Mapping Normative Frameworks of EThics and Integrity of REsearch

D.5.1: Protocol for Systematic Searches and Tagging of RE+RI Cases

D.5.1 Protocol for Systematic Searches and Tagging of RE+RI Cases

Project details	
Project:	Mapping Normative Frameworks of EThics and Integrity of REsearch
Project acronym:	EnTIRE
Project start date:	01.05.2017
Duration:	48 months
Project number:	741782
Project Coordinator:	Vrije Universiteit Medisch Centrum (VUmc) Amsterdam

Deliverable details	
Work Package:	WP 5 Cases, Casuistry and Scenarios
Deliverable description:	D.5.1 Protocol for Systematic Searches and Tagging of RE+RI Cases
Work package leader:	Dublin City University
Authors:	Jonathan Lewis, Mohammad Hosseini, Bert Gordijn, Péter Kakuk and Søren Holm

Table of Contents

1. About EnTIRE

The areas of Research Ethics and Research Integrity (RE+RI) are rapidly evolving. In the EU and internationally, new legislation, codes of conduct and good practices are constantly being developed. New technologies (e.g. gene editing), complex statistical methods (e.g. biostatistics), pressure to publish and obtain grants, and growing emphasis on stakeholder driven science (e.g. public-private partnerships) increase the complexity of conducting science. In this complex and dynamic environment, researchers cannot easily identify the correct rules and best tools for responsible conduct of research. This also increasingly constitutes a challenge for RE+RI experts.

Our aim is to create a platform that makes the normative framework governing RE+RI easily accessible, supports application in research and evaluation, and involves all stakeholders in a participatory way, thus achieving sustainability. The platform will foster uptake of ethical standards and responsible conduct of research, and ultimately support research excellence and strengthen society's confidence in research and its findings.

2. About 'Cases, Casuistry and Scenarios'

Work package 5, 'Cases, Casuistry and Scenarios', contains all tasks that provide for a comprehensive selection of RE+RI cases and case analysis methods as well as the presentation of actual case analyses. Moreover, a selection of prominent RE+RI cases will be analysed with different case analysis methods. Finally, a set of scenarios will be developed. The purpose of making the cases, case analysis methods, actual cases analyses and scenarios available is to foster structured analysis and thorough debate about RE+RI cases, which in its turn might serve as the bedrock for responsible future RE+RI regulation and practice.

The RE+RI cases to be uploaded onto the EnTIRE platform will result from searches in different potential sources, e.g. academic literature, reports of RE+RI committees, professional regulators, grey literature, media outlets and the blogosphere. In addition, the focus groups sessions, both the face-to-face and the online focus groups, in WP2 will be used to generate, as well as to reflect and deliberate on, cases from local practice. The cases resulting from the searches within these different sources will be tagged and categorised, with the purpose of enhancing the indexing of cases in the EnTIRE platform. Both traditional methods of categorisation – e.g. using well-known concepts such as misconduct, falsification, fabrication, plagiarism – as well as more innovative ways of categorisation, such as using the ethical principles within the RE+RI normative framework, will be applied and a thesaurus developed. In addition, case analysis methods suitable to the analysis of RE+RI cases will be identified through a systematic literature review and made available on the online platform. Moreover, a selection of prominent RE+RI cases will be analysed thereby showcasing relevant case analysis methods. Finally, a set of scenarios will be built for educational purposes and in order to stimulate strategic thinking about RE+RI.

3. Description of Work

*Task 5.1. Preparation of data collection on cases (M1-6, **DCU**, UNIDEB, Manchester)*

1. Identify potential sources of RE+RI cases, for example: a) academic literature; b) reports by RE+RI committees and regulatory bodies; c) grey literature such as government documents, white papers, theses and dissertations, conference proceedings and policy statements; d) media outlets such as newspapers and magazines; and e) the blogosphere.
2. Develop appropriate systematic methods to conduct searches within these sources so as to gather RE+RI cases, e.g. through a systematic literature review.
3. Develop a system of categories, a 'thesaurus', for tagging cases that enhances retrievability and orientation within the EnTIRE database. Tagging methods will involve traditional approaches, e.g. tagging according to issues such as misconduct, falsification, fabrication, plagiarism, fake peer- review, data management, as well as innovative methods, for example tagging according to the main ethical principles from the RE+RI normative framework that have been violated.

The output for task 5.1 is D.5.1: Protocol for systematic searches and tagging of RE+RI cases [M7].

Task 5.1 is the first of five tasks to be carried out by Work Package 5. The output of task 5.1 is a protocol for the retrieval and tagging of RE+RI cases in tasks 5.2 and 5.3. The deadline for submitting the protocol to the European Commission is 1st December 2017.

Task 5.1 (and its associated deliverable D.5.1) is intimately linked with tasks 5.2 and 5.3. Task 5.2 will be carried out between 1st November 2017 and 30th April 2018 (M6-12) of the EnTIRE project. Task 5.3 will occupy WP5 between 1st May 2018 and 30th April 2020 (M12-36).

Task 5.2 will put this protocol into practice by conducting pilot searches for RE+RI cases to see whether the search methodologies detailed in this protocol are adequate and feasible.

Furthermore, the tagging system detailed in this protocol will be adjusted according to the results of the pilot searches, the development of the structures and interface for the Wiki platform and input from the initial rounds of WP2's stakeholder consultation. It follows that, where necessary, adjustments will be made to this protocol during task 5.2, with details of, and justifications for, any amendments to the protocol provided in the sub-report that will be produced by 30th April 2018 in order to coincide with the conclusion of task 5.2.

The report for task 5.2 ('**Report on Task 5.2 Pilot Collection of Data on Cases**') should be read in conjunction with this protocol. It provides an overview of the pilot searches for Research Ethics and Research Integrity cases. The aim of this report is to make explicit and justify any substantial changes that have been made to the D.5.1 protocol during months 6 to 12 of the EnTIRE project.

Subsequently, as part of task 5.3, WP5 will conduct full-scale systematic searches for RE+RI cases between 1st May 2018 and 30th April 2020. At the same time, cases will be tagged according to the system finalised during task 5.2. Details of the full-scale systematic searches for RE+RI cases will be provided in the search report that will be produced by 30th April 2020 in order to coincide with the conclusion of task 5.3. As part of tasks 5.2 and 5.3, the first RE+RI cases will be uploaded to the online platform by 31st October 2018 (D.5.2).

3.1 Timeline for WP5 Tasks 5.1, 5.2 and 5.3

Description	Timeframe/Deadline
D.5.1: Protocol for systematic searches and tagging of RE+RI cases [M7]	30th November 2017
T.5.2. Pilot collection of data on cases (M6-12, DCU, UNIDEB, Manchester) 4. Conduct pilot searches in IE, HU and UK within each of the potential pools of RE+RI cases to see whether the initially chosen search methodologies are adequate and feasible. Make adjustments, if necessary. 5. Conduct structuring, posting and testing activities on the web platform.	1st November 2017 - 30th April 2018

6. Adjust the system of tagging based on the results of the pilot searches and the normative framework defined in the focus groups (WP2).	
Sub-Report for Task 5.2	30th April 2018
T.5.3. Scale up: Collecting and categorising RE+RI cases (M12-36, DCU, UNIDEB, VUmc) 7. Conduct full-scale systematic searches for RE+RI cases. 8. Add the RE+RI cases that have resulted from the face to face and the online focus groups sessions 9. Work through the RE+RI cases and tag them	1st May 2018 – 30th April 2020
D.5.2: Delivery of the first RE+RI cases as input for the platform [M18]	31st October 2018
Sub-Report for Task 5.3	30th April 2020

4. Abstract of Key Steps: A Practical Guide for Conducting Searches

This section provides a synopsis of the key steps involved in searching for and retrieving RE+RI cases. It also contains details of the preliminary processes involved in tagging cases. This section is meant as a quick, practical guide for WP5 members undertaking tasks 5.2. and 5.3. However, researchers searching for and tagging RE+RI cases should use this guide along with the pertinent sections of the protocol where full details and justifications are provided.

4.1 Timeframe: 1st Nov 2017 – 30th April 2018 (M6-12)

Task 5.2 will be conducted within this timeframe:

Task 5.2. Pilot collection of data on cases (M6-12, DCU, UNIDEB, Manchester)

4. Conduct pilot searches in IE, HU and UK within each of the potential pools of RE+RI cases to see whether the initially chosen search methodologies are adequate and feasible. Make adjustments, if necessary;
5. Conduct structuring, posting and testing activities on the web platform;
6. Adjust the system of tagging based on the results of the pilot searches and the normative framework defined in the focus groups (WP2).

To coincide with task 5.2, those conducting searches will store a list of references and details of retrieved cases within a shared database. The database will function as an index of available cases. These details will allow for cases to be sorted/filtered according to 'basic' criteria such as source (including web URLs, DOIs and standard academic citations), number of words, date, name and affiliations of parties involved.

4.1.1 Academic Literature

We will use this timeframe to carry out a pilot collection of data from academic literature based on a refined version of the Cochrane systematic review. The protocol requires that we:

I. Conduct a scoping search;
II. Create and refine a set of search terms using Boolean operators – one set of search terms should include 'research ethics' and another search should incorporate the term 'research integrity OR research misconduct';
III. Choose the bibliographic and citation databases to search;

IV. Search the databases;
V. Make adjustments to the methodology based on the results of the pilot search;
VI. Log search histories and input basic features of the extracted cases using appropriate data management software.
VII. Report the pilot search providing justifications for search terms and bibliographic databases and details of amendments to the protocol.

4.1.2 RE+RI Committees and Regulatory Bodies

In terms of gathering cases from RE+RI committees and regulatory bodies, we will use this timeframe to contact the organisations detailed in appendices 1, 2 and 3. The protocol for contacting these organisations includes:

I. Preparing an email template in English;
II. Sending an email to the organisations (University of Debrecen to contact Research Ethics Committees and Dublin City University to contact Research Integrity Offices and regulatory bodies);
III. In cases where no response is received, sending a reminder after two weeks;
IV. In cases where no response is received, sending a letter after two weeks;
V. In cases where no response is received, and drawing upon the EnTIRE project's institutional partners and individual contacts, an email will be drafted for our contacts to send to known individuals working within the organisation;
VI. Log the history of each correspondence using appropriate data management software;
VII. Where access to RE+RI cases is granted, an initial list of cases including their basic features will be produced using appropriate data management software;
VIII. Report the search providing details of amendments to the protocol.

4.1.3 Grey Literature

We will use this timeframe to carry out a pilot collection of data from grey literature based on a refined version of the Cochrane systematic review. The protocol requires that we:

I. Conduct a scoping search;
II. Create and refine a set of search terms using Boolean operators – one set of search terms should include 'research ethics' and another search should incorporate the term 'research integrity OR research misconduct';
III. Choose the grey literature databases and repositories to search;
IV. Search the databases;
V. Make adjustments to the methodology based on the results of the pilot search.
VI. Log search histories and input basic features of the extracted cases using appropriate data management software;
VII. Report the pilot search providing justifications for search terms and grey literature databases and details of amendments to the protocol.

4.1.4 Media Outlets

We will use this timeframe to search for news items using the pertinent databases identified in section 7 in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

4.1.5 The Blogosphere

We will use this timeframe to search the blogs identified in section 7 in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic

case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

4.1.6 Online Repositories

We will use this timeframe to search the online repositories identified in section 7 in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

4.1.7 WP2 Stakeholder Consultation

According to the WP2 Protocol, the outputs from the face-to-face focus groups have been timed to arrive before we conclude our initial data collection processes in task 5.2.

Based on the official transcripts of the face-to-face focus groups, EnTIRE researchers will:

I. Store the official transcripts using appropriate data management software;
II. Input basic features of cases using appropriate data management software;
III. Report the search with details of any amendments to the protocol.

4.1.8 Consolidation

After conducting the pilot searches, gathering references and cases and registering basic case features, we will compile a shared master document that contains all the references, associated cases and basic case information.

4.2 Timeframe: 1st May 2018 – 30th April 2020 [M12-36]

Task 5.3 will be conducted within this timeframe:

Task 5.3. Scale up: Collecting and categorising RE+RI cases (M12-36, DCU, UNIDEB, VUmc, EUREC)

7. Conduct full-scale systematic searches for RE+RI cases;
8. Add the RE+RI cases that have resulted from the face to face and the online focus groups sessions;
9. Work through the RE+RI cases and tag them.

To coincide with task 5.3, those conducting searches will store a list and details of retrieved cases within a shared database. The database will function as an index of available cases. These details will allow for cases to be sorted/filtered according to 'basic' criteria such as source (including web URLs, DOIs and standard academic citations), number of words, date, name and affiliations of parties involved. At this stage, the database will be expanded to include a register of keywords for each individual case based on the tagging system proposed in section 8.

4.2.1 Academic Literature

In this timeframe, a full-scale systematic search for, and retrieval of, RE+RI cases is planned:

I. Utilising the search terms and bibliographic and citation databases finalised in task 5.2, search the databases (one search should include the term 'research ethics' and another search should incorporate the term 'research integrity OR research misconduct');
II. Export references to bibliographic software;
III. Identify and remove duplicates (literature that expands upon or offers a different perspective of an already-identified case will not be classed as a duplicate);

IV. Read the literature in order to identify RE+RI cases that match the definition given in section 5;

V. At the same time as step IV, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;

VI. Produce a refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

VII. Search for publications that cite the literature on the refined list but are not included on the list;

VIII. Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;

IX. At the same time as step VIII, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;

X. Produce another refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

XI. Conduct a meta-analysis on the literature detailed in the refined list of references;

XII. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;

XIII. Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and bibliographic databases.

4.2.2 RE+RI Committees and Regulatory Bodies

Having received the cases from the various organisations during task 5.2, the 12-36-month timeframe will be concerned with evaluating the content of cases and attributing keywords as the primary steps of the tagging process:

I. Comparing the cases with those provided through the other sources listed within this protocol, duplicates will be identified and removed (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
II. The cases will be read to identify those that match the definition given in section 5;
III. At the same time as step II, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;
IV. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
V. The search will be reported, providing justifications for the exclusion of identified cases.

4.2.3 Grey Literature

In this timeframe, a full-scale systematic search for, and retrieval of, RE+RI cases is planned:

I. Utilising the search terms and grey literature databases finalised in task 5.2, search the databases (one search should include the term 'research ethics' and another search should incorporate the term 'research integrity OR research misconduct');
II. Export references to bibliographic software;
III. Identify and remove duplicates (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);

IV. Read the literature in order to identify RE+RI cases that match the definition given in section 5;

V. At the same time as step IV, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;

VI. Produce a refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

VII. Additional search for publications that cite the grey literature on the refined list but are not included on the list;

VIII. Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;

IX. At the same time as step VIII, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;

X. Produce another refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

XI. Conduct a meta-analysis on the literature detailed in the refined list of references;

XII. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;

XIII. Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and grey literature databases.

4.2.4 Media Outlets

In this timeframe, we will read each news item detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply

with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

4.2.5 The Blogosphere

In this timeframe, we will read each blog entry detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

4.2.6 Online Repositories

In this timeframe, we will read each RE+RI case detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

4.2.7 WP2 Stakeholder Consultation

According to the WP2 Protocol, the outputs from the face-to-face focus groups have been timed to arrive before we conclude our initial data collection processes in task 5.2 whilst the outputs from the online focus groups will arrive before we begin the scale-up phase of our data collection in task 5.3.

In months 6-12, researchers on WP5 will have used the official transcripts in order to extract the pertinent information from the face-to-face focus group sessions. Consequently, WP5 researchers will have stored the official transcripts using appropriate data management software and input basic features of cases.

During months 12-36, researchers on WP5 will use the information provided by WP2 from the online focus groups in order to:

I. Store the information provided by WP2 using appropriate data management software;
II. Input basic features of cases using appropriate data management software.

Subsequently:

I. The cases will be read to identify those that match the definition given in section 5;
II. At the same time as step I, we will produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, we will remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;
III. We will produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
IV. The search will be reported, providing justifications for the exclusion of identified cases.

5. Cases

The tasks of this protocol include identifying potential sources of RE+RI cases and developing appropriate systematic methods to conduct searches within these sources so as to gather RE+RI cases.

The proposal submitted to the European Commission provides the means to delineate the concept of a case as it relates to the EnTIRE project:

> Confidence in research is severely undermined by evidence of violations, misbehaviours and poor judgement. In recent years, high profile cases of scientists falsifying and fabricating data have put a spotlight on researcher behaviour and damaged public confidence in research findings. Thankfully, serious violations – such as falsification, fabrication and plagiarism (FFP) - are relatively rare, with an estimated 1 to 2% of scientists engaged in such practices. Less serious issues, generally known as questionable research practices (QRPs) - such as inappropriate authorship or research design and analysis - however, are more prevalent (undertaken by approximately a third of scientists) and arguably have a greater impact on the research process (Widdershoven et al. 2015, 13).

The proposal makes it clear that a case involves some sort of 'violation', 'misbehaviour', 'poor judgment' or 'questionable research practice'. It begs the question of what has been violated or how a research practice can be considered to be questionable. With the aim of making the normative frameworks that governs RE+RI easily accessible, the EnTIRE project suggests that violations, misbehaviours, poor judgments and questionable research practices can only be considered as such in relation to these normative frameworks. In turn, a framework 'consists of explicit rules, formulated in laws, regulations, codes, and guidelines, and implicit rules, which structure local RE+RI practice, and influence the application of explicitly formulated rules' (ibid., 3). Therefore, if a case, as it relates to RE+RI, involves a violation of, or misbehaviour, poor judgment or questionable research practice in relation to, a normative framework, then it does so on the basis of explicit and/or implicit rules governing RE+RI practice.

Utilising the work of Steneck (2006), the EnTIRE proposal posits a link between research ethics and the implicit aspects of the normative framework as well as between research integrity and the framework's explicit elements. Whereas 'research ethics is described as the "critical study of the moral problems associated with or that arise in the course of pursuing research"', 'research integrity is said to entail "possessing and steadfastly adhering to professional standards, as outlined by professional organisations, research institutions and, when relevant, the government and public"' (Widdershoven et al. 2015, 4). It follows that research integrity focuses on the *explicit* professional standards of scientific research *tout court*. That said, just as research ethics is also governed by explicit rules in the sense that the latter come face-to-face with implicit rules in the determination of local RE+RI practices, research integrity 'entails clear norms which should be obeyed (for example, the rules concerning scientific misconduct)' (ibid.). It follows that 'any effort at mapping the RE+RI normative framework must cover not just the [explicit] regulations that promote RE+RI, but also the [implicit] values and norms developed in practice and through processes of deliberation regarding responsible conduct of research' (ibid.).

Based on the fact that 'the normative framework…will focus on general values, norms and regulations in both RE and RI, as well as their mutual interdependence' (ibid), RE+RI cases are much more than just violations of, or questionable practices in relation to, explicit rules and regulations, they also consist of violations of, or misbehaviours, poor judgments or questionable research practices in relation to, implicit rules, standards and practices, that is, 'the [implicit] normative views and experiences of stakeholders and learning processes during the application of rules and regulations in research and evaluation practices' (ibid., 6-7). Indeed, as well as considering the application of explicit rules, these normative views and experiences, as already mentioned, concern both the moral challenges and implicit rules local rules governing research practices.

What can also be inferred from the EnTIRE proposal is that an RE+RI case is not solely a violation of, or questionable practice in relation to, the normative framework. Discussing the Office for Research Integrity's (ORI) online repository of misconduct case summaries, the proposal states

that the 'naming and shaming' approach, 'whilst making interesting reading, does not engender a discussion on RE+RI nor support a community of interest, whereas the negative content can repel rather than attract stakeholders' (ibid., 13). Indeed, one of the aims of developing an innovative, user-friendly open access and open source Wiki-platform as part of the EnTIRE project is to increase awareness 'of cases and scenarios embodying best practices' and create 'preconditions for scientific excellence' (ibid.,). As a result, the definition of a RE+RI case can be expanded to include acts demonstrating best practice and/or scientific excellence in relation to the normative framework governing RE+RI.

The question now turns to whether cases should be fictional or non-fictional, that is, should cases include hypothetical scenarios or should they solely be made up of time-tethered, verifiable and concrete acts, practices and experiences? In terms of the EnTIRE project's proposal, task 5.4 of WP5 demands that we 'develop a set of RE+RI scenarios for educational purposes and in order to stimulate strategic thinking about RE+RI'. Bearing in mind that RE+RI scenarios created for educational purposes are almost always fictional, the need to develop such scenarios later in the project is a good enough reason to include hypothetical RE+RI cases as well as non-fictional ones.

In what format should cases be presented? Although, as Donna Zucker (2009) has observed, 'the terms "case study", "case review" and "case report" are used loosely in the scientific and professional literature', it is clear that if we conceive a 'case study' as having 'scientific credentials' in terms of an 'evidence base for professional applications', a 'case review' as emphasising 'a critical reappraisal of a case' and a 'case report' as 'a summary of a case' or a 'document reporting a case, as in case law or medicine' (Zucker 2009, 2), then we are faced with documents that 'state the facts', as it were, with little or no critical engagement and those, such as case reviews, that necessarily involve analysis, judgment and interpretation. However, as anthropologists, researchers in the political, cultural and social sciences and those engaged with hermeneutics have observed, the problem with the simple reporting of facts is that descriptions of acts, practices and experiences intrinsically contain (implicit) interpretations, judgments and analyses. Furthermore, as Zucker observes, case studies in health care research 'often involve in-

depth interviews with participants and key informants, reviews of the medical records, observation, and excerpts from patients' personal writings and diaries' (ibid.). Indeed, it is within the remit of task 5.1 to 'embody implicit elements of the normative framework' (Widdershoven et al. 2015, 16). We will do this by engaging with WP2's stakeholder consultation, collecting cases that elucidate implicit 'experiences and cases from practice' (ibid., 16), that is, the 'normative experiences of stakeholders and learning processes during the application of rules and regulations' (ibid., 10). The fact that the EnTIRE project is predicated on such a 'bottom-up, participatory approach' (in order to make explicit the implicit experiences of individual and group dealings with RE+RI cases) means that we must include analyses of cases as well as 'mere' reports as a matter of necessity. It is also the case that cases considered by Research Ethics Committees and Regulatory Bodies are primarily evidenced by interviews, correspondence and observations, all of which will necessarily contain accounts of individual and group experiences determined by interpretations, analyses and judgments. Consequently, it is misleading to claim that RE+RI cases exist in documented form utterly devoid of critical engagement with the matters of a case. It follows that in order to do justice to task 5.1, cases will be gathered from 'fact-stating' case studies and case reports as well as from analysed and interpreted reviews and interviews.

