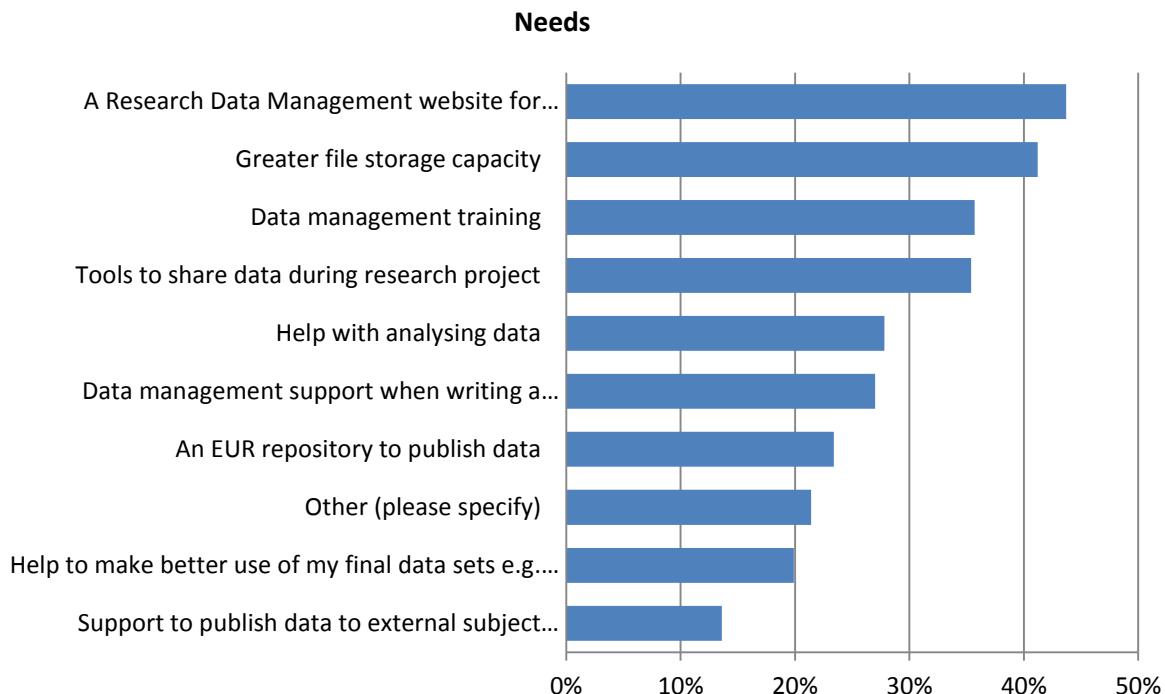


Figure 9 Researchers' needs concerning data management

If you are interested in data management training, in which aspects would you like to be trained?

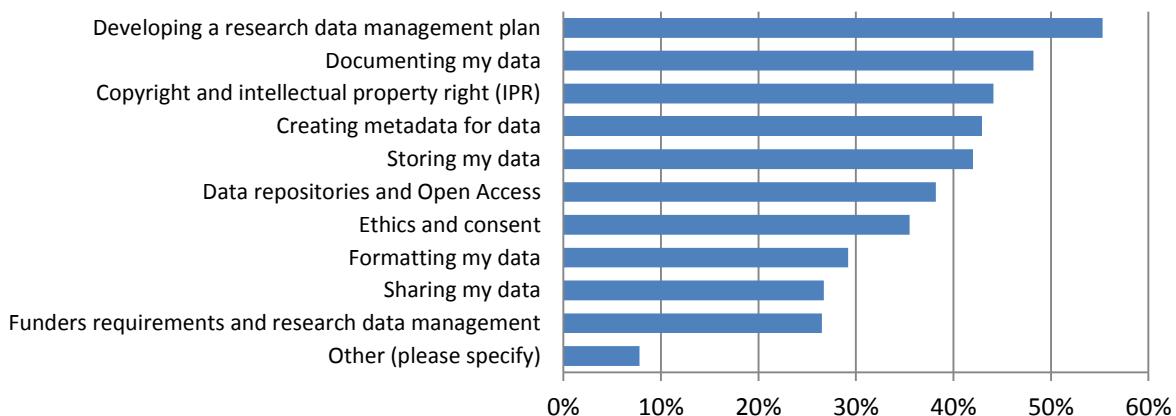


Figure 10 Interest in training in aspects of data management

3.3 Assessment and discussion

Researchers store or back up their data mainly on hard disks, either on their campus computer, computer at home, their notebook or on an external drive. A recent study at Pennsylvania State University found similar backup practices among scholars³. SSC ICT assesses this data storage practice as unsafe, even if scholars back up their data on different hard disks at the same time. Due to the fact that the average lifespan of a hard disk is three to five years, the risk of losing data is still

³ Personal Archiving And Scholarly Workflow: An Exploratory Study Of Pennsylvania State University Faculty. Retrieved from <http://www.cni.org/topics/scholarly-communication/personal-archiving-and-scholarly-workflow-an-exploratory-study-of-pennsylvania-state-university-faculty/>

considerable. The even shorter lifespan of external hard drives needs further attention. Especially the use of external hard drives by PhDs, due to a shortage of storage space, seems to be widespread.

Asked about the frequency with which respondents back up their data, almost a third answers to back up on an ad-hoc basis. This may be an artefact of the research process itself; researchers are not working with data every day, so may have no need to back up daily. However, broken down to school/institute level, data at the Erasmus Medical centre are backed up more systematically than in other schools or institutes (Figure 6). This at least suggests that the research process itself not the only factor determining backup behaviour.

It is alarming that on average 15% of all responding researchers do not know if their data is being backed up.

Researchers frequently use cloud-based storage services (22,6% for storage and 14,5% for backup). Especially Dropbox and Google Drive are mentioned by several researchers in the open answer categories as convenient applications to transport, synchronise, store or share data. However, there are risks to consider when using cloud storage services: the data might not be covered by EU data protection laws, for instance, or data might be altered without a user's knowledge⁴.

There seem to be two reasons for frequent use of internal and external hard disks and web-based services by researchers: insufficient storage space on the campus network and lacking technical infrastructure to share, transport or access data from outside the university network. 40% of the respondents ask for more storage space and more than one third is in need of better tools to share data during research projects. These conclusions are further supported by scholars' remarks in the open answer categories and personal communication resulting from the survey as well as in the interviews (see Chapter 5).

Many researchers (45%) add metadata, although not yet in standardised ways. About one third does use metadata standards or guidelines. Given the fact that a high percentage of scholars wish to be trained in describing data/creating metadata (more than 40%), we conclude that the introduction of protocols and guidelines to describe data in a more systematic way would be met with benevolence by researchers.

In general, researchers would like more information on and training in a wide variety of issues in data management. Many respondents mention especially writing a data management plan, intellectual property rights and storing data, but other issues need to be addressed as well.

3.4 Recommendations based on the survey results

- Discourage storage on hard disks, USB sticks and cloud-based services and make research staff aware of the dangers. Inform them via the data management web site with practical information about do's and don'ts when using such media.
- Create safe storage and backup facilities for individual researchers (see Chapter 4).
- Create safe ways to share, collaborate on and remotely access research data during research projects.
- Introduce a programme to stimulate the use of guidelines for annotating research data with metadata (see also Section 7.3).

4 Safe storage and collaboration

In evaluating possible solutions for safe storage of and collaboration on research data, we have to distinguish research in progress from finished or rather, from published research. During the first stage a collaboration repository or data lab is needed and in the latter a preservation (or publication) repository. "The collaboration repository is where researchers actively work with and analyse data. The publication repository is for research that is 'finished', with some of the results being available for

⁴ See e.g.: Using Dropbox and other cloud storage services at LSE. Retrieved from: <http://www.lse.ac.uk/intranet/LSEServices/IMT/guides/softwareGuides/other/usingDropboxCloudStorageServices.aspx>

public viewing.”⁵ Both kinds of storage must be secure, but they differ in time horizon (e.g. project duration versus long term, respectively) and very likely in access rights.

4.1 Criteria and candidate platforms

Several potential platforms and services for data storage, preservation and collaboration have been explored and compared in a quick scan, using a fixed set of criteria. The results are presented in Appendix 9. The comparison is intentionally limited in character, both given the platforms included and the criteria used. The list of potential storage solutions is based on current storage practices amongst researchers and on storage facilities under consideration by SSC ICT and the university library.

Platforms marked “OR” are considered relevant for storing data during ongoing research, whereas platforms marked “A” are intended for long-term archiving. Whether metadata other than basic information like the file name (see Section 3.2.2) can be assigned and whether a platform supports version control, is relatively straightforward. However, several criteria that were applied turn out to be rather ambiguous. The criterion “flexibility to accommodate various file types” for instance needs refinement and a definition, because it is currently non-discriminating. How broad is the range of file types that a particular platform supports? Does the (organisation that offers the) platform distinguish between preferred – sustainable – formats and accepted formats; does it offer a policy and advice?

The same need for proper definitions applies to “secure storage” and “secure file transmission”, both important issues in research data. Secure file transmission is a gradual feature: the https web protocol offers a certain security, but not the highest level possible, for instance. Therefore this cannot be a mere Yes-No-criterion. Secure storage is a very broad notion as well: does it mean secure from fire, from intentional misconduct, or from becoming inaccessible when under “unfriendly” jurisdiction? A promising approach to security issues would be to classify research data in terms of sensitivity. Sensitive data about individuals for instance requires stricter security measures than already published data sets. When such a classification is available researchers can be advised about best practices in using or refraining from specific platforms for their data storage⁶.

Also “international collaboration possibilities” should be clarified: it is not clear yet what EUR researchers require in this respect. Does it mean just making data available – for some or for all – on the web or does it also imply data analysis functionality or social media, for instance? Some platforms support a federated approach, such that an EUR researcher is internationally identified by his or her ERNA account; most platforms in the current selection don’t. Another feature that is relevant for academia is the assignment of so-called persistent identifiers. In contrast to a web URL a persistent identifier is a unique and persistent reference, which enables scholarly citation. In our selection most platforms don’t offer this, and among those that do one has a proprietary (brand) format, which might make it less sustainable than the well-known DOI (Digital Object Identifier), Handle and URN:NBN.

Two platforms have been explored in greater depth as potential solutions to accommodate researchers’ needs for safe storage and collaboration. One is a possible solution with on campus storage and backup that is suggested by SSC ICT in the context of the virtual work environment @wEURk. The other is the Dutch Dataverse Network, with which a pilot study has been carried out at the EUR.

4.2 @wEURk storage cluster

The virtual work environment @wEURk, to be implemented in 2014, presupposes a central storage cluster. Although @wEURk is essentially an office environment, the storage cluster behind it could be used to offer researchers storage space, which comes with backup facilities. This could solve problems with storage on hard disks and backups. Laptops, when connected to the network, could be automatically backed up.

⁵ H. Tjalsma, J. Rombouts: *Selection of Research Data, Guidelines for appraising and selecting research data*. DANS Studies in Digital Archiving 6. 2011, Den Haag en Delft. Stichting SURF, Data Archiving and Networked Services (DANS), 3TU.Datacentrum. ISBN 978-94-90531-06-5. Page 40.

⁶ For instance: Using Dropbox and other cloud storage services,
<http://www.lse.ac.uk/intranet/LSEServices/IMT/guides/softwareGuides/other/usingDropboxCloudStorageServices.aspx>

Researchers themselves can authorise colleagues from other Dutch academic institutions to access their data, using SURFconext facilities⁷. Data can be made citable with a URL. Collective use of these storage facilities would lead to lower prices per terabyte, but the ratio is as yet unknown. Archiving functionalities of the cluster and the connections between storage and archiving are to be explored.

4.3 Dutch Dataverse Network

The Dutch Dataverse Network (DDN)⁸ is an external platform used by six Dutch universities to store and share research data, based on the platform originally developed by Harvard University⁹. Using SURFconext researchers from participating universities can start their own online project environment – called a Dataverse – and store a variety of scientific data (texts and raw research data, but also audiovisual material and complete databases). Researchers can index the data they have stored and share their data with other scientists or interested parties. Researchers themselves determine who gets access to what data.

In DDN so-called collections can be defined, which in their turn contain studies. Studies can be created for separate subprojects, experiments or other parts of a research project or research career. Collections and studies from other Dataverses may be included (even from other owners, provided the studies are freely accessible).

In DDN it is possible to refer to stored data sets by means of a 'persistent URL' and link the data via this persistent identifier to publications, other data sets, personal webpages. The persistent identifier allows for standardised data citation, on a par with publication citation. Research Data Netherlands (the DANS/3TU.Datacentre consortium) is currently exploring technical connections between DDN and their archiving facilities.

DDN has been piloted at the EUR with several researchers and data types, with emphasis on storing and describing data (see Appendix 3 for detailed results from the pilot study). Through collaboration with other universities we also know about their user experiences. Other steps in the data life cycle like sharing data and archiving data for the long term need to be explored.

4.4 Recommendations

At the moment it is difficult to compare the suggested storage solutions. More in depth analysis of technical, organisational and financial aspects and consequences for institutes and schools is needed. The storage solution proposed by SSC ICT is not yet in place and could only be discussed tentatively.

This leads to the following recommendations:

- Ask the university library's front office (see Section 7.2) and SSC ICT to improve the comparison matrix and deliver an extensive comparison of relevant storage platforms. Publish "best practice" fact sheets for each major platform on the data support website as an awareness raising instrument. Updates of the matrix and fact sheets are a shared responsibility of front office and SSC ICT.
- Start a pilot study to further explore the possibilities and consequences of using the storage cluster behind @wEURk for data storage and sharing and the connection with archiving facilities. Ask SSC ICT to assess as precisely as possible financial consequences for schools and institutes.
- Find on campus projects to pilot the DDN with larger datasets taking into account the whole data cycle. If necessary pay out of pocket costs for participating research groups or organise a contest where extra data storage and sharing facilities can be won (including advice and support). The contest could be a building block in a broader awareness campaign on data management. Work out a cost model for DDN usage together with other universities.

⁷ With SURFconext users only have to log on once with their institution's account to safely access services at connected institutions. EUR researchers can use their ERNA account. See <http://www.surfnet.nl/en/Thema/coin/Pages/default.aspx>.

⁸ <https://www.dataverse.nl/dvn/>

⁹ <http://dvn.iq.harvard.edu/dvn/>

5 Data management workflows and protocols

In the light of scientific integrity it is of great importance that researchers take care of their data in a responsible fashion. When one is uncertain how to do this, not only the risk of scientifically improper behaviour arises, but also the risk of data loss and inefficiency in workflows. To reduce these risks it is useful to make use of a protocol for data management, which is related to a researcher's workflow. In this section the project team proposes a blueprint for a protocol. At the minimum level this blueprint contains a few EUR-wide elements that should be refined to institute-specific and/or discipline-specific minimum protocols. Suggestions for an advanced level are also made.

5.1 Approach taken

The project plan stated the following goal: “Per data type data storage protocols are available as well as workflows. We consider protocols as documents that prescribe actions and describe associated responsibilities during a workflow, in this case with respect to data storage.”

It is evident that the creation and imposition of one general data storage protocol for all research practices is undesirable and impossible. Faculties and research schools simply practice too many different types of research and also differ to a great extent in organisational culture. Therefore the approach was taken to interview representatives from EUR faculties, schools and institutes to collect common ingredients in research workflows, in order to design a blueprint for a data storage protocol. It is then the task of each organisation (i.e. faculty, research school or institute) to adapt and implement the components of this blueprint to their own local standards.

To answer the question ‘what research data and documentation is needed to replicate a study in the future?’, interviews were held with nine research directors or their representatives and with ten senior researchers and research group leaders across disciplines at the Erasmus University Rotterdam. The questions asked focussed on *what* should be stored for their particular kind of research *at which moment* in the research cycle. This focus was chosen to act upon the second conclusion in the report by the Schuyt-committee: “While the research cycle is subject to stringent and critical monitoring, there is one phase that is relatively free of external review and that leaves considerable scope for creativity. That phase can be found within the primary research process (i.e. the process of data acquisition, storage, organisation, processing, and accountability), in other words specifically after the start of the study and before the peer review. Depending on the field involved, this exploratory phase may represent a risk with respect to data management.”¹⁰

5.1.1 Timeline

For the sake of discussing this “relatively free” phase, the research cycle has been simplified and straightened to the timeline in Figure 11. This timeline is recognisable for most respondents, although often different in real life. For instance, the timeline is more easily applied to quantitative studies with explicit research questions than to qualitative research, where data processing and data analysis are rather intertwined, or to studies that start from rich data that originally was not collected for research purposes, be it in population health research or in economical contract research. Furthermore it should be noted that whereas grant research, contract research and PhD research typically do have a clearly marked start, this doesn't apply to all research. Also, research is often an iterative process. For instance, many respondents mention that the review of a submitted publication usually leads to re-analysis of the data. Furthermore, when a multiannual project is broken down into subsequent parts, the timeline applies to each of these.

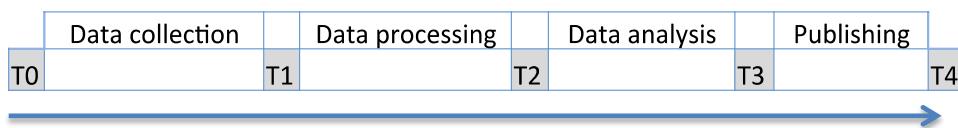


Figure 11 Abstracted research timeline

¹⁰ Royal Netherlands Academy of Arts and Sciences. *Responsible research data management and the prevention of scientific misconduct*. Amsterdam, 2013. Retrieved from http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20131009.pdf, p. 32.

Abstracting away from all kinds of preparation, it is at T0 that a research project starts, often but not always with an explicit, approved project plan or study protocol (terminology differs across disciplines). Now the data collection phase can begin. The diversity in practices is large in this phase, according to the respondents, and the types of data that have been collected at T1 are equally diverse. In essence, T1 is the moment where all the data that is to be used in the project is collected and one speaks of the “raw data”. Within several types of research it is necessary to process the raw data, since the raw data itself is not suitable for (statistical) analysis. This is for instance the case with interviews on tape that have to be transcribed, or with externally acquired data that requires changes in variable codings or anonymisation of confidential information. At T2, the data is fit for analysing. At T3, the diversity in practices diminishes: in all fields of research there is a specific data set present that has been used for analysis and there are results. During the publishing phase a publication is written and submitted to a journal; this process is very similar across research practices. T4 indicates the end of a project or a subproject.

5.2 What should be stored

Data storage is very specific to the organisation in question. What can or needs to be stored depends on the type of research, type of data, size of the data and the culture within an organisation. However, there is certainly common ground. All respondents are used to documenting – albeit sometimes in informal ways – certain decisions, study design considerations, changes to a previous plan et cetera, with an eye to the Methods section in the intended publication. Also, for most respondents it is common practice to provide the data that underlies a publication, as long as no legal or privacy-related restrictions apply. This situation leads to the following minimum data storage protocol with an EUR-wide core and organisation-specific refinements.

5.2.1 EUR-wide and local data management minimums

A few basic resources – data and documentation – are needed to make replication of a study possible. These resources are the same for most organisations and disciplines and are shown in Figure 12. *Each faculty, institute and school ought to commit to this EUR data management minimum and refine it to an organisation- or discipline-specific minimum.* This is the responsibility of the dean of research or the research director.

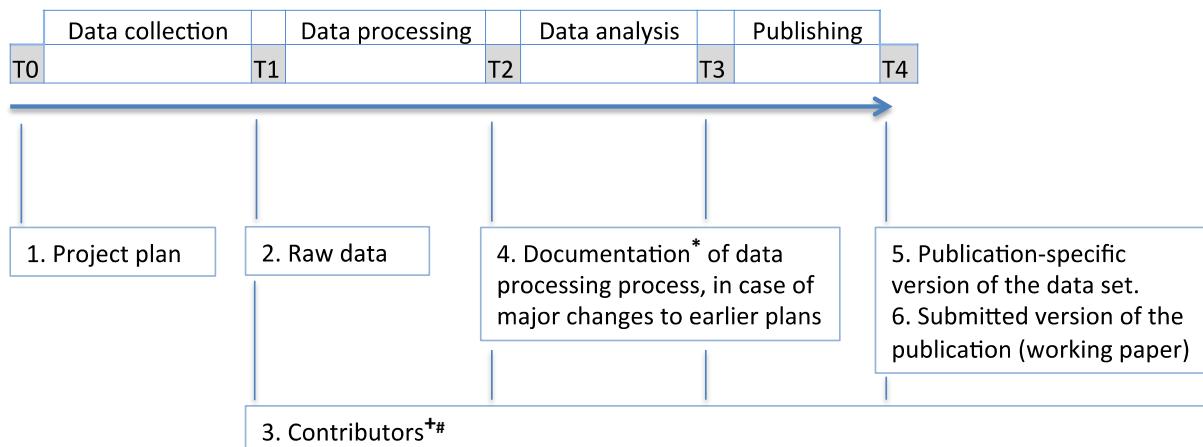


Figure 12 EUR data management minimum

In the following paragraphs several kinds of documentation and data are discussed which are important in at least some of these organisations and their disciplines. This can be used to answer the question “What does the EUR minimum entail for a specific organisation?”. Basically, the recommendation is to *store various kinds of documents when they exist anyway*. Furthermore, some respondents explicitly ask for practical, step-by-step instructions.

Ad 1 EUR minimum: always store the project plan

Local minimum: From the interviews it could be derived that nearly all the researchers within organisations make use of a project plan. For some types of research these plans are well defined and

detailed (e.g., externally funded studies and studies designed to test a theory), whereas for other types plans may evolve along the way (e.g., internally funded projects, exploratory research, and studies that aim to develop models or methodology). If a project plan is present at T0, it ought to be stored at that moment. Otherwise, it ought be stored as soon as it is finalised, for instance in contract research after sparring sessions with a client about the exact research question.

A project plan may include a data section and/or a data management plan¹¹, which is increasingly being requested by funding organisations. Also, in some disciplines (e.g., experimental research) research questions, hypotheses and the study design are part of the project plan. All this is worth storing.

When data from external data providers is used or in case of contract research, contracts or other legal documentation exist, often with agreements about (future) ownership, use and management of the data; this is worth storing. It should also be documented whether the researcher or author has any relationship with an external provider of data.

Ad 2 EUR minimum: always store raw data immediately after collection

Raw data storage plays an important role in accountability issues, but also allows the researcher to reuse his/her own data or to return to earlier stages of the research process when needed.

Local minimum: Raw data is very diverse, across and within organisations. When for instance interviews have been held, the interview transcripts (on paper and/or digital) should be stored. Ideally, interview transcripts have been authorised by the interviewees, but respondents mention that this is unfeasible if interviewees are for instance members of criminal gangs, illegal immigrants or unemployed.

In case of surveys, the digital version, the paper version, or maybe both can be stored. With (large) data sets a ‘snapshot’ (i.e. a point frozen in time) of the data should be stored, confidentiality and data ownership permitting.

Apart from the raw data some additional documentation should be stored with certain data and research types. It should be kept in mind to look closely at what information is already going to be in the publication, and what would make sense to document and store additionally. Examples could be: documentation about the sampling logic, sample size, boundaries of the population, the seed that was used in simulation research, informed consent forms, descriptives of data that cannot be published, metadata about the data collection process, etc.

In field research a field journal could be stored and analogously, lab journals are important sources of information in experimental environments. Obviously, digital-born journals yield easier access to information than paper or scanned versions.

In case of externally provided data with limited storage rights, relevant documentation such as a non-disclosure agreement should already have been stored at T0. In case of free public data, for instance provided by Statistics Netherlands (CBS), a description of the data and the selection criteria (database queries) should be stored.

Ad 3 EUR minimum: always document and store who is responsible for what

It is important to keep track of all persons who are involved in a study, be they researchers, assistants, programmers, readers, in short: all contributors. Moreover, it should be documented who is responsible for which part(s) of the study. The plus sign (+) indicates that this document must be updated at the subsequent crucial moments T2, T3 and T4.

Local minimum: Whenever relevant, add external contributors such as data providers, recruitment agents for test subjects, and referees.

Ad 4 EUR minimum: always document and store major deviations from the plan

¹¹ Examples can be found via <http://dataintelligence.3tu.nl/ii-data-management/data-management-plan/>

Documentation of the data processing process should be stored in case of major changes to earlier defined plans. Crucial changes made in the course of the project ought to be comparable to the initial design. The possibility to compare may encourage proper scientific behaviour and is essential when one is requested to account for one's results. This also holds for the subsequent analysis stage. The star sign (*) indicates zero or more versions of the document: no version when no major deviations occur at all, and versions at T2 and/or T3 whenever relevant.

Local minimum: Several respondents mentioned that while documenting changes makes sense when project plans exist, step by step documentation of all actions and changes is unnecessary. Rather, within each organisation the degree of relevant documentation must be specified such that it helps a researcher to return to previous decision points, allows for replication and is not too burdensome to the researcher.

Furthermore, at the end of the data processing stage (T2) one should store documentation about data screening methods, data pruning processes, handling of missing data and the syntax or code of any analysis software – whatever may apply to the discipline. When code books or data dictionaries are used, they should also be stored at this point. Storing intermediate data sets at T2 and/or T3 may be instrumental as well, especially if this can replace writing very detailed documentation about data transformation steps.

At the end of the analysis stage (T3), in several areas of research the most important thing to store is the syntax or code of the final analysis. Several respondents mentioned that statistical software stores the commands that have been executed¹², which makes these steps reproducible. Additionally, documentation about the configuration and version of the software can be relevant, including the date when the analysis was run.

Ad 5+6 EUR minimum: always store the publication-specific data set and the submitted publication

At T4 the submitted version of the publication should be stored; a reference suffices if the submitted publication is already stored in a trustworthy and accessible repository. Moreover, in all disciplines the data set that underlies any submitted or published publication ought to be stored. This version of the data set is important as it is the data from which conclusions of the study are drawn; it should be possible at all times to validate these conclusions with the use of this final data set.

Local minimum: Nowadays an increasing number of journals ask for the publication-specific data set when a paper is submitted. In general, however, this does not imply that the data is sustainably archived. Other resources that can be stored (or referred to) at this stage are so-called supplementary data.

A working paper might be a good alternative to the submitted paper. Several respondents publish working papers at conferences or for instance on the Social Science Research Network (SSRN¹³). For one respondent the working paper is also the version submitted to the journal. This means that, when the accepted version is published, comparison with the submitted version is easy.

5.2.2 Advanced data management

Beyond the minimum protocols there is room for an advanced level. Over time the EUR organisations can grow from the minimum level to this more mature level. Advanced data management involves more than storing what is necessary for replication purposes; it includes a variety of good practices that enhance scientific integrity and transparency.

The project team acknowledges that the respondents show many signs of advanced data management awareness. All good practices and suggestions listed in this section have been mentioned in the interviews. Because sharing knowledge and seeking consensus about researchers' ways of working are important instruments in embedding such practices in the organisation, examples of good practices in these areas are included, although they don't relate to data management per se.

¹² In the Stata™ software this is called a do-file. Other software packages are reported to have similar storage functionalities.

¹³ <http://www.ssrn.com/>

- There is an EUR organisation that carries out a plagiarism check after a PhD student's first year. The findings are delivered to the PhD student and his supervisor. The message is that one is allowed to make mistakes but that the thesis ought to be free of mistakes. At a later stage another check takes place. To make this process transparent the thesis supervisor needs a sparring partner. The respondent who mentioned this practice would like to extend the check to PhD proposals, because they are often written without the organisation's involvement.
- Issues with particular methodologies are discussed in so-called methodology interest groups consisting of five to ten researchers, which leads to consensus about ways of working. Experience shows that it is not easy to incorporate researchers from other disciplines, although they apply the same methodology.
- Another organisation internally audits publications. Authors of a randomly chosen publication are for instance asked to replicate a particular table. This requires the availability of the data files, the syntax of the calculations as well as clear explanations.
- This organisation has also experimented with keeping a log, like a diary. It turned out that it was unfeasible to record the huge number of small steps and that the log did not improve transparency of the working process.
- There is an organisation where researchers share grant proposals, in order to learn from accepted and rejected proposals.
- Many organisations have lab meetings, post work seminars and/or project meetings. One respondent mentioned a handbook for quality, available at the intranet.
- An organisation that introduced a working group for data management for the sake of replicability noted that this very fact already has a disciplining effect. Sharing data is also reported to have a disciplining effect. In a similar vein a respondent remarked that when he is the first author of a publication, he explicitly shares all data with his co-authors in order to be transparent.
- For the advancement of science it is valuable to store and publish studies that do not show an effect. A recent trend in the field of psychology is some journals' commitment to publish results – no matter which – if the journal has approved the study proposal at an earlier moment. Supporting this approach might bring about a change in the dominant tendency to publish only significant results.
- Given the confidential nature of the data, in one organisation researchers use a dedicated computer unconnected to the organisational network. The data is copied daily. When the project ends or the researcher leaves, the hard disk with the copy is archived for ten years and the computer is wiped clean.

Needs and suggestions also worth exploring are:

- It would be very helpful to be supported by a data manager who also keeps track of relevant developments and services (see Section 7.2 about the front office in the university library).
- It is common practice in health-related research to co-author a publication when one's data is used in it. When tissue samples are being used in research the tissue bank should at least receive a reference to the publication(s).

Since responsible data management, like science itself, is an ongoing process, it is recommended *that all EUR institutes evaluate periodically which of these good practices may be relevant and feasible to raise the level of their local protocol, and which of their own good practices could be interesting for others.*

5.3 Introduction and implementation of a data management protocol

Whereas the minimum blueprint is based upon a common denominator across disciplines and research strategies, the interviews also manifest many differences. These give rise to the following recommendations for introducing the protocols.

Recommendation Before introducing a protocol for data storage, organise a local debate about values in research and how e.g. collaboration, data sharing, peer consultation, and the publication of working papers support these values. Such workflow elements may not only contribute to the research

and to a professional attitude, but also yield documentation that can easily be stored. Professional research goes hand in hand with responsible data management.

Why this recommendation? The respondents agree that a certain level of data management, not just for one's own immediate needs, is or would be a good practice. They also report that researchers in general agree that a (more) structured approach to this is reasonable. Opinions differ, however, as to how directly a protocol for data storage should be introduced. On the one hand some respondents make a plea for data storage to be obligatory. On the other hand, however, there is concern about researchers avoiding the protocol when they would perceive it as a top-down administrative constraint on their workflow, without being related to improving research as such.

Recommendation When introducing and stimulating data management, take care to explain that this also includes documentation related to the data and to the research process.

Why this recommendation? Some respondents or researchers in their milieu associate the notion of "data" with numerical material and conclude that "data storage" and "data management" don't relate to them, for instance in ethnography. The same conclusion sometimes appears to be drawn in disciplines that predominantly analyse publicly available texts, such as legal texts.

Recommendation When advocating data management, be clear about limitations that are inherent in the discipline and define alternatives. If exact replication is known to be unfeasible, proper documentation becomes even more important.

Why this recommendation? Not all respondents subscribe to the view that perfect replication should be aspired, although they agree that it might be desirable both from a scientific perspective and from the perspective of investigating scientific integrity. A respondent at one end of the spectrum mentioned, regarding a large genetics study, that storing the raw data is more expensive than running the experiment again. In other words, when the process is well-documented and replicable, this kind of raw data is of lesser importance.

At the other end of the spectrum, however, it is for instance impossible to exactly replicate interviews or observational studies. Another example is provided by respondents who make use of data that underlies a non-disclosure agreement, for instance provided by a private company; the researchers are not allowed to store the data for other uses than the study itself. Although such data is in principle (commercially) available to others, they may change over time because the external data providers update them, without necessarily keeping track of earlier versions of the database. In such a case, a data description is the best approximation of the real data and this is probably what would be needed for publications as well.

The project team makes two other recommendations for the early stages of data management:

Recommendation Appoint a research data support officer at the central EUR level.

Why this recommendation? Support at a central level is crucial for continuity and for sharing good practices between institutes and disciplines. The data management support officer stimulates and supports the deans of research or research directors of the EUR faculties and graduate schools to introduce relevant data workflows and protocols and to maintain them over time¹⁴.

Recommendation Draft a covenant to the extent that the deans of research/research directors embed research data management in their institute.

Why this recommendation? A covenant of this kind is being introduced at Erasmus MC (see Appendix 5). A covenant is a good way to document commitment and responsibilities within the organisation. The deans of research/research directors should commit themselves to undertake certain data management steps and to evaluate them periodically. They should formulate the steps themselves, in a way that fits the kinds of research carried out at their institute. Furthermore, they should note obstacles and requirements (organisational, legal, technical) from other sections of the institute. When the central research data support officer is involved in the draft and in the periodical evaluation,

¹⁴ Note that this central function is not a data manager. A data manager is more hands-on involved in data curation and storage in a research institute or research group, and might for instance help with assigning metadata.

obstacles and requirements can be dealt with in generic ways and good practices can be shared. Considering the number of staff at Erasmus MC the heads of department would be better placed to acquire commitment from researchers than the dean and to commit themselves to the covenant. Therefore managing the (department) covenants could be delegated to the heads of department; the Erasmus MC coordinator for scientific integrity could centrally monitor this in lieu of the dean.

5.4 Stakeholders, responsibility and supervision

In the interviews about data management many stakeholders have been addressed: researchers, research group leaders, deans of research, the EUR, journals, funding organisations, and long-term archives in the back office.

Nearly all respondents see the *dean of research / the research director* as the person who should supervise and stimulate data management workflows. At Erasmus MC this could be delegated to the heads of department, supported by the coordinator for scientific integrity.

There is consensus among the respondents that *researchers* are responsible for the quality of their data and for properly storing the relevant data and documentation. It is also clear that young researchers should learn to do this, and here both PhD supervisors and the graduate schools are said to have a responsibility in socialising them. Several respondents mentioned that they already pay attention to the importance of data management when they teach students, for instance in courses on statistics and methodology.

It is considered to be important that *senior researchers and research group leaders* are role models. This appears to be delicate: some respondents mention that it can be hard to motivate senior staff to take up this role, and others mention that senior staff is not very eager to participate in training about scientific integrity and related topics.

In the context of the EBL pilot study (see Chapter 6) it was mentioned that part of a researcher's responsibility for storing data shifts towards the lab and the *organisation*, once the lab system stores information about an experiment for the purpose of safeguarding raw data. Although this is akin to the organisation's responsibility for making regular backups of all data that is stored on the EUR ICT network, it is advisable to make this explicit.

A certain responsibility is also attributed to *journals*, when storing data as well as information about the review and revision process is concerned. However, despite recent interest in so-called data availability policies in economic journals¹⁵, there is no general data management policy that journals adhere to and certainly not for the long term. Therefore the project team recommends to keep copies of the correspondence with the journal. For sustainable *archiving* of data and related documentation the team recommends to develop the front office – back office model (see Section 7.2).

Funding bodies increasingly require grant-holders to develop and implement Data Management Plans (DMPs). Such plans typically state what data will be created and how, and outline the plans for data sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied¹⁶.

5.5 Discussion

This task in the project has benefited greatly from two sets of recommendations for data storage: recommendations from ERIM (see Appendix 4) and recommendations from the coordinator for scientific integrity at Erasmus MC (see Appendix 5). These two approaches differ in scope, in the granularity of data types considered, and in the appropriate moment(s) in the research cycle for storing data and documentation. In preparing a concept checklist for data storage to discuss with respondents (see Appendix 6) the project team has used elements from both approaches. Although validating these earlier approaches was no objective of this project, it is interesting to note that the interviews provide

¹⁵ See <http://www.clariah.nl/clio-dap/samenvatting> (in Dutch).

¹⁶ See for example: <http://www.dcc.ac.uk/resources/data-management-plans>, <http://www.nwo.nl/subscription-documents/magw/data-archivering-en-beschikbaarstelling---formulier-2011>

support for both approaches and for all different positions in them. This goes to show the diversity of responses.

The project team concludes that the original goal, to deliver data storage protocols per data type, has not been met exactly. Instead, we recommend to introduce the blueprint as an EUR-wide minimum protocol – roughly along the lines of the Erasmus MC input – and to have this refined per discipline within the faculties and research schools, which shows more similarity with the ERIM input.

To return to the “relatively free” phase after the start of the study and before the peer review: there is no consensus among the respondents that researchers should or will store data and documentation between the moments T1 (raw data) and T4 (submitted publication). Rather, researchers should be trusted to store whenever, whatever they need later on in the process and what should be in or associated with the publication. Therefore it is recommended *to introduce the minimum data storage protocol including all crucial moments T0-T4 and to make explicit that it is the researcher’s responsibility to store data and documents at appropriate moments to prevent loss of data*. Administratively, it is recommended that *the researcher then adds a reference to major updates in the EUR’s research information system* (currently METIS, but in the near future to be replaced by another system).

5.6 Recommendations

In this chapter the following recommendations concerning protocols have been made:

- Each faculty, institute and school ought to commit to the minimal EUR data management protocol and to refine it to an organisation- or discipline-specific minimum protocol, under the responsibility of the dean of research, the research director or (at Erasmus MC) the head of department.
- Before introducing a protocol for data storage, it is advisable to organise some debate about values in research and how e.g. collaboration, data sharing, peer consultation, and the publication of working papers support these values. Professional research goes hand in hand with responsible data management.
- When introducing and stimulating data management, take care to explain that this also includes documentation related to the research process and other materials used. The term “data” runs the risk of being misunderstood.
- In refining the EUR-wide minimum protocol it is advisable to explicitly include various kinds of documents that exist anyway within the discipline or the organisation. The protocol should be practical and give step-by-step instructions. It is recommended as well to include all crucial moments T0-T4 and to make explicit that it is the researcher’s responsibility to store data and documents at appropriate moments to prevent loss of data. It is recommended to add a reference to major storage packages in the EUR’s research information system.
- When advocating data management, be clear about limitations that are inherent in the discipline and define alternatives. If exact replication is known to be unfeasible, proper documentation becomes even more important.
- Since responsible data management, like science itself, is an ongoing process, evaluate periodically which good practices developed elsewhere may be relevant and feasible to raise the level of the local protocol, and which local good practices might be interesting for others.
- For the sake of continuity the Executive Board should appoint a research data support officer at the central EUR level, to support deans of research/ research directors with the above.
- Also, the Executive Board should draft a covenant to the extent that the deans of research or research directors and – at Erasmus MC – the heads of department embed research data management in their institute. They should have room to describe how they will fulfill the data management minimum and what, if any, obstacles they see to achieve this.

6 Erasmus Behavioural Lab pilot

6.1 Goal

The EBL pilot assignment is to model and support a workflow that, as a main goal, underpins the genuineness of digital data gathered by means of psychological or behavioural research conducted in the Erasmus Behavioural Lab. In this context data consists of experimental executables or scripts, including digital stimulus assets as well as the resulting output files – and metadata – that contain measurements and/or recorded information. Normally “data” refers to multiple files as in a data set.

While searching for existing or similar systems, the term “data provenance” was found. Many definitions of “provenance” exist, but all center on documenting where, when and how something was acquired, be it information or a work of art¹⁷. This term suits the goal of these efforts well, as a new workflow and software system is targeted mainly at supporting the trustworthiness of the data that is generated within the EBL.

6.2 Setup

In the pilot proposal was mentioned that private suppliers are essential to extend the capacity of the EBL staff. So, the EBL pilot starts beginning of June with the negotiations with suppliers that can be hired for the pilot. Also during June, information and materials that contain existing and future requirements of the 4 existing administrative systems is gathered and organised. Various aspects of the pilot and the intended end result are discussed with other EUR representatives, in particular with SSC ICT.

After several meetings with suppliers the decision is made to hire 2 suppliers. A larger supplier (named Finalist) carries out the main analysis and inventory, while a smaller supplier (named Impart) is hired as a consultant to support the EBL in managing this pilot.

The EBL pilot is split into two main phases: the analysis phase and the Proof-of-Concept phase.

During the first phase, which approximately takes one month, existing administrative workflows and user requirements are inventoried and analysed in the context of the main project goal. This is done in order to develop a consistent organisational model and supporting software automation system (together they constitute the ‘EBL system’) to provide data provenance for research data that has its origin at the EBL computer systems.

To get user feedback during the analysis phase a user group from both Psychology and ERIM gives their input and feedback twice. This phase delivers a project analysis document, which contains an organisational and application design of a new EBL system.

During the second phase a Proof-of-Concept (PoC) workflow is built with two main goals:

- Detailed output of lab activities, reporting on the trustworthiness of the EBL-originating data(sets), a “Data Provenance Report”. This report contains a structured and categorized timeline of all recorded ‘lab-events’, like lab booking entries, data file generation information, subject signups and lab study sessions, lab key usage, et cetera, as well as the relation between these events.
- Provide the researcher (one variant of) a solid and clear workflow when using the EBL, in order to register and secure all research data that has its origin at the EBL computer systems.

Possible implementation risks determined during analysis are minimized by tests during the PoC. During this phase, two or three sessions provide user feedback.

¹⁷ Definition from taverna.org.uk: “Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness.” See also the information at http://www.w3.org/2005/Incubator/prov/wiki/What_Is_Provenance

6.3 Current workflow

The current EBL workflow matured over several years and is flexible to the researchers. However, the current workflow is sub-optimal and certainly needs improvement. It offers the following administrative and operational software systems to the researcher:

- OASE, EURO and ERPS subject pools
- EBL lab booking system
- EBL key management system
- EBL Assistant

The subject pools are used to recruit participants, compensated in different ways.

The lab booking system is used to book labs and the key management system was introduced in 2011 to better track key and lab usage.

The main use of the EBL Assistant —for the researchers— is the ‘copy and retrieve’ functionality to copy data to the EBL computer systems and vice versa.

More details on the current workflow and process can be found in chapter 4 of the analysis.

6.4 Analysis phase

Finalist started July 15th by building an inventory of the functional capabilities of the existing administrative systems of the EBL workflow.

As stated in the pilot proposal, the assumption is that [an integration of] the 4 types of administrative systems currently used in the EBL together provide a good basis for ‘lab data provenance’, but not a complete view.

After the first inventory, several elementary questions were discussed with the feedback groups during the first meeting, which took place end of July.

One question was whether it was allowed [by local legislation or for example APA guidelines] to even temporarily register the relation between subject and data file. From the discussion during the feedback sessions it was not clear if this could be done or not, so the analysis report presents different scenarios depending on the outcome. In general, data provenance quality will be better if more related events support the data file concerned.

Another question was whether participants should register when arriving at the EBL waiting room. Further analysis shows that it is important and also feasible to register participants’ arrival at the EBL, thereby registering this important event not only from the perspective of the researcher but also of the participant.

All feedback and suggestions from the first meeting were used in further building and detailing the new model. In addition to the functionality of the existing systems, a main new part of the workflow will be a lab-study intake system for each study conducted in the EBL.

The resulting, more mature model of the EBL system was presented in the second meeting with the feedback groups, on August 15th.

In general, the feedback group agrees on the identified functional and non-functional blocks, the workflow —including the new intake system— and further functionality that was proposed in the meeting.

6.5 The proposed new EBL system

During the second half of the analysis, existing and new requirements are brought into a new EBL workflow model and the software systems necessary to support it. The goals and boundaries of an initial implementation are determined, also in order to see which resources will be necessary.

The new workflow is similar to the current one; with the addition that every study conducted in the EBL will be registered by the intake system. Depending on study-requirements, sub-workflows like

Ethics Committee and technical feasibility are offered to the researcher. Much of the interaction the researcher has with the currently used systems shifts to the intake part of the new EBL system.

Information given by the researcher during the intake is used to guide lab usage and to securely and automatically store all experimental raw data in a storage system. Currently, the main responsibility for storage and backup of research data is with the researcher. A new workflow will shift these responsibilities to the EBL system. The safely archived raw data is accessible to the owning researcher through a web-interface.

The details on the new workflow and organisation can be found in chapter 5 of the analysis.

A proposal for the new technical software and hardware architecture can be found in chapter 6 of the analysis. This architecture is used and tested during the Proof-of-Concept phase.

Chapter 7 of the analysis describes which events in the system logging will support data provenance reporting.

6.6 Proof-of-Concept

During the Proof-of-Concept phase, actual functionality is tested and built to enable [demonstration of] the detailed reporting on usage of the EBL. As a part of this effort, a workflow for the researcher is built, visualising in which way the researcher interacts with the new EBL system.

During the PoC, several methods of using the Dell DX object storage, provided by the SSC-ICT, are tested.

The Proof-of-Concept runs during September 2013 and has three sessions in which the feedback group is informed on the status of development. Their input is used to further develop the PoC. This phase will also be used to see if development methods used [scrum¹⁸] are feasible for the initial implementation, further development and maintenance.

Reporting on the Proof-of-Concept phase will be an additional appendix or, when possible, during a demonstration session.

6.7 Recommendations

The proposed new ‘EBL system’ provides data provenance for research data originating at the computer system used in the EBL. It does this by using a formal lab-study intake system that enables automatic storage of the raw data in archival storage. In addition to that, the system automatically logs all related user events in a database.

It is recommended that before initial implementation, a clear statement is necessary on whether the relation between data file and subject may be stored by the system. The involved research community will have to look at local legislation and for example APA guidelines to determine this.

Furthermore, it is recommended to split implementation of a new EBL system in an initial phase and a ‘maturing’ phase. The initial phase will provide a solid but basic system, while the ‘maturing’ phase will further enrich data provenance and user experience.

In general, one could further explore whether the EBL system development can be aligned with the recommended extension of the DDN pilot study (see Section 4.3). In that case it should be modeled in such a way that manual or automatic connection to the DDN is possible.

¹⁸ Scrum is based on the theory of empirical process control and uses an iterative and incremental approach to get predictable results and control risks. From scrum.org: ‘read the scrum guide’.

7 Awareness raising: website, front office and workshop

Raising awareness of data management was a second goal of the survey and this has worked out well: during the survey several researchers contacted the university library for information and advice on data management issues like training and posed specific questions on intellectual property rights. The survey results show a clear need for an informative website on different aspects of data management, advice on data storage, an information point on intellectual property rights concerning research data and need for data management training (see Section 3.2.3). To some extent these needs are interconnected. Therefore the three tasks that in the project plan were listed under the heading Awareness Raising – website, virtual information desk, and workshops about data management – are discussed together in this chapter.

As a direct result of the survey a small network of PhDs was formed as sounding board for future workshops, on-line courses and a new web site on data management. The survey and the Taskforce Scientific Integrity in general were discussed in an issue of Erasmus Magazine (nr 18, 2013)¹⁹.

7.1 The data support web site

7.1.1 Results

Given the very short project period, the team opted for the practical approach to use another data management site as template. Instead of spending energy and time on extensive research on user requirements and blue prints, work has been done incrementally or agile²⁰ to deliver a basic site in the summer of 2013. The basic site is developed step-by-step based on user input, in connection with the evolution from virtual information desk to front office.

After inspecting other university websites, the project team has selected the site ‘Data management support for researchers’²¹ from the University of Glasgow as an example. The development of this site was funded through the JISC Managing Research Data programme²². Tone of voice and structure of the Glasgow site are directly targeted at supporting researchers in a practical and matter-of-fact way. Several PhD researchers at EUR rated the site positively on dimensions as clearness, informativeness and usefulness.

The already published content on Research Matters was remapped on the new structure. A first version of the site went live with basic content about aspects of data management, a virtual information desk, links to online courses and a news section.

The data support web site is a section of Research Matters and is structured around four stages in the research data cycle: creating, organising, accessing and looking after your data. Apart from the informative content, the site has sections on:

- Online courses and tools
- A Virtual Information Desk with a FAQ, a news section, a document library and a blog.
- News from external stakeholders like DANS/3TU and partner front offices at Leiden and Delft Universities.
- A share functionality, which allows researchers to share content pages with colleagues.

7.1.2 Recommendations: from stand-alone site to centre of expertise

- In Q3 and Q4 of 2013 the site has to be developed further from a stand-alone site on data management to a centre of expertise and advice in a virtual network of (local and national) stakeholders.
- Social media and smart connections (RSS feeds, APIs) should be used to connect with already existing local online communities centred around data management to ensure the site will

¹⁹ Retrieved from https://www.erasmusmagazine.nl/archief/jaargang_16_2012_2013/em_18_06_06_2013/ p. 8

²⁰ http://en.wikipedia.org/wiki/Agile_software_development

²¹ <http://www.gla.ac.uk/services/datamanagement/>

²² The site is based on findings in a scoping study:

http://www.lib.cam.ac.uk/preservation/incremental/documents/Incremental_Scoping_Report_170910.pdf

become the place to be for the latest information, expertise, tools and support on data-related issues.

- To trigger attention and traffic to and from the site, a campaign should be designed. This campaign should be part of a campus-wide data awareness campaign.
- The data support site will function as a tool(box) for the virtual information desk and the front-office that is to be developed (see next section). Editorial responsibility should be organised accordingly.
- In a follow-up project the relation between the data management site and the intranet at Erasmus MC needs to be addressed.

7.2 The virtual information desk

The website also provides a virtual information desk for all questions about data management. Researchers can contact the information desk by a pre-structured web form on the data support site, via the university library's Virtual Desk or by direct mail. Questions are answered by the EDSC Data Librarian or by the Copyright Information Point. The Data Librarian is responsible for maintaining a list of Frequently Asked Questions for the website.

Desk personnel will be instructed to refer questions sent through the Virtual Desk to the Data Librarian and/or the Copyright Information Point. The 1-hour instruction will address the very basics of data management and its context. A glossary of terms will be provided for future reference.

Eventually the virtual desk will evolve into a front office with a more elaborate set of services, including advice and instruction with regard to data management of current research, data storage and archiving, funder requirements, IPR, and the relation between publications and data. It will also initiate workshops and courses on data management (see Section 7.3).

The notion "front office" suggests the existence of a back office. The role of back office is played by archives such as those collaborating in Research Data Netherlands. They are responsible for long-term preservation of and access to data in a trusted digital repository. In this network researchers can seek information and advice locally from the front office, which in turn is informed about data management developments by the archives in the back office. Front office and back office might collaborate on training courses for the research community.

7.2.1 Organisation and training

Evolving from information point to front office will bring into play Liaison Librarians who actively promote data management and who even acquire data sets from researchers, with the aim of making the data sets available to others. As yet it is to be determined how the Data Librarian and Liaison Librarians will work together in the front office.

A first step will be to train Liaison Librarians in data management issues. The 3TU/DANS course "Data Intelligence 4 Librarians"²³ is especially designed for this purpose and not only addresses data management issues, but also skills like data acquisition and advice. The revised version of the course, scheduled for early 2014, will be called "Training 4 RDM Supporter", to emphasise the importance of research data management.

7.2.2 Recommendations

A follow-up project should:

- Explore the relation and the division of work between front and back office, in collaboration with Research Data Netherlands.
- Describe workflows concerning data management, describe the tasks of data librarian(s), liaison librarians and other UL personnel and describe further training needs.
- Implement an awareness program on data management and the services of the front office.

²³ <http://dataintelligence.3tu.nl/nl/home/>

- Describe editorial management and further development of the data support site including moderation of social media, blog content et cetera.
- Address the role of the front office for Erasmus MC.

7.3 Workshop and courses

In the project plan the goal of this component was expressed as follows (summarised): “A training module or a workshop about data management is available. The workshop is intended as an optional instrument that can help researchers and research support staff to deepen their knowledge of research data management.” As the survey results indicate (see Section 3.2.3) the need for training in data management is paramount.

7.3.1 Results

During the project period an out-line was written for an explorative data management workshop targeted at PhD students to be held at a seminar on Scientific Integrity organized by the EUR graduate school EGS3H. However, due to the emphasis on integrity and the integrity game in the parallel workshop, the workshop on data management was withdrawn after consultation with the organizers of the seminar; competition between these related issues seemed counterproductive.

Several PhDs were asked to test an open source online training module about data management called Mantra data management training²⁴, developed by the University of Edinburg. The course covers the main aspects of data management from storage and file types to IPR-issues. The online course was evaluated as very informative and useful as introduction to the basics of data management. This PhD sounding board suggested to combine the course with a face-to-face follow up workshop to address in depth questions and issues related to specific data types or research projects. In November 2013 EGS3H, in cooperation with the university library, will organise these follow up workshops.

The university library is exploring other courses on data management which could be offered to researchers as easy to access training modules, any place any time. An example is Epigeum’s course on Scientific Integrity²⁵, in which several aspects of data management, e.g. IPR, are taught in a blended learning environment (online and face-to-face). Delft University just started staff training with this course in July 2013.

7.3.2 Recommendations

- Basic training on data management using online training modules is a quick win. Advice about these modules should be a task of the virtual information desk, as is testing new courses.
- Developing courses or extensive workshops centred on different data types is time-consuming and costly and presupposes discipline specific knowledge. Therefore it should be explored if the EUR could share courses and training personnel with other universities, for instance in cooperation with Leiden and Delft or with the UKB (the consortium of the university libraries and the National Library).

8 Cost estimations

As yet, the available information about the costs for carrying out the recommended data management activities is limited. Therefore some scenarios are sketched, which have to be developed further.