Based on the demands of the EnTIRE project's proposal, the following definition of a case will be used for the basis of WP5's data collection:

A case:

1) Is either a fictional or non-fictional 'violation' of, or 'misbehaviour', 'poor judgment' or 'questionable research practice' in relation to, or act demonstrating best practice and/or scientific excellence relative to, the explicit and/or implicit rules governing RE+RI, whereby the explicit rules are formulated in laws, regulations, codes, and guidelines and the implicit rules are the normative views and experiences of stakeholders and learning processes during the application of explicit rules and regulations; AND

2) Is presented in documented form in the mode of ‘fact-stating’ case studies and case reports and interpreted, analysed and judged case reviews and interviews;
3) A case can be in any language. However, bearing in mind the resources available to WP5, only cases in English will be tagged. As the cases will be presented on the online platform in the form of a case name, basic case information, URL/hyperlink or academic citation to the case source on third-party websites/texts and a list of appropriate tags, cases not in English may still be present on the online platform in the form of a URL/academic citation. Taking into account the user-oriented, collaborative approach to the EnTIRE project as a whole, those cases not in English will be open to being tagged by the end users of the online platform.

Due to the fact that we will ‘use the outcome of the face-to-face focus groups to assess the relevant elements of the normative framework and define the boundaries of the data to be collected’ (ibid., 28), the definition of a case (as it relates, specifically, to the *implicit rules* governing RE+RI) is contingent upon the findings of the stakeholder consultation. As a result, the meaning of an RE+RI case as posited in this section should be considered as a ‘working definition’, one that is subject to change based on the deliverables, findings and results of other EnTIRE Work Packages.

6. Sources

The EnTIRE project's proposal identifies several categories of sources for the collection of cases:

a) Academic Literature;
b) Reports by RE+RI Committees and Regulatory Bodies;
c) Grey Literature such as Government Documents, White Papers, Theses and Dissertations, Conference Proceedings and Policy Statements;
d) Media Outlets such as Newspapers and Magazines;
e) The Blogosphere; and
f) WP2's Focus Group Sessions.

In addition to each of these categories, for which individual methodologies for searching for cases are detailed below, we have also included 'Online Repositories' as a separate entity. During the course of our scoping exercises, we identified organisations, some of which fall under one or more of the categories 'a' to 'e', that have already made RE+RI cases publicly available in single online repositories. Rather than include such organisations within the above categories, whose methodologies may not be applicable to online repositories, we have produced pertinent and more appropriate methodologies on the basis of their individual sources for RE+RI cases.

The sources we have identified for searching for and retrieving cases are:

a) Academic Literature;
b) Reports by RE+RI Committees and Regulatory Bodies;
c) Grey Literature;
d) Media Outlets;
e) The Blogosphere;
f) Online Repositories; and
g) WP2's Focus Group Sessions

7. Systematic Methods to Conduct Searches for RE+RI Cases

7.1 Academic Literature

The best method for conducting searches for RE+RI cases in academic literature is to use the *structure* of a refined systematic review. According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.1.0), a systematic review 'reduces the impact of authors' biases, promotes transparency of methods and processes, reduces the potential for duplication, and allows peer review of the planned methods' (Green and Higgins eds. 2011). The 'gold standard' for systematic reviews is detailed in the *Cochrane Handbook*, produced by The Cochrane Collaboration. The *Handbook* instructs researchers on the method of full-scale systematic reviews, which, as the Cochrane Collaboration acknowledges, can take several years to produce. Due to the time constraints of WP5, the *Cochrane Handbook* has been used to produce a refined systematic review for searching for and retrieving RE+RI cases in the academic literature. The systematic review structure is derived from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses ('PRISMA') guidelines, an example of which is provided in the *Cochrane Handbook* (see Figure 1). It is structured as follows:

- Conduct a scoping search;
- Create and refine a set of search terms using Boolean operators (two parallel searches will be undertaken, one utilising the search term 'research ethics', the other will include 'research integrity OR research misconduct');
- Choose the bibliographic and citation databases to search;
- Search the databases and export references to bibliographic software;
- Identify and remove duplicates (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
- Read the literature in order to identify RE+RI cases that match the definition in section 5;
- Remove references that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

- Produce a refined list of references using appropriate data management software;
- Additional search for publications that cite the literature on the refined list but are not included on the list;
- Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;
- Remove references that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;
- Produce another refined list of references using appropriate data management software;
- Conduct a meta-analysis on the literature detailed in the refined list of references;
- Produce a list and details of RE+RI cases using appropriate data management software;
- Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and bibliographic databases.

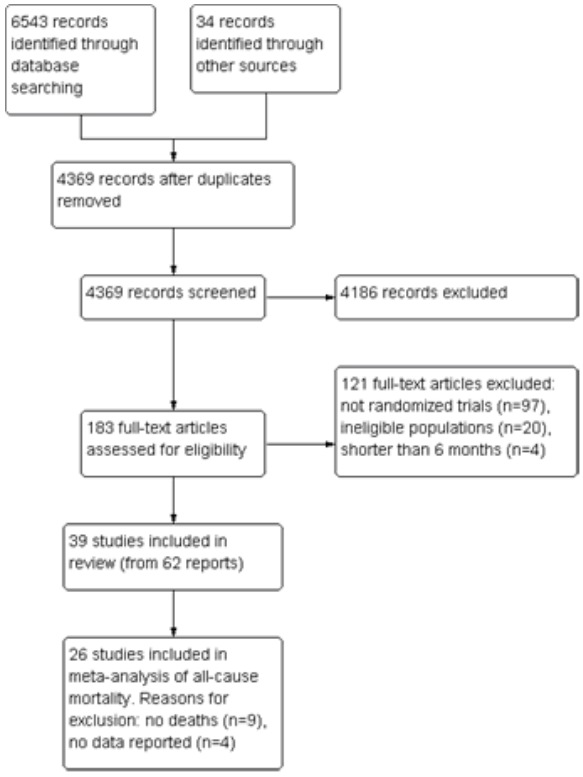

Figure 1. Systematic Review Flow Diagram (reproduced from the *Cochrane Handbook*, Fig. 11.2a)

Applying this structured systematic review to the tasks of WP5, the steps of the review can be rationalised in the following way:

*Task 5.2. Pilot collection of data on cases (M6-12, **DCU**, UNIDEB, Manchester)*

I. Conduct a scoping search;
II. Create and refine a set of search terms using Boolean operators (two parallel searches will be undertaken, one utilising the search term 'research ethics', the other will include 'research integrity OR research misconduct');
III. Choose the bibliographic and citation databases to search;
IV. Search the databases;
V. Make adjustments to the methodology based on the results of the pilot search;
VI. Log search histories and input basic features of the extracted cases using appropriate data management software.
VII. Report the pilot search providing justifications for search terms and bibliographic databases and details of amendments to the protocol.

*Task 5.3. Scale up: Collecting and categorising RE+RI cases (M12-36, **DCU**, UNIDEB, VUmc)*

I. Utilising the search terms and bibliographic and citation databases finalised in task 5.2, search the databases (two parallel searches will be undertaken, one utilising the search term 'research ethics', the other will include 'research integrity OR research misconduct');
II. Export references to bibliographic software;
III. Identify and remove duplicates (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
IV. Read the literature in order to identify RE+RI cases that match the definition given in section 5;

V. At the same time as step IV, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

VI. Produce a refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

VII. Search for publications that cite the literature on the refined list but are not included on the list;

VIII. Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;

IX. At the same time as step VIII, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

X. Produce another refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

XI. Conduct a meta-analysis on the literature detailed in the refined list of references;

XII. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;

XIII. Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and bibliographic databases.

7.1.1 Preliminary Risks

The objective of the systematic review method is to limit bias. Consequently, a comprehensive literature search is required to identify as much of the relevant literature as possible. The role of multiple database searching is twofold: firstly, it broadens coverage to include additional sources; secondly, it takes advantage of differences in indexing across databases to increase the chances

of retrieving relevant items. The bibliographic and citation databases to be used to search for and retrieve RE+RI cases will likely include:

- Academic Search Complete
- Biosis Citation Index
- CINAHL
- Directory of Open Access Journals
- EMBASE
- Google Scholar
- Heinonline
- HUDOC
- Justis
- Just Cite
- JSTOR
- Legal Source
- LexisNexis
- Lexis Library (Legal)
- MEDLINE (Ovid MEDLINE)
- PhilPapers
- Project Muse
- PsycINFO
- PubMed
- ScienceDirect
- SciFinder
- Scopus
- Social Services Abstracts
- Web of Science (incorporating Conference Proceedings Citation Index, Science Citation Index, Social Sciences Citation Index, Arts & Humanities Citation Index and Book Citation Index)

- Westlaw
- WorldCat

This list of bibliographic and citation databases is indicative of the resources WP5 will use to search for and retrieve RE+RI cases. They are listed here as they are considered to be the most widely-used and frequently-cited databases in academic research associated with the aims, purposes and content of this protocol. The actual resources we will use in the full-scale systematic search during task 5.3 will be determined by our pilot case searches during task 5.2. Justifications for the databases used during the systematic search will be provided in the sub-reports that emerge from tasks 5.2 and 5.3.

As several commentators have noted, dealing with the overlap between key bibliographic and citation databases is standard practice when it comes to a systematic review.[1] For example, there is approximately 85% overlap between EMBASE and MEDLINE/PubMed, while PubMed is slightly more up-to-date compared to MEDLINE. Of the 4,800 journals indexed in EMBASE, 1,800 are not indexed in MEDLINE. Similarly, of the 5,200 journals indexed in MEDLINE, 1,800 are not indexed in EMBASE. Differences in coverage vary with particular disciplines, with EMBASE indexing more European journals and more non-English journals. Scopus covers all the citations included in PubMed/MEDLINE and EMBASE from 1966 and some older PubMed citations from 1949-1965. Web of Science will cover some publications missed by Scopus, EMBASE and PubMed/MEDLINE. While there is a considerable degree of overlap between Web of Science, Scopus, EMBASE and PubMed/MEDLINE, there will often be articles found in one database that are not found in the other using the same search. However, when searching additional databases with overlapping coverage but fewer precision-enhancing features, one may encounter the problem of reintroducing irrelevant material that has already been eliminated from the retrieval in the database with the fullest feature set. Nevertheless, studies comparing searches of different databases have generally concluded that a comprehensive search requires that all key databases

[1] See, for example, Hood and Wilson (2003), Sampson et al. (2006), Yiva and Iselid (2008), Michigan State University (2017), University Library (2017) and The Francis A. Countway Library of Medicine (2017).

be searched (Suarez-Almazor et al. 2000). One may find that relevant items may be missed in one database when assigned indexing terms different than those used by the searcher only to find the same records retrieved from another database because the indexing in that database matches the terms selected by the searcher.

According to the EnTIRE project's proposal another potential risk 'is the number of RE+RI cases that are found, which may be either too small or too large. If the number is too small, further search strategies will be applied, making use of expert knowledge of partners and other RE+RI leaders. If the number is too large, methods of limiting the number of cases will be deployed in the search strategies such as exclusively focusing on a recent time period' (Widdershoven et al. 2015, 46). Baseline numbers and methods for including and excluding cases will be made explicit and justified in the search report.

7.2 Reports by RE+RI Committees and Regulatory Bodies

For the purposes of this study, a regulatory body is a government/public institution concerned with the governance of research ethics and/or research integrity. Such a definition includes (within this task's search parameters) regulatory bodies that are explicitly concerned with the governance of RE+RI whilst removing organisations that are only tangentially involved in RE+RI matters. As a government body mandated under the terms of a legislative act to ensure compliance with the act in the areas of research ethics and/or integrity, a regulatory body will either form its own Research Ethics Committees or set itself up as a formal conduit for applications to appropriate regional or local Research Ethics Committees and/or Research Integrity Offices. For example, when it comes to health research involving human participants, the UK's Health Research Authority – a Non-Departmental Public Body mandated by the Care Act (2014) – not only provides a national system for the governance of health research and its ethics, it acts as a formal gateway to NHS Research Ethics Committees.

Were we to search for and retrieve cases from Research Ethics Committees ('RECs'), Research Integrity Offices ('RIOs') and regulatory bodies on an international, regional and local scale, the task would be impossible within the timeframes of WP5. Consequently, as the EnTIRE project's proposal demands that WP3-5 'assemble the relevant normative elements, including RE+RI rules and procedures, educational materials, and illustrative casuistry, and identify relevant institutions *across EU countries*' [italics added], task 5.1 of WP5 will confine its search to RECs, RIOs and regulatory bodies within the EU.

7.2.1 Research Ethics Committees

In order to emphasise the delineation of research ethics cases and research integrity cases (something that has also been built into the methodologies for retrieving cases from academic literature and grey literature), the University of Debrecen will carry out searches for cases from Research Ethics Committees.

The European Network of Research Ethics Committees ('EUREC') brings together national REC associations, networks and relevant bodies, the Forum of National Ethics Committees ('NEC Forum') and the European Commission's Ethical Review System to form the necessary infrastructural basis to promote awareness of specific working practices of RECs across Europe, to enhance the shared knowledge base of European RECs, to support coherent reviews and the development of opinions and to meet new challenges and emerging ethical issues.

EUREC provides the most comprehensive and up-to-date (yet open and evolving) publicly available list of RECs in Europe.

EUREC (http://www.eurecnet.org) implicitly distinguishes or explicitly names over 1,100 Research Ethics Committees in Europe. There are also a significant number of RECs that are not explicitly numbered or detailed by EUREC. These tend to be local RECs housed in health care institutions, academic institutions and research centres.

Details of the RECs numbered or named by EUREC can be found in **APPENDIX 1**.

Taking into account the details provided in Appendix 1, each of the national, regional and local RECs in Europe answer to state-specific 'parent' organisation. As EUREC demonstrates, not only do these national bodies act as 'umbrella' organisations for regional and local RECs, laying down guidelines for the appropriate conduct of individual RECs, coordinating dialogues and partnerships between RE+RI committees, auditing the practices of RECs, receiving case details and considering appeals against localised case judgments, they also provide a focal point for enquiries regarding RECs within a particular country. Rather than retrieve cases from individual RECs, many of which do not have details provided by EUREC, WP5 will engage the 'umbrella' organisations listed in Appendix 1 in order to make the task manageable within the timeframes of the EnTIRE project.

The University of Debrecen's process of retrieving cases will be structured as follows:

1. Prepare an email template in English;
2. An initial email (in English) will be sent to each of the 'umbrella' organisations detailed in Appendix 1, providing details of the EnTIRE project and its plans for formatting, tagging disclosing and disseminating RE+RI cases. Access to their RE+RI cases will be requested and their agreement sought on the basis of their knowledge about our plans for making the cases publicly available;
3. Where no correspondence is forthcoming, a second email will be sent two weeks after the date of the first email;
4. If there is no response to the second email, a letter will be sent two weeks later;
5. In the event that we do not receive a reply, a final effort will be made to contact the organisation (in the appropriate local language where possible). Drawing upon the EnTIRE project's institutional partners and individual contacts, an email will be drafted for our contacts to send to known individuals working within the 'umbrella' organisation. We will request that the email, which will contain the same details as our initial correspondence, is forwarded from the known contact in the 'umbrella' organisation to the person(s) responsible for granting access to RE+RI cases;
6. Log the history of each correspondence using appropriate data management software;
7. Where access to RE+RI cases is granted, an initial list of cases, including their basic features, will be produced using appropriate data management software;
8. Comparing the cases with those provided through the other sources listed within this protocol, duplicates will be identified and removed (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
9. The cases will be read to identify those that match the definition given in section 5;
10. At the same time as step IX, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

11. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
12. The search will be reported, providing justifications for the exclusion of identified cases.

7.2.2 Research Integrity Offices

In order to emphasise the delineation of research ethics cases and research integrity cases (something that has also been built into the methodologies for retrieving cases from academic literature and grey literature), Dublin City University will carry out searches for cases from Research Integrity Offices

The European Network of Research Integrity Offices ('ENRIO') is an informal network that aims to enhance research integrity. It contains 28 member organisations within 22 European countries. The current members of ENRIO are all national offices or national organisations. Contact details for each member organisation are provided on the ENRIO website.

ENRIO runs parallel to EUREC. Together, ENRIO and EUREC are responsible for initiating the European Network of Research Ethics and Research Integrity ('ENERI'), which brings together the key players in Europe involved in research ethics and research integrity, allowing for more effective exchange of information between and among RECs and RIOs.

As is the case with EUREC, ENRIO provides the most comprehensive and up-to-date (yet open and evolving) publicly available list of RIOs in Europe.

ENRIO explicitly details 28 RIOs in Europe. Details of the RIOs can be found in **APPENDIX 2**.

Dublin City University's process of retrieving cases will be structured as follows:

1. Prepare an email template in English;

2. An initial email (in English) will be sent to each of the organisations detailed in Appendix 2, providing details of the EnTIRE project and its plans for formatting, tagging disclosing and disseminating RE+RI cases. Access to their RE+RI cases will be requested and their agreement sought on the basis of their knowledge about our plans for making the cases publicly available;
3. Where no correspondence is forthcoming, a second email will be sent two weeks after the date of the first email;
4. If there is no response to the second email, a letter will be sent two weeks later;
5. In the event that we do not receive a reply, a final effort will be made to contact the organisation (in the appropriate local language where possible). Drawing upon the EnTIRE project's institutional partners and individual contacts, an email will be drafted for our contacts to send to known individuals working within the 'umbrella' organisation. We will request that the email, which will contain the same details as our initial correspondence, is forwarded from the known contact to the person(s) responsible for granting access to RE+RI cases;
6. Log the history of each correspondence using appropriate data management software;
7. Where access to RE+RI cases is granted, an initial list of cases, including their basic features, will be produced using appropriate data management software;
8. Comparing the cases with those provided through the other sources listed within this protocol, duplicates will be identified and removed (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
9. The cases will be read to identify those that match the definition given in section 5;
10. At the same time as step IX, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;
11. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
12. The search will be reported, providing justifications for the exclusion of identified cases.

7.2.3 Regulatory Bodies

In order to emphasise the delineation of research ethics cases and research integrity cases (something that has also been built into the methodologies for retrieving cases from academic literature and grey literature), Dublin City University will carry out searches for cases from regulatory bodies.

The Office for Human Research Protections (https://www.hhs.gov/ohrp/) within the US Department of Health and Human Services produces an annual International Compilation of Human Research Standards, which enumerates over 1,000 laws, regulations, and guidelines that govern human subjects research in 126 countries, as well as standards from a number of international and regional organisations. This Compilation was developed for use by researchers, Institutional Review Boards/Research Ethics Committees, sponsors and others involved in human-subjects research around the world. These laws, regulations, and guidelines are classified into eight categories:

1. General
2. Drugs and Devices
3. Clinical Trial Registries
4. Research Injury
5. Privacy/Data Protection
6. Human Biological Materials
7. Genetic
8. Embryos, Stem Cells and Cloning

The International Compilation of Human Research Standards does not include standards from state, provincial or local levels. It also does not cover:

- Enabling legislation such as laws that authorise an agency to promulgate human-subjects' standards but do not direct the content of those regulations;
- Laws, regulations or guidelines specific to research integrity, clinical bioethics, product liability, clinical trial inspection procedures, intellectual property, good manufacturing practice, bioequivalence testing and informed consent in clinical practice;
- Codes of Ethics for academic, medical and professional organisations;
- Working papers, drafts, commentaries and discussion papers.

Based on the *International Compilation of Human Research Standards*, the regulatory bodies in Europe mandated to ensure governance of RE+RI are provided in **Appendix 3**.

The *International Compilation of Human Research Standards* provides the most comprehensive and up-to-date publicly-available, single-source list of European Regulatory Bodies for RE+RI.

Dublin City University's process of retrieving cases will be structured as follows:

1. Prepare an email template in English;
2. An initial email (in English) will be sent to each of the Regulatory Bodies detailed in Appendix 3, providing details of the EnTIRE project and its plans for formatting, tagging disclosing and disseminating RE+RI cases. Access to their RE+RI cases will be requested and their agreement sought on the basis of their knowledge about our plans for making the cases publicly available;
3. Where no correspondence is forthcoming, a second email will be sent two weeks after the date of the first email;
4. If there is no response to the second email, a letter will be sent two weeks later;
5. In the event that we do not receive a reply, a final effort will be made to contact the organisation (in the appropriate local language where possible). Drawing upon the EnTIRE project's institutional partners and individual contacts, an email will be drafted for our

contacts to send to known individuals working within the regulatory body. We will request that the email, which will contain the same details as those produced in for our initial correspondence, is forwarded from the known contact in the regulatory body to the person(s) responsible for granting access to RE+RI cases;

6. Log the history of each correspondence using appropriate data management software;
7. Where access to RE+RI cases is granted, an initial list of cases, including their basic features, will be produced using appropriate data management software;
8. Comparing the cases with those provided through the other sources listed within this protocol, duplicates will be identified and removed (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
9. The cases will be read to identify those that match the definition given in section 5;
10. At the same time as step IX, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;
11. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
12. The search will be reported, providing justifications for the exclusion of identified cases.

7.2.4 Preliminary Risks

It is the vision of the EnTIRE project to make the normative frameworks governing RE+RI publicly available on an online, Wiki-based platform. The aim is to include practical information on how to comply with EU, national and discipline-specific RE+RI standards and legislation, including information on rules and procedures, educational materials and illustrative cases and scenarios. Due to the fact the cases retrieved from RECs and regulatory bodies will be made publicly available via the online platform, our initial correspondence with these organisations needs to make the project's public disclosure of data explicit. Where appropriate, agreement will need to be sought from the organisations granting access to RE+RI cases if their cases have not already

been made publicly available. Consequently, prior to drafting the correspondence to be sent to the RECs and regulatory bodies, advice will be sought from the EnTIRE project's legal representatives in order to ensure that our plans to make third-party RE+RI cases available are legal and that the correct disclosure agreements and legal documents are supplied.

As is the case with academic literature, a potential risk 'is the number of RE+RI cases that are found, which may be either too small or too large. If the number is too small, further search strategies will be applied, making use of expert knowledge of partners and other RE+RI leaders. If the number is too large, methods of limiting the number of cases will be deployed in the search strategies such as exclusively focusing on a recent time period' (Widdershoven et al. 2015, 46). Baseline numbers and methods for including and excluding cases will be made explicit and justified in the search report.

7.3 Grey Literature

Described as 'unconventional literature' (Augur 1998), grey literature encompasses 'that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers' (*Grey Literature Report* 2017).

Grey literature publications may include, but are not limited to, reports (such as pre-prints, preliminary progress and advanced reports, technical reports, statistical reports, memoranda, state-of-the art reports and market research reports), theses, conference proceedings, technical specifications and standards, non-commercial translations, bibliographies, technical and commercial documentation, and official documents not published commercially (primarily government reports and documents) (Alberani 1990). A more comprehensive list of grey literature formats is provided by GreyNet.