Storing and archiving research data (Chapter 4)

The costs for storing and archiving research data depend to a large extent on the storage capacity that is required and on aspects such as data security and availability. The survey has shed some light on the storage capacity that researchers need. For instance, the self-reported estimated average of produced data circles around 50 GB per researcher, with extreme estimates (over 100 TB) by respondents from Erasmus MC and the Rotterdam School of Management. However, 1 in 7 respondents does not know the volume of their data, so these numbers should be used with care. Furthermore, extra security

²⁴ <http://datalib.edina.ac.uk/mantra/>

²⁵ <http://www.epigeum.com/>

measures might introduce extra costs, as does fast data recovery from backups. As yet the overall EUR requirements are unavailable for budgeting. Furthermore, in the Research Data Netherlands consortium the business models for archiving institutional data – i.e. large amounts from trusted data providers – are still under development. Another unknown is the amount of data that EUR researchers will eventually select for long-term archiving, because it will not necessarily be everything that is being stored during a study. The project team recommends to *carry out a pilot study with one faculty, the front office, the back office and SSC ICT over a sufficiently long period to collect answers to these unknowns*. This might be connected to the recommended extension of the Dataverse pilot study (see Section 4.4).

Protocols for data management (Chapter 5)

To refine, implement and maintain the minimum protocols for data management the project team estimates that at least the following effort is needed:

- At the EUR central level:
 - Initially: implementation of the EUR minimum and support for deans²⁶ and research support officers: research data support officer: 25 days
 - Structurally: support for faculties, collaboration with the university library front office, contact with the Executive Board: research data support officer: 16 days per year
- Locally, per faculty or school:
 - Initially: refinement of the EUR minimum to one or more local minimum protocols and local implementation:
 - the dean of research and the research support officer: 2 days together
 - per discipline 2 senior researchers and/or PhD supervisors: 4 days together
 - commitment building activities, possibly in the broader context of scientific integrity awareness: 1 day per person involved
 - Structurally: evaluation and adaptation of the protocol:
 - the dean of research and the research support officer: 2 days per year together
 - per discipline 2 senior researchers and/or PhD supervisors: 4 days per year together

Erasmus Behavioural Lab pilot study (Chapter 6)

For the pilot study at the Erasmus Behavioural Lab the estimate given in April 2013 still holds. It is attached as Appendix 7.

Awareness raising: front office, workshops and website (Chapter 7)

For a front office that also offers data management workshops and develops and maintains the associated website the following budget is needed for 2014:

Function	Initial costs
Question handling	17.040
Advice and support	6.800
Website/research data portal	11.200
Training/workshops for schools/institutes	33.442
Liaison and acquisition	9.700
Information point IPR and copyright	29.400
Project Management	8.440
Total	116.022

Table 1 Initial costs for implementing the front office

This total amount covers the initial costs plus the costs for training and courses in the second half of 2014 (€14.600). The estimated costs for “information point IPR and copyright” are based on

²⁶ In the context of responsibility for supervising data management protocols, for “deans of research” one can also read “research deans” of – at Erasmus MC – “heads of departments”.

experience with the copyright information point in Q1-Q3 2013. Following the year 2014 structural costs are estimated at roughly the same level, but with a significantly larger share for question handling and a correspondingly smaller one for development of training materials and the website. The underlying detailed budget can be found in Appendix 8.

8.1 Recommendation

Carry out a pilot study with one EUR faculty, the front office, the back office and SSC ICT over a sufficiently long period to collect the information needed to budget the costs of research data storage and archiving.

9 Appendix 1 – Respondents and project team

We gladly acknowledge contributions from many people, especially from deans of research, research group leaders and senior researchers who provided information and gave feedback on research workflows and concept protocols for data storage (see Chapter 5):

Prof. Dr V. Bekkers, EUR – FSW / Dean of Erasmus Graduate School of Social Sciences and the Humanities
Prof. Dr P. van Bergeijk, EUR – ISS / Professor of International Economics and Macroeconomics
Prof. Dr A. Burdorf, EUR – EMC / Professor of Determinants of Population Health
Prof. Dr G. Engbersen, EUR – FSW / Research Director of the Sociology Department
Dr J. van Erp, EUR – ESL / Associate Professor Criminology
Dr C. Festen, EUR – EMC / Adjunct Director Education and Research
Prof. Dr D. Fok, EUR – ESE / Professor of Applied Econometrics
Prof. Dr P.H. Franses, EUR – ESE / Dean ESE, Professor of Applied Econometrics and Professor of Marketing Research
Dr S. Groeneveld, EUR – FSW / Associate professor Public Administration
Prof. Dr K. Heine, EUR – ESL / Professor of Law and Economics
Prof. Dr A. de Jong, EUR – RSM / Professor of Corporate Finance and Corporate Governance
Prof. Dr J. van der Lei, EUR – EMC / Professor of Medical Informatics
Dr K. Redekop, EUR – iMTA / Associate Professor, Clinical epidemiologist
Dr P. Riegman, EUR – EMC / Head of the Erasmus MC Tissue Bank
Prof. Dr M. Salih, EUR – ISS / Professor of States, Societies and World Development
Prof. Dr E. Schut, EUR – iBMG / Director of Research & Professor of Health Economics
Prof. Dr J. van Strien, EUR – FSW / Director of Research & Professor Biological and Cognitive Psychology
Prof. Dr A. Uitterlinden, EUR – EMC / Head of Genetic Laboratory Internal Medicine
Prof. Dr M. Verbeek, EUR – RSM / Professor of Finance, Scientific Director ERIM, Dean of Research RSM

The following persons have provided valuable input for the EBL pilot (Chapter 6):

H. Houtgraaf, EUR – CIO Office on EUR Application Architecture recommendations
M. Scholten, EUR – SSC ICT on data storage
L. de Bruijn, EUR – SSC ICT on data storage and servers for PoC
P. Oost, EUR – SSC ICT on Data Security and Privacy
S. Kamp, EUR – SSC ICT on Data Storage and servers for PoC
M. Steenhuis, EUR – UL on the EUR’s Research Information System
W. Mijnhardt, EUR – ERIM on the EUR’s Research Information System

Project steering committee:

F. Wynstra, EUR – ERIM / Professor and NEVI Professor Purchasing and Supply Management
F. van der Veen, EUR – FSW / Assistant Professor
G. Goris, EUR – UL / Deputy Librarian
M. van Donzel, EUR – ABD / Senior Policy Advisor
R. Juttmann, EUR – Erasmus MC / Coordinator Scientific Integrity
P. de Jaegere, EUR – Erasmus MC / Professor of Intervention Cardiology
M. van Dijk, EUR – ESHCC / Assistant Professor
J. London, EUR – SSC ICT / Staff

Project team:

M. Grootveld, Data Archiving and Networked Services (DANS) / Project Manager
R. Otten, EUR – UL / Head of Academic Services and Work Package Leader
G. de Bie, EUR – EBL / Manager Erasmus Behavioural Lab
P. Plaatsman, EUR – UL / Liaison Librarian Data & Economics
E. Haaijer, EUR / Project Assistant

10 Appendix 2 – Survey questionnaire

Research Data Management at the Erasmus University Rotterdam

Aim of this survey

Welcome and thank you for agreeing to take part in Erasmus University Rotterdam (EUR) research data management survey.

The survey is intended for EUR researchers across all schools and institutes.

It consists of 19 questions and should take maximum 15 minutes to complete.

In 2012 a Taskforce Scientific Integrity was established at the EUR, by the Executive Board, with a charge to develop appropriate measures for guaranteeing the integrity of scientific research. Developing an EUR-wide infrastructure to support the research data lifecycle and to store data in a sustainable way in particular is one of the Taskforce's subprojects.

This survey is conducted by the University Library and aims to discover how research data is used and managed across the University.

The questionnaire is designed to:

- Assist the project team to understand the data held by researchers
- Discover the influences and barriers to managing research data
- Establish what advice and support you require
- Identify current levels of research data management practice in institutes and schools.

All your comments will be treated as confidential and only anonymised information will be included in our project reports.

Any personal and identifiable data we collect from your survey responses will be accessible to the project team members only.

Your answers may be used to inform and develop research data management tools, infrastructure, and policies at the EUR.

Personal Information

1. Name

*2. Faculty/School

- Erasmus School of Economics
- Erasmus School of Law
- Faculty of Social Sciences (FSW)
- Erasmus Medical Centre
- Faculty of Philosophy
- Erasmus School of History, Culture and Communication
- Rotterdam School of Management
- Institute of Health Policy & Management
- International Institute of Social Studies

Research Data Management at the Erasmus University Rotterdam

*3. Which of the following best describes your research role

- PhD researcher
- Post-doctoral researcher
- Research fellow
- Assistant Professor
- Associated Professor
- Professor
- Other (please specify)

About your research data

In this section we would like to find out how you create and manage your research data.

*4. What types of research data do you create or work with as part of your research?

Select all that apply:

- Survey data
- Free public data (e.g. CBS data)
- Public data not available in a researcher-friendly database, requiring manual collection (e.g. generating a database by combining various databases that may require scripts)
- Commercially available data (purchased, often with non-disclosure agreement (NDA))
- Firm proprietary data (protected, specifically available only for the research project)
- Case studies
- Interviews (audio, video and written transcripts)
- Ethnographic participant-observation including web-based observations/virtual ethnography
- Documents
- Patient data under the Medical Research Involving Human Subjects Act (WMO)
- Patient data not under the Medical Research Involving Human Subjects Act (WMO)
- Experimental data (except data from clinical trials)
- Simulated data
- Other (please specify)

*5. Is your dataset primarily static or dynamic?

(Static: not likely to be modified. Dynamic: periodically updated or expanded)

- Static
- Dynamic

Research Data Management at the Erasmus University Rotterdam

*6. Where is this research data stored? Select all that apply:

- Hard disk drive of campus computer
- Hard disk drive of off-campus computer provided by funder or other partner in an externally funded project
- Hard disk drive of off-campus computer elsewhere
- Hard disk drive of laptop/netbook
- Hard disk drive of instrument/sensor which generates data
- External hard drive
- USB/Flash drive
- Shared drive/university server
- An institutional repository (please specify in „Other“)
- Web based service (e.g. Dropbox, Flickr, Google Docs, Basecamp)
- DataVerse
- Blackboard
- LiveLink
- Sakai
- Sharepoint
- CD/DVD
- Email client/server
- VHS/Video Cassette
- Floppy Disk
- Audio Cassette Tape
- Microfiche
- On paper
- Other: (please specify)

Research Data Management at the Erasmus University Rotterdam

***7. Please estimate the volume of research data across all of your work.**

- 1GB
- 1-50 GB
- 50-100 GB
- 100-500 GB
- 500 GB - 1 TB
- 1-50 TB
- 50-100 TB
- >100 TBs
- I don't know

***8. How frequently is your research data backed up?**

- Daily
- Weekly
- Monthly
- Ad-hoc
- Never
- I don't know

Research Data Management at the Erasmus University Rotterdam

*9. Where is your data backed up?

- Hard disk drive of campus computer
- Hard disk drive of off-campus computer provided by funder or other partner in an externally funded project
- Hard disk drive of off-campus computer elsewhere
- Hard disk drive of laptop/netbook
- Hard disk drive of instrument/sensor which generates data
- External hard drive
- USB/Flash drive
- Shared drive/university server
- An institutional repository (please specify in „Other“)
- Web based service (e.g. Dropbox, Flickr, Google Docs, Basecamp)
- DataVerse
- Blackboard
- LiveLink
- Sakai
- Sharepoint
- CD/DVD
- Email client/server
- VHS/Video Cassette
- Floppy Disk
- Audio Cassette Tape
- Microfiche
- On paper
- Other (please specify)

*10. Do you document or record any metadata about your data? Metadata is to make data more meaningful or easier to search for; it can vary from the data creation date to variable definitions and complete lab notebooks.

- Yes
- No
- I don't know

Research Data Management at the Erasmus University Rotterdam

11. If Yes, do you use any standards or guidelines?

- Yes
- No
- I don't know

* 12. Are you aware of any policy or requirements from your (main) funder (or research group, school et cetera) regarding research data management? In particular:

	Yes	No	N/A
Should you have a data management plan?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Should you store research data at a specified location during the project?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Should you make the research data publicly discoverable at the end of the project?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Should you make the research data accessible via Open Access at the end of the project?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

* 13. Have you developed a research data management plan for your (main) project?

- Yes
- No
- I don't know

* 14. How long do you typically store the research data after your last publication concerning this research?

(Please specify the duration of data storage in numbers of years)

Research Data Management

Research Data Management at the Erasmus University Rotterdam

* 15. Who are the (main) funders of your research? (select all that apply)

- EUR
- NWO
- ZonMw
- Agentschap NL
- KNAW
- Health funds (e.g. KWF/Dutch cancer society)
- TNO
- EU
- Commercial organisations (please specify)
- Other (please specify)

Sharing your data

You may share your research data via data repositories, data banks and data centres, submission to a journal to support publication, and informal sharing between researchers.

* 16. Who should typically access the research data you are creating? (select all that apply)

	Would not share with anyone	Would share with my immediate collaborators	Would share with others in my research center or at my institution	Would share with others in my field	Would share with funders/publishers	Would share with anyone
Immediately after the data has been generated.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
After the data has been normalized and/or corrected for errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
After the data has been processed for analysis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
After the data has been analyzed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Immediately before publication.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Immediately after the findings derived from this data have been published.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Research Data Management at the Erasmus University Rotterdam

*17. Do you / Would you deposit your data in a public subject/disciplinary repository if available?

Yes, I am required to do so.

I deposit data in (name of repository)

Yes, I choose to do so. I deposit data in (name of repository)

No, because

Not sure, because

Support for Research Data Management

The Erasmus University Rotterdam is committed to supporting researchers across the research lifecycle. We would like to know where you require help.

*18. I would require:

- Greater file storage capacity
- Tools to share data during research project
- An EUR repository to publish data
- Data management support when writing a research proposal
- A Research Data Management website for guidance and support
- Support to publish data to external subject repositories
- Help with analysing data
- Help to make better use of my final data sets e.g. create a website to showcase data
- Data management training
- Other (please specify)

Research Data Management at the Erasmus University Rotterdam

19. If you are interested in data management training, in which aspects would you like to be trained?

- Developing a research data management plan
- Documenting my data
- Formatting my data
- Storing my data
- Sharing my data
- Creating metadata for data
- Ethics and consent
- Funders requirements and research data management
- Copyright and intellectual property right (IPR)
- Data repositories and Open Access
- Other (please specify)

Thank you

Thank you!

We appreciate the time you have taken. Your responses will help us understand how research data is managed at the EUR.

The Taskforce Scientific Integrity would like to get an insight into policies, requirements, and good practices that are in use across the EUR. Therefore it would be very helpful if you would upload or point us to documentation that you know about.

Please send any documents or comments you find relevant to: rdm@ubib.eur.nl

Should you require more information about this survey or services on research data management, please contact:
Drs. Roel Otten, Manager Academic Services UB, otten@ubib.eur.nl

11 Appendix 3 – Findings from the Dutch Dataverse Network pilot

Goals

The Dutch Dataverse Network²⁷ (DDN) is an external platform used by five Dutch universities. The EUR University Library will carry out small pilot studies to determine its usability for storing and sharing research data of various data types. The amount of Dataverse studies (in Dataverse jargon datasets are called “studies”) should suffice for a qualitative insight into workflows. Explicit attention is paid to metadata and access rights.

Set up

The pilot builds on earlier work done in 2012 by the ERIM and the University Library. They started pilots after the ERIM report and workshop on Management of Research Data²⁸. The aim was to explore how researchers could be supported in storing different data types during different phases of their research.

To make these pilots as ‘hands on’ as possible, The Dutch Dataverse Network (DDN) was introduced as platform to store the research data during the pilot. The choice for DDN was inspired with by its growing popularity within University Libraries in the Netherlands and earlier experience during an EU funded project *Network of European Economists Online (NEEO)*.

Contributions came from three researchers from ERIM for three different data types: Experimental, Survey and Qualitative data. In the spring of 2013 the Research Data Management at Erasmus University Rotterdam Project plan started. In this project context piloting with Dataverse was continued with other data types and more in depth testing together with four more researchers from ESHCC and Psychology.

During the interview the researcher and UL staff member filled out the Cards Data management plan²⁹³⁰. The UL staff member created a Dataverse for the researchers and checked with the researcher if the metadata is correct and explains the basics of the Dataverse. After the interview the researcher tested the Dataverse and provided feedback. During the pilot, researchers were in charge of their own Dataverses; they could store data sets, change access rights etc.

Current status

The Dutch Dataverse Network

The Dutch Dataverse Network is created for researchers and lecturers. This service makes it possible to store a wide variety of scientific data (texts and raw research data, but also video material and complete databases) in an online environment, safely and sustainably. These data include digital research material as well as material used for educational purposes.

Researchers can index the data they have stored in a user-friendly way, and share their data with other scientists or interested parties. Researchers themselves determine who gets access to what data. Another option which makes the Dutch Dataverse Network inviting to researchers, is the possibility to refer to stored datasets by means of a URL and to link them to, for instance (scientific) publications. The great advantage of this 'persistent URL' is that the data is permanently available via this link, even if they are moved to another location.

In addition, within a Dataverse so-called collections can be defined, which in their turn contain studies. In such a collection also subcollections can be defined, resulting in a multi-layered tree structure. Collections studies from other Dataverses may be included (even from other owners, provided the studies are freely accessible). In DDN so-called studies can be created for separate subprojects, experiments or other parts of a research project or research career.

²⁷ <https://www.dataverse.nl/dvn/>

²⁸ Management of Research Data: Minimum and best practices for data collection, analysis and archiving, ERIM September 2012.

²⁹ http://www.eur.nl/researchmatters/research_data/data_management/data_management_plan/data_management_plan_interview/

³⁰ For more information on the Cards project see: <http://www.surf.nl/en/projecten/pages/cards.aspx>

To a limited extent the graphic design can be branded by adding headers, footers and banners, a descriptive text and an organisation's own conditions of use.

Within the Netherlands five University Libraries use DDN as a data management platform. The Netherland Institute for Ecology, NIOO joined the Network in July 2013. The ULs meet regularly, both virtual and face-to-face. The network is professionalising and is in the process of formulating a SLA, writing application management documentation and revising the terms of use.

Main user comments

- Presently it is not possible to do a bulk upload in DDN, which by most researchers is seen as deal breaker. Bulk upload will probably be available in the next release of DDN. For now a temporary solution would be to upload a zip file. Other ways to upload files in bulk are tested by partners in the network.
- The persistent identifiers to refer to research data are one of the advantages of DDN, but this is only mentioned by one of the researchers. That researcher liked that option but had some difficulty sharing the identifier.
- Questions arose about responsibility for the metadata of the research data: is assigning metadata a task for the library or the researchers or a combination of both?
- Researchers found the basic DDN instruction by UL staff during the Pilot(s) too limited.

Comments from the University Library

- Given the present terms of use DDN is not suited for large files and sensitive data.
- Given the present terms of use research ownership of research data is at least ambivalent.

Recommendations

- Provide researcher with structure, support and guidelines about the metadata and standards. Some data explanation might be required from them if they want re-use of their data or careful reproducing their research outcomes. Metadata is provided by UL staff, but the researcher checks this and if needed corrects and add meta data.
- The terms of use of DDN need to be adjusted to several partners instead of just the University of Utrecht.

Remarks from the pilot participants (in Dutch)

The text below is the complete set of comments made by pilot participants; it has not been edited.

ERIM onderzoeker, Qualitatieve data, RSM Assistant Professor

Enkele observaties van mijn kant:

- Het datamanagementplan: hierbij wil ik louter opmerken dat dit prima is in te vullen, op enkele vragen na. Dit zijn precies die vragen die ik jouw stuurde en die rood of geel waren gemarkeerd. Juist daar zou dus wat nadere duidelijkheid kunnen worden gegeven om ervoor te zorgen dat elke onderzoeker dit zelfstandig en makkelijk kan invullen.
- Wbt DDN: we hebben twee punten gezamenlijk gedaan: eerst de algemene invuloeefening om een 'project' te starten. Hierbij kan ik me herinneren was niet overal duidelijk wat precies ingevuld moet worden. Deze pagina is momenteel dus niet zelfstandig en makkelijk in te vullen door een onderzoeker. Wat betreft het tweede punt: het opslaan van de bestanden: dit is in mijn geval (met veel bestanden achteraf), nauwelijks te doen: het kost teveel tijd om dit een voor een te doen. Stel dat je vanaf de start met DDN begint, dan is nog steeds de verantwoordelijkheid bij de onderzoeker om nieuwe versies op te slaan. En ook dan lijkt mij dat telkens een aparte file moet worden geupload en gekenmerkt in het systeem. Ook dat lijkt me niet zo handig. Het kan natuurlijk wel, maar gebruiks vriendelijk is anders.

ERIM onderzoeker, Survey data , RSM Assistant Professor

We hebben het programma natuurlijk nog niet echt gebruikt, maar mijn eerste indruk is dat het programma niet heel gebruikersvriendelijk is. In dat opzicht zou het wel beter kunnen.

ESHCC, assistent professor en deelnemer aan EUR RDM projectgroep

De indexering van de data kan beter. Waar denk ik aan:

- 1) Thesaurus met aantal voorstellen voor categorieën die je aan je data kan linken.
- 2) Mogelijkheid om bij data ook verschillende keywords toe te voegen waarop je later kan zoeken, maar die ook data in algemene termen beschrijven.

Tot slot vraag ik me af of je niet clusters van files kan uploaden die samen één bron uitmaken. Zo zou ik binnen mijn studie liefst een mapje bevolkingsschattingen hebben waaronder ik verschillende files kan uploaden van verschillende bronnen die allemaal indicaties bevatten voor de bevolkingsschattingen die ik heb gemaakt.

Wat ik nog wil opmerken:

- 1) De structuur van Dataverse is soms nog wat onduidelijk. Het blijft een wat onhandige versie van Dropbox.
- 2) Doordat je files één voor één moet toevoegen is dit systeem vooral geschikt voor nieuwe projecten. Als je al bezig bent met je onderzoek, zoals ik nu, of het is al afgerond, dan is het toch tijdsintensief omdat je echt alle handelingen gewoon opnieuw moet doen. Dus als het ingevoerd wordt, best voor nieuwe projecten volgens mij.
- 3) Nog over de structuur van de Dataverse: hoe gaat het als je data in verschillende studies wilt gebruiken? Moet je dan opnieuw uploaden en alles invoeren, of kan je dan verwijzen naar bestaande gegevens die al in de Dataverse staan? Dat laatste lijkt het handigste, maar is me niet helemaal duidelijk.

EURFSW Psychologie, assistent professor en deelnemer aan research data projectgroep

Ik heb tot nu toe al mijn data van het nieuwe experiment overgezet. Dat ging vrij soepel.

Positief vond ik de flexibiliteit van het systeem. Ook bij een in eerste instantie foute indeling was het vrij eenvoudig dit netjes recht te brengen. Positief vond ik ook de snelheid van het systeem. Ook bij het overzetten van relatief grote bestanden (60-100 MB) was er weinig wachttijd. Laatste positieve punt is de overzichtelijkheid van de opslag. Je ziet in een oogopslag wat er is opgeslagen.

Negatief vond ik de benadering in individuele files. Je moet 1 voor 1 je data overzetten. In mijn geval was dit niet zo'n probleem omdat ik nog relatief weinig data had; allemaal 'work in progress'. Ik zou bij mijn andere experimenten hier wel grote problemen mee hebben. Bij sommige van mijn fMRI experimenten gaat het werkelijk om duizenden files ruwe data. Dus voor echt ruwe data in die context is dat geen doen.

Een positief en negatief punt tegelijk vond ik de ongedwongen benadering waarbij je als onderzoeker vrij bent op te slaan zoals je dat zelf wilt. Vrijheid en flexibiliteit zijn natuurlijk meestal positief.

Negatief aspect hierbij is dat ik niet gedwongen wordt iets over de indeling van mijn data te vertellen, zodat voor een willekeurige buitenstaander nog steeds volstrekt onduidelijk is wat de data inhoudt en hoe deze gebruikt kan worden. Natuurlijk kan dit op initiatief van de onderzoeker zelf toegevoegd worden, maar de incentive bij mij hierbij is erg klein. Dit heeft voor mij tot gevolg dat ik nog steeds geen file heb bijgevoegd met een beschrijving van de data en zo de winst voor mij beperkt is. Ik heb nu extern de data staan wat natuurlijk een veilig gevoel geeft, maar als er iets met mij gebeurt dan heeft niemand wat aan de data. Ook precieze controle op mijn data is zo niet mogelijk.

ESHCC, onderzoeker

ihab ik de volgende opmerkingen:

- De usability is niet erg goed. Ik moest bijvoorbeeld erg zoeken om erachter te komen hoe ik nu precies data kan toevoegen. Kan me voorstellen dat jullie daar weinig aan kunnen veranderen maar denk dat het collega's wel zal afschrikken.
- Belangrijkste is dat voor mij nog onduidelijk is wat nu precies de toegevoegde waarde is van dit systeem boven DANS. Het is natuurlijk verleidelijk om te gaan vergelijken met dat wat je al kent maar tegelijk is het ook onvermijdelijk. Maar is het grote voordeel dat ik hier zelf data kan uploaden en updaten terwijl de persistent identifier hetzelfde blijft?

Voor mij blijft dat gewoon wat onduidelijk en icm de niet erg gebruiksvriendelijk ogende interface ben ik bang dat veel collega's nogal zullen schrikken.... Meer specifiek:

M.b.t. toevoegen:

- Ik moest behoorlijk zoeken hoe ik een studie aanmaak. Uiteindelijk gevonden dankzij handleiding maar lastig dat er nergens een knop zit als 'Create study' of iets dergelijks.
- Wel erg veel velden die je kunt invullen waarvan ik van veel niet weet wat het betekent, zoals topic classification, related material, related studies.

M.b.t. Uploaden

- Nadat ik een bestand heb geupload, krijg ik dit scherm maar snap niet goed wat ik kan invullen bij Category etc. Gelukkig kan ik gewoon op verder klikken maar je bent toch geneigd alles te willen invullen.

The screenshot shows a web browser window with the URL www.dataverse.nl/dvn/dv/eurshcc1/faces/study/AddFilesPage.xhtml?studyId=152. The page is titled 'Dutch Dataverse Network > MartijnKleppe Dataverse'. It displays a table for uploading files, with one row selected:

Category	File Name	Description	Size (bytes)	Remove
	1970-8 speciale onderwerpen		473479	Remove

Buttons at the bottom include 'Save' and 'Cancel'.

- Bulk upload zou voor mij wel een must zijn

Na accessible maken:

- Krijg ik nu automatisch een persistent identifier? Volgens mij wel, want geloof dat ik die heb kunnen afleiden maar werkt die ook meteen?
- En hoe kan ik andere nu attenderen op mijn dataset? Toen ik mijn collega de URL gaf, kon hij 'm niet openen ondanks dat ik 'm op public heb gezet.

ESHCC onderzoeker

Allereerst, ik heb de handleiding niet gelezen. Deels wegens tijdgebrek, maar ook om me meer in te kunnen leven in wat er uiteindelijk mee gaat gebeuren (mensen lezen handleidingen nou eenmaal niet).