As demonstrated by the sheer number of different grey literature formats on GreyNet, it can be difficult to locate pertinent literature and collate what exists in the realm of grey literature in a comprehensive way. Increasingly, material is made available online, but there is still a wide range of grey literature which is not. In addition, a lot of grey literature, such as government reports, is only made available online for a limited period of time. Indeed, some resources, including unpublished conference papers and certain dissertations, may only be available through personal contact with the authors.

Despite the problems of searching for and accounting for the appropriate grey literature in a comprehensive way, the *Cochrane Handbook for Systematic Reviews* (v. 5.1.0) has concluded that failure to identify references and studies reported in conference proceedings and other grey literature might affect the results of a systematic review. It is now standard practice to include grey literature within systematic literature reviews. Consequently, in order to search for and retrieve RE+RI cases from grey literature, WP5 will utilise the structure of the refined systematic review detailed in the 'Academic Literature' section above:

- Conduct a scoping search (two parallel searches will be undertaken, one utilising the search term ‘research ethics’, the other will include ‘research integrity OR research misconduct’);
- Create and refine a set of search terms;
- Choose the grey literature databases and repositories to search;
- Search the databases;
- Export references to bibliographic software;
- Identify and remove duplicates (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
- Read the literature in order to identify RE+RI cases that match the definition given in section 5;
- Remove references that are unsuitable or ‘off topic’ according to criteria that will be made explicit in the systematic review report;
- Produce a refined list of references using appropriate data management software;
- Additional search for publications that cite the grey literature on the refined list but are not included on the list;
- Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;
- Remove references that are unsuitable or ‘off topic’ according to criteria that will be made explicit in the systematic review report;
- Produce another refined list of references using appropriate data management software;
- Conduct a meta-analysis on the literature detailed in the refined list of references;
- Produce a list and details of RE+RI cases using appropriate data management software;
- Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and grey literature databases.

Applying this structured systematic review to the tasks of WP5, the steps of the review can be rationalised in the following way:

*Task 5.2. Pilot collection of data on cases (M6-12, **DCU**, UNIDEB, Manchester)*

I. Conduct a scoping search (two parallel searches will be undertaken, one utilising the search term 'research ethics', the other will include 'research integrity OR research misconduct');
II. Create and refine a set of search terms;
III. Choose the grey literature databases and repositories to search;
IV. Search the databases;
V. Make adjustments to the methodology based on the results of the pilot search;
VI. Log search histories and input basic features of the extracted cases using appropriate data management software;
VII. Report the pilot search providing justifications for search terms and bibliographic databases and details of amendments to the protocol.

*Task 5.3. Scale up: Collecting and categorising RE+RI cases (M12-36, **DCU**, UNIDEB, VUmc)*

I. Utilising the search terms and grey literature databases finalised in task 5.2, search the databases (two parallel searches will be undertaken, one utilising the search term 'research ethics', the other will include 'research integrity OR research misconduct');
II. Export references to bibliographic software;
III. Identify and remove duplicates (texts that expand upon or offer a different perspective of an already-identified case will not be classed as a duplicate);
IV. Read the literature in order to identify RE+RI cases that match the definition given in section 5;
V. At the same time as step IV, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in

section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

VI. Produce a refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

VII. Additional search for publications that cite the grey literature on the refined list but are not included on the list;

VIII. Read the additional literature in order to identify additional RE+RI cases that match the definition given in section 5;

IX. At the same time as step VIII, produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the systematic review report;

X. Produce another refined list of references and associated keywords (to be later used as tags) using appropriate data management software;

XI. Conduct a meta-analysis on the literature detailed in the refined list of references;

XII. Produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;

XIII. Report the search, providing justifications for the exclusion of identified literature/cases and reasons for the use of specific search terms and grey literature databases.

7.3.1 Preliminary Risks

As previously stated, the objective of a systematic review method is to limit bias. Consequently, a comprehensive literature search is required to identify as much of the relevant grey literature as possible. However, taking into account the number of grey literature formats detailed by GreyNet together with the diversity of the literature sources, it would be impossible (within the timeframes of WP5 and the EnTIRE project in general) to undertake a comprehensive search for RE+RI cases by engaging with each and every grey literature format emanating from individual

organisations on international, national, regional and local levels. Indeed, a considerable amount of bias would affect the search methodology were WP5 to narrow its search to specific grey literature sources produced by specific organisations. In order to mitigate methodological bias, WP5 will collate RE+RI cases by confining its search to readily searchable, publicly available online repositories and the most frequently cited, global, multidisciplinary, online grey literature databases.

Details of the pertinent grey literature repositories are included within the 'Online Repositories' section below.

In addition to specific online repositories, the grey literature databases to be used to search for and retrieve RE+RI cases will include:

- BIOSIS Previews
    - BIOSIS Previews includes proceedings of many meetings that may not be electronically available elsewhere.
- Dissertation Abstracts
    - Dissertations Abstracts is the largest single source of information about doctoral dissertations and master's theses available. It contains over 2 million entries. You can search two theses and dissertations databases: Index to Theses (UK and Ireland) and Dissertation Abstracts (international: mainly North American theses, with a smaller selection from around the world).
- DocuTicker
    - DocuTicker collects abstracts from grey literature: PDF reports published by government agencies, think tanks, NGOs, research institutes and other public interest groups.
- Grey Literature Report

    - A bi-monthly publication of the New York Academy of Medicine, the GLR includes listings of recently published reports in health science and public health. The archives are tagged with MeSH terms and are searchable.
- GreyNet
    - Includes a listing of grey literature sources via GreySource.
- NTIS
    - The National Technical Information Service ('NTIS') provides access to the results of both US and non-US government-sponsored research and can provide the full text of the technical report for most of the results retrieved.
- OpenGrey
    - A multidisciplinary database of technical reports, meetings, dissertations and official publications. Largely European in focus, it includes records from biomedicine and other sciences, economics and the humanities.
- OpenSIGLE
    - EAGLE (the European Association for Grey Literature Exploitation), has closed the SIGLE (System for Information on Grey Literature) database, which was one of the most widely-used databases of grey literature. INIST (Institute for Scientific and Technical Information) has launched OpenSIGLE, which provides access to all the former SIGLE records, new data added by EAGLE members and information from GreyNet.
- Policy Archive
    - A digital library of public policy research containing over 30,000 documents.
- PolicyFile
    - Offers public policy reports and studies published by think tanks, university research programs and research organisations including the OECD, IMF, World Bank, the Rand Corporation and a number of federal agencies.
- ProQuest Dissertations and Theses Global
    - Indexes more than 2 million doctoral dissertations and masters' theses worldwide.
- Think Tank Search

    - From the Kennedy School of Government, the tool provides a customised Google search of documents produced by think tanks.
- UNESCO Documents and Publications
    - Education, social and natural science, culture and communication database of bibliographic records which also provides access to the full text of United Nations Educational, Scientific and Cultural Organization (UNESCO) documents, publications and periodicals.
- Web of Science Core Collections
    - Includes Conference Proceedings Citation Indexes.

This list of Grey Literature databases is indicative of the resources WP5 will use to search for and retrieve RE+RI cases. They are listed here as they are considered to be the most widely used and frequently cited databases in academic research associated with the aims, purposes and content of this protocol. The actual resources we will use in the full-scale systematic search during task 5.3 will be determined by our pilot case searches during task 5.2. Justifications for the databases used during the systematic search will be provided in the sub-reports that emerge from tasks 5.2 and 5.3.

Finally, bearing in mind grey literature is made available outside traditional models of publishing, with new findings often appearing long before they are published in peer-reviewed publications, the immediacy of the material does mean that it will most likely not have been through any rigorous quality-assessment or peer-review process.

7.4 Media Outlets

Although believed to publish news concerned only with reports of ethical failure (Davis 1999), media outlets can also be a source for finding additional information such as interviews and commentaries, and help to fill in the gaps about cases that are retrieved from other sources. Furthermore, they can provide information about the extent of exposure given to RE and RI in the news.

We have chosen two sources to search:

- LexisNexis academic database;
- Google News search engine to include content that was only published digitally.

In order to search media outlets, we propose a two-step strategy. First, we will identify relevant keywords based on the scope of the research, and, subsequently, we will make various combinations with them. By juxtaposing mentioned keywords in order to produce compound terms, which we will link via the 'AND' operator; we narrow down our search to relevant results and avoid gathering news articles in other contexts. Our search strings will include:

- "case study" AND "research ethics"
- "case report" AND "research ethics"
- "case study" AND "research integrity"
- "case report" AND "research integrity"
- "case study" AND "research misconduct"
- "case report" AND "research misconduct"
- "case study" AND "research fraud"
- "case report" AND "research fraud"
- "case of" AND "research integrity"
- "case of" AND "research ethics"

- “case of” AND “research misconduct”
- “case of” AND “research fraud”

The table below shows the results of our scoping search. As well as registering the number of hits, we also checked the search results in terms of publication dates.

Search engine	Search strings	Number of hits	comments	Date of search
Google News	“case study” And “research integrity”	16	News items are available from 2012 until 2017	09/08/2017
	“case report” And “research integrity”	13	News items are available from 2010 until 2017	
	“case study” And “research ethics”	87	News items are available from 2010 until 2017	
	“case report” And “research ethics”	42	News items are available from 2013 until 2017	
	“case study” And “research misconduct”	43	News items available from 2012 until 2017	15/09/2017
	“case report” And “research misconduct”	5	News items available from 2013 until 2016	
	“case study” And “research fraud”	8	News items available from 2009 until 2017	
	“case report” And “research fraud”	2	News items available from 2011 until 2017	
	“case of” And “research integrity”	266	News items available from 2004 until 2017	26/10/2017

	“case of” And “research ethics”	442	News items available from 2003 until 2017	
	“case of” And “research misconduct”	152	News items available from 2004 until 2017	
	“case of” And “research fraud”	104	News items available from 2004 until 2017	
LexisNexis	“case study” And “research integrity”	89	News items are available from 1987 until 2017	10/08/2017
	“case report” And “research integrity”	13	News items are available from 1994 until 2017	
	“case study” And “research ethics”	340	News items are available from 1981 until 2017	
	“case report” And “research ethics”	63	News items are available from 1991 until 2017	
	“case study” And “research misconduct”	43	News items are available from 1998 until 2017	15/09/2017
	“case report” And “research misconduct”	9	News items are available from 1995 until 2015	
	“case study” And “research fraud”	18	News items are available from 1987 until 2017	
	“case report” And “research fraud”	7	News items are available from 1987 until 2014	
	“case of” And “research integrity”	990	News items are available from 1987 until 2017	26/10/2017
	“case of” And “research ethics”	996	News items are available from 1985 until 2017	

	“case of” And “research misconduct”	988	News items are available from 1989 until 2017	
	“case of” And “research fraud”	999	News items are available from 1980 until 2017	

A total of 5735 items emerged. During task 5.2 [M6-12], we will search the media databases identified above in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

During task 5.3 [M12-36], we will read each reference detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

7.5 The Blogosphere

Although the blogosphere may contain unwarranted information, due to their participatory nature and interaction with a small target audience, they enable dialogue and co-production in a unique manner (Hookway 2008). As such, the blogosphere, in addition to locating cases that may not be published through more official channels, can help to capture the implicit elements of the normative framework that may not have been mentioned in other sources.

To relevant blogs, we engaged two initiatives, both of which maintain several active blogs in the field:

- Ethics and Integrity in Research provides a list of useful blogs on scientific integrity, academic redaction and publication as well as other aspects of scientific research;
- A complimentary source is the website of the Netherlands Research Integrity Network ('NRIN'), which includes several blogs for accessing information on research ethics and integrity.

The table below provides a list of blogs compiled from these two sources. All the blogs contain cases in English from several disciplines. A short summary is made available about each blog, explaining their approach and highlighting the number of cases that could potentially be extracted.

Blog	Comments
Retractionwatch.com	It provides thousands of cases of misconduct and exemplary practice from different countries. Their database can be searched using different queries. The Retractionwatch platform can be used to find cases of misconduct that ended in retraction of scientific articles, or exemplary cases that resulted in the correction of a previously published paper. ➢ On 19/07/2017 contained information about 3153 retracted scientific articles.

Plagiarismwatch.org	A US-based blog that monitors plagiarism and aims to educate publishers and authors. They receive information from the whistle blowers and present the evidence without passing judgment. ➢ On 10/08/2017, the blog contained 91 cases of plagiarism received from February 2016 until November 2016.
ethics-and-integrity.net	A multilingual blog that provides information in English, German, French and Arabic, with the aim of centralising information and resources. ➢ On 11/08/17, the blog contained 19 historical cases of fraud from 1980 until 2016.
copy-shake-paste.blogspot.com	Initiated and maintained by Debora Weber-Wulff, who is also active in two other German-language blogs, namely "Portal Plagiat" and “VroniPlag Wiki”. ➢ On 10/08/207, the blog contained 191 entries about cases of misconduct or questionable practices since 2006.
Researchethicsblog.com	Initiated and maintained by Nancy Walton, the blog is mainly focused on human subjects in research. ➢ On 09/08/2017, it contained 10 historical cases.
statistically-funny.blogspot.com	Developed and maintained by Hilda Bastian, the blog stores a collection of studies that use statistically unsound methods. ➢ On 09/08/201, it contained 50 cases (2012-2016).

During task 5.2 [M6-12], we will search the blogs identified above in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

During task 5.3 [M12-36], we will read each blog detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

7.6 Online Repositories

Supported and maintained by national and international organisations, case study repositories store a wealth of fictional and non-fictional cases.

Gathering case studies across disciplines, two projects provide a useful list of available case-repositories from different countries:

- The Online Ethics Library project of the Illinois Institute of Technology (detailed in **Appendix 4**);
- The Levan Online Ethics Resource Center of the University of Southern California (detailed in **Appendix 5**).

During our scoping exercise, it was discovered that some of the repositories are no longer useable due to broken links or changes in their public availability policies. A total of 3305 cases emerged from 31 repositories. The table below provides a list of valid repositories including information about the cases:

Repository	Quantity and other features of cases
American anthropological association	Association's aim is to educate members on ethical issues within anthropology. As part of this mission, the Committee on Ethics provides case studies on how to approach ethical dilemmas in the discipline. ➢ On 15/08/17 contained two fictional cases including questions for discussions and comments. ➢ Cases are neither tagged nor categorised.
American medical association – Journal of ethics	The ethics journal of the largest association of physicians and medical students in the US. Based on the monthly theme case and commentary are gathered from experienced physicians and other experts in the field who can help readers think productively about that topic.

	➢ On 15/08/17 contained 1570 cases. Mostly based on historical scenarios but rewritten for publication. Cases are not accompanied with questions but include experts' commentary. ➢ Cases are categorised under several headings. Lots of duplicates and rather chaotic. Cases are neither tagged nor searchable.
American physics society	Society provides a series of case studies on ethical issues that can arise in the course of doing physics research, and intended to be an educational resource for researchers, mentors, and students. ➢ On 14/08/17, contained 39 fictional cases that were specifically designed for doing physics research. Cases come in two versions, one for students and another including possible assignments and discussion guides for teachers. ➢ Cases are classified under 9 topics in a report. Cases are not tagged.
American sociological association	The association has 21000 active members and aims to articulate policy and implement programs to impact sociology on a national and international level. ➢ On 14/08/2017, contained 103 fictional cases on ethical standards that should be followed in research. Cases are followed by questions and discussions. ➢ Cases are neither categorised nor tagged.
Austrian Agency for Research Integrity	The agency is responsible for investigating alleged cases of scientific misconduct in Austria, evaluating the severity of each violation and proposing consequential measures with the first two tasks being assigned to the Commission for Research Integrity. A selection of cases and inquiries from 2009 until 2015 is made available in English. ➢ On 18/08/2017, contained 38 anonymised case summaries incorporated in the annual reports. ➢ Cases are neither categorised nor tagged.
Bioethics outreach - Iowa state university	Cases are developed by the office of biotechnology as a method for improving bioethics outreach.

	➢ On 14/08/2017, contained 45 cases with questions, moral issues and some discussion points highlighted. ➢ In addition to the English cases, there are also two cases in Spanish and two in Portuguese. ➢ Cases are not tagged and are not searchable.
Centre for ethics in professions – University of Puerto Rico	Centre provides cases for different fields of applied ethics aiming to emphasise the ethical dimensions involved in the science based professions. ➢ On 14/08/17, contained 12 cases on research ethics and plagiarism, including discussion questions. ➢ Cases are not tagged and are not searchable.
Center for the Study of Ethics in the Professions - Illinois Institute of Technology	The centre provides a large library of various resources including case studies and codes of conducts. The 'Research ethics case studies' collection provides cases from several disciplines. ➢ On 14/08/2017, contained links to 110 fictional cases that are available in several other online sources. ➢ Cases are tagged and can be searched and filtered based on subject and discipline
Committee on Publication ethics (COPE)	COPE provides advice to editors and publishers on all aspects of publication ethics and, in particular, how to handle cases of research and publication misconduct. The case repository stores anonymised cases of real scenarios that were discussed in COPE forums. ➢ On 18/08/2017, contained 570 cases together with the advice given by COPE. More recent cases include follow-up information and outcome too. ➢ Based on the type of violations or issue at hand, cases are classified under 10 categories. Cases are also tagged and searchable.
CTSPEDIA (A knowledge base for clinical and translational researchers)	CTSPEDIA is a knowledgebase that provides useful resources for clinical and translational researchers. Fictional cases relevant to clinical research are submitted by field experts. ➢ On 14/08/2017, contained 112 short cases, some with and some without questions for discussion.

	➢ Although some other resources are tagged, but cases are neither categorised nor tagged, and therefore, are not searchable.
Danish Committee on Research Misconduct	The Committee handles Danish cases of research misconduct. As a general rule a case is raised via a complaint handed in at the research institution where the research in question was carried out. All anonymised case summaries are available in Danish, including a few in English. ➢ On 18/08/2017, contained 8 case summaries in English. Cases are anonymised and differ in the provided context ➢ Cases are neither categorised nor tagged.
Economic and social research council of the UK	A non-departmental public body for funding research on economic and social issues in the UK. The repository aims to raise awareness about some of the ethical issues that can arise in research. ➢ On 14/08/17 contained 21 research ethics cases. Case studies are inspired by real issues that happen during the process of research, and were rewritten for educational purposes, and highlight the ethical process and lessons to be learned. ➢ Cases are subsumed under several categories but are neither tagged nor searchable.
Hale chair in applied ethics	The Hale Chair in Applied Ethics is devoted to research and teaching in ethics within and about the professional disciplines. ➢ On 15/08/17 contained 23 fictional cases developed for pedagogical purposes, including questions for discussions. ➢ Cases are neither tagged nor categorised.
Markkula Center for Applied Ethics -at Santa Clara University	The repository includes resources of different types on applied ethics across a variety of fields. ➢ On 14/08/2017, contained real fictional cases on educational integrity (25), bioethics (21) and 191 cases from other contexts (e.g. business ethics, journalism ethics, etc.). Cases also include questions for discussions. ➢ Within the categories of educational integrity and bioethics, cases are subsumed under several other categories, but they are not tagged, nor searchable.

National academy of engineering (online ethics centre)	The online repository of the US national academy of engineering, that is regularly updated and well maintained. In addition to engineering cases, the platform provides hundreds of cases on life and environmental science, social and behavioural sciences, computer, maths and physical sciences. ➢ On 14/08/17, contained 565 cases including 117 historical cases, 216 fictional cases, 183 open-ended scenarios, 37 mini-cases, and 12 role-plays. ➢ Cases are tagged and categorised by type of resource, topic, field, author and other publication information. Easy to search and filter based on various criteria.
National Center for Case Study Teaching in Science - State University of New York in Buffalo	A peer-reviewed collection containing cases in all areas of science. Under the 'ethics' topical area, a collection of cases that are all hosted on the platform can be found. ➢ On 14/08/2017 contained 135 curated cases. Most cases include questions for discussions and relevant illustrations/images. ➢ Cases are tagged, filtered and searchable based on several attributes. ➢ More resources including teaching notes and answer keys are under paywall.
National institute of health (NIH)	The sourcebook of NIH provides case studies to use by the medical community to help staff involved in research define or refine their own standards. Each year a theme is specified and, fictional cases about that topic are developed and presented (except for 2005 & 2006). ➢ On 14/08/2017, contained 65 cases on 15 topics. Cases are designed for medical settings and include questions. ➢ Cases are subsumed under each year's theme, but they are not tagged.
National center for professional and research ethics – Ethics CORE project	The repository is the final deliverable of a project that was founded by the US National Science Foundation (NSF) to help the scientific community share ethics resources and work together. ➢ On 14/08/2017, contained 195 fictional cases with a short abstract. Some are hosted on the platform and are

	easy to download as a PDF file; some others are linked to other online sources. ➢ Cases are tagged and the repository can be easily searched.
National science foundation	The US National Science foundation's office of inspector general provides independent oversight and aims at preventing and detecting fraud, waste, and abuse. ➢ On 18/08/2017 contained 3245 records of investigated cases from 1989 until present. Cases are strictly anonymised. ➢ Records are neither categorised nor tagged.
Office for Human research protection (OHRP)	The OHRP provides leadership in the protection of the rights, welfare, and wellbeing of subjects involved in research conducted or supported by the U.S. Department of Health and Human Services. ➢ On 18/08/2017, contained 34 fictional cases explaining a situation, without further questions or discussions. ➢ Cases are organised based on host entities (the context) and issues, but cases are not searchable.
Office of Research Integrity (ORI)	The ORI oversees and directs Public Health Service (PHS) research integrity activities on behalf of the Secretary of Health and Human Services (HHS) with the exception of the regulatory research integrity activities of the Food and Drug Administration. In addition to fictional case studies for educational purposes, ORI stores information about publicly known cases of misconduct that were adjudicated by HHS since 2008. ➢ On 18/08/2017 contained 35 historical cases of misconduct including names, affiliations and the decision of the court. ➢ Additionally, ORI's casebook contains 24 fictional cases including 8 role-plays. These cases are categorised in 8 categories based on the involved issue, and include questions and discussions questions for facilitators.
Research involving children	Financed by the UNISEF office of research and few academic institutions, the International Research Involving Children (ERIC) project provides a repository of information and resources to

	assist researchers, and guide and improve research involving children. ➢ On 14/08/2017, contained 27 cases inspired by historical events and rewritten by experts. Ethical issues are highlighted and motivations for the decisions are explained. ➢ Cases are categorised under four categories (Harms and Benefits, Informed Consent, Privacy and Confidentiality, Payment and Compensation) but they are not tagged.
The center for bioethics and human dignity studies, Trinity international university	A repository to analyse biomedical cases from a Christian ethics perspective. ➢ On 14/08/17 contained 16 case studies based on historical cases including recommendations and comments from several experts and an audio podcast. ➢ Cases are tagged based on the topic. ➢ Four of the published cases are also available in Chinese.
The global health network - WHO	The online platform made to present cases that originally appeared in a casebook published by the World Health Organisation (WHO) in 2009. ➢ On 14/08/2017, contained 64 fictional cases about ethical issues raised in health research. Cases include study questions. ➢ Cases are subsumed under eight categories but are not tagged.
The graduate school of Michigan state university	A collection of research integrity resources for graduate students. ➢ On 14/08/2017, contained 17 cases and 8 video case studies, all of which present fictional cases. PDF files also include questions for discussion. ➢ Cases are categorised under seven topics but are not tagged.
The MC Graw Hill Bioethics case studies	An educational source in the context of general human biology. ➢ On 14/08/17 contained 23 fictional cases including questions for discussions. ➢ Cases are neither categorised nor tagged.