Daarnaast, ik denk dat het goed is om na te denken voor wie deze tool nou bedoeld gaat zijn. Het vergt toch wel wat nadenken en overwegingen om data op deze manier aan te bieden, waardoor het kan zijn dat het vooral interessant is voor onderzoekers die toch al bewust bezig zijn met data open access aanbieden. Daarbij ook de vraag dan in welke fase van onderzoek een onderzoeker deze tool zou gaan gebruiken: in mijn eigen use case zou dit toch meer richting einde van onderzoek zijn, met als grootste voordeel tov DANS dat je de dataset nog kunt aanpassen, en verschillen kunt uitleggen.

Overige opmerkingen:

1. Hoe de Handle aan de praat te krijgen is erg onduidelijk, geeft geen feedback over wat er veranderd moet worden. Na lang spelen bleek dit pas te werken na release van dataset, wat jammer is want wat mij betreft zou je best mogen linken naar een dataset die nog in Draft status is
2. Het verschil tussen een Study en een Dataverse is vrij onduidelijk. Kan een onderzoeker meerdere Dataverses hebben? Waarom? Daarnaast heb je ook nog Collections, wat zijn dat dan?
3. Keywords: vocabulary & URL onduidelijke meerwaarde
4. Co-author van dataset: hier zou ik eigenlijk een persoon willen kunnen kiezen (via DAI, of via SURFconext), zodat het ook onder zijn/haar dataverse terecht komt
5. Onder "edit cataloging information" en "create study template" zitten heel andere metadata-velden. Is "create study template" een aparte pagina voor beschrijving tbv replicatie-onderzoek? Als

het een template is om snel Studies te maken in Dataverse, dan is mij niet duidelijk waarom het extra metadata-velden bevat

6. Version history voor release ziet er netjes uit, maar toont interessant genoeg niet de version notes, die worden hierna pas gevraagd in een pop-up

7. Waarom staat in de referentie explicet [Distributor] en [Version]? Dat zou ik er meteen uitslpen als het eenmaal in Word staat.

a. Daarbij ook, ik mis knoppen om het te importeren in Mendeley, RefWorks, of om export te maken naar een Bibtex file

8. Wat is deaccession?

9. Jammer dat bestanden niet overschreven kunnen worden, maar moeten worden gedelete en opnieuw geupload

10. Heb study gereleased en op public gezet, alle bestanden op 1 na restricted. Maar zie nu mijn study niet staan op homepage met Released Dataverses

Ik zou mijzelf in ieder geval Dataverse wel zien gebruiken!

Technisch netwerkbeheerder UB

- samenwerking met Utrecht verloopt soepel
- beheermogelijkheden (permisies) totnutoe toereikend
- performance/stabiliteit onduidelijk, soms snel, soms traag tot onbereikbaar
- inrichting moet worden aangepast voor gebruik door meerdere universiteiten
- geen SLA voor applicatie, wel voor hosting (hardware)
- er moet nog een upgrade van de DDN-software plaatsvinden, december
- koppeling ERNA-id komt er aan, december
- basisfunctionaliteit inmiddels redelijk onder controle
- Utrecht en Tilburg hanteren verschillende policies m.b.t. gebruik
- onze eigen policy inmiddels bijgesteld (study -> dataverse), dient nog verder te worden uitgewerkt
- standaard documentatie van IQSS niet toereikend
- er moeten nog verschillende legal statements komen
- ‘gebruikskader’ verdient m.i. verduidelijking/aanscherping, b.v. waarvoor wel/niet, wanneer wel/niet, hoe, en dit alles in relatie tot omliggende en alternatieve opslag- en publicatieoplossingen als PC-harddisk, netwerk-schijf, RePub, DANS, clouddiensten, etc.
- leuk project!

Onderbibliotheekaris

- TU Twente gaat ook meedoen. Heeft in Escape project geleerd hoe datasets aan publicaties te koppelen. Zij brengen die ervaring mee binnen DDN
- DDN niet handig voor grote datasets, communiceer dit goed
- Oefen ook peer review process binnen DDN
- Kies enkele vervolg Pilot(s) maar manage de verwachtingen. Mag specialiseren binnen één data type voor het bedenken van een workflow, misschien behavioral lab.

EDSC coordinator

- we hebben nu een DDN met één file en vele gebruikers en een DDN met één gebruiker en heel veel files.
- onderzoekers hebben Dataverse nodig en geen study. In study kun je namelijk permisies niet aangeven.
- contributor heeft geen permisies, curator wel
- voorlopig library specialist ook als admin aanwezig.
- Lastig hoe je files te ordenen binnen je DDN. Weet nu dat het via categorieën en studies kan
- telkens terugkerende study version notes popup is erg hinderlijk. Kan die ook uitgezet worden?
- eigen look and feel
- DDN Utrecht nog erg in ontwikkeling, Dataverses: 27 | Studies: 31 | Files: 434 waarvan 3DDN, 3 studies en 7 files van de EUR.
- spss files worden geconverteerd, dit is handig voor analyse programma van DDN maar wil je dat niet dan is het juist lastig want conversie duurt lang. Kan via zip of rar worden omzeild. Opgelost in nieuwe versie door other format te kiezen.

- Komen nog nieuwe versies van DDN voor verschillende type files en bulk upload
- Alternatief voor DDN Oxford's Dataflow <http://www.dataflow.ox.ac.uk/index.php/about/about-datastage>
- Lopend onderzoek in datastage daarna opslag in databank
- Inloggen via Surfconnect geeft soms moeilijkheden in IE. Pagina reloaden en dan lukt het wel.

12 Appendix 4 – General recommendations for storing research data

- 1 Always maintain copies of the original, “raw” research data. In case of paper and pencil questionnaires, this means storing the actual forms. In case of electronic data, it means the original completed electronic forms. In case of qualitative research, it means the original audio files or transcripts of interviews, or field notes. In case of secondary data or data collected by others, it means the originally obtained data (data ownership issues permitting). Thus, while the nature and form of the actual raw data may vary, the basic principle applies that the researcher should be able to convincingly demonstrate that this original version of the raw research data has not yet undergone any selection, purification or transformation steps.
- 2 We recommend tying the original data to the identity of the research informant/participant, even in the case of confidential data. Confidentiality can be maintained by separately storing a key, controlled by the (lead) researcher. At a minimum, the identity of individual research informants/participants should be recorded (without necessarily relating this to a specific response). The key principle here is that anonymity can be guaranteed (if necessary) with respect to published data, without sacrificing the identity of research participants for the original data collected.
- 3 The data collection process should be clearly described. This includes the names and roles of the researchers involved and/or the organisations providing the data (such as research agencies). The descriptions should be detailed to the extent that the process can fully be traced back.
- 4 The data input and analysis procedure should be documented in detail, so that the analysis can be replicated exactly. This includes major analysis steps that may in the end not be reported in further publications, but which have been instrumental in steering the analysis process. All substantial files should be stored, including for instance specific software syntax, diagrams, graphical presentations, etcetera. Again, the names and roles of the researchers involved should be provided.
- 5 For each crucial data compilation, purification or transformation step, it is recommended that clearly identifiable and described data sets are stored. (Crucial steps transform data such that it is impossible to revert to the rawer data when only the transformed data is available.)
- 6 All original, “raw” data and the documentation of the data collection and analysis process should be stored for a minimum of five years after publication of the most recent publication using this data. This applies as long as specific professional or journal policies do not require a longer storage period.
- 7 All (electronic) “raw” data and the documentation of the data collection, input and analysis process should be stored in duplicate. At least one set of data should be stored on a university or external network, with appropriate safeguards regarding anonymity and data ownership. Data collected on paper should be transferred to an electronic medium in its entirety and stored electronically (where possible by scanning the entire documents).
- 8 In the case of co-authored papers where another person is executing data collection, input and/or analysis, we recommend storing a copy of the raw data yourself as well (confidentiality and data ownership issues permitting), and storing the documentation of data collection, input and data analysis procedures.

Source: letter about scientific integrity from Professor Verbeek to ERIM fellows, members and doctoral students, 04/07/2012 (reference MV/tv 0012.003840).

13 Appendix 5 – Digitalisering en opslag van onderzoekdata

Doelstellingen:

1. Van alle onderzoekdata bestaat een definitieve gedigitaliseerde versie die geschikt is voor opslag in een digitaal depot.
 2. Elke afdeling maakt gebruik van één of meerdere depots, die voldoen aan hoge eisen met betrekking tot bestendigheid, toegankelijkheid en beveiliging. Bij elke publicatie worden op het moment van acceptatie ten minste de volgende data opgeslagen in één of meerdere van deze depots:
 - de (concept) publicatie zelf
 - het betreffende onderzoeksplan/protocol,
 - ingevulde en ondertekende informed consent formulieren (indien van toepassing)
 - de verzamelde “ruwe data”,
 - een (statistisch) analyse bestand en
 - de analyse resultaten.
 3. De gedeponeerde data per publicatie zijn onveranderlijk. Data m.b.t. afzonderlijke publicaties die gebaseerd zijn op dezelfde dataset, worden niettemin afzonderlijk gedeponeerd.
 4. Onderzoeksgegevens worden zodanig gedeponeerd dat deze snel en eenvoudig kunnen worden geraadpleegd. Er wordt een register aangelegd waarin voor elke publicatie staat geregistreerd in welk depot, op welke plaats welke documentatie is gedeponeerd.
 5. Alle gedeponeerde data zijn toegankelijk voor (vertegenwoordigers van) de Raad van Bestuur.
- a. *Beschrijf de logistieke en technische knelpunten die de afdeling voorziet bij het nastreven van deze doelstellingen*
- b. *Beschrijf de door de afdeling gewenste centrale ondersteuning bij het oplossen van deze knelpunten*
- c. *Beschrijf, met in achtneming van a. en b. de stappen die de afdeling neemt om deze doelstellingen te bereiken.*

Source: memorandum about action plan scientific integrity (WI) from dr. Juttmann, coordinator WI, to the theme chairs in Erasmus MC, 15/05/2013. It should be noted that the action plan also addresses organisational culture and education, alongside (an infrastructure for) research data.

14 Appendix 6 – Concept data checklist for various disciplines

The following concept list was used in the data storage interviews with senior researchers and research group leaders.

Research Data Management:

What should be stored for the sake of replication and accountability?

Research process:

Data collection	Data processing	Data analysis	Publish	
T0 = project start	T1	T2	T3	T4

Research Data Management: What should be stored for the sake of replication and accountability?					Respondent information: (Name, affiliation, research area)										
Research process:															
<table border="1"> <tr> <td>Data collection</td> <td>Data processing</td> <td>Data analysis</td> <td>Publish</td> <td></td> </tr> <tr> <td>T0 = project start</td> <td>T1</td> <td>T2</td> <td>T3</td> <td>T4</td> </tr> </table> 					Data collection	Data processing	Data analysis	Publish		T0 = project start	T1	T2	T3	T4	
Data collection	Data processing	Data analysis	Publish												
T0 = project start	T1	T2	T3	T4											
What would make sense to store at T0 when the project starts?															
Suggestions:															
0.1 Project plan, study design or study protocol															
0.2 "Dataparagraaf" ("Data section" in e.g. a grant proposal)															
0.3 Data Management Plan															
0.4 Unique project identifier															
0.5 Future owner of the research data															
0.6 Agreements about the future management and use of the data. E.g., when you are not the first author, how will data storage and archiving be arranged?															
0.7 Research questions															
0.8 Hypotheses															
0.9 ?															
What would make sense to store at T1 at the end of the data collection stage?															
Suggestions:															
1.1 Raw data															
1.1.1 A snapshot of the data															
1.1.2 Interview transcripts with member review and respondent validation															
1.1.3 Interview transcripts without such authorisation															
1.1.4 Notes from the interview sessions															
1.1.5 A field journal															
1.1.6 Digital scans of "paper" versions															
1.1.7 First digital version of "paper" data, i.e. entered in a software application															
1.1.8 Printed out or saved PDFs of electronic Internet documents															
1.1.9 ?															
1.2 Documentation about															
1.2.1 Sampling logic; sample size et cetera to prevent "additional sampling"															
1.2.2 The boundaries of the population to which the study applies															
1.2.3 Stimuli used															
1.2.4 The seed, in case of simulated data															
1.2.5 Invitation letters															
1.2.6 The original questionnaire															
1.2.7 The data items/identifiers/codes/labels from public or commercially available data in the description of the variable definitions; the definitions and criteria used by the data provider															
1.2.8 Researchers and others such as readers who are involved															
1.2.9 Informants, respondents, interviewees, participants															
xxx <i>Suitably anonymised, but checks ought to be possible.</i>															
1.2.10 Signed informed consent forms															
1.2.11 Elaborate descriptives of data that cannot be published (i.e., restricted by the terms of a Non-Disclosure Agreement)															
1.2.12 The data collection process, including:															
1.2.12.1 Data collection protocol or method, if different from what is in the project plan/study design															
1.2.12.2 Database queries that were used, including when the query was carried out															
1.2.12.3 In case of externally collected and coded data (e.g. from a market research company): maximum information about the provenance of the data and about the definitions and criteria used by the external data provider															
1.2.12.4 Ditto: Any relationship the researchers/authors may have with the external provider of the data															
1.2.13 ?															

What would make sense to store at T2 = at the end of the data processing stage?	
Suggestions:	
2.1 Step by step documentation of the data processing process, including all changes to the earlier plans and protocols: What was changed and why?	
2.2 Data screening methods	
2.3 Documentation of any imputation, selection, and data pruning processes	
2.4 Documentation of deleted outliers, missing data report, and how labels are linked to the survey items.	
2.5 The syntax or other code used in data processing.	
2.6 Variables, code books	
2.7 Lab journals	
2.8 Intermediate results, such as:	
2.8.1 ?	
What would make sense to store at T3 = at the end of the data analysis stage?	
Suggestions:	
3.1 Step by step documentation of the analysis process, including all changes to the earlier plans and protocols: What was changed and why?	
3.2 Current version of the data	
3.3 "Everything that is not in the paper to be submitted"	
3.4 Method, variables and underlying items (including labels) used in the final model.	
3.5 Robustness checks, sensitivity analyses, summary statistics and graphs	
3.6 Syntax or code of your final data analysis, plus documentation of configuration and version of the software used, including when the analysis was run	
3.7 Syntax and output log files of the statistical software, including when the analysis was run	
3.8 Lab journals	
3.9 Anything that you have tried and didn't work or didn't yield significant results	
3.10 A detailed log of all the "trials" of an experiment that you run	
3.11 ?	
What would make sense to store at T4 = publication?	
Suggestions:	
4.1 The publication-specific version of the dataset	
4.2 The submitted version of the publication	
4.3 The response from the editor (even in case of a rejection) plus the reviews	
4.4 Your answers to the reviewers' questions	
4.5 The reviewed and revised version of the publication	
4.6 The final version of the publication	
4.7 ?	

15 Appendix 7 – Raming structurele kosten na de EBL-pilot

This information will be available by the formal conclusion of the EBL-pilot, early October 2013.

16 Appendix 8 – Front office budget

Function group	Timing	Tasks	Function scale	Hours *	Costs ** / ***	
					Initial	Structural/year
Virtual Desk /Front office	Initial Q1/Q2/ 2014	Project management development Front Office: organisation/workflows/etc (manager AS) designing workflows front office	12	40	5.400	
	Q1/Q2/ 2014 Structural	Management Front office (Data librarian EDSC)	11	60	3.040	
Question handling	Initial Q1/Q2 2014	Basic question handling Virtual information point / building FAQ, organising coordination with other Frontoffices (Leiden/Delft) and back offices (Dans/3TU, DDN)	11	80/year		6.800
	Structural	Question handling and advice on data collection/ data processing/ data storage/ data archiving issues/data reuse/ security/citing data/ publishing data/, data mining;	average 10	200	14.000	
Advice and Support	Initial Q1/Q2 2014 Structural	Participation/advice in pilots Support on Metadata / generating persistent identifiers/Advice on storage possibilities	11	40	3.040	
	Initial Q3 2013 Q1/Q2 2014 Structural	Basic data portal Development data portal content/ soc media/ relation with sites from schools and institutes Editorial Coordination and Quality control	10	160	11.200	
			10	80/year	5.600	

Training / workshops					
Initial					
Q1	Inventory training needs schools and institutes / evaluating existing on and off line training material	11	80	6.700	
Q1/Q2 2014	Development training program/ workshops pilots included	11	160	12.160	
Q3/Q4	executing training program (in/external trainers, licences, other material) based on 8 faculties X 3 workshops X 8 hours ***	11	192	14.582	
Structural					
	Coordination training program	11	50/year	3.800	
	Executing training program (in/external trainers, licences, other material) based on 8 faculties X 3 workshops/year X 6 hours	11	144/year	10.944	
Faculty liaison and acquisition					
Initial					
Q1 t/m Q2 2014	Training 5 data librarianship LL's Training EDSC personnel (2) Training desk personnel (UL intern) Facility Liason and acquisition 5LL's	n/a n/a n/a 11	100	1.500 600 PM 7.600	
Q2/Q3	structural contacts schools/institutes 5 LL's	11	100/year	7.600	
	schooling/congres 5 LL's	11	100/year	7.600	
	schooling/congres data librarian (EDSC)	11	100/year	7.600	
Initial					
Q1	Organising back office (on/off campus)	8	420	29.400	
Q1/Q4 2014	Question handling /FAQ	8	420	29.400	
Structural	Question handling /FAQ		Total	116.022	119.594

* Hours based on 42 effective working weeks per year

** cost per hour based on : "Handleiding Overheidstarieven 2013", Ministerie van Financieen. Kolom 1: kostendekkend tarief per uur
 *** (licence) fees for storage services are not included

**** cost for training/workshops in 2014 are booked as initial cost.

17 Appendix 9 – Storage criteria matrix

This matrix is discussed in Section 4.1.

Platform	Type	Metadata information	Version administration	International collaboration possibilities	Storage up to 5 GB data	Flexibility to accommodate various file types	Persistent citation	Jurisdiction	Support organisation	Secure storage	Secure file transmission	Costs
@wEURk	OR	Needs to be configured	Needs to be configured	N	Y	Y	?	NL	N	Y	Y	€2,57/GB/year
3TU.Datacenter	A/ OR	Y	Y	?	Y	Y	Y	NL	Y	Y	Y	?
Basecamp	OR	N (tags)	N	Y	Y	Y	N	US	N	Y	Y	\$3,000/year 500 GB
Blackboard	OR	N	N	N	N	Y	N	NL	N	Y	Y	At EUR cost for storage over 1GB is for user
DANS EASY	A	Y	Y	N	Y	Y	Y	NL	Y	Y	Y	approx. €60 / GB for the long term
Dropbox	OR	N	Y	Y	Y	Y	N	US	N	N	Y	\$120/year 500 GB
Dutch Dataverse Network	OR	Y	Y	Y	Y	Y	Y	NL	Y	Y	N	Fixed (€2675 with 6 partners) and variable costs per institution (€ 5,40/ GB)
Figshare	OR	Y	Y	N	Y (250 mb per file)	Y	Y	US?	N	Y	?	No costs
Google Drive	OR	N (description fields)	Y	Y	Y	Y	N	US	N	N	Y	\$300/year 500 GB
Icebox	OR	N	Y	Y	Y	Y	N	US	N	Y	?	Starting price ranging from 3300 to \$34,995 and \$0.4/GB
Livelink	OR	Y	Y	Y	Y	Y	N	NL	N	Y	N	?
Sharepoint	OR	Y	Y	Y	Y	Y	?	NL	N	Y	Y	?



Besrijpen en begrepen worden.

Met openheid van zaken communiceren.

*Waardoor een intense en
hechte samenwerking
mensen en organisaties.
ontstaat.*

Informeren. Implementeren. Instrueren.

Overzicht is niets.

Inzicht is alles.

Samen.

Altijd en Overal.

Analyse voor project
Vestigium

Erasmus Behavioural Lab

T.a.v. Gerrit Jan de Bie



Versiebeheer

Datum	Versie	Commentaar	Auteur
29 juli '13	0.1	Eerste versie	Robin Roestenburg
20 augustus '13	0.2	Review door Gerrit Jan de Bie	Marco Plaisier
22 augustus '13	0.3	Review door Gerrit Jan de Bie	Marco Plaisier
28 augustus '13	0.4	Hoofdstuk 7 toegevoegd	Marco Plaisier
30 augustus '13	1.0	Laatste review en definitief gemaakt	Marco Plaisier

Titel	Analyse voor project Vestigium
Status	Definitief
Project	Erasmus Behavioural Lab analyse
Datum	25 juli '13
Auteurs	Robin Roestenburg, Marco Plaisier
Document eigenaar	Gerrit Jan de Bie
Distributielijst	Gerrit Jan de Bie
Goedkeuring	Gerrit Jan de Bie

Contactpersoon

Naam	Robin Roestenburg
Functie	Senior Software Engineer
Kantooradres	Stationsplein 45, A4.205 3013 AK Rotterdam
Telefoonnummer	06-526 382 46
E-mail adres	robin@finalist.com

Naam	Marco Plaisier
Functie	Informatie analist
Kantooradres	Stationsplein 45, A4.205 3013 AK Rotterdam
Telefoonnummer	06-196 00 497
E-mail adres	marco@finalist.com

Inhoudsopgave

1.	Inleiding	5
1.1.	Samenvatting.....	5
1.2.	Het EBL	5
1.3.	Probleemomschrijving	5
1.4.	Doelgroep.....	5
1.5.	Afbakening.....	5
1.6.	Leeswijzer.....	5
1.7.	Aannames en afhankelijkheden	6
2.	Visie	7
2.1.	Stakeholders	7
3.	Concepten en begrippenlijst	8
4.	Inventarisatie huidige situatie	11
4.1.	Registreren onderzoek.....	11
4.2.	Uitvoeren onderzoeken	11
4.3.	Verwerken data en publiceren onderzoek	16
5.	Requirements nieuwe situatie	17
5.1.	Doelstellingen.....	17
5.2.	Nieuwe situatie	17
5.3.	Scope.....	31
5.4.	Non-functionele requirements.....	32
6.	Voorstel architectuur nieuwe situatie	34
6.1.	Vestigium	34
6.2.	Data Copy Agent.....	34
6.3.	Dell DX Storage.....	37
7.	Integriteitsrapportage.....	38
7.1.	Onderzoeker	38
7.2.	Proefpersoon	41
7.3.	Beheerder	42

1. Inleiding

1.1. Samenvatting

Als een onderzoeker de integriteit van zijn onderzoek wil aantonen, dan is dat op dit moment zeer lastig. De benodigde informatie is aanwezig, maar niet of slechts deels vastgelegd of verspreid over verschillende bronnen. Voor het EBL is een analyse uitgevoerd om de voorwaarden van een nieuw systeem te bepalen die het aantonen van de integriteit vereenvoudigt en completer maakt.

Een onderzoek binnen het EBL volgt 3 fases: het registreren van een onderzoek ten behoeve van werving van proefpersonen, het uitvoeren van het onderzoek en het verwerken van de data. Daarbij bieden beheerders ondersteuning aan onderzoekers. Proefpersonen nemen deel aan het onderzoek en worden hiervoor gecompenseerd. Dit proces wordt ondersteund door een half dozijn losse systemen.

In de nieuwe situatie wordt de registratie van onderzoek een verplicht onderdeel van het proces. Hierdoor wordt het mogelijk om alle registreerbare gebeurtenissen in het lab te koppelen aan een van de onderzoeken. Met behulp van de integriteitsrapportage, die de gebeurtenissen in het lab inzichtelijk maakt, kan de onderzoeker aantonen dat zijn onderzoek integer is uitgevoerd. Hiervoor moeten de huidige losstaande systemen worden geïntegreerd in één centraal systeem. Deze integratie biedt verder mogelijkheden om het beheer van het lab effectiever en efficiënter uit te voeren.

1.2. Het EBL

Het Erasmus Behavioural Lab is bedoeld voor onderzoekers van de Institute of Psychology (IOP) en Erasmus Research Institute of Management (ERIM). Er worden onderzoeken uitgevoerd op het gebied van psychologie, bedrijfskunde en economie. Hiervoor heeft het EBL beschikking over verschillende labs, zoals cubicles, video, EEG en eyetracking labs.

1.3. Probleemomschrijving

In het recente verleden is bij verschillende wetenschappelijke onderzoeken twijfel ontstaan over de aan het onderzoek ten grondslag liggende data. Dit heeft uitgebreid aandacht gehad van de media. Om te garanderen dat de onderzoekdata niet ten behoeve van gewenste uitkomsten gemanipuleerd wordt, dient er een systeem ontwikkeld te worden voor de registratie, opslag en beveiliging van deze data. Dit systeem moet de gehele workflow dekken vanaf de registratie van het onderzoek tot het genereren van *compliancy* rapporten.

1.4. Doelgroep

De doelgroep van dit document zijn alle betrokkenen bij het RDM Pilot EBL, genaamd Vestigium.

1.5. Afbakening

Dit document beschrijft op een hoog niveau de voorwaarden voor een oplossing van het probleem. Het bevat een omschrijving van de huidige situatie en de requirements voor de nieuwe situatie, inclusief een voorstel voor de nieuwe architectuur.

Dit document is geen functionele of technische specificatie. Deze specificaties volgen in de realisatie van het project.

1.6. Leeswijzer

Dit document bestaat uit 6 hoofdstukken. Hoofdstuk 1 is de inleiding. In hoofdstuk 2 wordt de visie op het nieuwe systeem geformuleerd, gecombineerd met een beschrijving van de stakeholders. Hoofdstuk 3 bevat een overzicht van alle relevante concepten en begrippen die gebruikt worden in dit document, inclusief de definitie. Hoofdstuk 4 beschrijft de huidige processen vanuit het perspectief van een onderzoek. Hoofdstuk 5 bevat een beschrijving van de nieuwe situatie, bestaand uit de doelstellingen, taak analyse, scope en niet-functionele requirements. Hoofdstuk 6 geeft een voorstel voor de architectuur van het nieuwe systeem. Dit hoofdstuk is technisch georiënteerd. Hoofdstuk 7 omvat de onderbouwing voor de integriteitsrapportage.

1.7. Aannames en afhankelijkheden

Bij het schrijven van dit document zijn de volgende aannames gedaan:

- Koppeling met een Research Information System of andere werkomgeving voor onderzoekers is geen onderdeel van dit document. We nemen aan dat data wordt opgeslagen in de Dell DX Storage en niet gekopieerd hoeft te worden naar een RIS.
- De infrastructuur voor de productieomgeving van het nieuwe systeem wordt gehost bij het Shared Service Centre ICT.
- Alle onderzoekdata wordt van en naar computers gestuurd die uitgerust zijn met Windows XP, Windows 7 of Windows 8.
- De sleutelkast waarmee gekoppeld moet worden blijft ongewijzigd.
- De andere systemen worden herbouwd.

2. Visie

De visie schetst een toekomstbeeld van het nieuwe systeem. Voor dit project is de volgende visie geformuleerd:

Vestigium biedt onderzoekers de middelen om de integriteit van hun wetenschappelijk onderzoek te onderbouwen door vastlegging van de oorsprong van de onderzoekdata en middelen om de faciliteiten van het EBL beter te benutten.

2.1. Stakeholders

Stakeholders zijn die personen, groepen of instanties die actief betrokken zijn bij het project, te maken zullen krijgen met het de gevolgen van het project of het project kunnen beïnvloeden. Er wordt onderscheid gemaakt tussen primaire stakeholders en secundaire stakeholders. Primaire stakeholders gebruiken het systeem direct. Secundaire stakeholders hebben er op een of andere manier invloed op het systeem of worden er door beïnvloed.

Voor dit project zijn de volgende primaire stakeholders geïdentificeerd:

- Onderzoek eigenaren,
Dit is de persoon die opdracht heeft gegeven tot het onderzoek. Dit hoeft niet de onderzoeker te zijn die het daadwerkelijk in het lab uitvoert. Vaak wordt het onderzoek bedacht door een professor en uitgevoerd door een PhD of student-assistent.
- Onderzoekers,
De persoon die het onderzoek uitvoert in het lab.
- Proefpersonen,
De persoon die het onderzoek ondergaat.
- Lab beheerders.
De beheerders van het lab.

De volgende secundaire stakeholders zijn geïdentificeerd:

- College van Bestuur van de Erasmus Universiteit,
- opleiding Psychologie van FSW,
- Bedrijfskunde en Economie (ERIM),
- Commissie Wetenschappelijke Integriteit (CWI),
- Medisch-Ethische Toetsings Commissie (METC),
- Ethische Commissie,
- American Psychology Association (APA).

3. Concepten en begrippenlijst

Term	Omschrijving
Beheerder	De beheerders van het lab. Verantwoordelijk voor de apparatuur, technische begeleiding van onderzoek, toegang, etc.
Boekingssysteem	In het Boekingssysteem worden de ruimtes gereserveerd en kunnen ze ingezien worden. Een gebruiker heeft standaard de mogelijkheid om het Boekingssysteem in te zien en kan de autorisatie om ruimtes te reserveren aanvragen bij de beheerders van het lab. Het systeem is gebouwd in PHP en slaat zijn data op in een MySQL database.
CWI	Commissie voor Wetenschappelijke Integriteit Deze commissie waakt over de integriteit van onderzoeken uitgevoerd binnen, of in naam van, de Erasmus Universiteit. Integriteit heeft in dit kader te maken met de wetenschappelijke eerlijkheid van de onderzoekers.
Data	Term die gebruikt wordt voor het bestand dat stimuli genereert tijdens het onderzoek (de operationalisatie) of voor de onderzoekresultaten die in een bestand worden opgeslagen.
EBL Assistant	De EBL Assistant is een applicatie die op lab, support room en staf computers draait. De applicatie wordt voor meerdere doeleinden gebruikt. De belangrijkste zijn: <ul style="list-style-type: none">• installeren van software,• kopiëren data naar het lab vanuit staf netwerk,• kopiëren data van het lab naar het staf netwerk. De EBL Assistant is geschreven in VB.net.
Eigenaar	Hiermee wordt de eigenaar van het onderzoek bedoeld. De eigenaar heeft het onderzoek geregistreerd en is eindverantwoordelijk. Een eigenaar kan ook een onderzoeker zijn, maar dat hoeft niet.
Erasmus Behavioural Lab	Het psychologisch en gedragskundige lab van de Erasmus Universiteit. Hier worden onderzoeken gedaan naar psychologische aspecten van menselijk gedrag.
ERNA	Het unieke studentnummer of stafnummer van een persoon. Aangezien er ook externen proefpersoon kunnen zijn, heeft niet iedereen een ERNA. Alle onderzoekers krijgen een ERNA als ze op de universiteit komen werken.
ERPS	Proefpersoon wervingssysteem. Dit systeem is een instantie van het door SonaSystems geleverde 'web-based human subject pool management software'.
EURO	Proefpersoon wervingssysteem. Dit systeem is een instantie van het door SonaSystems geleverde 'web-based human subject pool management software'.
Experiment	Deze term niet gebruiken i.v.m. negatieve connotaties. Zie 'Onderzoek'
Janno	Server in het EBL die zowel in het netwerk van de universiteit als in het netwerk van het EBL staat. Janno wordt ook wel de buffer genoemd. Data die een onderzoeker naar het lab toe kopieert, zet hij impliciet eerst op Janno en vervolgens op de computers van het lab. Bij het kopiëren van data vanuit het lab naar een staf computers verloopt dit in omgekeerde richting.

Lab computer	Een computer in het netwerk van het EBL. Dit netwerk is op VLAN-niveau gescheiden van het netwerk van de universiteit. De computers van het EBL kennen maar twee type gebruikers: <ul style="list-style-type: none"> • <i>ebluser</i> (voor alle gebruikers), • <i>ebladmin</i> (gebruikt door beheerders) Alle gebruikers die onder <i>ebluser</i> de lab computer gebruiken kunnen elkaar data inzien.
LDAP	Lightweight Directory Access Protocol. Een directory is in dit verband informatie die op een hiërarchische manier, gegroepeerd naar een bepaald attribuut, is opgeslagen. Denk aan een telefoonboek waarin telefoonnummers en adressen van personen per bedrijf worden opgeslagen. Een directorynaam komt overeen met de bedrijfsnaam. Iedere directory bevat dan alle personen binnen dat bedrijf als objecten, met contactgegevens zoals telefoonnummer en e-mailadres als attributen. Authenticatie in decentraal ontwikkelde en beheerde systeem vindt plaats tegen de centrale LDAP server die is gekoppeld aan de administratie van HRF. De lokale systemen bepalen de autorisatie van een bepaalde gebruiker, in de LDAP liggen geen autorisatie gegevens vast.
METC	Medisch-Ethische Toetsing Commissie. Deze commissie toetst de noodzaak of onderzoek beoordeeld moet worden door de Ethische Commissie. Als het onderzoek aan bepaalde criteria voldoet, dan zal de METC adviseren dat het onderzoek door de Ethische Commissie beoordeeld wordt.
OASE	Proefpersoon wervingssysteem. Dit systeem is gebouwd in PHP en slaat zijn data op in een MySQL database.
Onderzoek	Met onderzoek wordt een wetenschappelijk onderzoek binnen de EBL bedoeld.
Onderzoeker	De onderzoeker voert het onderzoek uit in het lab.
Proefleider	Zie Onderzoeker.
Proefpersoon	Iemand die een onderzoek ondergaat, bijvoorbeeld een eye-tracking onderzoek of een EEG. Proefpersonen zijn over het algemeen studenten van de Erasmus Universiteit, maar ze kunnen ook extern worden geworven.
Salvia	Server in het EBL die o.a. gebruikt wordt om de nachtelijke synchronisatie van alle lab computers naar de buffer (zie Janno) uit te voeren.
Sleutelkast	Deze kast hangt bij de ingang van het lab. Onderzoekers kunnen hier met een pincode sleutels ophalen voor specifieke ruimtes in het lab.
Sleutel cabinet	Onderzoekers moeten een geautoriseerd zijn, voordat ze een pincode krijgen en sleutels kunnen ophalen. Het volledige systeem is geleverd door Deister. Het systeem heeft een op SOAP-gebaseerde service waarmee andere applicaties met het systeem kunnen werken. Ook is er een web applicatie waarmee een beheerder handmatig wijzigingen kan doorvoeren in de gegevens van het Sleutelkastsysteem (bijv. toevoegen van gebruikers die niet in het Boekingssysteem vastliggen). Zowel de SOAP service als de web applicatie zijn gebouwd in Java. Het systeem slaat zijn data op in een SQLServer database.

Slot	Een tijdvak waarop een proefpersoon zichzelf kan inschrijven om deel te nemen aan een onderzoek.
Staf computer	Een computer in een van de VLAN's van de universiteit.
Support room	Ruimte in het EBL waar de onderzoeker zijn onderzoek kan voorbereiden en zijn resultaten kan verwerken.
Support room computer	Een computer in een support room van het EBL. Deze computers staan in het netwerk van de universiteit, net als de staf computers.

4. Inventarisatie huidige situatie

We kunnen de huidige situatie het beste beschrijven door het proces te volgen dat een onderzoek doorloopt in het Erasmus Behavioural Lab (EBL).

Om een onderzoek uit te voeren in het EBL moet een onderzoeker een aantal stappen doorlopen:

- Registreren van een onderzoek ten behoeve van het werven van proefpersonen,
- Uitvoeren van het onderzoek,
- Verwerken data en publiceren onderzoek.

De beschrijving in volgende paragrafen heeft als doel om alleen die systemen en functies te identificeren die invloed hebben op eventuele aanpassingen in de toekomstige situatie.

4.1. Registreren onderzoek

Op een gegeven moment is de onderzoeker klaar met de voorbereiding en wil hij zijn onderzoek uitvoeren in het EBL. Als hij proefpersonen nodig heeft in het onderzoek, dan zal hij zijn onderzoek registreren in een van de proefpersoonwervingsystemen. Dit is de enige plaats waar het onderzoek nu wordt geregistreerd.

Als de onderzoeker geen proefpersonen wil werven via een van de wervingssystemen, dan zal het onderzoek niet worden geregistreerd. Het onderzoek is dan alleen bekend bij de betrokken proefleiders.

4.2. Uitvoeren onderzoeken

Alvorens de onderzoeker onderzoeken kan uitvoeren moet hij:

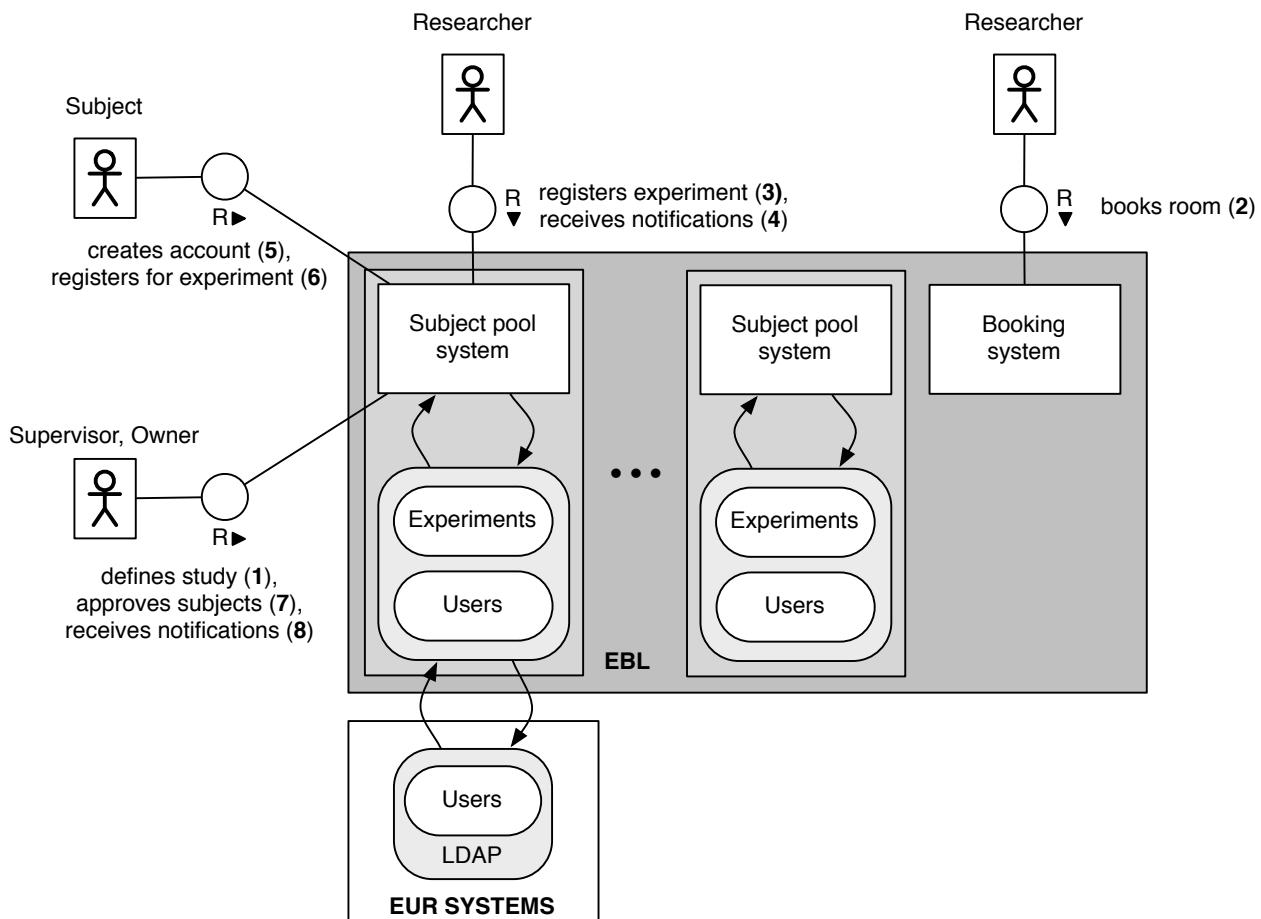
- proefpersonen werven voor de onderzoeken, en
- ruimtes in het EBL boeken waar de onderzoeker de onderzoeken wil uitvoeren, en
- het operationalisatie-bestand plaatsen op de lab-computers.

Deze drie acties kunnen in willekeurige volgorde plaatsvinden en het hangt van het type onderzoek af hoe de onderzoeker dit daadwerkelijk uitvoert. Vervolgens kan de onderzoeker het onderzoek daadwerkelijk uitvoeren. In de volgende paragrafen zullen we de huidige situatie voor deze stappen in detail bespreken.

4.2.1. Werven proefpersonen

Een onderzoeker kan zijn onderzoek registreren in Proefpersoon-wervingssystemen. Proefpersonen kunnen zich via deze systemen inschrijven om mee te doen aan een onderzoek.

Onderstaand figuur laat op hoog niveau de systemen en gebruikers zien die betrokken zijn bij het werven van proefpersonen. We zullen dit proces verder bespreken en verwijzen naar de verschillende nummers in onderstaand figuur.



Afbeelding 1 Werven van proefpersonen

De onderzoeker kan op een aantal manieren proefpersonen werven voor zijn onderzoek, de volgende drie zijn de meest voorkomende:

1. Individuele afspraken maken met proefpersonen,
2. Handmatig rekruteren van proefpersonen,
3. Aanbieden van een aantal tijdvakken (slots) waarop door proefpersonen in een Proefpersoon-wervingssysteem ingeschreven kan worden.

Als eerste moet de eigenaar van het onderzoek het registreren in het wervingssysteem (1). De onderzoeker moet bij elk van deze manieren er voor zorgen dat hij tijdig een kamer in het lab geboekt heeft waar het onderzoek in plaats kan vinden (2). Zie paragraaf 4.2.2 voor meer informatie over het boeken van een kamer in het lab.

Bij manieren 1 en 3 zal de onderzoeker zijn onderzoek moeten registreren in één van de Proefpersoon-wervingsystemen (3):

- OASE voor Psychologie,
- ERPS voor RSM of,
- EURO pool voor onderzoeken waarbij de proefpersoon betaald worden.

Het handmatig rekruteren van proefpersonen, manier 2, loopt buiten de Proefpersoon-wervingssystemen om.

Het verschil tussen deze systemen is de compensatie die een proefpersoon krijgt voor het meedoen aan een onderzoek.

De Proefpersoon-wervingssystemen, met uitzondering van OASE, zijn gekoppeld aan de centrale LDAP van de EUR. De Proefpersoon-wervingssystemen hebben voor authenticatie toegang tot alle gebruikers in de centrale LDAP.

Omdat het Proefpersoon-wervingssysteem voor PSY in de leeromgeving PsyWeb draait heeft dit systeem naast de gebruikersgegevens t.b.v. authenticatie ook toegang tot meta-data als volledige namen. OASE heeft eigen authenticatie en autorisatie.

Nadat het onderzoek is geregistreerd in een Proefpersoon-wervingssysteem zal het systeem op verschillende momenten via e-mail notificaties naar de onderzoeker sturen (4). De Proefpersoon-wervingssystemen sturen echter geen notificatie naar de onderzoeker op het moment dat er voldoende proefpersonen voor zijn onderzoek zijn, de onderzoeker moet dit zelf controleren.

Proefpersonen kunnen een gebruikersaccount aanmaken in het Proefpersoon-wervingssysteem (5). Zij hoeven dus niet bekend te zijn in de centrale LDAP database van de EUR.

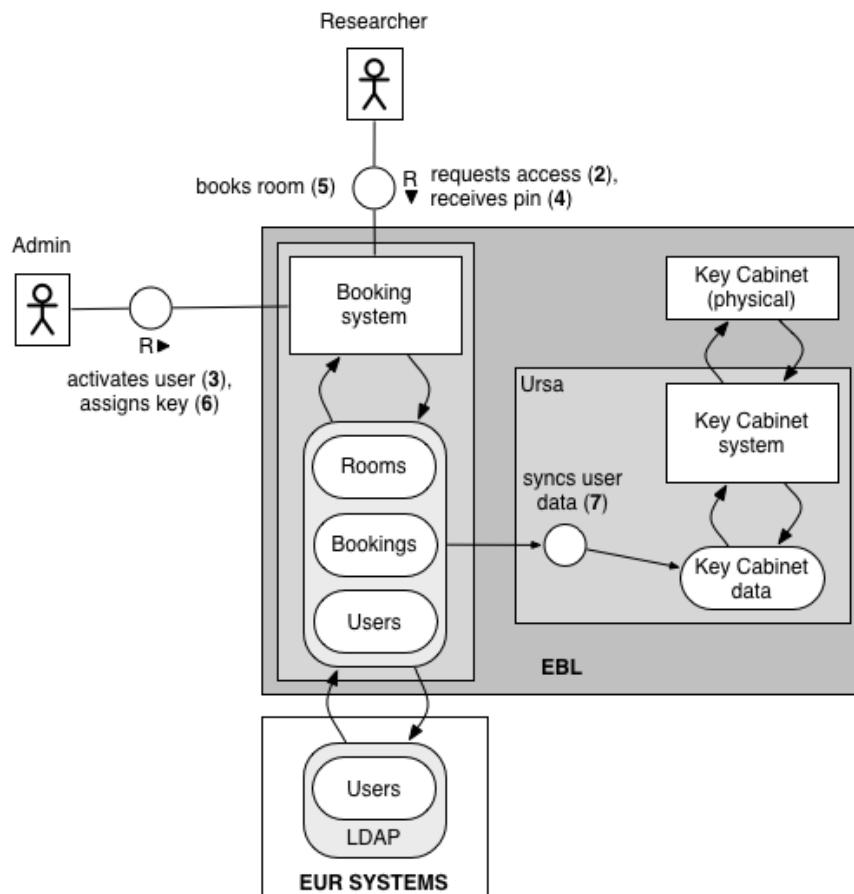
Een proefpersoon kan zich vervolgens in het Proefpersoon-wervingssysteem aanmelden voor een onderzoek (6). De proefleider of de eigenaar moet elke proefpersoon die zich anmeldt voor het onderzoek goedkeuren (7).

Na verloop van tijd zijn er voldoende proefpersonen geworven om het onderzoek uit te voeren. Vaak zal een onderzoeker al starten met het onderzoek voordat alle proefpersonen geworven zijn.

4.2.2. Boeken kamers in het lab

In het Boekingssysteem kan een onderzoeker kamers van het lab boeken. Een geboekte kamer kan de onderzoeker gebruiken om een onderzoek met proefpersonen in uit te voeren.

Onderstaand figuur laat op hoog niveau de systemen en gebruikers zien die betrokken zijn bij het boeken van een kamer. We zullen dit proces verder bespreken en verwijzen naar de verschillende nummers in onderstaand figuur.



Afbeelding 2 Boeken kamers in het lab

Het Boekingssysteem is gekoppeld aan de centrale LDAP van de EUR (1). Het Boekingssysteem heeft voor authenticatie toegang tot alle gebruikers in de centrale LDAP. Omdat het Boekingssysteem voor PSY toegang heeft tot de leeromgeving PsyWeb, heeft dit systeem naast de gebruikersgegevens t.b.v. authenticatie ook toegang tot meta-data als volledige namen.

Een onderzoeker kan standaard het Boekingssysteem alleen inzien. Voor het boeken van ruimtes dient de onderzoeker zijn account te laten activeren door een beheerder van het EBL (2). Dit kan de onderzoeker op de volgende manieren doen:

- invullen van registratie formulier (dit is een onderdeel van het Boekingssysteem), of
- (Indien de onderzoeker tot de faculteit Psychologie behoort) aanvragen van gebruikersaccount vanuit de leeromgeving van Psychologie,
- sturen van e-mail naar de beheerder.

De beheerder van het EBL krijgt de aanvragen voor activatie van een account via e-mail binnen. Er zijn nu twee mogelijkheden:

- De beheerder keurt de aanvraag af. De onderzoeker krijgt geen toegang tot het Boekingssysteem en kan geen kamers reserveren. Dit is een uitzonderlijke situatie en zal in praktijk niet vaak voorkomen.
- De beheerder keert de aanvraag goed (3). Het Boekingssysteem stuurt de gebruiker een e-mail met daarin de bevestiging dat hij toegang heeft tot het Boekingssysteem. In deze e-mail staat ook de pincode waarmee hij toegang kan krijgen tot het Sleutelkastsysteem (4).

De onderzoeker kan nu in het Boekingssysteem een kamer boeken voor zijn onderzoek (5).

Naast de geboekte kamer (de boeking) heeft de onderzoeker ook een sleutel voor de kamer nodig. Hij krijgt deze sleutel niet automatisch bij de boeking toegewezen, maar hij moet deze via zijn proefleider aanvragen. De proefleider vraagt mondeling, via e-mail of via de EBL Assistant de sleutel voor de boeking aan bij een beheerder van het EBL. De beheerder kent in het Boekingssysteem één of meerdere sleutels toe aan het account van de onderzoeker. (6).

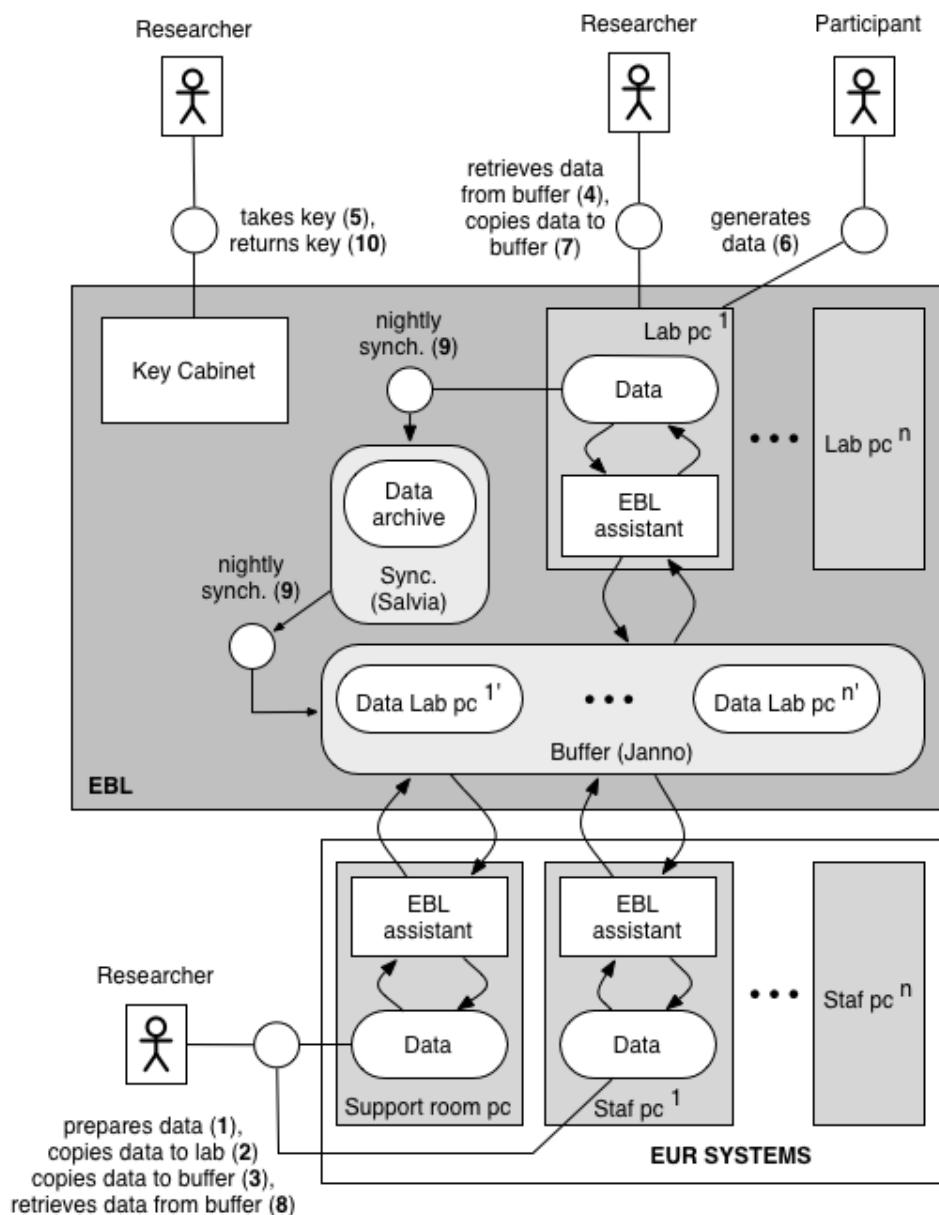
Stafleden (o.a. proefleiders en profeigenaren) hebben met hun pincode overigens standaard toegang tot alle sleutels en hoeven dus geen sleutels aan te vragen.

Het Boekingssysteem synchroniseert met het Sleutelkastsysteem, zodat alle gebruikers (inclusief bijbehorende pin codes & sleutels) in dit systeem beschikbaar zijn (7).

4.2.3. Uitvoeren onderzoek

De onderzoeker voert het onderzoek uit in het lab. De computers in het lab bevinden zich in een afgeschermde omgeving. De onderzoeker moet daarom eventueel benodigde data voor de start van onderzoek naar het lab toe kopiëren. Tijdens het uitvoeren van het onderzoek ontstaat er data die de onderzoeker uit het lab ophaalt en (eventueel) verwerkt in zijn publicatie.

Onderstaand figuur laat op hoog niveau de systemen en gebruikers zien die betrokken zijn bij het uitvoeren van een onderzoek. We zullen dit proces verder bespreken en verwijzen naar de verschillende nummers in onderstaand figuur.



Afbeelding 3 Uitvoeren onderzoek

Om een onderzoek uit te voeren heeft een onderzoeker data voorbereid. Voordat hij een onderzoek gaat uitvoeren kopieert de onderzoeker zijn data naar de computers in de kamer van het lab waar het onderzoek zal plaatsvinden. Deze kamers heeft hij eerder geboekt in het Boekingsysteem (zie paragraaf 4.2.2).