The Netherlands Board on Research Integrity (LOWI)	LOWI is an independent advisory body, established in 2003. Anonymised cases are stored on the website from 2007 until 2017 (in Dutch). ➢ On 18/08/2017, contained 24 anonymised case summaries in English including an explanation of relevant considerations. ➢ Cases are neither categorised nor tagged.
UK clinical ethics network - UKCEN	The network aims at providing up to date and reliable information on ethical issues that are commonly presented to clinical ethics committees or arise in clinical practice. ➢ On 15/08/2017 contained 3 fictional cases based on common scenarios in clinical practice. Cases are followed by questions for discussions and highlight several principles for moral deliberation. ➢ Cases are neither tagged nor categorised.
UK research integrity office	The UK Research Integrity Office (UKRIO) is an independent charity, offering support to further good practice in academic, scientific and medical research. Sample case study packs are provided for education and training purposes. ➢ On 18/08/2017, contained one free case study pack including 4 Fictional scenarios based on real life situations. Some suggested points for discussion accompany each case study. ➢ Cases are neither categorised nor tagged.
Washington university of saint Louise – centre for clinical and research ethics	A library of Research Ethics Case Studies for use in small group case discussions, and train-the-trainer support for units. ➢ On 14/08/2017, contained 106 case studies mainly on ethical issues raised in clinical research. Cases also include questions for group discussions. ➢ Cases are neither categorised nor tagged.
Yale Interdisciplinary Centre for Bioethics	The centre provides a series of cases rooted in the work of hospital IRBs and drafted by their members. ➢ On 14/08/2017, contained 6 fictional cases. They are very elaborate and highly contextualised cases in

	research ethics. Each case comes with several expert comments that shed light on various aspects of the case. ➢ Cases are neither categorised nor tagged.

During task 5.2 [M6-12], we will search the repositories identified above in order to locate RE+RI cases. We will produce a list of references and associated cases, ensuring that we include basic case features using appropriate data management software. We will also provide a report of the search with details of any amendments to the protocol.

During task 5.3 [M12-36], we will read each case detailed in the master document that was created during months 6-12. We will manually search through the content to include cases that comply with the definition of a case in section 5. Simultaneously, we will develop and expand the list of references, associated cases and basic case features by registering keywords (that can later be used in the tagging process) according to the tagging system posited in section 8. The search will be reported, providing justifications for the exclusion of identified cases.

7.7 WP2's Focus Group Sessions

According to sub-objective (2) of the WP2 'Stakeholder Consultation: Protocol', the stakeholder consultation will 'explore stakeholders' experiences and perspectives regarding RE+RI, including implicit rules and practices and local cases' (Evans 2017, 5).

Drawn from the Netherlands, Spain and Croatia, stakeholders for the pilot, face-to-face focus groups will include 'researchers from various disciplines [biomedical, social sciences, natural sciences, applied sciences] (4), journal editors (2), members of national and RE+RI committees (4), policy makers (2), representatives from industry, including pharmaceutical companies (2), and representatives from research funding organisations (2)' (ibid., 7). Consequently, 'sixteen stakeholders (two groups of eight participants) from each of the three countries (48 in total) will participate in two rounds of focus groups' (Figure 2: 'WP2 Stakeholder Consultation: Protocol'). Additionally, 'a selection of participants from each country will be selected to take part in a third, multi-country focus group in Amsterdam to discuss similarities and differences between countries' (ibid., 10).

Furthermore, the scale-up phase of WP2 will take the form of online focus groups, covering similar content as the face-to-face focus groups but in an online format. The online focus groups will either be made up of a diversity of stakeholders from each EU country or take the form of 'multi-country online community discussions' (ibid., 12).

Based on the three guides for each of the face-to-face focus groups (Appendix 5: 'WP2 Stakeholder Consultation: Protocol'), which will also be used as the basis for the scale-up, online focus group sessions, stakeholders will be asked the following questions:

- "Tell me about a time when you witnessed research that did not meet what you consider good research practice. Why was that?" (Focus Group Round 1)

- "Do you have good examples for the categories/cases from your experience?" (Focus Group Round 1)

WP5 will use these explicit questions, together with any additional details provided throughout the course of the stakeholder consultation, to gather data relating to specific RE+RI cases.

According to the WP2 Protocol, the outputs from the far-to-face focus groups have been timed to arrive before we conclude our initial data collection processes in task 5.2 whilst the outputs from the online focus groups will arrive before we begin the scale-up phase of our data collection in task 5.3.

During task 5.2 [M6-12], WP5 researchers will use the official transcripts in order to extract the pertinent information from the face-to-face focus group sessions. Based on the official transcripts, WP5 researchers will:

I. Store the official transcripts using appropriate data management software;
II. Input basic features of cases using appropriate data management software;
III. Report the search with details of any amendments to the protocol.

During task 5.3 [M12-36], WP5 researchers will use the information provided by WP2 from the online focus groups in order to:

I. Store the information provided by WP2 using appropriate data management software;
II. Input basic features of cases using appropriate data management software.

Subsequently:

I. The cases will be read to identify those that match the definition given in section 5;

II. At the same time as step I, we will produce a register of keywords for each case utilising the tagging system detailed in section 8 (for cases that match the definition given in section 5). Simultaneously, we will remove cases that are unsuitable or 'off topic' according to criteria that will be made explicit in the search report;
III. We will produce a finalised list of references and associated keywords (to be later used as tags) using appropriate data management software;
IV. The search will be reported, providing justifications for the exclusion of identified cases.

7.7.1 Preliminary Risks

The collection of case details is contingent upon the questions asked and the information provided through WP2's stakeholder consultation.

According to the 'WP2 Stakeholder Consultation: Protocol', 'one risk associated with the focus group is the possibility of participants' personal knowledge of deviant cases being exposed to others or even made public. Efforts to mitigate this risk include asking all participants to return confidentiality agreements and to minimise the use of identifying characteristics. Participants will also be reminded to respect privacy and confidentiality at the beginning of each focus group (both face-to-face and online)' (Evans 2017, 16). All data will be anonymised for analysis. Furthermore, summaries of the focus groups will be send to the participants and written consent will be sought before any anonymised case or scenario based on the data will be published on the EnTIRE platform.

Ethics approval has been applied for separately in the Netherlands, Spain and Croatia.

According to WP2's protocol, the burden of responsibility for data protection lies with the Dutch partner (VUmc). Any sensitive data collected will be stored electronically in 'Dark Storage', a maximum-security data storage facility at VUmc. Audio recordings of face-to-face focus groups will be destroyed after they have been transcribed and quality checks conducted. Only the

transcripts will be archived. These focus group transcripts will have identifying information removed as much as possible. In addition, they will only be accessible to authorised personnel. Data from the online discussions will be collected through third party software WP2 will choose a suitable party in the next months, selected on the basis of their compliance with EU data protection acts and their ability to guarantee anonymity. A data processing agreement with this party will be produced.

According to WP2's protocol, the characteristics of the sample recruited will be described using descriptive statistics. Analysis of data generated during the face-to-face and online focus groups will be thematic. The preliminary data collection categories (guidelines, codes, legislations, and standards; committees, training courses and expert advice and contacts and cases, casuistry and scenarios) will be used as a deductive coding scheme in initial line-by-line coding of the transcripts of the face-to-face focus groups and the text from the online discussions. Any topics that fall outside of this coding scheme will allow the coding scheme to be developed further and, subsequently, inform the both the focus group topic guide for subsequent discussions and data collection categories. According to WP2, data analysis will be conducted using MAXQDA qualitative data analysis software.

8. System of Categories for Tagging Cases

The EnTIRE Project Proposal demands that WP5:

- Develop a system of categories, a 'thesaurus', for tagging cases that enhances retrievability and orientation within the EnTIRE database. Tagging methods will involve traditional approaches, e.g. tagging according to issues such as misconduct, falsification, fabrication, plagiarism, fake peer-review, data management, as well as innovative methods, for example tagging according to the main ethical principles from the RE+RI normative framework that have been violated (Widdershoven et al. 2015, 32).

In tagging and categorising RE+RI cases, the aim is 'to enhance the indexing and retrievability of cases in the EnTIRE database' (ibid.).

8.1 Introduction to Tagging Cases

A tag, otherwise known as a 'knowledge tag', is a type of metadata that, as a keyword or term, captures knowledge of a phenomenon in the form of descriptions, categorisations, classifications, semantics, hyperdata, hyperlinks and references. In turn, these metadata can be collected into an 'ontology' that describes and defines the types, entities and relationships existing in a specific realm of tagging discourse.

With the development of Web 2.0 technologies and its emphasis on user contribution and participation, interest has grown in the use of tagging as a potential means of indexing online content. As Choi and Syn observe, 'tagging allows users to add their own keywords or tags to online documents and images so that they can organise resources for themselves, share them with others, and find resources that others have tagged' (Choi and Syn 2012). That said, Bischoff

et al. (2008) have shown that there are substantial differences in tag types among resource types. For example, there are more location- and time-related tags for images, and more author/owner- and type-related tags for music resources.

8.2 Hierarchical Categorisation and Linear Tagging

If we understand that tags are metadata, then the relationships between tags can be hierarchical, linear/'flat' or a combination of the two. A hierarchical scheme, or 'top-down taxonomy', is constituted by 'parent-child' relationships between the metadata elements, which allows for classification of phenomena in terms of categories and subcategories. The hierarchy allows users of the category system to easily search for and retrieve information. Non-hierarchical schemata are one-dimensional (otherwise known as 'linear' or 'flat'). Consequently, each element within a flat scheme is completely discrete from other elements and classified according to a single dimension. Whereas top-down taxonomies tend to be created by an authorised, 'expert' group of designers/researchers (sometimes in the form of a controlled vocabulary), non-hierarchical approaches to categorisation (such as collaborative, social tagging), which are often called 'folksonomies', are created by all users.

Hierarchical, top-down, classifications are widely used to sort and structure. A popular example is the Dewey Decimal Classification ('DDC'). In many cases, strict semantics are assigned to the simple hierarchy of a classification. By contrast, as Jakob Voß observes, 'tagging with uncontrolled keywords is gaining popularity on the web in collaborative tagging (or folksonomy) systems. The tags in these systems are not connected (beside their correlation) and they are used without specific rules – nevertheless powerful structures evolve from collaborative action' (Voß 2006). Voß demonstrates that a third system, which combines hierarchical relations with collaborative tagging, is the 'thesaurus' system, according to which 'tags are connected more flexibly with less strict semantics' (ibid.).

An example of hierarchical categorisation can be found in the Categories for the Description of Works of Art ('CDWA'), developed by the Art Information Task Force ('AITF') and funded by the J. Paul Getty Trust, the National Endowment for the Humanities ('NEH') and College Art Association ('CAA'). According to the CDWA, the categories 'are a set of guidelines for best practice in cataloguing and describing works of art, architecture, other material culture, groups and collections of works and related images, arranged in a conceptual framework that may be used for designing databases and accessing information. CDWA includes around 540 categories and subcategories of information. A small subset of categories is considered *core* in that they represent the minimum information necessary to identify and describe a work' (CDWA 2015). Despite acknowledging that the designation of 'core' categories 'should vary depending upon the end-users whom the particular art information system is intended to serve, the mission of the specific institution and a number of other factors', the CDWA 'recommends a relational data structure, where records for objects/works are linked to each other in hierarchical relationships, where necessary' (ibid.). Details of the hierarchical system for the CDWA can be found at:

http://www.getty.edu/research/publications/electronic_publications/cdwa/categories.html

An example of linear tagging has been employed by Wiley for its Discover the Future of Research project. As this schema demonstrates, there is no indexing of tags according to supercategories and subcategories. The tags are presented in one dimension, albeit with visual cues to highlight the most frequently accessed tags.

In the fields of Research Ethics and Research Integrity, both top-down taxonomies and linear schemata have been employed. On the one hand, Ethics Core: Collaborative Resource Environment, developed by the National Center for Professional & Research Ethics, utilises a flat scheme of 2,247 tags. The tags are collaboratively developed, implemented, maintained and amended by end-users, taking the control away from an authorised, 'expert' group of categorisers commonly employed in hierarchical taxonomies. The individual tags can be browsed

based on the number of objects they tag. In addition, the tags are searchable with the most 'recently used' and 'top 100' presented on the website for users to access.

On the other hand, the Online Ethics Center for Engineering and Science utilises a hierarchical system of categorising content determined by the OEC Editorial Boards combined with a flat scheme for individual fields. OEC resources and announcements are categorised in six fields, including Engineering, Life and Environmental Sciences, Computer, Math, and Physical Sciences, Social and Behavioural Sciences, Research Ethics and International Ethics. In turn, the 'Research Ethics' field, for example, is subcategorised using 24 keywords, which, to the user, do not appear to rely upon more fundamental categories. The resources tagged by these keywords can be found within several fields, suggesting that more complex, hierarchical relationships exist than can be made explicit by the user.

Bearing in mind that the EnTIRE project is concerned with adopting 'the Wiki approach' for the online platform, it is pertinent to see how Wikipedia approach the tension between hierarchical categorisation and flat, collaborative tagging. Indeed, the category system of Wikipedia is a thesaurus that combines collaborative tagging and hierarchical subject indexing. The process of assigning categories to Wikipedia articles is a form of collaborative tagging. Furthermore, the introduction and removal of tags and categories is carried out by end-users. However, as Jakob Voß observes, 'an essential difference between known collaborative tagging systems and Wikipedia's categories is that one can also assign categories to other categories. This way hierarchical relationships with supercategories and subcategories are defined. From these hierarchies, one can derive tree structures like those of known classifications. Most of the categories are connected to a selected main category that is superordinated to all other categories' (Voß 2006). That said, while professional classification systems ensure that every document is classified with only one class, tagging allows multiple descriptors to be assigned. Consequently, 'there is no directive on how many descriptors to assign to a document in collaborative tagging' (ibid.). For Wikipedia, the result is that every category can be assigned to many other categories. It follows that 'the category system of Wikipedia is not a classification

[like the Dewey Decimal Classification, for example] but a thesaurus' – 'a controlled vocabulary of terms that can be used as keywords', thereby ensuring that 'every tag ("descriptor") has a marked meaning' (ibid.).

In practice, Wikipedia's category system evolves from bottom to top. It is automatically generated from the category tags introduced by end-users at the bottom of articles. By tagging articles with existing or new categories, end-users automatically and implicitly build up a hierarchical structure. The scheme terminates with 'Wikipedia's major topic classifications' – root descriptors without broader terms ('top terms' in a thesaurus) – which, although not top of Wikipedia's overall category system (Category:Contents is technically at the top of the category hierarchy, but contains many categories of no use to readers), groups major topic classifications in one place from which one can reach all other descriptors. Currently, the controlled topic classifications are Reference, Culture, Geography, Health, History, Mathematics, Nature, People, Philosophy, Religion, Society and Technology.

8.3 Problems with Hierarchical Classifications

Hierarchies are useful for a major type of information retrieval task – browsing. When we do not know exactly what we are looking for, it is easier to be able to broaden and narrow our search areas than to make random connections between ideas. The top few categories of a traditional hierarchy provide users with a better overview of the contents of a data collection than thousands of individual tags, even if these tags are ranked by their frequency in the collection. However, being rooted in the culture, society and time in which they were conceived, hierarchical classification systems can be slow to change. By contrast, the flexibility of tagging allows users to classify their collections of items in the ways that they find useful and that can change as their needs adapt. The problem is that, as Prokofyev and his co-authors observe, 'the ontology of scientific knowledge *per se* is very complex and vaguely defined at any given point in time' (Prokofyev et al. 2012, 326). Consequently, as Voß argues, 'hierarchy seems to have a strict semantic that does not fit with the vagueness of the world. In practice, there are always several

ways to classify an object (for instance diseases by region of the body, dangerousness, type of treatment etc.)' (Voß 2006). When it comes to scientific ontologies in the realm of hierarchical categorisation, Prokofyev suggests that the limits to the hierarchical system can only be overcome by taking the management of metadata away from 'the heads of the experts', replacing it with a renewed focus on 'human-machine collaboration' and 'user-friendliness' (Prokofyev 2012, 326.)

The respective claims of Prokofyev and Voß have been supported by evidence that discloses a gap between end-users and data categorisers in the realm of information retrieval and data collection. By looking at identical or near identical tags and keywords, Heckner, Mühlbacher and Wolff 'found an overlap ratio of 60% relative to the minimum of the number of tags and keywords per document. The coverage of tags with respect to all authors' keywords reaches only 30% which means that almost two thirds of author keywords are not reflected in (user) tag contents' (Heckner, Mühlbacher and Wolff 2008, 13). Elaborating, they show that 'in most cases, tags tend to be more general, which is to be expected as tags are shorter (less multiword terms). Additionally, in some cases taggers tend to use faceted tags where authors employ (more specific) multi-word terms' (ibid.). Consequently, compared with expertly-derived hierarchical categories, social tags tend to introduce fewer and simpler concepts. Furthermore, Heckner's research shows that 'user tags considerably add to the lexical space of the tagged resource' (ibid., 16). It follows, according to Furnas (1987), that in order to support successful retrieval in complex environments, it is necessary to index an object with a variety of aliases. Foskett (1997) and Lancaster (1993), respectively, claim that a thesaurus may assist users in achieving a better match between their search query and the indexing terms provided by indexers and authors. In this spirit, Heckner, Mühlbacher and Wolff argue that 'social tagging enhances the possibilities of traditional (author or expert) indexing by adding user-created retrieval vocabularies which could bridge the gap between users and authors or indexers without the expensive process of creating a thesaurus or crossvocabulary concordance' (Heckner, Mühlbacher and Wolff 2008).

On a pragmatic level, Trant (2006) demonstrates that museums could utilise folksonomies to open collections up to unexpected and more personal meanings, and that the content elements found in folksonomies were missing from formal museum documentation. He goes on to suggest that user tags might help to bridge the gap between professional and public discourse by providing a source of terms not in museum documentation (Trant 2009). If hierarchical classification systems drive a pragmatic wedge between end-users and metadata experts and indexers, then, according to Choi and Syn, 'user tags help reduce the gap as tags represent users' needs and understanding of resources' (Choi and Sun 2012). According to Trant, 'tagging begins as a personal information management and re-discovery tool. New concepts often emerge in personal tags that are then shared in social systems, where social information discovery leads new users to content. This is where social tagging seems to offer a number of affordances, primarily related to the use of the resulting folksonomy by others for information retrieval, browsing, searching or current awareness' (Trant 2009, 23). Additional benefits of collaborative tagging include: 'promoting a sense of ownership of content', 'developing social cohesion in a group' and conquering the 'economics of indexing' through a 'broad distribution of effort' (ibid.).

8.4 Problems with Collaborative Tagging

Despite the move away from formal hierarchical classification schemes towards linear, collaborative tagging, there is a concern that user-generated tagging will lead to an overly subjective approach to categorisation. In other words, tags will emerge that are more disclosive of individual sense-making, desires, intentions and motivations than they are of the phenomenon to be tagged. On the contrary, a number of studies have shown that most collaborative tags are topic-related, in the sense that that they are oriented primarily towards either the phenomenon being tagged in terms of its form and content or the social normative framework governing the use of specific tags.[2] Indeed, such studies have demonstrated that tags related to personal opinions or personal tasks are very much in the minority. According to Choi and Syn, 'many user

[2] See, for example, Bischoff et al. (2008) and Choi and Syn (2012).

tags were related to the nature or physical characteristics of resources. This represents that in doing research and educational activities in digital humanities domains, not only subject content aspects of resources, but also the nature of the work itself was important for users' (Choi and Sun 2012).

It is more likely that a personalised variety of terms can present challenges when searching and browsing. For example, it has been shown that tagging systems lead to semantic difficulties which may hinder the precision and recall of information.[3] In addition, when users choose tags, the resulting metadata can include homonyms and synonyms, which may lead to inappropriate connections between items, thereby exacerbating searching inefficiencies. Users can also choose tags that suggest different inflections of words, which can contribute to navigation difficulties if the system does not include stem tags when searching or browsing. Furthermore, as tags are an uncontrolled form of keywords, tag structures and grammatical forms can include abbreviations, initialised proper names ('MLK' instead of Martin Luther King), acronyms ('OEC' instead of Online Ethics Center), singular/plural terms and spelling errors. In addition, as Choi and Syn (2012) have observed, due to inconsistent use of, or missing, hyphens and underscores, related keywords are listed as separate tags without an explanation or meaning-defining context. As a whole, 'there is concern over the relativistic nature of socially created vocabularies, the lack of term consistency, problems with synonymy and polysemy, and the inherent inconsistency of a user-generated vocabulary. All of these characteristics may limit the value of a folksonomy as an indexing language and retrieval tool' (Trant 2009, 23).

It has been argued that larger-scale folksonomies can address some of these problems associated with tagging. Specifically, users of tagging systems tend to notice the current use of 'tag terms', thereby, 'correctly' appropriating tags according to socially normative structures in order to easily form connections to related items. However, even though users tend to use tags correctly, research has shown that 'agreement on tags among categorisers is significantly lower compared to agreement among describers' (Strohmaier, Körner and Kern 2012). In other words, there tends

[3] See, for example, Marlow 2006, Lakoff 2005, Golder and Huberman 2006 and Shirky 2005.

to be more agreement concerning what tags are used to accurately and precisely refer to or represent the contents of phenomena compared to those tags that are used as 'a means to categorise resources according to some shared high-level characteristics', which act as 'a navigational aid to the resources for later browsing' (ibid.). Not only is there greater disagreement amongst tag categorisers, large-scale collaborative tagging exercises have also shown that the larger the sample of users, the more the tagging motivation of individuals varies (ibid.). Furthermore, the more the tagging motivation varies, the greater the influence on the resulting tags and folksonomies (ibid.).