De onderzoeker heeft twee mogelijkheden voor het kopiëren van zijn data. Hij kan:

- de data eerst op de computers in support rooms voorbereiden (1) en vandaar kopiëren naar de computers in de kamer van het lab (2), of
- de data voor te bereiden op zijn staf computer (1) en vandaar te kopiëren naar de computers in de kamer van het lab (2)

In beide gevallen voert de onderzoeker deze kopieer actie uit met behulp van de EBL Assistant. De Assistant kopieert eerst de data naar een buffer (3) (de Janno server) en plaatst de data dan naar de geselecteerde computers (4).

Om toegang te krijgen tot het lab haalt de onderzoeker de sleutel van de door hem gereserveerde kamer in het lab uit de sleutelkast (5). De onderzoeker moet dan de eerder naar hem verstuurde pincode invoeren.

Vervolgens wacht de onderzoeker tot de proefpersonen aanwezig zijn en gaat hij met hun naar het lab. De proefpersonen voeren het onderzoek uit en genereren daarmee data (6).

Om de gegenereerde data van het onderzoek veilig te stellen heeft de onderzoeker wederom twee mogelijkheden. De onderzoeker kan:

- van elke computer die in het onderzoek gebruikt is de data door middel van de EBL Assistant kopiëren naar de buffer (7) en vervolgens de data vanaf de buffer naar een computer in de support rooms of een computer in het staf netwerk te kopiëren (8), of
- een dag wachten en kopieert vervolgens de data met de EBL Assistant vanaf de buffer naar een computer in de support rooms of een computer in het staf netwerk (8).

Het tweede punt behoeft enige uitleg. Een achtergrondproces synchroniseert elke nacht de data van alle computers in het EBL naar de buffer (9). Wanneer de onderzoeker een dag later zijn data wil kopiëren, kopieert hij het in feite vanaf de buffer en niet van de computers in het lab.

Het achtergrondproces draait op de zgn. Salvia server. Het proces bepaalt per te synchroniseren computer de delta van:

- de data die op dit moment op de computer staat, en
- de data zoals die de vorige nacht op de computer stond.

Het achtergrondproces slaat de delta op waarmee, op een later tijdstip, de beheerders van het lab tot op zekere hoogte kunnen bepalen welke data er op een zeker moment op een computer in het lab heeft gestaan.

Om de proefpersonen te compenseren voor hun deelname aan het onderzoek (bijv. door middel van een verhoging van hun cijfer) moet de proefleider in het Proefpersoon-wervingssysteem accorderen dat een proefpersoon daadwerkelijk aan het onderzoek heeft meegedaan.

Indien de proefpersoon na een week nog niet goedgekeurd is stuurt het Proefpersoon-wervingssysteem een mail met daarin de uitstaande accordering(en).

Aan het einde van de sessie of dag hangt de onderzoeker de sleutel terug in de sleutelkast (10). De onderzoeker krijgt een e-mail indien hij dit vergeet.

4.3. Verwerken data en publiceren onderzoek

De onderzoeker verwerkt de data van zijn onderzoeken in zijn onderzoek met als uiteindelijke doel zijn onderzoek te publiceren. In dit vervolgproces maakt de onderzoeker in principe geen gebruik van het EBL en zijn systemen. Eventueel gebruikt hij de computers in de support rooms om zijn data te verwerken.

5. Requirements nieuwe situatie

Requirements beschrijven wat het systeem moet doen om aan de wensen en eisen van de stakeholders te voldoen. Requirements worden opgesplitst in functionele requirements die beschrijven wat het systeem moet doen en niet-functionele requirements, die beschrijven hoe het systeem moet werken.

De functionele requirements van de nieuwe situatie zijn gebaseerd op de workshops, interviews met betrokkenen en beschikbare documentatie. Uit deze informatie zijn eerst vijf doelstellingen geformuleerd voor het nieuwe systeem. Op basis van de input in combinatie met de doelstellingen is een taakanalyse uitgevoerd. Deze taakanalyse beschrijft de functionaliteit die het nieuwe systeem moet bieden. Hierbij is er voor gekozen om de taakanalyse te maken vanuit het perspectief van de volgende gebruikersgroepen: de onderzoeker, de proefpersoon en de beheerder. Daarnaast worden de rapportages kort toegelicht. De integriteitsrapportac

Voor iedere gebruikersgroep is de taakanalyses opgesplitst in features. Deze features zijn geprioriteerd, zodat duidelijk is wat de meest waardevolle of belangrijke elementen van het systeem zijn.

Als laatste zijn de non-functionele requirements uitgewerkt. Deze omvatten onderdelen zoals privacy, beveiliging, performance en gebruiksvriendelijkheid.

5.1. Doelstellingen

5.1.1. Doelstellingen Vestigium

Een nieuw systeem moet aan de volgende doelstellingen voldoen om succesvol te zijn:

1. Alle beschikbare aspecten van onderzoek dat uitgevoerd wordt in het EBL worden vastgelegd, zodat er een rapportage gemaakt kan worden waarmee onderbouwd kan worden dat het onderzoek integer is uitgevoerd,
2. Rapportages ten bate van het CWI en aangesloten faculteiten kunnen met minimale inspanning worden gegenereerd,
3. Onderzoekers kunnen altijd en overal bij hun onderzoekdata,
4. De huidige workflow van onderzoeken wordt gestroomlijnd, zodat nieuwe onderzoekers binnen maximaal 4 uur werktijd hun onderzoek kunnen invoeren, inplannen, de juiste ruimtes kunnen reserveren en open kunnen stellen voor proefpersonen, mits de onderzoeker al de juiste autorisaties heeft,
5. De benutting en planning van de EBL-faciliteiten en gegeven ondersteuning door beheerders bij het gebruik van het EBL wordt met minimale inspanning inzichtelijk.

5.1.2. Grenzen

Naast bovenstaande doelstellingen, zijn er doelstellingen die voor de hand liggen om mee te nemen in (het ontwerp van) het nieuwe systeem. Echter, er is voor gekozen om deze explicet buiten beschouwing te laten.

1. Vastleggen van onderzoeken buiten het EBL,
Onderzoekers gebruiken bijvoorbeeld vragenlijsten die aan proefpersonen thuis worden gestuurd of voeren hun onderzoek deels uit in een andere omgeving dan het EBL, bijvoorbeeld in het Erasmus MC. Deze data en de ontstaansgeschiedenis daarvan worden niet vastgelegd in het nieuwe systeem.
2. Het systeem zal geen mogelijkheden bieden aan onderzoekers, beheerders of anderen om binnen het systeem de validiteit van de gebruikte onderzoeksopzet of –methode te beoordelen,
3. Idem voor het analyseren van onderzoekdata,
4. Onderzoekers de mogelijkheid geven om direct, of met behulp van beheerders, (groepen) proefpersonen te selecteren en uit te nodigen voor een onderzoek,
5. Buiten de preselectie van proefpersonen extra hulpmiddelen aanbieden voor het bepalen van geschiktheid van de proefpersonen voor het onderzoek.

5.2. Nieuwe situatie

De nieuwe situatie komt in grote lijnen overeen met de huidige situatie. Het belangrijkste verschil is dat de onderzoeker verplicht wordt zijn onderzoek te registreren. Zonder deze stap kan niet verder in het proces. Deze stap is verplicht, omdat alle gebeurtenissen die in het kader van het onderzoek plaatsvinden te herleiden moeten zijn naar een geregistreerd onderzoek.

Voor onderzoekers en proefpersonen zal er buiten de registratie weinig veranderen. Na registratie kan het onderzoek worden getoetst. Na de registratie en optionele toetsing volgt de wervingsfase van proefpersonen, hoewel deze fase in de praktijk deels parallel zal verlopen aan het uitvoeren van het onderzoek. Tijdens de uitvoerfase voert de onderzoeker het onderzoek uit samen met de proefpersonen. De resultaten worden vastgelegd en blijven beschikbaar voor de onderzoeker. Nadat alle proefpersonen zijn geweest, wordt het onderzoek afgesloten.

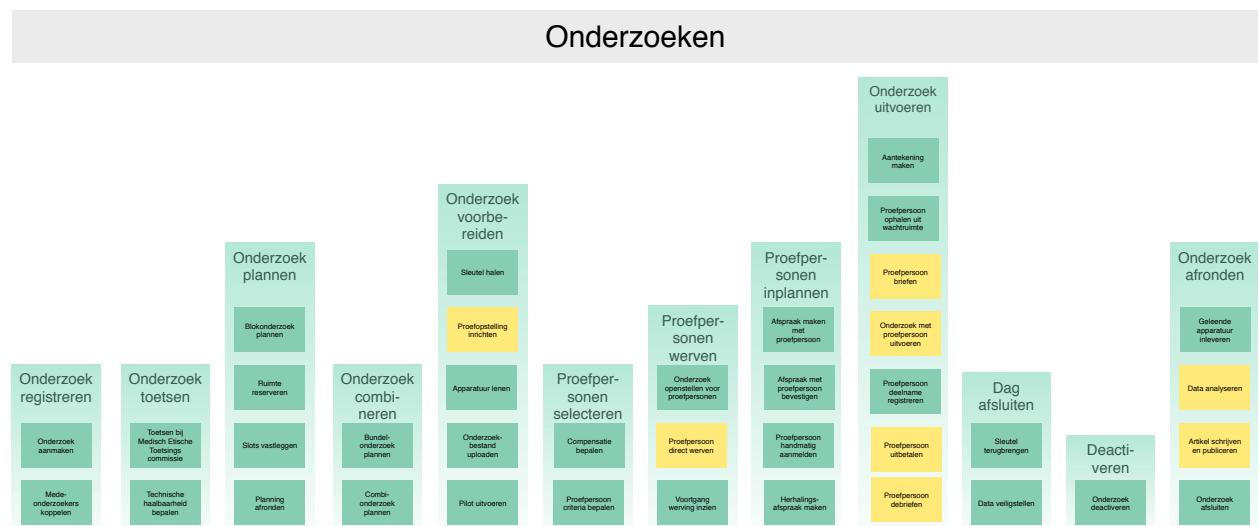
Het belangrijkste verschil is dat alle betrokkenen met één systeem te maken krijgen in plaats van met een aantal verschillende. Hierdoor wordt de informatie gecentraliseerd en wordt het makkelijker om een onderbouwing te geven van de integriteit van het onderzoek en om inzicht te krijgen in de benutting en planning van de lab-faciliteiten.

Het systeem is opgedeeld 4 deelgebieden. De drie belangrijkste deelgebieden zijn de functies van het systeem voor de onderzoeker, de proefpersoon en de beheerder. Het laatste deelgebied omvat de rapportages.

Voor ieder deelgebied is een diagram gemaakt. Ieder diagram bestaat uit torens die een functioneel gebied afbakenen, zoals *onderzoek registreren* of *proefpersonen werven*. Iedere toren is opgebouwd uit de high-level functionele requirements. Zo bestaat de toren *onderzoek registreren* uit de features *Onderzoek aanmaken* en *Mede-onderzoekers koppelen*. Deze blokken worden in de sub-paragrafen toegelicht. De gele blokken in het diagram zijn acties of werkzaamheden waar het systeem niet direct voor gebruikt zal worden, maar wel als informatiebron kan dienen.

5.2.1. De onderzoeker

Vanuit het perspectief van de onderzoeker zijn er elf functionele gebieden geïdentificeerd. De onderzoeker start bij het registreren van zijn onderzoek en eindigt uiteindelijk bij het afronden van het onderzoek. In de praktijk zal de volgorde onregelmatig zijn en zullen taken parallel uitgevoerd worden. Sommige gebieden en taken zijn ook optioneel, afhankelijk van de aard van het onderzoek, de manier van werving en de deelname van proefpersonen.



5.2.1.1. Onderzoek registreren



Doel: het onderzoek registreren, zodat het bekend wordt in het systeem.

De onderzoeker start met het registreren van het onderzoek in het systeem. Dit is een verplichte stap, want zonder zal de onderzoeker geen onderzoek kunnen uitvoeren in het lab. Ook wordt het onmogelijk om de integriteit van het onderzoek via het systeem te onderbouwen, omdat andere gebeurtenissen niet herleid kunnen worden naar het onderzoek.

Het registreren bestaat uit het vastleggen van de karakteristieken van het onderzoek die nodig zijn, zoals de omschrijving, naam, verwachte looptijd, aantal proefpersonen, wel/geen pilot, etc.

Daarnaast kan de onderzoeker die het onderzoek registreert ook andere onderzoekers koppelen als het onderzoek door meerdere personen wordt uitgevoerd. Als er mede-onderzoekers

gekoppeld worden, dan krijgen deze ook rechten om voor dat onderzoek bijvoorbeeld ruimtes te reserveren, sleutels op te halen en proefpersonen te selecteren. De precieze rechtenstructuur is nog niet uitgewerkt.

5.2.1.2. Onderzoek toetsen

Onderzoek toetsen

Toetsen bij Medisch Ethische Toetsings commissie

Technische haalbaarheid bepalen

Doel: laten controleren of het onderzoek voldoet aan ethische en technische voorwaarden, voordat het uitgevoerd wordt.

Sommige onderzoeken zullen eerst getoetst worden door de METC. Op basis van de methode van het onderzoek kan de commissie bepalen of toetsing door de Ethische Commissie wellicht nodig is en kunnen de vervolgstappen gezet worden. De beslissing van de METC en Ethische Commissie moet worden vastgelegd bij het onderzoek.

Daarnaast moet voorkomen worden dat onervaren onderzoekers zonder begeleiding in het lab werken. Op basis van de karakteristieken van het onderzoek en de gekoppelde onderzoekers bepaalt het systeem dat een instructie door de lab-beheerders gewenst is. Dit resulteert in een notificatie naar de beheerders en betrokken onderzoeker. De beheerder moet dit onderzoek eerst goedkeuren, voordat het verder in het proces mag.

5.2.1.3. Onderzoek plannen

Onderzoek plannen

Blokonderzoek plannen

Ruimte reserveren

Slots vastleggen

Planning afronden

Doel: het plannen en reserveren van faciliteiten.

Het onderzoek moet gepland worden. Het belangrijkste is dat de ruimte gereserveerd wordt, zodat er geen twee onderzoeken tegelijkertijd in dezelfde ruimte plaatsvinden. De onderzoeker moet kunnen zien welke labs hij kan reserveren op basis van de benodigde faciliteiten en de bestaande reserveringen.

Het komt in de huidige situatie voor dat bepaalde ruimtes zeer lang van te voren worden gereserveerd, zonder dat er een uitgewerkt onderzoek is. Het gebeurt dan regelmatig dat de ruimte toch niet wordt gebruikt. Andere onderzoekers worden dan gedupeerd. Het nieuwe systeem zal deze situatie al deels voorkomen, doordat het niet mogelijk is een ruimte te reserveren zonder een geregistreerd onderzoek.

Daarnaast kan het systeem dit probleem op andere manieren oplossen:

1. Door reserveringen niet te accepteren als die meer dan een bepaalde periode van te voren zijn gemaakt.
2. Door overboekingen toe te staan. Het onderzoek waar de eerste proefpersoon zich voor inschrijft, krijgt dan de reservering toegewezen.
3. Door arbitrage van een beheerder.

De precieze oplossing zal in overleg met de primaire stakeholders gekozen moeten worden.

Een andere situatie die regelmatig voorkomt is dat een onderzoeker een ruimte voor langere tijd reserveert voor afspraken. Als er binnen een periode geen afspraken worden gemaakt, dan wordt dat deel van de reservering tijdig weer vrijgegeven, zodat andere onderzoekers de ruimte kunnen gebruiken. Het nieuwe systeem moet deze situatie ook ondersteunen.

Een van de manieren waarop een proefpersoon zich kan inschrijven is via slots (tijdvakken). Als een onderzoek 30 minuten duurt en een ruimte met 4 computers is voor 2 uur gereserveerd, dan zijn er 16 slots voor proefpersonen beschikbaar. De onderzoeker moet kunnen aangeven welke slots er zijn.

Een blokonderzoek is een speciaal soort onderzoek waar Psychologie-minors als proefpersoon kunnen deelnemen. Er worden speciale dagen aangewezen waarop meerdere onderzoeken voor minors worden uitgevoerd. Zo kunnen minor-studenten binnen hun minor de verplichte deelname-uren behalen. Dit type onderzoek kan waarschijnlijk gebruik maken van de mogelijkheid om slots vast te leggen.

Als de planning van het onderzoek gereed is, dan kan de onderzoeker de planning afronden. Dit betekent dat het onderzoek zichtbaar wordt voor andere onderzoekers die het systeem gebruiken. Hierdoor wordt het mogelijk om bijvoorbeeld onderzoeken te combineren. Het onderzoek wordt nog niet zichtbaar in het wervingssysteem.

5.2.1.4. Onderzoeken combineren

Onderzoek combineren

Bundel-onderzoek plannen

Combi-onderzoek plannen

Doele: meerdere onderzoeken combineren, zodat proefpersonen voor een langere tijd in het lab zijn en een hogere compensatie krijgen.

Als een onderzoek kort duurt (minder dan 30 minuten) dan is het handig om twee of meer onderzoeken te combineren, zodat een proefpersoon zich op 1 onderzoek inschrijft en deeltneemt aan meerdere. Dit wordt een bundel-onderzoek genoemd. De reden hiervoor is dat proefpersonen niet graag deelnemen aan korte onderzoekjes, omdat ze dan relatief veel tijd kwijt zijn en weinig krijgen uitbetaald. Het systeem kan suggesties geven voor het bundelen van onderzoeken en maakt de synchrone planning van beide onderzoeken mogelijk. Dit betekent dat van de gebundelde onderzoeken één inschrijving zichtbaar wordt in het wervingssysteem.

Daarnaast zijn er onderzoeken die bestaan uit meerdere onderdelen in verschillende labs. Een voorbeeld is een onderzoek waar eerst een instructie en enquête ingevuld wordt achter een computer waarna er een EEG-onderzoek volgt. Hiervoor moet een combi-onderzoek ingepland worden.

5.2.1.5. Onderzoek voorbereiden

Onderzoek voorbereiden

Sleutel halen

Proefopstelling inrichten

Apparatuur lenen

Onderzoeksbestand uploaden

Pilot uitvoeren

Doele: de voorbereidingen treffen voor het onderzoek, zodat het onderzoek zelf snel en probleemloos uitgevoerd kan worden.

Voordat het onderzoek daadwerkelijk kan plaatsvinden, moet alles in gereedheid gebracht worden. Allereerst moet de onderzoeker de sleutel van het gereserveerde lab ophalen in de sleutelkast die bij de ingang van het lab hangt. Hiervoor heeft hij een pincode nodig en moet hij geautoriseerd zijn in het systeem.

Daarna kan de onderzoeker of de beheerder de proefopstelling inrichten. Het systeem zal hier ondersteunende documentatie voor aanbieden, maar verder zal de onderzoeker geen informatie in het systeem registreren.

Voor sommige onderzoeken is apparatuur nodig (zoals koptelefoon of reactieknoppen) die geleend kunnen worden bij de beheerders. De onderzoeker kan vooraf aangeven dat hij bepaalde apparatuur wil lenen. Het systeem stelt de beheerder hiervan op de hoogte.

De belangrijkste stap is het kopiëren van het bestand met de stimuli (operationalisatie-bestand) naar de onderzoeks-computer. De onderzoeker gebruikt hiervoor het systeem. Hij kan een bestand op een gedeelde schijf kiezen en opdracht geven om deze naar een of meer computers in het lab te kopiëren. Het systeem voert deze opdracht verder zelfstandig uit (op de achtergrond) en stuurt een notificatie naar de onderzoeker als de kopieer actie is voltooid.

De onderzoeker kan er voor kiezen om een pilot uit te voeren, vaak met ongeregistreerde proefpersonen, om te controleren of alle apparatuur werkt en om de laatste problemen op te lossen. De data die hieruit voortkomt, moet geregistreerd worden in het systeem met het kenmerk *pilot*.

5.2.1.6. Proefpersonen selecteren

Proefpersonen selecteren

Compensatie bepalen

Proefpersoon criteria bepalen

Doele: alleen die proefpersonen aan het onderzoek laten deelnemen die mogen deelnemen en de validiteit van het onderzoek niet verminderen.

Dit functioneel gebied bestaat uit twee onderdelen: het bepalen van de compensatie en het bepalen van de criteria waaraan een proefpersoon moet voldoen om deel te nemen aan het onderzoek.

De compensatie bepaalt deels welke studenten mee mogen doen aan het onderzoek. Als deelname aan het onderzoek beloont wordt met geld of met kans op een prijs, dan kan in principe iedereen meedoen aan het onderzoek. Als de proefpersoon gecompenseerd wordt met verplichte uren, dan zal het onderzoek alleen beschikbaar zijn voor psychologiestudenten. Als de compensatie meetelt op het eindcijfer van een vak, dan zal het alleen voor de bedrijfskunde studenten beschikbaar zijn.

De onderzoeker kan daarnaast aangeven aan welke criteria proefpersonen moeten voldoen om aan het onderzoek mee kunnen doen. In het algemeen zijn dat criteria als geslacht, leeftijd of studiejaar, maar proefpersonen kunnen ook uitgesloten worden omdat ze eerder aan een vergelijkbaar onderzoek hebben meegedaan (en dus al getraind zijn of de truc achter het onderzoek al kennen). Een volledig overzicht van de criteria is nog niet bekend.

5.2.1.7. Proefpersonen werven

Proefpersonen werven

Onderzoek openstellen voor proefpersonen

Proefpersoon direct werven

Voortgang werving inzien

Doel: proefpersonen werven voor deelname aan het onderzoek.

Nadat de voorwaarden van deelname zijn bepaald, kan de onderzoeker het onderzoek openstellen. Proefpersonen kunnen nu het onderzoek zien in het wervingssysteem en zichzelf er voor inschrijven.

Er zijn drie niveaus van zichtbaarheid: openbaar, EUR en privé. Openbare onderzoeken zijn zichtbaar zonder dat er ingelogd hoeft te worden. Alleen als een proefpersoon zich wil inschrijven, zal er een account nodig zijn. EUR onderzoeken zijn alleen zichtbaar voor studenten en staf die ingelogd zijn op het wervingssysteem. Privé onderzoeken zijn nooit zichtbaar in het wervingssysteem. Dit zijn onderzoeken die over gevoelige onderwerpen gaan en waar proefpersonen via andere wegen worden geworven.

Een onderzoeker kan voor een openbaar of EUR onderzoek ook reclame maken door bijvoorbeeld de link te posten op Facebook of Twitter.

Natuurlijk kan de onderzoeker nog altijd proefpersonen direct werven. Hiervoor is geen interactie met het systeem nodig. Tijdens de uitvoer van het onderzoek moet de onderzoeker wel de deelname van de proefpersonen registreren (zie 5.2.1.9).

Tijdens de werving kan de onderzoeker ook zien wat de voortgang van de werving is.

5.2.1.8. Proefpersonen inplannen

Proefpersonen inplannen

Afspraak maken met proefpersoon

Afspraak met proefpersoon bevestigen

Proefpersoon handmatig aanmelden

Herhalingsafspraak maken

Doel: afspraken met proefpersonen inplannen.

Het systeem faciliteert een aantal manieren voor een onderzoeker om afspraken te maken met proefpersonen. Ten eerste kan een onderzoeker slots openstellen, ten tweede een directe afspraak maken met de proefpersoon en ten derde proefpersonen werven en direct inschrijven.

Het openstellen van slots is onderdeel van het plannen van het onderzoek (5.2.1.3).

Voor het maken van een afspraak zal de onderzoeker contact opnemen met de proefpersoon om de mogelijkheden te bespreken. Het systeem faciliteert hierbij de onderzoeker door inzicht te geven in de beschikbaarheid van de faciliteiten en de contactgegevens van de proefpersoon.

Sommige proefpersonen zijn slordiger met het bijhouden van hun afspraken en daarom is het gebruikelijk dat onderzoekers de afspraak op bevestigen ter herinnering. Het systeem faciliteert dit door automatische herinnering-e-mails te sturen of een overzicht van contactgegevens te tonen, zodat onderzoekers de afspraken kunnen nabellen.

Als de onderzoeker proefpersonen werft, dan betekent dat de onderzoeker ergens op de campus mensen aanspreekt en, als ze interesse hebben om mee te doen, meeneemt naar het lab. De onderzoeker moet dan de proefpersoon handmatig aanmelden, zodat de compensatie geregeld kan worden en de deelname wordt vastgelegd.

In sommige onderzoeken moet de proefpersoon een onderzoek ondergaan, gevolgd door een follow-up enkele dagen daarna. Het systeem faciliteert het maken van deze herhaalspraak.

5.2.1.9. Onderzoek uitvoeren

Onderzoek uitvoeren

Proefpersoon ophalen uit wachtruimte

Proefpersoon briefen

Onderzoek met proefpersoon uitvoeren

Proefpersoon deelname registreren

Proefpersoon uitbetalen

Proefpersoon debriefen

Doel: het onderzoek uitvoeren en het verkrijgen van onderzoekdata.

Voor het uitvoeren van het onderzoek is het systeem voornamelijk ondersteunend.

Als een proefpersoon in de wachtkamer komt, dan kan hij zich daar aanmelden. De onderzoeker kan deze aanmeldingen zien in het systeem en dan de proefpersoon ophalen uit de wachtkamer. Daarna zal de onderzoeker de proefpersoon briefen met uitleg over het onderzoek en eventuele instructies. Daarna volgt het onderzoek zelf. Het systeem biedt voor deze drie taken geen faciliteiten voor de onderzoeker met uitzondering van het inzien van de wachtruimte aanmeldingen.

Na het onderzoek moet de deelname van de proefpersoon geregistreerd worden door de onderzoeker. Dit is het controlemoment waarop bepaald wordt of de proefpersoon daadwerkelijk heeft deelgenomen en bepaalt dus ook of de proefpersoon wordt gecompenseerd. De onderzoeker zal de deelname registreren in het systeem.

In alle gevallen krijgt de proefpersoon een bewijs van deelname mee, zodat hij zijn deelname kan bewijzen. Bij onderzoeken waar een proefpersoon voor krijgt betaald, zal de proefpersoon voor ontvangst moeten tekenen.

Uiteindelijk moet de onderzoeker de proefpersoon ook debriefen. Veel onderzoeken hebben een bepaalde truc of mechanisme dat andere deelnemers niet mogen weten om de validiteit van het onderzoek niet in gevaar te brengen. In de debriefing wordt dit toegelicht en uitgelegd wat het doel van het onderzoek was. Het systeem biedt hier geen faciliteiten voor.

5.2.1.10. Dag afsluiten

Dag afsluiten

Sleutel terugbrengen

Data veiligstellen

Doel: aan het einde van de dag het lab afsluiten en de data veiligstellen

Tijdens het onderzoek, vaak aan het einde van de dag, moet de data die ontstaan is tijdens het onderzoek veilig worden gesteld. De onderzoeker kan deze actie gedurende de dag zelf in het systeem initiëren, of het systeem doet dit op eigen initiatief aan het einde van de dag. De data wordt dan veilig weggeschreven in de centrale opslag en beschikbaar gemaakt voor de onderzoeker.

Het wegschrijven naar de centrale opslag kan op twee manieren, door middel van een kopieeractie of door middel van een verplaatsactie. Met een kopieer-actie blijft de data beschikbaar op de labcomputer. Met de verplaatsactie wordt de data ook verwijderd van de labcomputer. Dit is vooral belangrijk voor onderzoeken naar gevoelige onderwerpen.

Aan het eind van de dag moet de sleutel van het lab weer ingeleverd worden in het sleutelkastsysteem. Dit wordt geregistreerd in het systeem.

5.2.1.11. Deactiveren

Deactiveren

Onderzoek deactiveren

Doel: tijdelijk een onderzoek deactiveren.

Als een onderzoek om een bepaalde reden niet kan doorgaan, dan kan de onderzoeker het onderzoek tijdelijk deactiveren. Het onderzoek zal wel zichtbaar blijven voor de onderzoeker en zijn collega's, maar proefpersonen kunnen zich niet inschrijven, de onderzoeker kan geen ruimtes boeken of proefpersonen werven.

5.2.1.12. **Onderzoek afronden****Onderzoek
afronden**Geleende
apparatuur
inleveren

Data analyseren

Artikel schrijven
en publicerenOnderzoek
afsluiten

Doel: na afloop van het onderzoek het onderzoek in het systeem afronden.

Als alle proefpersonen het onderzoek hebben ondergaan, dan kan het onderzoek worden afgerond. Dit betekent dat de onderzoeker de proefopstelling kan opruimen en de geleende apparatuur weer inlevert bij de labbeheerders.

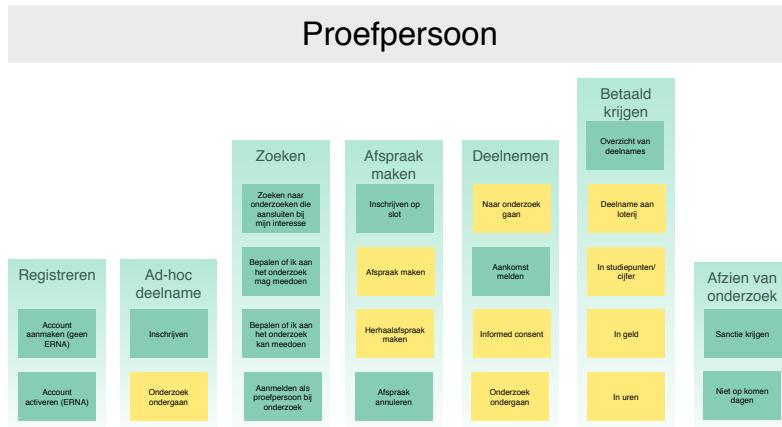
De onderzoeker zal ook de data downloaden uit het systeem en de data analyseren en verwerken in een artikel of in andere wetenschappelijke documenten. Hierbij zal het systeem zorgen dat de data beschikbaar blijft.

Uiteindelijk kan de onderzoeker het onderzoek afsluiten. Het onderzoek en de data blijft beschikbaar in het systeem, maar kan niet meer worden gewijzigd.

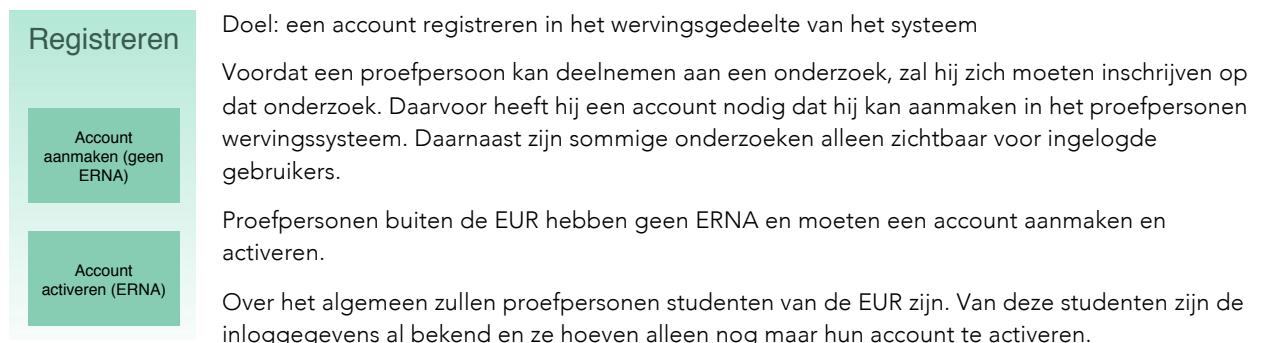
5.2.2. De proefpersoon

Een proefpersoon neemt deel aan een onderzoek, maar zal zich daarvoor eerst moeten registreren en een afspraak moeten maken. Hiervoor wordt de proefpersoon gecompenseerd door geld, studiepunten of met een kans op een prijs. Daarnaast krijgt een proefpersoon een sanctie als hij (een aantal keer) niet komt opdagen.

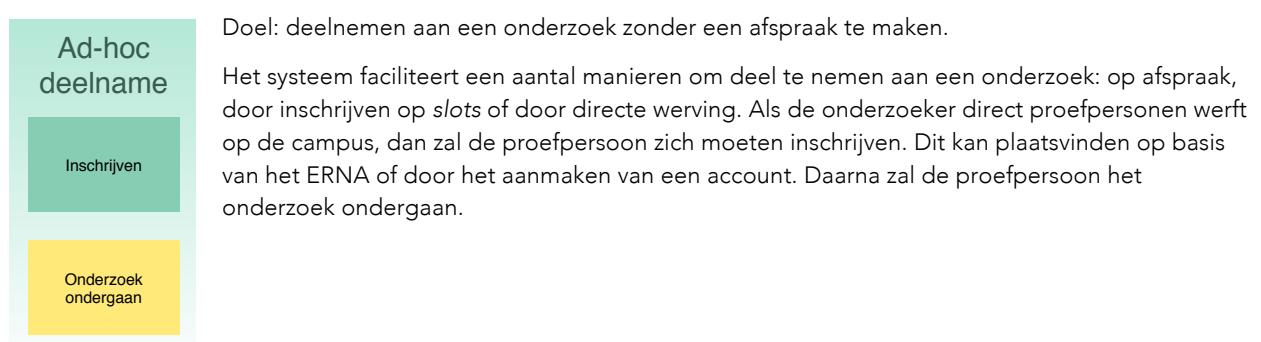
De functionele gebieden waar een proefpersoon het systeem voor nodig heeft, zijn in het diagram hieronder weergegeven. Ieder gebied wordt behandeld in de volgende paragrafen. De gele blokken tonen acties die de proefpersoon wel uitvoert, maar waar hij geen interactie met het systeem heeft.



5.2.2.1. Registreren



5.2.2.2. Ad-hoc deelname



5.2.2.3. Zoeken

Zoeken

- Zoeken naar onderzoeken die aansluiten bij mijn interesse
- Bepalen of ik aan het onderzoek mag meedoen
- Bepalen of ik aan het onderzoek kan meedoen
- Aanmelden als proefpersoon bij onderzoek

Doel: een proefpersoon wil een onderzoek vinden om aan deel te nemen

Een onderzoeker kan zijn onderzoek publiceren naar het wervingsgedeelte van het systeem als het onderzoek openbaar of zichtbaar is voor de EUR. De proefpersoon kan op het publieke gedeelte zoeken naar alle openbare onderzoeken en na inloggen zoeken naar alle EUR onderzoeken die voor hem beschikbaar zijn.

Het systeem biedt een aantal zoekfilters die een proefpersoon kan gebruiken om te zoeken naar specifieke onderzoeken. De precieze filters zijn op dit moment nog niet bekend.

Als de proefpersoon een onderzoek vindt dat aanspreekt, dan kan hij de details van het onderzoek bekijken om te zien of hij kan en mag deelnemen aan het onderzoek. Veel onderzoeken kennen randvoorwaarden voor deelname die betrekking kunnen hebben op de eigenschappen van de proefpersoon (geslacht, lengte, roken of niet-roken, etc.), op eerdere deelnames (sommige onderzoeken gebruiken dezelfde truc) of op de kennis van de proefpersoon (studiejaar). De precieze criteria zijn nog niet bekend.

Daarna kan de proefpersoon bepalen of er mogelijkheden zijn voor deelname (planning) en zich aanmelden voor het onderzoek.

5.2.2.4. Afspraak maken

Afspraak maken

- Inschrijven op slot
- Afspraak maken
- Herhaalspraak maken
- Afspraak annuleren

Doel: afspraak maken met de onderzoeker wanneer het onderzoek zal plaatsvinden

Als de onderzoeker slots heeft gedefinieerd, dan kan de proefpersoon zien welke slots er nog beschikbaar zijn en zich op een geschikt slot inschrijven. De afspraak is dan volledig automatisch gemaakt en de proefpersoon krijgt een notificatie.

Als voor het onderzoek geen slots beschikbaar zijn, dan zullen de onderzoeker en proefpersoon een afspraak moeten maken wanneer het onderzoek uitgevoerd kan worden. Alleen de onderzoeker zal hiervoor het systeem gebruiken, maar de proefpersoon kan inzien welke afspraken hij heeft staan. Hetzelfde geldt voor de herhaalspraken.

Natuurlijk kan het voorkomen dat een proefpersoon de afspraak wil annuleren. Dit wordt gefaciliteerd door het systeem. Hierbij wordt de onderzoeker ook op de hoogte gesteld dat een afspraak is geannuleerd.

5.2.2.5. Deelnemen

Deelnemen

- Naar onderzoek gaan
- Aankomst melden
- Informed consent
- Onderzoek ondergaan

Doel: deelnemen aan het onderzoek

Als de proefpersoon gaat deelnemen aan het onderzoek, dan zal hij eerst naar de wachtruimte van het EBL gaan. Daar kan hij zijn aankomst melden via een aanmeldapplicatie die beschikbaar is in de wachtruimte. Hij zal dan opgehaald worden door de onderzoeker.

Nadat de proefpersoon is gebriefd door de onderzoeker en informed consent heeft geaccordeerd, zal hij het onderzoek ondergaan. Het is mogelijk dat informed consent al op een eerder moment gevraagd wordt (bijv. aan het begin van het studiejaar).

Daarna zal de proefpersoon het onderzoek ondergaan.

5.2.2.6. Betaald krijgen

Betaald krijgen

Overzicht van deelnames

Deelname aan loterij

In studiepunten/ cijfer

In geld

In uren

Doel: de proefpersoon wil gecompenseerd worden voor zijn deelname.

Voor proefpersonen biedt het systeem een overzicht van deelname. Met dit overzicht kan de proefpersoon zien aan welke onderzoeken hij heeft deelgenomen en wat de compensatie daarvoor was. Hiermee kunnen psychologie studenten altijd precies inzien hoeveel verplichte uren ze nog moeten deelnemen aan onderzoeken.

Onderzoekers kunnen op verschillende manieren de proefpersonen uitbetalen. Dit is afhankelijk van de omstandigheden. Voor psychologiestudenten is er bijvoorbeeld een verplichting om een aantal uur per jaar als proefpersoon aan onderzoeken deel te nemen. Bedrijfskunde studenten kunnen voor bepaalde vakken extra toeslag op hun cijfer krijgen als ze aan een onderzoek hebben deelgenomen. Natuurlijk kunnen proefpersonen ook in geld of natura uitbetaald worden.

5.2.2.7. Afzien van onderzoek

Afzien van onderzoek

Sanctie krijgen

Niet op komen dagen

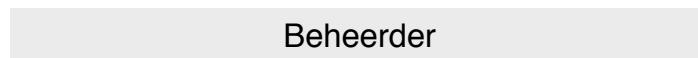
Doel: een proefpersoon een sanctie opleggen na misdragingen tijdens het onderzoek of het niet komen opdagen.

Hoewel het overgrote deel van de proefpersonen met plezier deelneemt aan onderzoeken en op tijd aanwezig is, zijn er proefpersonen die niet komen opdagen of zich tijdens het onderzoek misdragen. Deze kunnen gestraft worden door uitsluiting voor een bepaalde periode, aftrek van gemaakte uren (voor psychologie studenten) of andere sancties.

5.2.3. Beheerder

De beheerder van het lab heeft een ondersteunende rol binnen het lab. Hij zal de onderzoekers, en eventueel proefpersonen, helpen bij het uitvoeren van hun onderzoek. Aan de ene kant door kennis en expertise en aan de andere kant door praktische zaken bij het uitlenen van apparatuur en het onderhoud van de faciliteiten.

De functionele gebieden waar een beheerder het systeem voor nodig heeft, zijn in het diagram hieronder weergegeven. Ieder gebied wordt behandeld in de volgende paragrafen. De gele blokken tonen acties die de beheerder wel uitvoert, maar waar hij geen interactie met het systeem heeft.



5.2.3.1. Labruimte beheren



Doeleind: zorgen dat alle labruimtes up-to-date zijn en gebruikt kunnen worden voor experimenten.

De beheerders zorgen er voor dat onderzoekers de labruimtes kunnen gebruiken voor hun onderzoek. Soms heeft een onderzoeker een specifieke proefopstelling of inrichting van de ruimte nodig. Als de beheerders deze opstelling neerzetten, dan moeten ze bij die ruimte aangeven dat deze tijdelijk anders is ingericht. Andere onderzoekers kunnen dit inzien en bepalen of ze de ruimte wel of niet kunnen gebruiken.

Er zijn dagen waarop een, meer of alle labruimtes gesloten zijn voor gebruik. Bijvoorbeeld op nationale feestdagen of als er onderhoud aan een proefopstelling wordt gepleegd. Dan kan de beheerder die ruimtes blokkeren, zodat er geen reserveringen op gemaakt kunnen worden.

5.2.3.2. Begeleiden onderzoekers



Doeleind: onderzoekers helpen bij het uitvoeren van hun onderzoeken

De beheerders van het lab hebben veel kennis en ervaring met de apparatuur en de werking van de labfaciliteiten. Daarom kunnen ze advies geven of een onderzoek wel of niet uitgevoerd kan of mag worden.

Daarnaast kan hij technische begeleiding geven aan de onderzoekers en een intake gesprek voeren met nieuwe onderzoekers om de werking van het lab toe te lichten. Het systeem faciliteert de beheerder door een overzicht te geven van onderzoeken waar extra begeleiding gewenst is.

5.2.3.3. Resource planning

Resource planning <ul style="list-style-type: none"> Overzicht in gebruik zijnde labs Overzicht boekingen Tijd/uren schrijven op onderzoek Stamdata beheren Proefpersoon uren vrijgeven 	<p>Doel: de beheerder moet overzicht hebben van de beschikbare faciliteiten en planning</p> <p>Voor de beheerder is het belangrijk om een tijdig en accuraat inzicht te hebben in de planning van de faciliteiten en de gemaakte kosten voor onderzoeken. Hierdoor kan er op tijd ingegrepen worden als er teveel onderzoeken tegelijk worden gepland en kan het onderhoud gepland worden in rustigere periodes. Het overzicht van in gebruik zijnde labs gecombineerd met een overzicht van boekingen, stelt de beheerder in staat om dit te doen.</p> <p>Als een beheerder ondersteuning biedt aan een onderzoeker, dan kan de tijd daarvoor bij het onderzoek worden geregistreerd. Uiteindelijk kan dan een goede verdeling van de kosten bepaald worden per onderzoek.</p> <p>Een ander belangrijk aspect van het lab zijn de proefpersoon uren. Voor verschillende groepen proefpersonen kan vooraf bepaald worden hoeveel uren ze als proefpersoon zullen besteden. Dit kan inzichtelijk gemaakt worden in het systeem, zodat de beheerder van het lab hiermee rekening kan houden in de planning.</p> <p>Als laatste kan de beheerder stamdata beheren die van belang is voor de planning. Voorbeelden zijn het beheren van stamdata van ruimtes, apparatuur, type onderzoek, etc.</p>
--	---

5.2.3.4. Account beheer

Account beheer <ul style="list-style-type: none"> Accounts verifiëren Toegang tot bepaalde labs goedkeuren 	<p>Doel: onderzoekers toegang geven tot de juiste labs</p> <p>De beheerders zullen verantwoordelijk zijn voor het verifiëren van accounts van onderzoekers en het beheren van de accounts uitvoeren.</p> <p>Om te voorkomen dat onderzoekers labs reserveren met apparatuur waarvan ze onvoldoende kennis hebben, is er een extra verificatiestap. De beheerder moet eerst het onderzoek dat gebruik maakt van de labs goedkeuren, voordat de onderzoekers toegang krijgen tot het lab (zie ook § 5.2.1.2).</p>
---	--

5.2.3.5. Uitlenen/innemen

Uitlenen/innemen <ul style="list-style-type: none"> Apparatuur beheren Apparatuur uitlenen Apparatuur innemen 	<p>Doel: niet-standaard apparatuur uitlenen aan onderzoekers</p> <p>Voor sommige onderzoeken heeft een onderzoeker extra apparatuur nodig, zoals koptelefoons of specifieke software. De beheerder kan deze apparatuur uitlenen en innemen.</p> <p>De beheerder kan apparatuur beheren, door nieuwe apparatuur te registreren en oude te verwijderen uit het systeem. Ook kan de beheerder de kalibratie- en onderhoudsdatums van de apparatuur vastleggen.</p>
--	--

5.2.3.6. Kwaliteitscontrole

Kwaliteits-controle	Doel: kwaliteit van de processen rondom het onderzoek verhogen
No-show criteria beheren	De beheerder is verantwoordelijk voor de manier waarop onderzoek uitgevoerd wordt. Hij zal bijvoorbeeld bepalen welke sanctie een proefpersoon krijgt als hij niet komt opdagen bij een onderzoek.
Follow-up uitvoeren	Daarnaast zal de beheerder ook ervoor zorgen dat de onderzoekers goed onderzoek uitvoeren door regelmatig een follow-up vragenlijst te sturen naar proefpersonen die recent hebben meegewerkt aan een onderzoek.
Onderzoekers berichten	Als laatste kan de beheerder een bericht sturen naar een groep onderzoekers, gebaseerd op hun kenmerken om bepaalde zaken extra onder de aandacht te brengen.

5.2.4. Rapportages

Het systeem biedt standaard een aantal rapportages aan die voor verschillende doeleinden gebruikt kunnen worden. In het algemeen zal de beheerder de rapportages uit het systeem halen en doorsturen, maar in deze analyse wordt er vanuit gegaan dat ieder beoogd ontvanger dat zelf doet.

Rapportages



5.2.4.1. CvB



Doel: rapportage over de integriteit van 1 onderzoek

Voor het College van Bestuur of de Commissie Wetenschappelijke Integriteit is er een rapportage beschikbaar met daarin een overzicht van 1 onderzoek en alle gebeurtenissen en wijzigingen die plaats hebben gevonden in het kader van dat onderzoek. Hiermee is een gecategoriseerde tijdlijn te maken vanaf de registratie van het onderzoek in het systeem, naar de voorbereiding en werving van proefpersonen tot aan het ontstaan van de data.

5.2.4.2. Labbeheer



Doel: overzicht van verbruikte uren en financiën

De labbeheerders hebben beschikking over rapportages waarmee inzicht te verkrijgen is in de tijd die besteed is bij de ondersteuning van onderzoekers en het gebruik van de labs per onderzoek. Daarnaast is er een overzicht van de verbruikte proefpersoon uren die gebruikt kan worden voor het spreiden van de uren door het hele jaar.

5.2.4.3. Faculteit/bureau onderwijs



Doel: overzicht van gemaakte uren per proefpersoon voor compensatie

Voor de faculteiten is een rapportage beschikbaar waarmee de gemaakte proefpersoon uren inzichtelijk gemaakt kan worden. Dit kan gebruikt worden door het bureau onderwijs om de studenten de juiste punten te geven op hun tentamencijfer of het behalen van de uren norm te controleren.

5.2.4.4. Onderzoeker



Doel: verloop van eigen onderzoek analyseren

De onderzoeker kan op een zeker moment zijn eigen onderzoek analyseren als basis voor het artikel of andere publicatie. Hiervoor kan hij een rapportage uitdraaien die in principe dezelfde informatie toont als de integriteitsrapportage.

5.3. Scope

ID	Functioneel gebied	Context
FE-1	Labruimte beheren	Beheren
FE-2	Begeleiden onderzoekers	Beheren
FE-3	Resource planning	Beheren
FE-4	Accountmanagement	Beheren
FE-5	Stamdata beheren	Beheren
FE-6	Onderzoek registreren	Onderzoeken
FE-7	Onderzoek toetsen	Onderzoeken
FE-8	Onderzoek plannen	Onderzoeken
FE-9	Onderzoeken combineren	Onderzoeken
FE-10	Onderzoek voorbereiden	Onderzoeken
FE-11	Proefpersonen selecteren	Onderzoeken
FE-12	Proefpersonen werven	Onderzoeken
FE-13	Proefpersonen inplannen	Onderzoeken
FE-14	Onderzoek uitvoeren	Onderzoeken
FE-15	Administratie	Onderzoeken
FE-16	Onderzoek afronden	Onderzoeken
FE-17	Hulp en ondersteuning	Onderzoeken
FE-18	Uploaden onderzoek en veiligstellen data	Onderzoeken
FE-19	Registreren als proefpersoon	Proefpersonen
FE-20	Ad-hoc deelname	Proefpersonen
FE-21	Zoeken naar onderzoek	Proefpersonen
FE-22	Afspraak maken	Proefpersonen
FE-23	Deelnemen aan onderzoek	Proefpersonen
FE-24	Betaald krijgen	Proefpersonen
FE-25	Afzien van onderzoek	Proefpersonen
FE-26	CvB Rapportage	Rapportage
FE-27	Labbeheer rapportage	Rapportage
FE-28	Faculteit/bureau onderwijs rapportage	Rapportage
FE-29	Notificaties	Rapportage
FE-30	Inloggen	Algemeen

5.4. Non-functionele requirements

Naast functionele requirements zijn er non-functionele requirements. Deze beschrijven hoe het systeem moet werken. Voorbeelden zijn: hoe snel moet een pagina getoond worden, hoe veilig moet de data opgeslagen worden en hoeveel gebruikers kunnen tegelijkertijd van het systeem gebruik maken?

Naast de gebieden die hier explicet zijn benoemd, zijn er geen specifieke eisen ten aanzien van de volgende punten: betrouwbaarheid, onderhoudbaarheid. Naast deze requirements zijn er ook veel requirements die impliciet zijn, bijvoorbeeld op het gebied van performance en betrouwbaarheid. Gezien de gekozen architectuur zien wij geen aanleiding om deze requirements explicet te maken. Het uitgangspunt is dat het systeem prettig werkt voor de gebruikers.

5.4.1. Privacy en anonimiteit

Op dit moment is niet bekend in welke mate de privacy en anonimiteit van proefpersonen beschermd moet worden. Ook zijn er een aantal functionele eisen aan het systeem die enige mate van privacy verlies met zich meebrengen. Daarom formuleren we een aantal lagen die steeds strenger worden. In het uitvoering van het project kan een definitieve keuze gemaakt worden. Een zeer privacybescherming betekent wel dat een aantal requirements niet gerealiseerd kunnen worden.

1. Het systeem moet vastleggen aan welke onderzoeken een proefpersoon heeft meegewerkt en welk bestand dat tot gevolg heeft gehad. Deze informatie is alleen beschikbaar voor beheerders en de eigenaren van het onderzoek.
2. De koppeling tussen proefpersoon en databestand wordt na het afsluiten van het onderzoek verwijderd en kan niet meer automatisch herleid worden. De onderzoeken waar een proefpersoon aan mee heeft gewerkt, mogen wel bekend blijven.
3. De koppeling tussen proefpersoon en het onderzoek waar hij aan mee heeft gedaan en de resulterende bestanden mag nergens in het systeem worden gelegd.
4. Het systeem mag op geen enkele wijze informatie verstrekken of vastleggen over de deelname van een proefpersoon aan een onderzoek (zoals tijdstip van deelname, het onderzoek, resulterende bestanden), buiten het feit dat de proefpersoon gecompenseerd moet worden voor deelname aan een onderzoek of een sanctie moet krijgen voor het niet nakomen van een afspraak.

5.4.2. Audit en logging

Hieronder staat een overzicht van de non-functionele requirements voor audit en logging. In hoofdstuk 7 staat een uitgebreid overzicht van alle non-functionele requirements die specifiek nodig zijn voor de integriteitsrapportage.

- De volgende informatie wordt minimaal gelogd:
 - Basisinformatie over het onderzoek, inclusief meta-data en gekoppelde onderzoekers,
 - Deelname van een proefpersoon aan een onderzoek, inclusief de inschrijving en afspraak (als dat via het systeem gebeurt) en de goedkeuring na deelname,
 - Meta-data van data die gegenereerd is per onderzoek, inclusief creation date en bron (lab-computer), datum wanneer de data is verplaatst van de bron naar een andere computer,
 - De ruimtes die geboekt zijn voor het onderzoek, inclusief de meta-data zoals de datum waarop de boeking is gemaakt, en de datum en duur van de boeking,
 - Sleutels gebruikt door onderzoekers per onderzoek, inclusief de tijdstippen van uitgave en terugval,
 - Loginacties en -pogingen op het systeem.
- Alle wijzigingen in bovenstaande elementen inclusief het tijdstip waarop de wijziging heeft plaatsgevonden.
- De audit moet onbeperkt bewaard blijven.
- Het logging gedeelte moet ook logging vanuit andere systemen kunnen verwerken, mits deze logging direct betrekking heeft op een onderzoek.

5.4.3. Beveiliging

- Het systeem kent de volgende rollen:
 - Beheerder,
 - Eigenaar,
 - Onderzoeker,
 - Proefpersoon.
- Bestanden zijn alleen beschikbaar voor medewerkers aan een onderzoek.

- Onderzoeken kunnen publiek zijn, maar ook privé. Privé onderzoeken zijn wel geregistreerd in het systeem, maar proefpersonen kunnen zich er niet voor inschrijven. De werving verloopt volledig via de onderzoekers.

5.4.4. Performance

Het doel van het systeem is dat alle gebruikers er gemakkelijk mee kunnen werken. Dat betekent dat het systeem als snel ervaren moet worden. De specifieke requirement die voor het EBL van belang is, is de volgende:

- Systeem moet met bestanden kunnen omgaan van maximaal 1GB.

5.4.5. Capaciteit

- Het systeem moet minimaal 10 000 proefpersoon deelnames per jaar kunnen afhandelen.
- De verwachte capaciteit aan data storage per jaar is 2TB.

5.4.6. Beschikbaarheid

- Het systeem is web-based en beschikbaar via het Internet.

5.4.7. Integriteit

- Onderzoek bestanden dienen 1-op-1 vastgelegd te worden. Er mag dus geen *lossy* compressie worden toegepast.