8.5 Conclusion

A decision on the appropriate system of categories for tagging RE+RI cases must be informed by the broader demands and aims of the EnTIRE project. According to the project's proposal, 'our vision is that in order to make the normative framework governing RE+RI accessible, a dynamic online Wiki-platform, owned by the community of RE+RI stakeholders, is needed.' (Widdershoven et al. 2015, 2). In addition, that platform should be 'customer-tailored, up-to-date and self-sustainable' (ibid., 7). It follows that 'the key unique feature of this proposal is the iterative, "bottom up" approach, making explicit normative experiences of local stakeholders and principles embedded in local rules and practices, and enabling the structuring of data in a way that fits in with research and evaluation practice, providing useful, accessible information for local users' (ibid., 5).

In terms of the functionality of the online platform, 'the current situation of researchers needing accessible information whereas websites are too static and limited to serve researchers' needs can only be solved by creating an interactive platform, enabling users to navigate *quickly and intuitively* to appropriate content, developed and kept up-to-date by a community of active users' [italics added] (ibid., 14). it will combine 'data mining features, which allow for searchability' 'with more conventional, hierarchical data organisation approaches (e.g. by geography or discipline), allowing researchers to perform analyses more easily – greatly enhancing their ability to perform

cross-country comparative research with low effort and cost' (ibid., 11). Consequently, 'we will adopt an open source approach to address the challenges of maintaining the relevance of platform data and minimising recurrent costs. The Wiki-platform is open access and will be adapted to the specific needs of the project' (ibid., 8). Therefore, in terms of the aims of the EnTIRE project, the online platform should be:

- Self-sustainable
- Dynamic
- Community-owned
- Bottom-up structured
- Localised
- Up-to-date
- Organised hierarchically
- Searchable
- With accessible information
- Quick and intuitive to navigate
- Low cost

In terms of the information already provided in section seven, the following characteristics would support a linear, collaborative approach to tagging RE+RI cases:

- Self-sustainable
- Dynamic
- Community-owned
- Bottom-up structured
- Localised
- Up-to-date
- Low cost

However, the following features would be better supported by a hierarchical approach to categorisation:

- Organised hierarchically
- Searchable
- With accessible information
- Quick and intuitive to navigate

The EnTIRE project's aims for an online platform that blends the sustainable, community-owned, bottom-up-structured and low-cost approach of social, collaborative tagging with the searchable, accessible and intuitive features of a professional categorisation system, suggests that we should adopt the Wikipedia approach to categorisation. As we have seen, Wikipedia's category system is a thesaurus that combines collaborative tagging and hierarchical subject indexing. With end-users not only assigning categories to articles, but introducing and removing tags and categories, Wikipedia's category system evolves from bottom to top in the sense that it is automatically generated from the category tags introduced by end-users at the bottom of articles. It follows that by tagging articles with existing or new categories, end-users automatically and implicitly build up hierarchical structures.

Implicitly supporting the approach taken by Wikimedia, Prokofyev and his co-authors have developed The ScienceWISE System, which allows 'a community of scientists, working in a specific domain, to generate dynamically as part of their daily work an interactive semantic environment, i.e., a field-specific ontology with direct connections to research artefacts (e.g., research papers) and scientific data management services' (Prokofyev et al. 2012, 327). According to Prokofyev, 'since the underlying domain of the ontology is often rapidly changing and only loosely-defined, the best way to keep it up to date is to crowdsource its construction through the community of expert scientists' (ibid., 328). After creating a very loose, informal and contingent set of categories, 'ScienceWISE users (who are domain experts) are allowed to edit elements of the ontology (e.g., adding new definitions or new relations) in order to improve both

its quality and coverage' (ibid.). It follows that 'the proposed techniques [for The ScienceWise System] select tags directly from a collaborative, user-driven ontology' (ibid., 335).

If the act of assigning, introducing and removing tags is one carried out by the end-users of the EnTIRE platform and if the introduction and removal of tags by end-users determines the hierarchical structures that develop over time and with use and if the controlled vocabulary of terms with which the hierarchy terminates is, itself, contingent on tags and tag-structures developed by end-users, then we should avoid formalising explicit tags, terms, rules and ontologies at this stage unless we wish to undermine the aims espoused in the EnTIRE project's proposal and detailed above. Indeed, the aims for the online platform are similar to the results of end-user collaborative tagging systems, specifically, the development of social cohesion in a group, the promotion of a sense of content ownership and an overcoming of the high costs of professional indexing.

On a practical level, we cannot expect uploaded RE+RI cases to be tagged by end-users without a certain amount of instruction or demonstration, no matter how informal and contingent. Consequently, it is advisable that WP5 begins the process of tagging RE+RI cases outside of the dimensions of a formal, rule-governed system.

As already discussed, users of tagging systems tend to notice the current use of "tag terms" on the basis of socially normative structures. That said, and as Strohmaier, Körner and Kern (2012) have demonstrated, there is a considerable amount of disagreement concerning what categories, as high-level characteristics, should be employed within a tagging system. Bearing in mind that the processes and structures for information retrieval, browsing and searching will develop through the hierarchical structures derived from end-user collaborative tagging, it is necessary, at this stage, to demonstrate a system of end-user-oriented tagging that does not formalise a system of categories (which are open to disagreement), that does not posit a directive for how many descriptors to assign to an RE+RI case, that leaves open the possibility for end-users to

introduce new tags and remove tags and that allows users to assign tags in the ways that they find useful and that can change as their needs adapt.

Taking into account these conditions and bearing in mind that any demonstration of RE+RI case tagging is open to the contingencies of end-user engagement later in the project and, indeed, after the project has officially come to an end, we propose the following tagging 'system' for RE+RI cases:

- Bearing in mind that the explicit and implicit dimensions of the normative frameworks governing RE+RI vary according to research disciplines, we will begin tagging cases according to the discipline(s) that contextualise cases. These tags may be 'broad', such as 'science' or 'humanities', or 'narrow', such as 'physics' or 'philosophy'. Voß states that 'to build a formal ontology that can be fully processed automatically must be avoided by all means. But human beings can judge in individual cases. This is why "broader term" and "narrower term" do not need to be defined more precisely in many cases' (Voß 2006). In order to ensure that we do not prescribe rules that could be undermined by user engagement with the online platform, we will not define the meaning of 'broad' and 'narrow' with regards to research disciplines. As already stated, and as will apply to the majority of the aspects of the proposed 'system', the meanings, use and hierarchies of terms will develop through end-user tagging;
- Taking into account users' needs, 'the substantial and essential diversity of practices within and between countries and disciplines' (Widdershoven et al. 6), the fact that ethical frameworks, professional values and norms and pertinent legislation 'are strongly influenced by socio-cultural, political, economic and institutional contextual factors' (ibid.) and the demand for the platform to enable 'cross-country comparative research' (ibid.), we will begin tagging cases according to the geographical setting of the case. Again, these tags may be 'broad', such as 'Europe' or 'North America', or 'narrow', such as 'Netherlands', 'County Kerry' and 'London';

- Again, bearing in mind that the ethical frameworks, professional values and norms and pertinent legislation 'are strongly influenced by socio-cultural, political, economic and institutional contextual factors' (ibid.), we will begin tagging cases according to the time period that contextualises a case. Again, these tags may be 'broad', such as 'Nineteenth Century', or 'narrow', such as '2012' and 'August 2017';
- On the basis of the definition of a case in section 5, we will begin tagging RE+RI cases according to whether they are 'fictional' or 'non-fictional';
- On the basis of the definition of a case in section 5, we will begin tagging RE+RI cases according to whether they violate the normative frameworks governing RE+RI or demonstrate excellent practice in relation to the normative frameworks;
- On the basis of the definition of a case in section 5, and if the cases are violations of the normative frameworks governing RE+RI, we will begin tagging RE+RI cases according to whether they are violations of the explicit dimension of RE+RI in terms of rules, regulations and legislation and/or the implicit dimension in terms of those norms developed in practice and through processes of deliberation regarding research misconduct;
- On the basis of the definition of a case in section 5, and if the cases are violations of the normative frameworks governing RE+RI, we will begin tagging RE+RI cases according to the type(s) of violation, for example, 'misconduct', 'falsification', 'fabrication', 'plagiarism', 'fake peer-review', 'data management' etc.;
- On the basis of the definition of a case in section 5, and if the cases are violations of the normative frameworks governing RE+RI, we will begin tagging RE+RI cases according to the main ethical principles from the RE+RI normative framework that have been violated;
- On the basis of the definition of a case in section 5, and if the cases are examples of excellent practice in relation to the normative frameworks governing RE+RI, we will begin tagging RE+RI cases according to the type(s) of best practice demonstrated by the case;
- On the basis of the definition of a case in section 5, we will begin tagging RE+RI cases according to the form in which a case has been documented, for example, 'case study',

'case report', 'case review', 'newspaper article', 'blog', 'interview', 'case law', 'legislation' etc.;

- On the basis of the categorisation of sources used to search for and retrieve RE+RI cases in this protocol, we will begin tagging RE+RI cases according to the source(s) in which the case was located, for example, 'Academic Literature', 'Reports by RE+RI Committees and Regulatory Bodies', 'Grey Literature', 'Media', 'Blogosphere', 'Online Repository', 'User Consultation' etc.;
- Taking into account WP2's stakeholder consultation and the fact that the EnTIRE project's members and contacts will be end-users of the online platform, we will include recommendations for categories and tags not included in the 'system' currently proposed;
- Once the online platform is available for public use, users will be able to assign descriptors to RE+RI cases and introduce/remove tags with the result that the hierarchical tagging structure will begin to take form allowing for information retrieval, searchability and browsing to become more efficient and intuitive as the rate of collaborative tagging increases.

8.5.1 Preliminary Risks

What is proposed here is an idealised model for how case tagging should be carried out based on the theoretical advantages and disadvantages of both hierarchical and collaborative tagging provided in the scholarly literature. However, the proposed 'system' for tagging is, ultimately, dependent on the online platform's structures to be developed by WP6 in dialogue with the EnTIRE project's executive committee, consultants, stakeholders and Work Package Leaders. These IT structures are still in development. Consequently, WP5 awaits a final decision on the feasibility of the proposed 'system'. Task 5.2. demands that we 'adjust the system of tagging based on the results of the pilot searches and the normative framework defined in the focus groups (WP2)' (ibid., 32-3). As a result, the proposed tagging 'system', as posited above, should remain open to adaptation and innovation. Any changes to the 'system' will be detailed in the sub-report that will emerge at the conclusion of task 5.2.

As well as being contingent upon the tagging practices of the online platform's end-users, the proposed 'system' is dependent upon feedback from WP2's stakeholder consultation during the early stages of the project (as a result of the face-to-face focus groups) and throughout its development (through the online consultation). According to the EnTIRE project's proposal, 'the findings from the consultation will help define the boundaries of content to be collected and structure the information on the platform according to stakeholders' concerns; enabling the collection, provision and presentation of data sensitive to stakeholder needs. The process will also foster an ongoing dialogue on the content, priorities, data structure, and acceptability and usability of the platform by the stakeholders' (Widdershoven et al. 2015, 5). Not only that, by exploring experiences and perspectives regarding RE+RI (including implicit rules and practices), the consultation will likely have an effect on the content of the terms and types of keywords used to tag RE+RI cases on the basis of specific violations, excellent practices and ethical principles.

By 'defining the boundaries of data to be collected', 'developing a mapping structure adapted to user needs' and being 'involved in the evaluation and improvement of the ease of use of the platform' (ibid., 3; 5), the project's stakeholders have the potential and opportunity to affirm, challenge, undermine and alter our initial RE+RI case tags as well as introduce new and innovative tags both relatively early in the project and throughout its development.

WP3 is tasked with producing 'a detailed mapping and analysis of the normative documents on research ethics and research integrity that are available within the European Union' (ibid., 29). In addition, WP3 'will analyse and prepare the normative documents for integration in the wiki-platform of the project (e.g. distinction between legislation that must be applied and the soft laws and best practices that must be taken into account)' (ibid.). By collecting and making explicit the normative documents governing RE+RI practice, the tasks and outcomes of WP3 will likely have an effect on the content of the terms and types of keywords used to tag RE+RI cases on the basis of specific violations, excellent practices, ethical principles and the forms of documented cases. In addition, by analysing the normative documents, WP3 will introduce new tags for

incorporation within the online platform, such as, according to the EnTIRE project proposal, ‘legislation’, ‘soft law’ and ‘best practice’.

9. Interaction with other Work Packages

We expect the content of this proposal to be (more or less) reciprocally related to certain tasks, findings and deliverables of Work Packages 2, 3, 4 and 6, in particular. The interaction with WP2's Stakeholder Consultation has already been made explicit in sections 7.7 and 8.5.1 and will not be detailed here.

9.1 Interaction with WP3

WP3 is tasked with producing 'a detailed mapping and analysis of the normative documents on research ethics and research integrity that are available within the European Union' (Widdershoven et al. 2015, 29). In addition, WP3 'will analyse and prepare the normative documents for integration in the wiki-platform of the project (e.g. distinction between legislation that must be applied and the soft laws and best practices that must be taken into account)' (ibid.). By collecting and making explicit the normative documents governing RE+RI practice, the tasks and outcomes of WP3 will likely have an effect on the content of the terms and types of keywords used to tag RE+RI cases on the basis of specific violations, excellent practices, ethical principles and the forms of documented cases. In addition, by analysing the normative documents, WP3 will introduce new tags for incorporation within the online platform, such as, according to the EnTIRE project proposal, 'legislation', 'soft law' and 'best practice'.

9.2 Interaction with WP4

WP4 'is responsible for collecting and synthesising the information about: 1) RE+RI committees in different European countries and for different research domains; 2) RE+RI training courses for researchers on; and 3) RE+RI experts' advice and contact details' (ibid., 30). Given that one of the primary sources of gathering cases in WP5 is RE+RI committees, the master document that logs our attempts at contacting different RE+RI committees will be made available to WP4 representatives. Similarly, if WP4 uncover alternative and more reliable communication channels

with RECs and RIOs, sharing these details with WP5 will improve both the efficiency and the outcomes of tasks 5.2 and 5.3.

9.3 Interaction with WP6

WP6 is responsible for developing 'a user-friendly platform, including a website and online resources, to facilitate access to RE+RI knowledge and experience and support application in research and evaluation, thus fostering uptake of ethical standards and responsible conduct of research' (ibid., 33). Given the importance of providing a user-friendly environment for searching for, retrieving and tagging RE+RI cases, WP5 has appointed a representative to the community task force to ensure that what is proposed in this protocol is feasible from an IT point of view. Changes to this protocol resulting from discussions with WP6 will be made explicit in the reports emerging from tasks 5.2 and 5.3.

10. References

Alberani, V., Pietrangeli, P. D. C. and A. M. R. Mazza (1990) 'The Use of Grey Literature in Health Sciences: A Preliminary Survey', *Bulletin of the Medical Library Association* 78 (4): 358-63

Augur, C. P. (1998) *Information Sources in Grey Literature*, London and New Providence, NJ: Bowker-Saur

Bischoff, K et al. (2008) 'Can all tags be used for searching?', *Proceedings of the 17th ACM conference on Information and knowledge management*, CIKM '08, ACM, New York, NY, USA: 203-12

CDWA (2015) *Categories for the Description of Works of Art: Introduction.* CDWA Website. [Online]. Available at: http://www.getty.edu/research/publications/electronic_publications/cdwa/introduction.html. Accessed 25th August 2017

Choi, Y. and S. Y. Syn (2012) 'An examination of user tags in a digitized humanities online collection', *ASIST 2012*, Baltimore, MD, USA. [Online]. Available at: http://onlinelibrary.wiley.com/store/10.1002/meet.14504901294/asset/14504901294_ftp.pdf?v=1&t=j6ryendx&s=b35cc90cbb5a563436204ef43febc83693ba26ef. Accessed 29th August 2017

Davis, M. (1999) *Ethics and the University*, London and New York: Routledge

European Network of Research Ethics Committees. (2017). [Online]. Available at: http://www.eurecnet.org/index.html. Accessed on 25th August 2017

European Network of Research Integrity Offices (2017). [Online]. Available at: http://www.enrio.eu/. Accessed on 17th September 2017

Evans, N. (2017) *WP2 Stakeholder Consultation: Protocol.* [Unpublished Research Paper]. EnTIRE Project, Vrije Universiteit Medisch Centrum (VUmc), Amsterdam, Netherlands

Foskett, A. C. (1997) *The subject approach to information*, London: Library Association Publishing

Furnas, G. W. et al. (1987) 'The vocabulary problem in human-system communication', *Commun. ACM.* 30: 964-71

Gavel, Y. and L. Iselid (2008) 'Web of Science and Scopus: A Journal Title Overlap Study', *Online Information Review* 32 (1): 8-21

Golder, S., and B. A. Huberman (2006) 'The Structure of Collaborative Tagging Systems', *Journal of Information Science* 32: 198-208

GreyNet (2017) *Document Types in Grey Literature*. GreyNet Website. [Online]. Available at: http://www.greynet.org/greysourceindex/documenttypes.html. Accessed 25th August 2017.

Harvard University (2017) *Systematic Reviews and Meta-Analysis*, The Francis A. Countway Library of Medicine, Harvard Medical School Website. [Online]. Available at: http://guides.library.harvard.edu/c.php?g=309982&p=2070465. Accessed 14th August 2017

Heckner, M., Mühlbacher, S., and C. Wolff (2008). 'Tagging Tagging. Analysing User Keywords in Scientific Bibliography Management Systems', *Journal of Digital Information* 9 (2). [Online]. Available at: https://journals.tdl.org/jodi/index.php/jodi/article/view/246. Accessed 25th August 2017

Higgins J. P. T. and S. Green (eds.) (2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0, The Cochrane Collaboration. [Online]. Available at: www.handbook.cochrane.org. Accessed 25th August 2017

Hood, W. W. and C. S. Wilson (2003) 'Overlap in Bibliographic Databases', *Journal of the American Society for Information Science and Technology* 54 (12): 1091-103

Hookway, N. (2008) 'Entering the Blogosphere: Some Strategies for Using Blogs in Social Research', *Qualitative Research* 8 (1): 91-113

Lakoff, G. (2005) *Women, Fire and Dangerous Things*, Chicago, Il: University of Chicago Press

Lancaster, W. (1993) *Information Retrieval Today*, Arlington: Information Resources Press

Marlow, C. et al. (2006) 'HT06, tagging paper, taxonomy, Flickr, academic article, to read' *Proceedings of the seventeenth conference on Hypertext and hypermedia*, ACM, New York, NY: 31-40

Michigan State University (2017) *PubMed, Web of Science or Google Scholar? A behind-the-scenes guide for life scientists: So which is better: PubMed, Web of Science, or Google Scholar?* Michigan State University Website. [Online]. Available at: http://libguides.lib.msu.edu/pubmedvsgooglescholar. Accessed 25th August 2017

Prokofyev R. et al. (2012) 'Tag Recommendation for Large-Scale Ontology-Based Information Systems' in P. Cudré-Mauroux et al. (eds.) *The Semantic Web – ISWC 2012*, *Lecture Notes in Computer Science* 7650, Berlin and Heidelberg: Springer: 325-36

Sampson, M. et al. (2006) 'Managing Database Overlap in Systematic Reviews Using Batch Citation Matcher: Case Studies Using Scopus', *Journal of Medical Library Association* 94 (4): 461-3

Shirky, C. (2005) 'Ontology is overrated: categories, links and tags'. [Online]. Available at http://shirky.com/writings/ontology_overrated.html. Accessed 29th August 2017

Steneck, N. H. (2006) 'Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions', *Science and Engineering Ethics* 12 (1): 53-74

Strohmaier, M., Körner, C. and R. Kern (2012) 'Understanding Why Users Tag: A Survey of Tagging Motivation Literature and Results from an Empirical Study', *Web Semantics: Science, Services and Agents on the World Wide Web* 17: 1–11

Suarez-Almazor, M. E. et al. (2000) 'Identifying Clinical Trials in the Medical Literature with Electronic Databases: MEDLINE Alone is Not Enough', *Control Clin. Trials* 21 (5): 476-87

The Office for Human Research Protection (2017) *International Compilation of Human Research Standards*. The Office for Human Research Protection Website. [Online]. Available at: https://www.hhs.gov/ohrp/international/compilation-human-research-standards/index.html. Accessed 25th August 2017

The University of Adelaide (2017) *MBBS Resources: Which Databases Should I Use?* University Library Website. [Online]. Available at: http://libguides.adelaide.edu.au/mbbs/home/wd. Accessed 14th August 2017

Trant, J. (2006) 'Exploring the Potential for Social Tagging and Folksonomy in Art Museums: Proof of Concept', *New Review of Hypermedia and Multimedia* 12 (1): 83-105

Trant, J. (2009) 'Studying Social Tagging and Folksonomy: A Review and Framework', *Journal of Digital Information* 10 (1): 1-44

Voß, Jakob (2006) 'Collaborative thesaurus tagging the Wikipedia way', *CoRR*. [Online]. Available at: http://arxiv.org/abs/cs.IR/0604036. Accessed 28th August 2017

Widdershoven, G. et al. (2015) *EnTIRE Mapping Normative Frameworks for EThics and Integrity of Research*. [Unpublished Research Proposal]. H2020-SwafS-2016-17, European Commission Research and Innovation Proposal Submission Forms, European Commission

Zucker, D. M. (2009) 'How to Do Case Study Research', *Teaching Research Methods in the Humanities and Social Sciences* 2. [Online]. Available at: http://scholarworks.umass.edu/nursing_faculty_pubs/2. Accessed 25th August 2017

11. Appendix 1: Research Ethics Committees

Country/Region	Number of RECs Listed by EUREC	Keys RECs Listed by EUREC	Additional RECs	Umbrella Organisations for RECs	Links
Armenia	No Information	No Information	No Information	No Information	No Information
Austria	26 -	9 Federal State Regional RECs 7 Leading RECs	Local RECs -	Federal Office for Safety in Health Care Forum of the Austrian Ethics Committees	http://www.basg.gv.at/en/austrian-federal-office-for-safety-in-health-care/ http://www.ethikkommissionen.at/
Belarus	No Information	No Information	No Information	No Information	No Information
Belgium	215 -	Advisory Committee on Bioethics -	Local RECs in hospitals -	Advisory Committee on Bioethics -	http://jme.bmj.com/content/31/6/318.long https://www.health.belgium.be/language_selection?destination=bioeth
Bosnia and Herzegovina	No Information	No Information	No Information	No Information	No Information
Bulgaria	2 -	Ethics Committee for Multi-Centre Trials Central Ethics Committee	Regional Health RECs Academic/Research RECs	Bulgarian Drug Agency Central Ethics Committee	
Croatia	No Information	No Information	No Information	No Information	No Information
Cyprus	3	The Review Bioethics Committee for Biomedical Research on Human Beings and their biological substances; The Review Bioethics Committee for the Clinical trials on Medicinal Products of Human Use; The Review Bioethics Committee for Biomedical Research on Human Beings and their biological substances and the clinical trials on Medicinal Products of Human Use.		Cyprus National Bioethics Committee-CNBC	http://www.bioethics.gov.cy/moh/cnbc/cnbc.nsf/index_gr/index_gr?opendocument
Czech Republic	No Information	See here	Academic/Research RECs	Forum of Czech Ethics Committees	http://www.forumek.cz/