5.4.8. Compatibiliteit met andere systemen

Het systeem moet samen kunnen werken met de systemen zoals genoemd in hoofdstuk 4. Daarnaast moet het systeem goed werken op de volgende systemen:

- Andere systemen moet de logging functionaliteit kunnen gebruiken om eigen logregels aan toe te voegen
- Moderne webbrowsers (IE9, Chrome, Firefox, Safari)
- Tablets, smartphone

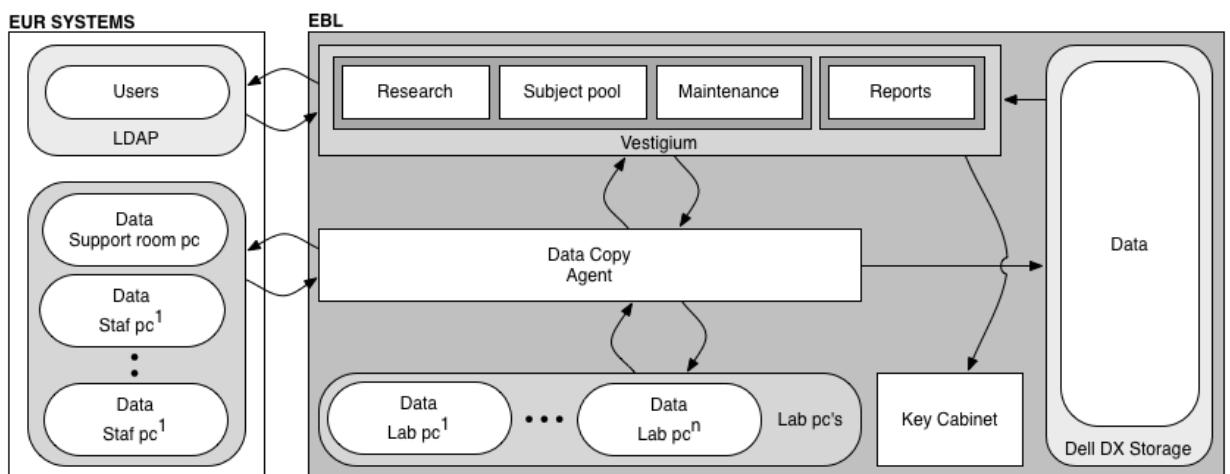
5.4.9. Gebruiksvriendelijkheid

- Een proefpersoon kan zich zonder uitleg voor een onderzoek inschrijven
- Het systeem is volledig Engelstalig.

6. Voorstel architectuur nieuwe situatie

Voorgaand hoofdstuk beschrijft de functionele en niet-functionele eisen waaraan de nieuwe situatie moet voldoen. Dit hoofdstuk beschrijft de wijzigingen in de onderliggende technische architectuur die nodig zijn om aan deze eisen te kunnen voldoen.

Onderstaand figuur laat op hoog niveau de systemen en gebruikers zien zoals die in de nieuwe situatie betrokken zijn bij het uitvoeren en studie met onderzoeken in het lab. De verschillende actoren zijn niet opgenomen, omdat de figuur hiermee te complex zou worden.



Afbeelding 4 Overzicht architectuur nieuwe situatie

Ten opzichte van de in de huidige situatie getoonde plaatjes (zie hoofdstuk 4) zijn de volgende onderdelen nieuw:

- de Vestigium applicatie waarin functionaliteiten van een aantal huidige applicaties samenkommen,
- de Data Copy Agent voor het kopiëren van data tussen het staf en lab netwerk, en
- de Dell DX Storage als opslag van data.

We zullen deze nieuwe onderdelen in onderstaande paragrafen verder bespreken.

6.1. Vestigium

De applicatie dient ter vervanging van de drie proefpersoon-wervingssystemen en het boekingssysteem. Ook bevat deze applicatie de nieuwe functionaliteiten ten behoeve van:

- intake,
- beheer, en
- rapportage.

6.2. Data Copy Agent

Op dit moment zorgt de EBL Assistant applicatie in combinatie met Janno en Salvia voor het kopiëren van data tussen het staf en lab netwerk, zie ook paragraaf 4.2.3 'Uitvoeren onderzoeken'. De Data Copy Agent gaat in de nieuwe situatie zorg dragen voor het uitvoeren van deze kopieer acties.

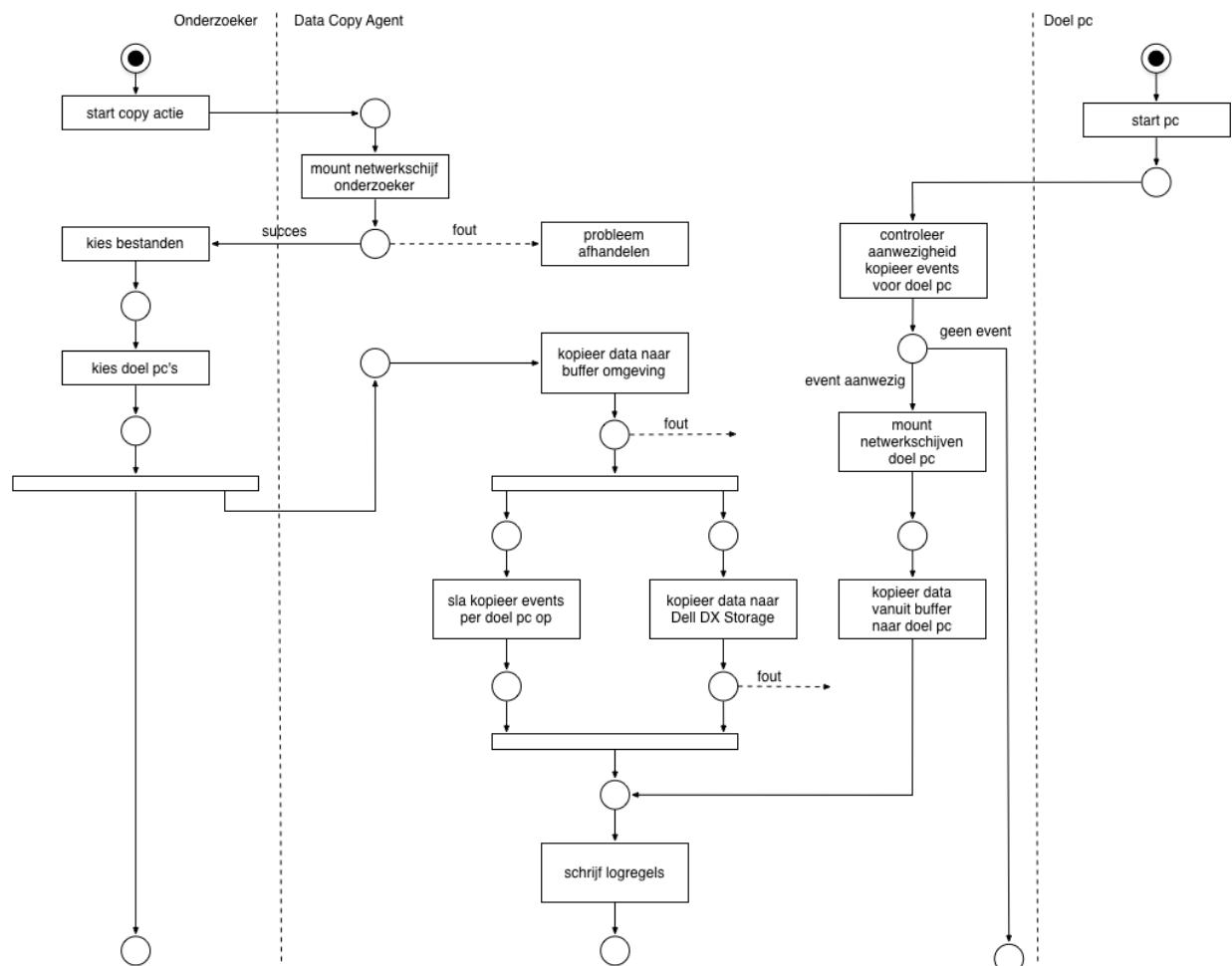
Er zijn drie momenten dat de Data Copy Agent data kopiëert:

- handmatige kopieer actie van staf- naar lab-netwerk door onderzoeker (de zgn. Copy),
- handmatige kopieer actie van lab- naar staf-netwerk door onderzoeker (de zgn. Retrieve),
- batch kopieer acties van lab- naar staf-netwerk door achtergrondproces.

6.2.1. Copy

Voordat hij een onderzoek gaat uitvoeren kopiëert de onderzoeker zijn data naar de computers in de kamer van het lab waar het onderzoek zal plaatsvinden. De onderzoeker initieert de kopieer actie van ofwel een computer in een van de support rooms van het lab, ofwel van zijn staf computer.

Onderstaand figuur toont de verschillende stappen die de onderzoeker en de Data Copy Agent in het kopieer proces volgen.



Afbeelding 5 Overzicht van Copy proces

De onderzoeker start vanuit de Vestigium applicatie een kopieer actie.

De Data Copy Agent mount vervolgens de netwerk schijf van de onderzoeker en stuurt en lijst van bestanden op de netwerkschijf naar de Vestigium applicatie. De Vestigium applicatie toont de onderzoeker de bestanden.

De onderzoeker kiest de te kopiëren bestanden.

De Vestigium applicatie toont de onderzoeker de mogelijke doel computers waar de Data Copy Agent de data naar toe kan kopiëren. De onderzoeker kiest de doel computers.

De Data Copy Agent zal vervolgens de data in een tijdelijk buffer opslaan. Van daaruit zal de Data Copy Agent de data uiteindelijk kopiëren naar:

- de geselecteerde doel computers, en
- de Dell DX Storage (ten behoeve van de v0 opslag).

Omdat er naar de lab computers geen kopieer acties mogen plaatsvinden terwijl ze in gebruik zijn, slaat de Data Copy Agent een kopieer event voor elk van de geselecteerde doel computers op in een event log.

Bij het (her-)starten van de doel computer zal deze de Data Copy Agent vragen om eventuele openstaande kopieer acties uit te voeren. De Data Copy Agent controleert vervolgens het event log om te zien of er kopieer acties aanwezig zijn. Er zijn nu twee mogelijkheden:

- *er is geen kopieer event aanwezig*
De Data Copy Agent hoeft verder niets uit te voeren en de doel computer zal verder opstarten.
- *er is wel een kopieer event aanwezig*
De Data Copy Agent mount de netwerk schijven van de doel computer en kopieert de in de buffer opgeslagen data naar de doel computer.

De Data Copy Agent stuurt van alle uitgevoerde acties logregels naar de Vestigium applicatie.

Op verschillende moment in het proces kunnen er fouten optreden die verder door de Data Copy Agent (en eventueel de Vestigium applicatie) verder afhandelt.

6.2.2. Retrieve

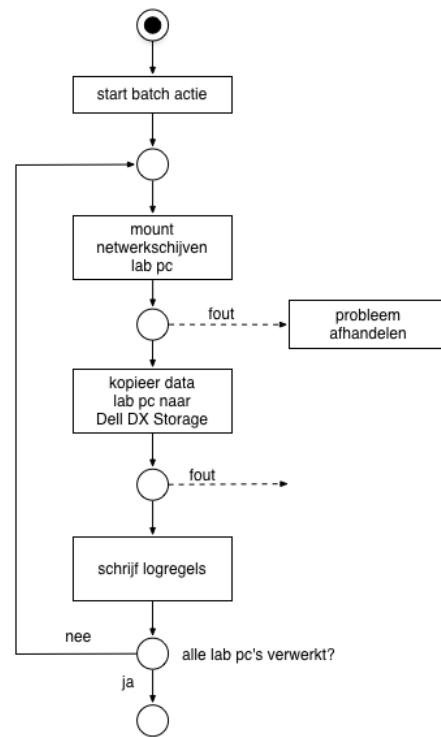
De handmatige kopieer actie van data vanaf het lab netwerk naar het staf netwerk (ook wel de *retrieve* actie genoemd) verloopt in tegenovergestelde richting aan het in paragraaf 6.2.1 beschreven proces.

6.2.3. Move

Het systeem biedt ook een move actie. Deze actie is identiek aan de retrieve actie (6.2.2), maar verwijdert tevens de data van de lab computer.

6.2.4. Batch

Data die de onderzoeker niet handmatig overgezet heeft zal de Data Copy Agent elke nacht alsnog veilig stellen. vergeleken met de handmatige *retrieve* actie is dit een vrij eenvoudig proces, omdat de Data Copy Agent de data alleen naar de Dell DX Storage hoeft te kopiëren. Het proces ziet er als volgt uit:



Afbeelding 6 Overzicht van Batch proces

De Data Copy Agent start elke nacht de *retrieve* actie voor alle lab computers.

De Data Copy Agent mount de netwerk schijven van de gekozen doel computers en kopieert de data naar de Dell DX Storage. Het bepalen van de te kopiëren data kan op diverse manieren plaatsvinden, het uiteindelijke algoritme kan tijdens het project bepaald worden.

Uiteindelijk stuurt de Data Copy Agent logregels van de uitgevoerde acties naar de Vestigium applicatie, zodat de data van het onderzoek van de onderzoeker de volgende dag beschikbaar is.

6.3. Dell DX Storage

Om aan een van de doelstellingen van het project, onderzoekdata opslaan in centraal systeem dat via het web bereikbaar is, te kunnen voldoen slaan we data op in het Dell DX Storage systeem. De wijze waarop dit gebeurd is beschreven in paragraaf 6.2.

Vanuit de Vestigium applicatie kan een onderzoeker de data van zijn onderzoek downloaden door middel van directe hyperlinks naar de Dell DX Storage server. Indien dit security technisch niet mogelijk blijkt (i.v.m. directe toegang tot de Dell DX Storage), kan er ook een alternatieve oplossing ingezet worden. Hierbij kopieert de Data Copy Agent de gegevens eerst naar de Vestigium applicatie en biedt het vanuit die applicatie ter download aan.

7. Integriteitsrapportage

De architectuur van het systeem (hoofdstuk 6) beschrijft dat de modules waarmee de onderzoekers, beheerders en proefpersonen werken, los staan van de logging module. De modules registreren een gebeurtenis en sturen deze naar de logging module. De logging module ontvangt de relevante gebeurtenissen en slaat deze als log-regels op. Dit hoofdstuk bevat een analyse van de log-regels die nodig zijn om de integriteit van een onderzoek te onderbouwen. Per gebied (onderzoeker, proefpersoon, beheerder) zijn de features opgesomd. Per feature is aangegeven of er een log-regel resulteert en is er een aanzet gegeven voor de inhoud van de log-regel.

Dit hoofdstuk geeft geen compleet overzicht van alle gebeurtenissen en acties die gelogd moeten worden. Vanuit de visie en de doelstelling is het belangrijkste doel dat het systeem een rapportage levert die de integriteit van onderzoek kan onderbouwen. Daarom is de onderbouwing hiervan in meer detail uitgewerkt. De features die in de tabel hieronder met een NVT zijn gemarkerd leveren geen (in)directe onderbouwing van de integriteit op, maar zullen wel in een of andere vorm gelogd worden.

Een voorbeeld: een onderzoeker is bezig in het systeem om mede-onderzoekers te koppelen aan het onderzoek dat hij zojuist heeft aangemaakt. Zodra hij een onderzoeker koppelt aan het onderzoek, zal de logging-module een regel wegschrijven in het log met de volgende informatie:

- de datum en het tijdstip waarop de mede-onderzoeker is gekoppeld aan het onderzoek,
- de user account van de onderzoeker die de koppeling heeft gelegd,
- het identificatienummer van het onderzoek waaraan de mede-onderzoeker is gekoppeld,
- de user account van de onderzoeker die het onderzoek heeft aangemaakt,
- de user account van de onderzoeker die gekoppeld is.

Met deze informatie is precies het wie, wat en waar te herleiden van deze actie. Hoewel dit niet expliciet wordt genoemd, wordt er ook een log-regel geschreven als er een onderzoeker wordt ontkoppeld van een onderzoek.

Door alle log-regels van een onderzoek samen te nemen, wordt het mogelijk om een chronologisch overzicht te maken van alle acties die voor een onderzoek zijn uitgevoerd en door wie. Dit wordt mogelijk gemaakt doordat iedere log-regel minimaal een datum en tijdstip bevat, de user account van de persoon die de actie uitvoert en het identificatienummer van het onderzoek.

Een interessant aspect is dat log-regels de bewijslast van elkaar versterken. Als er voor een bepaald onderzoek een ruimte-reservering in het log is geschreven en het afhalen van de sleutel is gelogd, dan is dat een sterke aanwijzing dat de ruimte daadwerkelijk is gebruikt. Als er daarnaast afspraken met proefpersonen in de periode van de reservering in de logs zijn opgenomen, dan wordt het bewijs weer sterker. Als uiteindelijk ook data is ontstaan in de periode van de reservering in het lab dat is gereserveerd en de deelname van proefpersonen is goedgekeurd, dan is het bewijs zeer sterk dat het onderzoek daadwerkelijk is uitgevoerd.

Alle log-regels zoals beschreven in de onderstaande tabellen zijn voorlopig en gebaseerd op het huidige inzicht. Tijdens de realisatie van het project, zullen hier velden aan toegevoegd of afgehaald worden.

De laatste kolom bevat specifieke opmerkingen bij de log-regel:

1. Onderdeel audi-trail van onderzoek zelf
Hiermee wordt bedoeld dat niet alleen de onderbouwing van het onderzoek gedaan kan worden aan de hand van de log-regels, maar dat ook de wijzigingen aan het onderzoek worden vastgelegd. Hiermee is te achterhalen wanneer de kenmerken van het onderzoek zijn gewijzigd.
2. Deze log-regel wordt ondersteund door ...
In veel gevallen biedt een log-regel op zichzelf niet veel bewijs, maar als de log-regel gecombineerd wordt met andere, dan wordt het bewijs sterker.

7.1. Onderzoeker

Functie	Features	Log-regel	Opmerking
Onderzoek registreren			

	Onderzoek aanmaken	Datum, tijdstip, user account, onderzoek-id, naam, omschrijving, startdatum, einddatum, aantal proefpersonen, onderzoeks-voorstel (link naar document), type onderzoek, methode	Onderdeel audit-trail van onderzoek zelf
	Mede-onderzoekers koppelen	Datum, tijdstip, user account, onderzoek-id, account eigenaar, account onderzoeker	Onderdeel audit-trail van onderzoek zelf
Onderzoek toetsen			
	Toetsen bij Medisch Ethische commissie	Datum, tijdstip, user account, onderzoek-id, beslissing	
	Technische haalbaarheid bepalen	NVT	Logregel ontstaat in <i>Onderzoek goedkeuren voor uitvoering</i>
Onderzoek plannen			
	Blokonderzoek plannen	Datum, tijdstip, user account, onderzoek-id, ruimte, startdatum en -tijd, einddatum en -tijd, slotduur, aantal simultane proefpersonen	Logregel identiek aan <i>Slots vastleggen</i>
	Ruimte reserveren	Datum, tijdstip, user account, onderzoek-id, ruimte, startdatum en -tijd, einddatum en -tijd	Deze log-regel wordt ondersteund door <i>Sleutel halen</i> en <i>Data veiligstellen</i>
	Slots vastleggen	Datum, tijdstip, user account, onderzoek-id, ruimte, startdatum en -tijd, einddatum en -tijd, slotduur, aantal simultane proefpersonen	Deze log-regel wordt ondersteund door <i>Sleutel halen</i> en <i>Data veiligstellen</i>
	Planning afronden	Datum, tijdstip, user account, onderzoek-id, statuswijziging	
Onderzoek combineren			
	Bundelonderzoek plannen	Datum, tijdstip, user account, onderzoek-id, type onderzoek	
	Combi-onderzoek plannen	Datum, tijdstip, user account, onderzoek-id, type onderzoek, onderzoekid's	Onderzoek-id's zijn koppelingen naar de andere onderzoeken (zie ook 5.2.1.4)
Onderzoek voorbereiden			
	Sleutel halen	Datum, tijdstip, user account, onderzoek-id, sleutel, tijdstip	Het is misschien niet mogelijk om een sleutel direct te herleiden naar een onderzoek i.v.m. algemene toegang tot sleutelkast voor stafleden
	Proefopstelling inrichten	NVT	
	Apparatuur lenen	NVT	Logregel onstaat in <i>Apparatuur uitlenen</i> van de beheerder
	Onderzoekbestand uploaden	Datum, tijdstip, user account, onderzoek-id, bron computer, bron directory, doel computer, doel directory, aantal bestanden, bestandsgrootte, tijdsduur	Deze log-regel wordt ondersteund door <i>Ruimte reserveren</i> en <i>Data veiligstellen</i>
	Pilot uitvoeren	Datum, tijdstip, user account, onderzoek-id, type onderzoek	Onderdeel audit-trail van onderzoek zelf
Proefpersonen selecteren			
	Compensatie bepalen	Datum, tijdstip, user account, onderzoek-id, type compensatie, hoogte compensatie	Onderdeel audit-trail van onderzoek zelf
	Proefpersoon criteria bepalen	Datum, tijdstip, user account, onderzoek-id, criteria	Onderdeel audit-trail van onderzoek zelf
Proefpersonen werven			

	Onderzoek openstellen voor proefpersonen	Datum, tijdstip, user account, onderzoek-id, statuswijziging	Onderdeel audit-trail van onderzoek zelf
	Proefpersoon direct werven	NVT	Voor loggen deelname, zie <i>Proefpersoon handmatig aanmelden of proefpersoon deelname registreren</i>
	Voortgang werving inzien	NVT	
Proefpersonen inplannen			
	Afspraak maken met proefpersoon	Datum, tijdstip, user account, onderzoek-id, proefpersoonid, startdatum en -tijd, einddatum en -tijd	Log-regel wordt ondersteund door <i>Ruimte reserveren</i>
	Afspraak met proefpersoon bevestigen	NVT	Dit wordt buiten het systeem om gedaan. Zie 5.2.1.8.
	Proefpersoon handmatig aanmelden	Datum, tijdstip, user account, onderzoek-id, proefpersoonid	
	Herhalingsafspraak maken	Datum, tijdstip, user account, onderzoek-id, proefpersoonid, startdatum en -tijd, einddatum en -tijd	Identiek aan de logregel bij <i>Afspraak maken met proefpersoon</i>
Onderzoek uitvoeren			
	Aantekening maken	Datum, tijdstip, user account, onderzoek-id, notitie	Kan later gebruikt worden door de onderzoeker als lab-notebook
	Proefpersoon ophalen uit wachtruimte	NVT	
	Proefpersoon briefen	NVT	
	Onderzoek met proefpersoon uitvoeren	NVT	
	Proefpersoon deelname registreren	Datum, tijdstip, user account, onderzoek-id, proefpersoonid	Deze log-regel is erg belangrijk, want ook de basis voor uitbetaling proefpersonen
	Proefpersoon uitbetalen	NVT	
	Proefpersoon debriefen	NVT	
Dag afsluiten			
	Sleutel terugbrengen	Datum, tijdstip, user account, onderzoek-id, sleutel, tijdstip	Het is misschien niet mogelijk om een sleutel direct te herleiden naar een onderzoek.
	Data veiligstellen	Datum, tijdstip, user account, onderzoek-id, bron computer, bron directory, doel computer, doel directory, aantal bestanden, bestandsgrootte, tijdsduur	Belangrijkste log-regel, want deze is sterke indicatie dat een onderzoek daadwerkelijk is uitgevoerd
Deactiveren			
	Onderzoek deactiveren	Datum, tijdstip, user account, onderzoek-id, statuswijziging	Onderdeel audit-trail van onderzoek zelf
Onderzoek afronden			
	Geleende apparatuur inleveren	NVT	
	Data analyseren	NVT	
	Artikel schrijven en publiceren	NVT	
	Onderzoek afsluiten	Datum, tijdstip, user account, onderzoek-id, statuswijziging	Onderdeel audit-trail van onderzoek zelf

7.2. Proefpersoon

Functie	Features	Logregel	Opmerking
Registreren			
	Account aanmaken (geen ERNA)	NVT	
	Account aanmaken (ERNA)	NVT	
Ad-hoc deelname			
	Inschrijven	NVT	Zie <i>Proefpersoon handmatig aanmelden</i>
	Onderzoek ondergaan	NVT	
Zoeken			
	Zoeken naar onderzoeken die aansluiten bij mijn interesse	NVT	
	Bepalen of ik aan het onderzoek mag meedoen	NVT	
	Bepalen of ik aan het onderzoek kan meedoen	NVT	
	Aanmelden als proefpersoon bij onderzoek	Datum, tijdstip, user account, onderzoek-id,	
Afspraak maken			
	Inschrijven op slot	Datum, tijdstip, user account, onderzoek-id, startdatum en -tijd	
	Afspraak maken	NVT	
	Herhaalspraak maken	NVT	
	Afspraak annuleren	Datum, tijdstip, user account, onderzoek-id, afspraakdatum en -tijd, reden van annuleren	
Deelnemen			
	Naar onderzoek gaan	NVT	
	Aankomst melden	Datum, tijdstip, user account, onderzoek-id, afspraak	Vormt samen met <i>Aanmelden als proefpersoon bij onderzoek</i> , <i>Proefpersoon deelname registreren</i> bewijs dat proefpersoon aanwezig was bij het onderzoek
	Informed consent	NVT	
	Onderzoek ondergaan	NVT	
Betaald krijgen			
	Overzicht van deelnames	NVT	
	Deelname aan loterij	NVT	
	In studiepunten/cijfer	NVT	
	In geld	NVT	
	In uren	NVT	
Afzien van onderzoek			
	Sanctie krijgen	NVT	
	Niet op komen dagen	NVT	

7.3. Beheerder

Functie	Features	Logregel	Opmerking
Labruimte beheren			
	Opstellingen inrichten	NVT	
	Onderhoud plannen	NVT	
	Vakantiedagen plannen	NVT	
Begeleiden onderzoekers			
	Onderzoek goedkeuren voor uitvoering	Datum, tijdstip, user account, onderzoek-id, statuswijziging	
	Technische begeleiding	NVT	
	Intake gesprek voeren met nieuwe onderzoeker	NVT	
Resource planning			
	Overzicht in gebruik zijnde labs	NVT	
	Overzicht boekingen	NVT	
	Tijd/uren schrijven op onderzoek	NVT	
	Stamdata beheren	NVT	
	Proefpersoon uren vrijgeven	NVT	
Account beheer			
	Accounts beheren	NVT	
	Toegang tot bepaalde labs goedkeuren	NVT	
Uitlenen/innemen			
	Apparatuur beheren	NVT	
	Apparatuur uitlenen	Datum, tijdstip, user account, onderzoek-id, apparaat, uitgiftedatum en -tijd	
	Apparatuur innemen	Datum, tijdstip, user account, onderzoek-id, apparaat, innamedatum en -tijd	
Kwaliteitscontrole			
	No-show criteria beheren	NVT	
	Follow-up uitvoeren	NVT	
	Onderzoekers berichten	NVT	