	-	-	Local Health Care RECs	-	-
Denmark	11	11 Regional RECs	-	Danish National Committee on Biomedical Research Ethics	http://www.cvk.sum.dk/CVK/Home/English.aspx
Estonia	2	Research Ethics Committee of the University of Tartu	-	-	http://www.ut.ee/en/research/ethics-review-committee
	-	Tallinn Ethics Committee On Medical Research	-	-	http://www.tai.ee/?id=1960
Finland	9	9 Regional RECs	Academic/Research RECs	National Committee on Medical Research Ethics (TUKIJA)	http://www.tukija.fi/en/
	-	-	-	National Advisory Board on Research Ethics (TENK) (For Academic/Research RECs)	http://www.tenk.fi/en/
France	39	39 Committees of Protection of Persons	-	-	
	-	National Consultative Ethics Committee	-	-	
Georgia	No Information	No Information	No Information	No Information	No Information
Germany	53	33 RECs at Universities	-	The National Council for Ethics	http://www.ethikrat.org/
	-	17 RECs at Medical Associations	-	Central Ethics Committee of the German Medical Association	
	-	3 RECs attached to the State	-	Permanent Working Party of Research Ethics Committees in Germany	http://www.ak-med-ethik-komm.de
Greece	1	National Ethics Committee of the National Organization for Medicines	Local Health Care RECs	National Bioethics Commission	http://www.bioethics.gr/index.php?category_id=3
	-	-	Academic/Research RECs	-	http://www.eof.gr/web/guest;jsessionid=9d7e0d5626859518193ac6d34c6f
Hungary	13	National Ethics Committee for clinical Pharmacology (KFEB)	Academic/Research RECs	Medical Research Council (Egészségügyi Tudományos Tanács, ETT)	
	-	National Ethics Committee for Human Reproduction (HRB)	-	-	
	-	National Scientific and Ethical Committee (TUKEB)	-	-	
	-	10 Regional Scientific and Ethical Committees (RKEBs)	-	-	
Iceland	No Information	No Information	No Information	No Information	No Information

Ireland	12	See here	Academic RECs	-	http://health.gov.ie/wp-content/uploads/2016/07/European-Communities-Clinical-Trials-on-Medicinal-Products-for-Human-Use-Regulations-2004.pdf
Italy	271 - - -	See here	- - - -	National Bioethics Committee (NBC) The National Ethics Committees Federation (FNaCe) - -	http://www.governo.it/bioetica/eng/index.html http://www.comitatietici.it/elenco/default.html http://www.iss.it/sibi/ http://www.unich.it/fnace/
Latvia	10 - -	Central Medical Ethics committee of Latvia 4 Clinical Drug Trial committees 5 Medical Ethics Committees	- - -	- - -	
Lithuania	3	Lithuanian Bioethics Committee 2 Regional Biomedical RECs	- -	Lithuanian Bioethics Committee -	http://bioetika.sam.lt/index.php?-277184495 -
Luxembourg	1	Comité National d'Ethique de Recherche	-	Comité National d'Ethique de Recherche	http://www.cner.lu/
Macedonia	No Information	No Information	No Information	No Information	No Information
Malta	No Information	No Information	No Information	No Information	No Information
Moldova	No Information	No Information	No Information	No Information	No Information
Montenegro	No Information	No Information	No Information	No Information	No Information
Netherlands	58 -	MREC (24 accredited / 34 non-accredited) -	- -	Central Committee (CCMO) Dutch Association of Medical Research Ethics Committees	http://www.ccmo.nl http://www.nvmetc.nl
Norway	7 -	7 Regional RECs -	- -	National Committee for Medical Research Ethics (NEM) -	https://www.etikkom.no/en/our-work/about-us/the-national-committee-for-medical-and-health-research-ethics-nem/ https://www.etikkom.no/en/library/
Poland	53	13 Bioethics Committees appointed by medical universities	-	Center of Bioethics of the Supreme Medical Council	www.bioetyka.net

	-	17 Bioethics Committees appointed by the medical research and development centers	-	-	http://www.eurecnet.org/information/poland.html
	-	23 Bioethics Committees appointed by the Regional Chambers of Physicians and Dentists	-	-	-
Portugal	1	National Ethics Committee for Clinical Research (CEIC)	Local Health Care RECs	National Ethics Committees Network coordinated by CEIC	http://www.ceic.pt.
Romania	2	National Committee of Medicines and Medical Devices (NCMMD)	Academic/Research RECs	National Ethics Council (NEC)	http://en.bioetica-medicala.ro/obiectivele-comisiei/
	-	National Agency for Medicines and Medical Devices (NAMMD)	-	-	http://www.anm.ro/anmdm/en/_/STRATEGII/NAMMD%20Communication%20Strategy%202013-2015_EL.pdf
	-	-	-	-	http://cne.ancs.ro/
Russia	No Information	No Information	No Information	No Information	No Information
San Marino	No Information	No Information	No Information	No Information	No Information
Serbia	No Information	No Information	No Information	No Information	No Information
Slovakia	76	Ethics Committee of the Ministry of Health	-	Ethics Committee of the Ministry of Health	http://www.health.gov.sk
	-	8 Regional RECs	-	-	http://www.eurecnet.org/information/slovakia.html
	-	40-50 Local RECs	-	-	-
Slovenia	No Information	No Information	No Information	No Information	No Information
Spain	140	Clinical Research Ethics Committees (CEIC)	Academic RECs	Ministerio de Sanidad, Servicios Sociales e Igualdad (for CEICs)	http://www.msc.es/profesionales/farmacia/ceic/home.htm
	-	-	Public Research Organisation RECs	Network of Ethics Committees in Universities and Public Research Centres in Spain (45 RECs)	http://www.ub.edu/rceue/index2.htm
	-	-	-	National Association of Research Ethics Committees	www.ancei.es
Sweden	7	Central Ethical Review Board	-	Central Ethical Review Board	www.epn.se
	-	6 Regional RECs in Universities	-	-	-
Switzerland	8	8 Ethics Committees		Federal Office of Public Health and Swissethics	http://www.swissethics.ch/index_e.html

Ukraine	No Information	No Information	No Information	No Information	No Information
United Kingdom	104	NHS RECs	Academic RECs	United Kingdom Ethics Committee Authority (UKECA)	http://www.hra.nhs.uk/resources/research-legislation-and-governance/four-nations/
	-	-	-	Health Research Authority for research relating to NHS patients	http://www.hra.nhs.uk/

12. Appendix 2: Research Integrity Offices

Country/Region	Number of RIOs Listed by ENRIO	Key RIOs Listed by ENRIO	Links
Armenia	N/A	N/A	N/A
Austria	1	Austrian Agency for Research Integrity	http://www.oeawi.at/
Belarus	N/A	N/A	N/A
Belgium	1	Flemish Commission for Research Integrity (VCWI)	http://www.kvab.be/vcwi/
Bosnia and Herzegovina	N/A	N/A	N/A
Bulgaria	N/A	N/A	N/A
Croatia	1	Croatian Committee on Ethics in Science and Higher Education (CESHE)	vilibic@izor.hr
Cyprus	N/A	N/A	N/A
Czech Republic	1	Commission for the Scientific Integrity of the Czech Academy of Sciences (CAS)	http://www.avcr.cz/en/about-us/cas-structure/academy-council/advisory-committees/
Denmark	1	Danish Committees on Scientific Dishonesty (DCSD)	http://ufm.dk/en/research-and-innovation/councils-and-commissions/the-danish-committees-on-scientific-dishonesty?set_language=en&cl=en
Estonia	1	Estonian Research Council (ETAg)	http://www.etag.ee/en/
Finland	1	Finnish Advisory Board on Research Integrity (TENK)	http://www.tenk.fi/en
France	2	French Agricultural Research Centre for International Development (CIRAD) Internal Office of the French National Institute of Health and Medical Research (INSERM)	http://www.cirad.fr/ http://www.inserm.fr/qu-est-ce-que-l-inserm/organigramme/comites/dis
Georgia	N/A	N/A	N/A

Germany	3	German Research Ombudsman (Ombudsman für die Wissenschaft) Geschäftsstelle für Ombudsangelegenheiten der Universität Hamburg Team Scientific Integrity (Team SciInt)	http://www.ombudsman-fuer-die-wissenschaft.de/ https://www.uni-hamburg.de/en/forschung/service/gute-wissenschaftliche-praxis.html http://www.scientificintegrity.de/en-index.html
Greece	2	Ethical Aspects in Research and Technology for Human (EARTHnet) Network of Responsible Conduct of Research in Greece (RCR-Greece)	http://earthnet.ntua.gr/?lang=en http://www.rcr.gr/index.php/en/
Hungary	N/A	N/A	N/A
Iceland	N/A	N/A	N/A
Ireland	2	Health Research Board (HRB) Royal Irish Academy (RIA)	http://www.hrb.ie/home/ https://www.ria.ie/
Italy	1	National Research Council (CNR)	https://www.cnr.it/it/ethics
Latvia	N/A	N/A	N/A
Lithuania	N/A	N/A	N/A
Luxembourg	1	Luxembourg National Research Fund (FNR)	https://www.fnr.lu/
Macedonia	N/A	N/A	N/A
Malta	N/A	N/A	N/A
Moldova	N/A	N/A	N/A
Montenegro	N/A	N/A	N/A
Netherlands	2	Netherlands Board on Research Integrity (LOWI) Netherlands Research Integrity Network (NRIN)	http://www.lowi.nl/en/netherlands-board-on-research-integrity-lowi?set_language=en https://www.nrin.nl/
Norway	1	National Research Ethics Committees (Etikkom)	https://www.etikkom.no/en/
Poland	1	Commission for Ethics in Science	https://institution.pan.pl/index.php/institution/science-ethics-committee

Portugal	1	Foundation for Science and Technology	http://www.fct.pt/index.phtml.en
Romania	N/A	N/A	N/A
Russia	N/A	N/A	N/A
San Marino	N/A	N/A	N/A
Serbia	N/A	N/A	N/A
Slovakia	N/A	N/A	N/A
Slovenia	1	Commission for Women in Science (CWS)	ursa.ok@gmail.com
Spain	1	Ethics Committee of the Spanish National Research Council (CSIC)	http://www.csic.es/etica-en-la-investigacion
Sweden	1	Central Ethical Review Board (CEPN)	http://www.epn.se/en/start/
Switzerland	1	Swiss Academies of Arts and Sciences (SAMW)	http://www.akademien-schweiz.ch/en/index/Portrait/Kommissionen-AG/Wissenschaftliche-Integritaet.html
Ukraine	N/A	N/A	N/A
United Kingdom	1	UK Research Integrity Office (UKRIO)	http://ukrio.org/

13. Appendix 3: Regulatory Bodies

Country/Region	Regulatory Body	Links
Europe (Region-wide)	European Commission: DG SANTE: Directorate-General for Health and Food Safety	http://ec.europa.eu/health/index_en.htm
	European Commission: DG GROWTH: Internal Market, Industry, Entrepreneurship, SMEs	https://ec.europa.eu/growth/sectors/medic al-devices_en
	European Commission: Directorate-General for Justice and Consumers	http://ec.europa.eu/justice/mission/index_ en.htm
	European Commission: European Group on Ethics in Science and New Technologies	http://ec.europa.eu//research/ege/index.cf m
	European Medicines Agency	http://www.ema.europa.eu/
Armenia	Drug and Medical Technology Agency	http://www.pharm.am/
	Ethics Committee of the Ministry of Health	
	Ethical Committee of the National Center for AIDS Prevention	http://www.armaids.am/main/index.php?l ng=1
Austria	Ministry of Health	http://www.bmg.gv.at
	Forum of Austrian Ethics Committees	http://www.ethikkommissionen.at
	Bioethics Commission	http://www.bundeskanzleramt.at/site/3575 /default.aspx
	Austrian Agency for Health and Food Safety	http://www.ages.at/ages/en/ages-austrianagency-for-health-and-food-safety/
	Austrian Federal Office for Safety in Health Care	http://www.basg.at/en/austrian-federaloffice-for-safety-in-health-care/
	Austrian Data Protection Authority	https://www.dsb.gv.at/DesktopDefault.asp x?alias=dsken
Belarus	Ministry of Health (MOH)	http://minzdrav.by/en/
	National Bioethics Committee	
	State Pharmacological Committee	
	Centre for Expertise and Testing in Health Care	http://rceth.by/
	State Service of Forensic Medicine (SSFM)	
Belgium	Belgium Advisory Committee on Bioethics (BACB)	http://www.health.belgium.be/en
	Commission for the Protection of Privacy	http://www.privacycommission.be/
	Superior Health Council (CSS)	http://www.health.belgium.be/eportal/Abo utus/relatedinstitutions/SuperiorHealthCo uncil/index.htm
	Federal Public Service	www.health.fgov.be
	Federal Commission for Medical and Scientific Research on Embryos in Vitro	http://health.belgium.be/eportal/Healthcar e/Consultativebodies/Commissions/Embr yoinvitro/19076630?ie2Term=research&i e2section=83
Bosnia and Herzegovina	Ministry of Health	http://www.fmoh.gov.ba/

	Medicines and Medical Devices Agency of Bosnia and Herzegovina:	http://www.almbih.gov.ba/
	Personal Data Protection Agency of Bosnia and Herzegovina	http://www.azlp.gov.ba/Default.aspx?lang Tag=enUS&template_id=147&pageIndex=1
Bulgaria	Ministry of Healthcare	http://www.mh.government.bg/
	Bulgarian Drug Agency (BDA)	http://en.bda.bg/
	Bulgarian Commission for Personal Data Protection	https://www.cpdp.bg/en/index.php?p=rubr ic&aid=2
	Executive Agency for Transplantation	http://bgtransplant.bg/
	Council of Ministers, Ethics Committee for Transplantation	
Croatia	Ministry of Health (MZSS)	http://www.zdravlje.hr/
	Agency for Medicinal Products and Medical Devices	http://www.almp.hr/#
	Agency for Medicinal Products and Medical Devices	http://www.almp.hr/#
	Croatian Health Insurance Fund	http://www.hzzo.hr/en/
	Croatian Personal Data Protection Agency	http://www.azop.hr/
Cyprus	Ministry of Health, Pharmaceutical Services	http://www.moh.gov.cy/Moh/phs/phs.nsf/dmlindex_en/dmlindex_en?opendocument
	Ministry of Health, National Bioethics Committee	http://www.bioethics.gov.cy/moh/cnbc/cn bc.nsf/index_en/index_en?OpenDocumen t
	Commissioner's Office for the Protection of Personal Data	http://www.dataprotection.gov.cy/dataprot ection/dataprotection.nsf/index_en/index_ en?opendocument
Czech Republic	Ministry of Health, Central Ethics Committee	http://www.mzcr.cz
	State Institute for Drug Control	http://www.sukl.cz/index.php?lchan=1&lred=1
	Office for Personal Data Protection	http://www.uoou.cz/uoou.aspx
	Ministry of Education, Youth, and Sport	http://www.msmt.cz/index.php?lchan=1&lred=1
	Research and Development Council, Bioethical Commission	http://www.vyzkum.cz/FrontClanek.aspx? idsekce=15908
Denmark	National Committee on Health Research Ethics	http://dnvk.dk/CVK/Home/English.aspx
	Danish Medicines Agency	http://www.dkma.dk
	Patient Compensation Association	http://patienterstatningen.dk/en.aspx
	Danish Data Protection Agency	http://www.datatilsynet.dk/english/
	Danish Council of Ethics	http://www.etiskraad.dk/daDK.aspx?sc_lang=en
Estonia	Estonian Council on Bioethics	http://www.eetikakeskus.ut.ee/en
	State Agency of Medicines	http://www.sam.ee/en/clinical-trialsmedicinal-products-estonia
	Minister of Social Affairs	https://www.sm.ee/en
	Estonian Data Protection Inspectorate	http://www.aki.ee/en/inspectorate
Finland	Ministry of Social Affairs and Health	http://www.stm.fi/en/frontpage

	National Committee on Medical Research Ethics (TUKIJA)	http://www.tukija.fi/en
	National Advisory Board on Research Ethics (TENK)	http://www.tenk.fi/en
	Finnish Medicines Agency (FIMEA)	http://www.fimea.fi/frontpage
	National Supervisory Authority for Welfare and Health (VALVIRA)	http://www.valvira.fi/en/licensing/medical _devices
	Finnish Patient Insurance Centre	http://www.potilasvakuutuskeskus.fi/ww w/page/pvk_www_2181
	Office of the Data Protection Ombudsman	http://www.tietosuoja.fi/1560.htm
	Board for Gene Technology	http://www.geenitekniikanlautakunta.fi/en
	National Advisory Board on Social Welfare and Health Care Ethics	http://www.etene.fi/en
France	Ministry of Social affairs and Health	http://www.sante.gouv.fr/
	National Consultative Bioethic Committee for Health and Life Sciences (CCNE)	http://www.ccneethique.fr/en
	National Commission for Informatics and Freedoms (CNIL)	http://www.cnil.fr/english/the-cnil/
	National Halth Products Safety Agency (ANSM)	http://ansm.sante.fr/
Georgia	Bioethics and Health Law Studies Society	http://www.patientsrights.ge/index.php?pa ge=385(=geo
	State Regulation Agency for Medical Activities (LEPL) of the Ministry of Labor, Health, and Social Affairs	http://www.moh.gov.ge/index.php?sec_id =10(_id=ENG
	Office of the Personal Data Protection Inspector	https://personaldata.ge/en/data-protectionday-event-2014/177
Germany	Federal Institute for Drugs and Medical Devices (BfArM)	http://www.bfarm.de/EN/Home/home_no de.html
	Federal Ministry of Education and Research (BMBF)	http://www.bmbf.de/en/index.php 3.
	Paul Ehrlich Institute	http://www.pei.de/EN/home/node.html;jsessionid=8A56CBB11CA133D70C010434A47D96B7.1_cid329
	Federal Ministry of Health (BMG)	http://www.bmg.bund.de/ministerium/eng lish-version.html
	Federal Commissioner for Data Protection and Freedom of Information	http://www.bfdi.bund.de/DE/Home/home _node.html
	German Ethics Council (DER)	http://www.ethikrat.org/welcome?set_lan guage=en
Greece	National Organization for Medicines (NOM):	http://www.eof.gr/web/guest/home
	National Bioethics Commission (NBC):	http://www.bioethics.gr/index.php?category_id=3
	Hellenic Data Protection Authority	http://www.dpa.gr/
Hungary	Ministry of Human Resources (EMMI)	http://www.kormany.hu/hu/emberieroforrasok-miniszterium
	Medical Research Council, Research Ethics Committees (KFEB, TUKEB, HRB)	http://www.ett.hu/
	National Institute of Pharmacy and Nutrition	http://www.ogyei.gov.hu
	Medical Research Council, Ethics Clinical Pharmacology Ethics Committee (KFEB)	http://www.ett.hu/kfeb.htm
	Authority for Medical Devices, Health Registration and Training Center	http://www.enkk.hu/index.php/hun/

	Hungarian National Authority for Data Protection and Freedom of Information	http://www.naih.hu/generalinformation.html
	Medical Research Council, Committee for Human Reproduction (HRB)	http://www.ett.hu/hrb.htm
Iceland	Ministry of Welfare (MOW): http://eng.velferdarraduneyti.is	http://eng.velferdarraduneyti.is/
	National Bioethics Committee (NBC)	http://www.vsn.is/en
	Icelandic Medicines Agency (MCA)	http://www.ima.is/
	Icelandic Health Insurance Agency (MCA)	http://www.sjukra.is/english
	Data Protection Authority	http://www.personuvernd.is/informationin-english/
Ireland	Department of Health	http://health.gov.ie/
	Health Products and Regulatory Authority	https://www.hpra.ie/
	Data Protection Commissioner	http://www.dataprotection.ie/docs/Home/ 4.htm
Italy	National Monitoring Center for	http://www.agenziafarmaco.com/en/conte nt/national-monitoring-centre-clinicaltrials
	Clinical Trials	
	Italian Medicines Agency	http://www.agenziafarmaco.it/
	Ministry of Health (MOH)	http://www.ministerosalute.it
	Italian Data Protection Independent Authority	http://www.garanteprivacy.it/garante/navi g/jsp/index.jsp?solotesto=N
Latvia	State Agency of Medicines	http://www.zva.gov.lv/?setlang=en&large
	Central Medical Ethics Committee	
	Data State Inspectorate	http://www.dvi.gov.lv/en/
Lithuania	Ministry of Health (MOH)	http://www.sam.lt/go.php/lit/IMG
	State Medicines Control Agency (SMCA)	http://www.vvkt.lt/lit/English
	State Health Care Accreditation Agency Under the Ministry of Health (SHCA)	http://www.vaspvt.gov.lt/en
	State Data Protection Inspectorate	https://www.ada.lt/go.php/lit/English
Luxembourg	Ministry of Health	http://www.ms.public.lu
	National Research Ethics Committee	http://www.cner.lu
	Division of Pharmacy and Medicines of the Ministry of Health	
	National Commission for Data Protection	http://www.cnpd.public.lu/fr/index.html
Macedonia	Ministry of Health of Republic of Macedonia	http://moh.gov.mk/
	Drug and Devices Register	https://lekovi.zdravstvo.gov.mk/
	Drug Agency	http://malmed.gov.mk/
	Directorate for Personal Data Protection	www.dzlp.mk

Malta	Medicines Authority: http://medicinesauthority.gov.mt/	
	Malta Competition and Consumer Affairs Authority, Technical Regulations Division, Regulatory Affairs Directorate	http://www.mccaa.org.mt/en/regulatoryaffairs-directorate
	Office of the Information and Data Protection Commissioner	http://idpc.gov.mt/index.aspx
Moldova	Ministry of Health, National Committee for Ethical Expertise of Clinical Trials	http://ms.gov.md/?q=comitetul-nationaletica
	Medicines and Medical Devices Agency	http://www.amed.md/
	National Center for Personal Data Protection of the Republic of Moldova	http://www.datepersonale.md/en/start/
	National Commission on Biological Security	http://lex.justice.md/index.php?action=vie w&view=doc(=1&id=303353
Montenegro	Ministry of Health of Montenegro	http://www.mzd.gov.me/en/ministry?alph abet=lat
	Agency for Medicines and Medical Devices:	https://www.calims.me/Portal/faces/glavn a?_adf.ctrl-state=rsbe35pln_83
	National Security Agency	http://www.anb.gov.me/en/Home?alphabe t=lat
Netherlands	Central Committee for Research Involving Human Subjects (CCMO)	http://www.ccmo.nl/en/
	Ministry of Health, Welfare, and Sport (VWS)	http://www.government.nl/ministries/vws #ref-minvws
	Netherlands Trial Register	http://www.trialregister.nl/trialreg/index.a sp
	Federation of Biomedical Scientific Societies (FMWV)	http://www.federa.org/
	Dutch Data Protection Authority	https://cbpweb.nl/en
	Ministry of Infrastructure and the Environment (IenM)	http://www.government.nl/ministries/ien m
	Dutch Health Care Inspectorate (IGZ)	http://www.igz.nl/english/
Norway	National Committee for Medical and Health Research Ethics (NEM)	http://www.etikkom.no/en/InEnglish/Committee-for-Medical-andHealth-Research/
	Norwegian Directorate of Health	http://www.helsedirektoratet.no/kvalitetplanlegging/medisinsk-utstyr/kliniskutprovning/Sider/default.aspx
	Regional Committees for Medical and Health Research Ethics	https://helseforskning.etikkom.no/ikbVie wer/page/forside
	Data Inspectorate	http://www.datatilsynet.no/English
	Ministry of Health and Care Services (MHCS)	https://www.regjeringen.no/en/dep/hod/id 421/
	Ministry of Education and Research (MER)	http://www.regjeringen.no/en/dep/kd.html ?id=586
	Norwegian Biotechnology Advisory Board	http://www.bion.no/english/
	Regional Committees for Medical Research Ethics (REK)	https://www.etikkom.no/en/InEnglish/Committee-for-Medical-andHealth-Research/
	Directorate for Health and Social Affairs	http://www.helsedirektoratet.no/kvalitetplanlegging/biogenteknologi/Sider/default.aspx
Poland	Ministry of Health, Bioethics Appeals Commission (MOH)	http://www.mz.gov.pl/en
	Center of Bioethics, Polish Chamber of Physicians and Dentists (NIL)	http://www.nil.org.pl/dzialalnosc/oro dek-bioetyki
	Office for Registration of Medicinal Products, Medical Devices, and Biocidal Products	http://en.urpl.gov.pl/general-information

	Inspector General for the Protection of Personal Data	http://www.giodo.gov.pl/168/j/en/
Portugal	National Institute of Pharmacy and Medicines	http://www.infarmed.pt/portal/page/portal/ INFARMED/ENGLISH
	Ethics Commission for Clinical Research (CEIC)	http://www.infarmed.pt/portal/page/portal/ INFARMED/MEDICAMENTOS_USO_H UMANO/CEIC
	National Data Protection Commission	http://www.cnpd.pt/english/index_en.htm
	Ministry of Health	http://www.portugal.gov.pt/en/theministries/ministry-of-health.aspx
	National Council of Ethics for the Life Sciences	http://www.cnecv.gov.pt/cnecv/en/
Romania	jon	http://www.ms.ro/
	National Agency for Medicines and Medical Devices	http://www.anm.ro/anmdm/en/index.html
	National Bioethics Committee for Medicines and Medical Devices	http://www.bioetica-medicala.ro/
	National Supervisory Authority for Personal Data Processing	http://www.dataprotection.ro/index.jsp?pa ge=documents(=en
Russia	Ministry of Healthcare of the Russian Federation (MOH)	http://www.rosminzdrav.ru
	Federal Service on Surveillance in Healthcare (Roszdravnadzor)	http://www.roszdravnadzor.ru/
	Russian Committee for Bioethics	http://www.bioethics.ru/eng/
	Association of Clinical Trials Organizations	http://acto-russia.org/en/
	Federal Agency for Technical Regulation and Metrology (GOST)	http://www.gost.ru/wps/portal/pages.en.Main
	Interdepartmental Commission on Genetic-Engineering Activity	
San Marino	San Marino Bioethics Committee	http://www.sanita.sm/online/home/comitato-bioetica/comitatosammarinese-di-bioetica.html
Serbia	Ministry of Health (MOH)	http://www.zdravlje.gov.rs/
	Serbian Drug Agency	http://www.alims.gov.rs/eng/
	Commissioner for Information of Public Importance and Personal Data Protection	http://www.poverenik.rs/en/thecommissioners-authority-di.html
Slovakia	Ministry of Health	http://www.health.gov.sk/
	Institute of Medical Ethics and Bioethics	http://www.bioethics.sk/
	State Institute for Drug Control:	http://www.sukl.sk/en
	Office for Personal Data Protection	https://dataprotection.gov.sk/uoou/en
Slovenia	Republic of Slovenia National Medical Ethics Committee (NMEC)	http://www.kme-nmec.si/
	Agency for Medicinal Products and Medical Devices (JAZMP)	http://www.jazmp.si/
	Information Commissioner of the Republic of Slovenia	http://www.ip-rs.si/
Spain	Spanish Bioethics Committee	http://www.comitedebioetica.es/?lang=en_ US
	Coordinating Center for Ethical Committees on Clinical Research	http://www.msc.es/profesionales/farmacia/ ceic/home.htm

	Institute of Health Carlos III, Ministry of Science and Innovation	http://www.isciii.es/htdocs/en/index.jsp
	Spanish Agency of Medicines and Medical Devices	http://www.aemps.gob.es/en/investigacion Clinica/medicamentos/home.htm
	Ministry of Health and Consumption	http://www.msc.es/en/home.htm
	National Commission for the Donation and Use of Embryos, Cells, and Human Tissues for Biomedical Research	http://www.isciii.es/ISCIII/es/contenidos/f d-el-instituto/organizacion.shtml
	National Biobank Register	http://www.isciii.es/ISCIII/es/contenidos/f d-el-instituto/organizacion.shtml
	National Stem Cell Bank	http://www.isciii.es/ISCIII/es/contenidos/f d-el-instituto/fd-organizacion/fdestructura-directiva/fd-subdirecciongeneral-investigacion-terapia-celularmedicina-regenerativa/fd-centrosunidades/banco-nacional-lineascelulares.shtm
Sweden	Central Ethical Review Board	http://www.epn.se/en/start/
	Medical Products Agency	https://lakemedelsverket.se/english/
	National Board of Health and Welfare (SOS)	http://www.socialstyrelsen.se/english
	Swedish Research Council (SRC)	http://www.vr.se/english
	BBMRI Sweden	http://bbmri.se/en/
	Ministry of Health and Social Affairs	http://www.sweden.gov.se/sb/d/2061
Switzerland	Federal Office of Public Health (FOPH)	http://www.bag.admin.ch/index.html?lang =en
	National Advisory Commission on Biomedical Ethics (NEK-CNE)	http://www.nek-cne.ch/en/homepage/
	Swiss Ethics Committees on Research Involving Humans	http://www.swissethics.ch/index_e.html
	Swiss Agency for Therapeutic Products	http://www.swissmedic.ch/index.html?lang=en
	Swiss National Clinical Trials Portal	http://kofam.ch/en/swiss-clinical-trialsportal/
	Federal Data Protection and Information Commissioner (FDPIC)	http://www.edoeb.admin.ch/index.html?la ng=en
	Swiss Academy of Medical Sciences (SAMS)	http://www.samw.ch/en/News/News.html
Ukraine	Ministry of Health of Ukraine State Expert Center	http://www.dec.gov.ua
	National Academy of Sciences Bioethics Committee	http://biomed.nas.gov.ua/index- en/bioethics-committee
	Ukrainian Parliament Commissioner for Human Rights:	www.ombudsman.gov.ua
	Ukrainian Ministry of Health	http://www.moz.gov.ua/en/
United Kingdom	Health Research Authority (HRA)	http://www.hra.nhs.uk/
	Department of Health (DH)	https://www.gov.uk/government/organisations/department-of-health
	NHSScotland, Chief Scientist Office (CSO)	http://www.cso.scot.nhs.uk/
	NHS Research Scotland	http://www.nhsresearchscotland.org.uk/
	Medicines and Healthcare Products Regulatory Agency (MHRA)	https://www.gov.uk/government/organisati ons/medicines-and-healthcare-productsregulatory-agency
	Administration of Radioactive Substances Advisory Committee (ARSAC)	https://www.gov.uk/government/organisati ons/administration-of-radioactivesubstances-advisory-committee

	Information Commissioner's Office:	https://ico.org.uk/
	Confidentiality Advisory Group (CAG)	http://www.hra.nhs.uk/about-the-hra/ourcommittees/section-251
	Human Tissue Authority (HTA)	http://www.hta.gov.uk/
	Healthcare Improvement Scotland	http://www.healthcareimprovementscotlan d.org/our_work/governance_and_assuranc e/human_tissue_banks.aspx
	Human Fertilisation and Embryology Authority	Human Fertilisation and Embryology Authority: http://www.hfea.gov.uk/

14. Appendix 4

Multi-disciplinary cases:
Cases from the Markkula Center for Applied Ethics --at Santa Clara University. Includes cases on business ethics, medical ethics, governmental, and technology ethics.
Dilemma Database, Institute for Global Ethics--Includes cases on business ethics, educational ethics , medical ethics, as well as philanthropic and personal ethics. Cases presented without resolutions.
Ethics Across the Curriculum - Case Studies-- Utah Valley State College. Emerging from their Ethics Across the Curriculum program in 1993, this page has a large collection of cases dealing with business, technology, medical ethics, etc.
Ethics Case Studies Database --(limited selection) from Philosophical & Religious Studies program at Leeds University, UK. Mostly dealing with medical and end-of-life ethics. Database searchable by keyword.
Ethics Education Library - database of cases from engineering, medicine, science, business, etc. developed by CSEP
Ethics Updates' Ethics Case Studies -- hosted by the University of San Diego. A large collection of cases covering everything from academic to social ethics.
Hale Chair in Applied Ethics Resources-- Rochester Institute of Technology. Twenty three cases detailed cases cover a wide variety of ethical topics by Lisa Newton, Director of the Program in Applied Ethics at Fairfield University.
The Center for Ethics in the Professions - Case Studies --at the University of Puerto Rico. Cases divided by topic including engineering ethics, science and research ethics, business ethics and computer ethics.
<p >Vanderbilt University Center for Ethics--Cases divided by topic including bioethics, journalism ethics, and military ethics.
Archaeology
Ethics in Anthropology - a large collection of case studies, codes of ethics, and other resources on ethics in anthropology put together by the University of Illinois at Urbana Champaign
Architecture Ethics Cases
American Institute of Architects National Ethics Council Advisory Opinions - a small collection of cases presented to the AIA's National Ethics Council, and the discussion and opinions of the Council. These cases include references to sections in the 2004 AIA's Code of Ethics.
Business & Accounting Ethics Cases
AICPA's Accounting and Business Ethics case studies--American Institute of Certified Accountants. Includes pdf file of case and teaching notes. Case on accounting fraud and corporate governance.
AICPA's cases & commentaries, "Ethics and Fraud in Business" --Much like the previous site, this collection of cases covers professional ethics, corporate leadership, regulatory, legal and financial analysis in the context of ethical dilemmas. Most cases include commentary.
Arthur Andersen - Ethics case studies: 1987-1994 -Carnegie Mellon Tepper School of Business. This collection of 90 case studies is a product of program sponsored by Author Anderson to promote ethical awareness in business schools. Cases cover a wide range of topics.
BusinessEthics.ca -a list of links to other cases. Site also includes a bibliography of articles and books on business ethics.
CasePlace.org-- Aspen Institute Business and Society Program. This is one of the largest collections of business ethics cases freely available on the internet. Cases can be searched by discipline, topic, or keyword.
Emerson Center for Business Ethics cases--John Cook School of Business, St. Louis University. Includes Center-sponsered case studies and case studies written by students.
Institute of Business Ethics Index to Cases- a small but interesting collection of business ethics cases. Many cases include teaching guides and bibliographies.
Web Miner's Guide to Business Ethics Cases-a comprehensive list of links to other business ethics case sources.

Computer Application/Programming Cases
Cases about Computers and Software-Online Ethics Center for Engineering and Science,. Small collection of cases.
ComputerCases.org- site contains three well-researched historical computer cases including accompanying teaching materials and supporting documents. The coases covered are Therac-25 (a radiation therapy device that delivered fatal overdoses), Machado (the first successfully procecuted case of an online hate crime), and Huges Aircraft (a whistleblowing case).
Education Cases
Case Studies in Education-Utah Valley State College. A collection of cases covering all aspects of education ethics from plagiarism to tenure.
Research Ethics and Academic Integrity- a collection of videos showing case studies dealing with academic integrity. Produced by Syracuse University. Videos available online.
Engineering Ethics Cases
ASCE's Journal of Professional Issues in Engineering Education and Practice-- each issue includes a case study. Full text of journal available.
Design Failure Lessons--from University of Texas at Austin. Covers four design flaw cases including the Ford Pinto Case and the Challenger space shuttle explosion. Each case includes links to a number of background articles and references.
Civil Engineering Cases -Carleton University, Rose-Hulman Institute of Technology. Not the best design, but find the section Contents of this Catalogue and follow link to the list of cases, and you can find the full text of a number of engineering cases. Cases can also be ordered.
Engineering.com Ethics Case Studies- a collection of 18 cases including a timeline of events, a detailed examination of the incident, ethical points for discussion, and a bibliography for each case.
Engineering Ethics: Concepts and Cases-cases from the book, Engineering Ethics: Concepts and Cases by Charles E. Harris Jr.
Learning from Failure Engineering Disasters --Materials Science and Engineering at University of New York, Stonybrook. An introduction to engineering failures, also includes a link to cases written by faculty.
NSPE Board of Ethical Review Cases - NIEE.org- A list of real cases presented to the National Society of Professional Engineers review board from 1976-2001. Includes summary of case and decisions made in relation to the NSPE Code of Ethics.
Online Ethics Center Cases About Engineering Professional Practice-an extensive list of cases provided by the National Academy of Engineering .
Penn State College of Engineering ethics cases a list of cases available on the web, as well as a collection of articles on using cases to teach engineering ethics.
Teaching Engineering Ethics with Cases -essay & resources by Michael Pritchard
Texas A&M University Engineering Ethics Cases -case studies taken from a variety of sources. .
UBC Library - Examples of Engineering Failures- University of British Columbia. A collection of reports, press releases, and statements relating to famous engineering failures.
Journalism Cases
Ethics Case Studies- Society of Professional Journalists. A small collection of case studies. Site also includes link to SPJ Code of Ethics.
Journalism Ethics Cases Online- School of Journalism, Indiana University-Bloomington. A large list of cases covering all aspects of journalism. Indexed by topics.
Poynter Online Journalistic Case Studies- an impressive collection of journalism cases. Page also includes a number of 'tip sheets' that describe how to deal with cases of plagiarism, undisclosed sources, etc.
Military Cases

Ethics of War Peace and Terrorism- University of San Diego. Small collection of cases. Site also has a number of articles and other resources on military ethics.
Military cases-Air University, Maxwell-Gunter Air Force Base. A small collection of summarized case studies. Site also includes links to articles on military ethics.
Science & Medicine/Bioethics Cases
Bioethics Outreach--Case Studies - Iowa State University. A large collection of case studies covering a wide range of subjects from biotechnology to agriculture.
Case Study Collection -National Center for Case Study Teaching in Science, State University of New York in Buffalo. This site has cases dealing with all different branches of biology and medicine. Cases are divided by discipline and each case includes teaching notes.
Center for Bioethics and Human Dignity case studies- a small collection of case studies put out by the American Medical Association and the Center for Bioethics and Human Dignity . Each case includes commentary from a Christian perspective.
Ethics Case Studies in Biodesign- Stanford University. A collection of 15 short cases dealing with biodesign.
Ethics in the Science Classroom -Western Michigan University. A small collection of cases dealing with genetic testing, racism in science, and research ethics.
Moral Reasoning in Scientific Research- Poynter Center for the Study of Ethics and American Institutions. Indiana University Bloomington. Cases taken from book Moral Reasoning in Scientific Research.
UK Clinical Ethics Network--Case Studies-a small collection of medical ethics cases.
Other Disciplines (education, funeral, etc)
American Anthropological Association -Ethical Currents Case Studies
Cyber ethics--case studies -a small collection of cases dealing with censorship on the internet and recent policies such as the Communications Decency Act. Some links are broken.
Funeral Ethics Organization- a collection of short cases exploring the professional ethics of the funeral industry.
Ethics for Trades and Technology- Utah Valley State College. A collection of cases dealing with general ethical questions in workplace.

15. Appendix 5

Cases and case studies
Anthropology:
Cases and Solutions from Ch 3 of the Handbook on Ethical Issues in Anthropology Cases and Comments from Ch 4 of the Handbook on Ethical Issues in Anthropology Smithsonian Institution: Anthropology Outreach Office--Ethical Dilemmas List of Ethics in Anthropology cases by Sharon Stoerger MLS, MBA Anthropology Ethics Cases and Resources from the Kutak Center for the Teaching and Study of Applied Ethics
Art:
List of Art Ethics cases by Sharon Stoerger MLS, MBA
Bioethics and Medical ethics:
Case Studies at the Center for Bioethics and Human Dignity (skip the AMA links - use "our links" instead) and go down to "Other Case Studies Produced by CBHD Fellows and Friends" The CBHD resource center also offers many proactive articles on issues in bioethics that can serve as discussion focus-points. McGraw Hill Bioethics Case Studies Bioethics at Iowa State University – Case Studies for the classroom
Business
Caseplace.org A free service of The Aspen Institute's Business and Society Program, casePlace.org is a practical and dynamic resource for up-to-date case studies, syllabi and innovative teaching materials on business and sustainability— from corporate governance to sustainable development. Arthur Andersen Case Studies in Business Ethics During the period 1987-94 Arthur Andersen funded a $5 million joint project with 525 universities to raise awareness of ethical issues in business. This collection of 90 case studies is one product of that effort. All participating universities, including USC, have license to use these materials and reproduce them as needed for instructional purposes. Other users are advised to consult an attorney regarding copyright issues. Harvard Business School Publishing Good ethics cases available for purchase. Restrict your search by "case" and search for your desired topic. Darden Case Collection (updated link coming) Ethics Cases on i-Case Website: Interactive cases; pdf files linked to audio and video segments. European Case Clearinghouse at Cranfield University A comprehensive source of management case studies from around the world. It distributes cases produced by the world's best-known management teaching establishments, as well as case studies in many languages produced by individual authors from almost every corner of the globe. Resource: Ethics in Small Business
Engineering
Cases and Scenarios from the Online Ethics Center for Engineering and Research Engineering Professional Practice cases from the Online Ethics Center for Engineering and Research

Engineering Ethics Case studies from the Ethics Education Library

Case Studies from the book Engineering Ethics: Concepts and Cases

Cases from Dr. Mike Rabins, Dr. Charles Harris, Dr. Michael Pritchard, Dr. Lee L. Lowery, Jr. and others on a NSF grant at Texas A&M University

Engineering Ethics Cases from Engineering.com

From Penn State College of Engineering

Engineering Ethics Cases from the Markkula Center for Applied Ethics

Environmental Ethics

Domestic and international cases Environmental Justice Case Studies written by University of Michigan Students

Cases from James Madison University College of Business

Cases from the Center for Ethics in the Professions at the University of Puerto Rico

International Relations

International Relations Ethics Case Studies

The Carnegie Council offers this series of 22 case studies for use in college and university classrooms. Each case presents and analyzes an historical example of an ethical dilemma in international affairs.

Journalism

Society of Professional Journalists Ethics Case Studies

Poynter Center Ethics Tip Sheet Case Studies

Indiana University School of Journalism Ethics Cases Online

Cases from Media Ethics: Issues and Cases 5th Edition (McGraw Hill)

Military Ethics

Cases from Air University: the intellecutal and leadership center of the Air Force

Cases from the Ethics Education Library

Cases from Texas A&M

Interactive Dillemas from the United States Naval Academy Stockdale Center for Ethical Leadership

Physics

American Physical Society: Ethics Case Studies

Sociology

Resources for Teaching Ethics in Sociology from the ASA

- By undergrad class
- By topic

Various Subjects

The Markkula Center for Applied Ethics

Cases on virtually all applied ethics topics, including campus ethics by one of the largest centers for applied ethics on the web.

Ethics Bowl Cases

Short cases (1-2 pages) used in Ethics Bowl competitions on a large variety of subjects (not organized by subjects)

CQ Researcher

Many well-prepared journalistic reports on a variety of ethics-related topics, including topic overview, pro/con positions on key questions, and situation analysis. You can also create email alerts that will let you know when new reports are posted on your favorite topics.

Mapping Normative Frameworks of Ethics and Integrity of Research

Report on Task 5.2 Pilot Collection of Data on Cases

Report on Task 5.2 Pilot Collection of Data on Cases

Project details	
Project:	Mapping Normative Frameworks of Ethics and Integrity of Research
Project acronym:	EnTIRE
Project start date:	01.05.2017
Duration:	48 months
Project number:	741782
Project Coordinator:	Vrije Universiteit Medisch Centrum (VUmc) Amsterdam

Deliverable details	
Work Package:	WP 5 Cases, Casuistry and Scenarios
Work package leader:	Dublin City University
Authors:	Anna Catharina Armond, Kristóf János Bodnár, Bert Gordijn, Søren Holm, Mohammad Hosseini, Péter Kakuk, Afroditi Kalambouka and Jonathan Lewis

Table of Contents

1. Summary

This report provides an overview of the pilot searches for Research Ethics and Research Integrity cases. Details of the methods employed to search for, retrieve and tag RE+RI cases are provided in 'D.5.1. Protocol for Systematic Searches and Tagging of RE+RI Cases'. The aim of this report is to make explicit and justify any substantial changes that have been made to the D.5.1 protocol during months 6 to 12 of the EnTIRE project.

2. Description of Work

T.5.2. Pilot collection of data on cases (M6-12, DCU, UNIDEB, Manchester)

1. Conduct pilot searches in IE, HU and UK within each of the potential pools of RE+RI cases to see whether the initially chosen search methodologies are adequate and feasible. Make adjustments, if necessary.
2. Conduct structuring, posting and testing activities on the web platform.
3. Adjust the system of tagging based on the results of the pilot searches and the normative framework defined in the focus groups (WP2).

Task 5.2 was carried out between 1st November 2017 and 30th April 2018 (M6-12) of the EnTIRE project. The task put the protocol into practice by conducting pilot searches for RE+RI cases to see whether the search methodologies and tagging system detailed in the initial protocol were adequate and feasible.

This report explains how the protocol has been adjusted according to the results of the pilot searches, the development of the online platform and the input from WP2's stakeholder consultation.

3. Changes to the Search Methodologies for RE+RI Cases

3.1 Academic Literature

Details of the search methodology employed to search for RE+RI cases in academic literature are provided on pp. 27-33 of the protocol (D.5.1). This section specifies the changes made to that protocol and provides justification for the proposed alterations.

3.1.1 Scoping Search

The first phase of the pilot was the scoping search, which helped to define the best keywords for the search. According to the protocol, we shall 'create and refine a set of search terms using Boolean operators, one set of search terms should include ''research ethics" and another search should incorporate the term "research integrity" OR "research misconduct"'. At this stage, we tested the search-term combinations suggested in the protocol using PubMed and Science Direct. Two researchers independently checked the results and screened the content (titles, journal names, genres, abstracts) of the first hundred results. As a team, we discussed the feasibility and precision of the keywords suggested in the protocol. The protocol suggests that we should develop a parallel search methodology, using two sets of keywords in order to generate two separate collections, one for research ethics (RE) and one for research integrity (RI). The two separate searches could enable us to analyse the overlaps and differences between, and evolution of the representation of, RE and RI in the academic literature.

3.1.2 Refinement of Search Terms

The original keywords suggested in the protocol need refinement. Executing a search with the original set of keywords ("research ethics" and "case") resulted in a huge number of results that could not be adequately accounted for within in the remits of the project due to a lack of time and resources. Limiting the number of findings by date restrictions (e.g. results after 1990) is inconsistent with the PRISMA statement. Consequently, we needed to look for other terms and variations of terms. Another important observation emerged at this stage of the

piloting. The term “case” or “case study” is broad and its inclusion (using the Boolean operator “AND”) both excluded important results and included many results unrelated to the topic of RE+RI. For example, using (“research ethics” AND “case”) as a keyword set in SCOPUS, PubMed, Science Direct, Web of Science, JSTOR and Ovid, 57,954 results were retrieved (see *Table 1*). For this reason, we decided not to use the term “case”. However, we recognise that the retrieval of cases is the key component of this task. On that basis, the search for “cases” will take place in two phases: i) by reading the titles and abstracts of retrieved articles; ii) by reading the articles.

Subsequently, we tested different keywords in order to generate manageable results. We decided to continue to refine keywords by testing candidate keywords and their variations in SCOPUS, PubMed, Science Direct, Web of Science, JSTOR and Ovid. Based on our experiences of conducting systematic reviews according to PRISMA guidelines, and based on the fact that the results would need to be screened, read and cross-checked by three reviewers, we estimated that between 10,000 and 15,000 articles would be a manageable size of aggregate results for the search. If the keyword variation generated results that significantly exceeded or fell short of this aggregate number, then we modified the keywords accordingly. We also tested two hundred results for sensitivity. If the initial two hundred results included a high number of irrelevant articles, then we had suitable justification to modify the search terms.

In order to identify research ethics (RE) cases in the academic literature, we developed the search term ((“research ethics”) AND (“violation” OR “unethical” OR “misconduct”)). It generated a manageable number of precise and relevant results.

For research integrity (RI) cases, the protocol suggested (“research integrity” AND “case”). We tested this search term in SCOPUS, PubMed, Science Direct, Web of Science, JSTOR and Ovid. Although the results were of a manageable size, they were, on the whole, too unrelated to the topic of RI (see *Table 2*). Since we decided to undertake two parallel searches for comparative purposes, we tested the same keywords employed in the RE search and screened the first two hundred findings for precision. Employing the search term ((“research integrity”) AND (“violation” OR “unethical” OR “misconduct”)) generated 3,077 titles.

3.1.3 Selection of Databases

The second phase of the pilot determined the databases that will be used for the up-scaled searches of task 5.3. Using all of the databases listed in the protocol would generate a huge number of results that could not be adequately accounted for within the remits of the project due to a lack of time and resources. Furthermore, due to access restrictions, we can feasibly only use those databases that are available through our institutional subscriptions. We decided to select the largest and the most commonly-used databases for systematic reviews and those that would provide a comprehensive set of results.

The databases are PubMed, Web of Science, SCOPUS, JSTOR, Ovid and Science Direct. The overlaps between the databases will be removed using EndNote citation software. The titles and abstracts will be read and those that possibly contain cases will read in full.

3.1.4 Inclusion and Exclusion Criteria

Three researchers will read the aggregated results (titles, abstracts) in order select the relevant articles. Articles will be included if it contains either a RE or a RI case, whereby a case, according to the protocol, 'is either a fictional or non-fictional "violation" of, or "misbehaviour", "poor judgment" or "questionable research practice" in relation to, or act demonstrating best practice and/or scientific excellence relative to, the explicit and/or implicit rules governing RE+RI'. Non-fictional and fictional cases will be included without any language or date restriction. Exclusion criteria will be: cases unrelated to academia, scientific activities and institutions or academic and industrial research and publication.

After selection processes have been completed, the three researchers will compare and cross-check their individual results. After performing the search, the reference list will be checked for the inclusion of any additional articles that have not been found in the online database search.

Table 1: Refinement of Search Keywords for Research Ethics (RE) Cases		
Databases	**Keywords**	**Total number**
PubMed Ovid Science Direct Web of Science JSTOR SCOPUS	"research ethics" AND case	57,954
PubMed Ovid Science Direct Web of Science JSTOR SCOPUS	("research ethics") AND (violation OR unethical)	6,181
Final Results for Research Ethics Keywords		
Databases	**Keywords**	**Total number**
PubMed Ovid Science Direct Web of Science JSTOR SCOPUS	("**research ethics**") AND (violation OR unethical OR misconduct)	10,548

Table 2: Refinement of Search Keywords for Research integrity (RI) Cases		
Databases	**Keywords**	**Total number**
PubMed Ovid Science Direct Web of Science JSTOR SCOPUS	"research integrity" AND case	10,492
PubMed Ovid Science Direct Web of Science JSTOR SCOPUS	("research integrity") AND (violation OR unethical)	2,356
Final Results for Research Integrity Keywords		
Databases	**Keywords**	**Total number**
PubMed Ovid Science Direct	("**research integrity**") AND (violation OR unethical OR misconduct)	3,077

Web of Science JSTOR SCOPUS		

3.1.5 Summary

As a result of the pilot search, the initial protocol was fine-tuned according to project needs and resources in accordance with PRISMA guidelines. The resulting changes can be summed up as follows:

1) The protocol keywords have been refined. For research ethics cases, the search term will be ((“research ethics”) AND (violation OR unethical OR misconduct)). For research integrity cases, the search term will be ((“research integrity”) AND (violation OR unethical OR misconduct));

2) The reference and citation databases have been limited to six of the most available and commonly-used databases: PubMed, Ovid, Science Direct, Web of Science, JSTOR and SCOPUS.

3.2 Reports by RE+RI Committees and Regulatory Bodies

Details of the search methodologies employed to search for RE+RI cases in reports by RECs, RIOs and Regulatory Bodies are provided on pp. 34-42 of the protocol (D.5.1).

Since the protocol was completed, several issues have arisen that have the potential to affect the methodologies detailed in the protocol:

I. The language of any request being sent to RECs, RIOs and Regulatory Bodies;
II. The legalities of making non-public RE+RI cases publically available on the EnTIRE platform;
III. The lack of resources available to transcribe cases from hard-copy to electronic format for the purposes of upload to the online platform;
IV. The need for specifically electronic cases that are already publically available and that can be located via a URL.

In terms of the language issue, this matter was discussed in a meeting with the secretary of EUREC. Since it was made clear in the recent EnTIRE Work Package Leaders' Meeting in Bonn that the official language of the project is English, and since the resources available to WP5 limit the opportunity to engage with RECs and RIOs in multiple languages, it has been decided that our requests for cases can only be made in English. WP5 recognise that this may affect the response rate.

Following discussions with other work package leaders in Bonn, and with the legalities of the platform's data protection policy, publication policy and general terms and conditions still to be finalised, it has been decided that WP5 can only request cases from RECs and RIOs that are already available to the public. Furthermore, because the legalities concerning the publication of case details are still to be finalised and due to a lack of resources available to WP5, it has been decided that these publically-available cases should be locatable online via a URL link.

In order to produce a revised methodology for retrieving cases from RECs, RIOs and Regulatory Bodies, the president of a central REC and the head of an academic REC were

interviewed. In addition, a virtual meeting was organised with the secretary of EUREC in order to discuss the challenges of case retrieval. WP5 received the following guidance:

I. Response rates from the various RECs, RIOs and Regulatory Bodies could be influenced by timing, language, ease of response and form of request (email, online, user friendly);
II. Because EUREC is frequently asked to conduct or support surveys, WP5 should avoid duplicating previous requests;
III. The period May-June is suggested as an appropriate timeframe for case requests;
IV. EUREC have offered to comment on our request for cases when it is drafted and it will assist WP5 in the dissemination of the request to its members via its central group email.
V. The request should be drafted before the 1st May 2018. Comments from WP5 should be provide and changes made to the request prior to 15th May 2018. EUREC will disseminate the request around 15th May 2018.

3.2.1 Summary

Taking into account the proposed methodology detailed in section 7.2 of the protocol, we propose the following amended case retrieval strategy for RECs (see *Table 3*):

Table 3: Revised Retrieval Strategy for RECs

Description	Timeframe/Deadline
1. Debrecen to send to DCU a list of their RIOs- and Regulatory-Body-related sources retrieved through the former's Grey Literature search;	15th April 2018
2. Debrecen to undertake an internet-based search for publically-available, online RE+RI cases using the list of 'umbrella' RECs detailed in Appendix 1 of the protocol. Where there is duplication of sources/cases already	15th April 2018

retrieved though Debrecen's Grey Literature search, the duplicated sources/cases will be reported in the Grey Literature systematic review and NOT in the RECs, RIOs and Regulatory Bodies searches;	
3. Debrecen to prepare a list of those 'umbrella' RECs detailed in Appendix 1 for whom they cannot identify online repositories of cases;	30th April 2018
4. Where cases can be found online for the organisations detailed in Appendix 1 of the protocol, cases will be entered into the online data template on the EnTIRE Wiki staging area. Where cases have been retrieved through the Grey Literature search, the same cases will NOT be reported in the RECs, RIOs and Regulatory Bodies searches;	1st May 2018-30th November 2018
5. WP5 will prepare an electronic request in English for those RECs, RIOs and Regulatory Bodies for whom it cannot identify online repositories of cases. The electronic request will also provide details of the EnTIRE project and its plans for tagging, disclosing and disseminating RE+RI cases. Bearing in mind that the RE+RI cases need to be in English and open to the public and taking into account the fact that we are legally constrained by the case details we make available on the online platform, WP5 should state that they can only accept publically-available cases in English and in a digital format;	1st May 2018
6. WP5 will provide comments and changes to the draft electronic request;	15th May 2018
7. Debrecen will send the list of 'umbrella' RECs to EUREC, requesting that EUREC distribute the electronic request to the organisations on the list.	15th May 2018

8. Where no correspondence is forthcoming, a second electronic request will be sent two weeks after the date of the first request directly to 'umbrella' RECs detailed in Appendix 1 for whom Debrecen cannot identify online repositories of cases;	1st June 2018
9. If there is no response to the second electronic request, a final request will be sent two weeks later.	15th June 2018

Taking into account the proposed methodology detailed in section 7.2 of the protocol, we propose the following amended case retrieval strategy for RIOs and Regulatory Bodies (see *Table 4):*

Table 4: Revised Retrieval Strategy for RIOs and Regulatory Bodies

Description	Timeframe/Deadline
1. Debrecen to send to DCU a list of their RIOs- and Regulatory-Body-related sources retrieved through the former's Grey Literature search;	15th April 2018
2. DCU to contact ENRIO to discuss the best way to disseminate WP5's request for cases to ENRIO's members;	15th April 2018
3. DCU to undertake an internet-based search for publically-available, online RE+RI cases using the list of those RIOs and Regulatory Bodies detailed in Appendices 2 and 3 of the protocol. Where there is duplication of sources/cases already retrieved though Debrecen's Grey Literature search, the duplicated sources/cases will be reported in the Grey Literature systematic review and NOT in the RECs, RIOs and Regulatory Bodies searches;	15th April 2018

4. DCU will prepare a list of those RIOs and Regulatory Bodies detailed in Appendices 2 and 3 for whom they cannot identify online repositories of cases;	30th April 2018
5. Where cases can be found online for the organisations detailed in Appendices 2 and 3 of the protocol, cases will be entered into the online data template on the EnTIRE Wiki staging area. Where cases have been retrieved through the Grey Literature search, the same cases will NOT be reported in the RECs, RIOs and Regulatory Bodies searches;	1st May 2018-30th November 2018
6. WP5 will prepare an electronic request in English for those RECs, RIOs and Regulatory Bodies for whom it cannot identify online repositories of cases. The electronic request will also provide details of the EnTIRE project and its plans for tagging, disclosing and disseminating RE+RI cases. Bearing in mind that the RE+RI cases need to be in English and open to the public and taking into account the fact that we are legally constrained by the case details we make available on the online platform, WP5 should state that they can only accept publically-available cases in English and in a digital format;	1st May 2018
7. WP5 will provide comments and changes to the draft electronic request;	15th May 2018
8. DCU will send the list of RIOs to ENRIO, requesting that ENRIO distribute the electronic request to the organisations on the list. With regards to Regulatory Bodies, DCU will send the electronic request directly to the organisations listed on Appendix 3;	15th May 2018
9. Where no correspondence is forthcoming, a second electronic request will be sent two weeks after the date	1st June 2018

of the first request directly to the RIOs and Regulatory Bodies detailed in Appendices 2 and 3 for whom DCU cannot identify online repositories of cases;	
10. If there is no response to the second email, a final request will be sent two weeks later.	15th June 2018

3.3 Grey Literature

Details of the search methodology employed to search for RE+RI cases in grey literature are provided on pp. 43-49 of the protocol (D.5.1). This section specifies the changes made to that protocol and provides justification for the proposed alterations.

3.3.1 Scoping Search

The first phase of the pilot was the scoping search, which helped to define the best keywords for the search. After testing a series of keywords, the most adequate search terms were chosen on the basis that they generated an optimal number of results that could be adequately accounted for within the timeframe of the project and with the resources available to WP5. Broader search terms would result in an unmanageable number of results while narrower search terms would likely exclude useful results.

3.3.2 Database Selection

The second phase determined the databases that will use for the up-scaled searches of task 5.3. Some of the databases listed in the protocol are either unavailable or have merged with larger databases. Unavailable databases have been removed from the list. For those databases that have been subsumed by a larger database, the 'parent' database will be searched. Consequently, in order to decide which databases should be used, the original list detailed in the protocol has been updated by removing those that are now unavailable and identifying those databases that have merged (see *Table 5*).

3.3.3 Refinement of Search Terms

Utilising the sets of search terms employed for searching for cases in academic literature, similar results were yielded during the pilot search. The proposed methodology, therefore, seems feasible in the cases of academic and grey literature. Bearing in mind that the systematic review method being employed will encompass and account for both academic

and grey literature, it was decided that the search terms used for academic literature should be employed for grey literature.

The terms that will be used for the up-scaled searches of task 5.3 are ((“research ethics”) AND (“violation” OR “unethical” OR “misconduct”)) and ((“research integrity”) AND (“violation” OR “unethical” OR “misconduct”)).

Table 5: Details of the Scoping Search

Database	Search Terms	Number of hits	Comments
Biosis	N/A	N/A	Subscription Request Pending
Dissertation Abstracts	N/A	N/A	Merged with ProQuest Dissertations and Theses Global
DocuTicker	N/A	N/A	No Longer Exists
Grey Literature Report	“research integrity” AND (violation OR unethical OR misconduct)	0	
	"research ethics" AND (violation OR unethical OR misconduct OR fraud)	0	
Greynet/OpenGrey	“research integrity” AND (violation OR unethical OR misconduct)	3	
	"research ethics" AND (violation OR unethical OR misconduct)	1	
NTIS/NTRL	“research ethics” AND (violation OR unethical OR misconduct)	953	
	“research integrity” AND (violation OR unethical OR misconduct)	1,527	
OPENSigle	N/A	N/A	Merged with OpenGrey
PolicyFile/ Dissertation and Thesis Abstracts/ ProQuest	“research ethics” AND (violation OR unethical OR misconduct)	13,501	
	“research integrity” AND (violation OR unethical OR misconduct)	10,599	

3.3.4 Summary

As a result of the pilot search, the initial protocol was fine-tuned according to project needs and resources in accordance with PRISMA guidelines. The resulting changes can be summed up as follows:

1) The protocol keywords have been refined. For research ethics cases, the search term will be ((“research ethics”) AND (violation OR unethical OR misconduct)). For research integrity cases, the search term will be ((“research integrity”) AND (violation OR unethical OR misconduct));

2) The reference and citation databases that will be searched will be Biosis, Grey Literature Report, Greynet/OpenGrey, NTIS/NTRL and PolicyFile/Dissertation and Thesis Abstracts/ProQuest.

3.4 Media Outlets

Details of the search methodology employed to search for RE+RI cases in media outlets are provided on pp. 50-53 of the protocol (D.5.1).

A scoping search was undertaken at the time of producing the original protocol (see *Table 6*). Using the methodology put forward in the protocol, we have been able to retrieve and register a large number of cases from different media outlets using both Google News and LexisNexis. All of these cases comply with the definition of the case mentioned in section 5 of the protocol.

Currently, five of the cases that have been tagged using the wiki data management template have been identified through Google News using one of the 12 suggested search strings ("Case Study" AND "Research Integrity"). The registered cases have been retrieved from *The Tennessean, In-pharmatechnologist, Nature-News, Inside Higher Ed* and *The New York Times*.

Additionally, we have tagged one case from the LexisNexis database using the search string ("Case Study" AND "Research Ethics"). The registered case is from *The Australian*. The six cases that have been tagged are factual cases.

Table 6: Details of the Scoping Search

Search engine	Search strings	Number of hits	comments	Date of search
Google News	"case study" And "research integrity"	16	News items are available from 2012 until 2017	09/08/2017
	"case report" And "research integrity"	13	News items are available from 2010 until 2017	
	"case study" And "research ethics"	87	News items are available from 2010 until 2017	
	"case report" And "research ethics"	42	News items are available from 2013 until 2017	
	"case study" And "research misconduct"	43	News items available from 2012 until 2017	15/09/2017

	“case report” And “research misconduct”	5	News items available from 2013 until 2016	
	“case study” And “research fraud”	8	News items available from 2009 until 2017	
	“case report” And “research fraud”	2	News items available from 2011 until 2017	
	“case of” And “research integrity”	266	News items available from 2004 until 2017	
	“case of” And “research ethics”	442	News items available from 2003 until 2017	
	“case of” And “research misconduct”	152	News items available from 2004 until 2017	26/10/2017
	“case of” And “research fraud”	104	News items available from 2004 until 2017	
LexisNexis	“case study” And “research integrity”	89	News items are available from 1987 until 2017	
	“case report” And “research integrity”	13	News items are available from 1994 until 2017	
	“case study” And “research ethics”	340	News items are available from 1981 until 2017	10/08/2017
	“case report” And “research ethics”	63	News items are available from 1991 until 2017	
	“case study” And “research misconduct”	43	News items are available from 1998 until 2017	
	“case report” And “research misconduct”	9	News items are available from 1995 until 2015	
	“case study” And “research fraud”	18	News items are available from 1987 until 2017	15/09/2017
	“case report” And “research fraud”	7	News items are available from 1987 until 2014	
	“case of” And “research integrity”	990	News items are available from 1987 until 2017	
	“case of” And “research ethics”	996	News items are available from 1985 until 2017	
	“case of” And “research misconduct”	988	News items are available from 1989 until 2017	26/10/2017
	“case of” And “research fraud”	999	News items are available from 1980 until 2017	

3.4.1 Summary

On the basis that the scoping search has retrieved a large (yet manageable) number of pertinent RE+RI cases utilising the methodology employed in the protocol, no changes have been made to this section of the protocol.

It should be noted, however, that, in order to avoid duplication of work and duplication of cases on the online platform, where sources have been retrieved through the Grey Literature search, the same sources will not be reported in the media outlets searches.

3.5 The Blogosphere

Details of the search methodology employed to search for RE+RI cases in the blogosphere are provided on pp. 54-55 of the protocol (D.5.1). The protocol also identifies a number of existing blogs focusing on RE+RI issues ('RE+RI blogs').

Two scoping searches have been performed: 1) a search in the already identified RE+RI blogs (see *Table 7*); and 2) a search in the general blogosphere.

Table 7: Details of Scoping Search (1)

Blog	Comments on scoping search
Retractionwatch.com	Contains a continually expanding, comprehensive set of cases involving full or partial retractions of papers from the academic literature since 2010. Currently contains just under 4,000 cases. A searchable database of all of the retraction cased identified is available (see Figure 1 below) which allows for the identification of cases on the basis of a number of parameters including subject, reason for retraction and country.
Plagiarismwatch.org	Contains a limited number of cases which can be 'hand searched', non-duplicate cases identified and original source identified and consulted (if necessary to provide a full account of the case)
ethics-and-integrity.net	Ditto
copy-shake-paste.blogspot.com	Ditto
Researchethicsblog.com	Ditto
statistically-funny.blogspot.com	Ditto

The scoping search shows that two different approaches for case selection are necessary, one for Retractionwatch and one for all the other RE+RI blogs. The reasons for this difference are two-fold:

1) Retractionwatch has a searchable database which enables detailed search queries, but similar functionality is not available for the other blogs. The cases in the other

blogs are, by the mere fact of inclusion in these blogs, already identified as potential RE+RI cases and do not need to be identified in a search. Rather, they need to be read, analysed and tagged (and checked for duplication);

2) Whereas it is (in principle) possible to include all of the cases in the other blogs, WP5 are unable to analyse and tag all relevant cases in Retractionwatch, which runs into the 1000s, due to time restrictions and resource limitations.

In relation to Retractionwatch, the very large number of available cases in the database, with many raising fairly similar issues (e.g. same scientific discipline, same type of misconduct, same geographical region, and so on), and the quantitative bias towards the life sciences, entail that selection has to be deliberate if all areas of academic research, all types of RE+RI issues and all geographical regions are to be represented in the case corpus. This will be achieved by initially searching according to type of misconduct and identifying a case for which there is sufficient information to allow in-depth tagging. A 'case-matching approach' will then be applied to this index case, identifying matching cases on the basis of scientific discipline and geographical area. For example, if the index case is a medical plagiarism case from Europe, a search will be performed for cases that differ on one of the two parameters so the matching set would be plagiarism cases from other areas of science and from other geographical regions.

Details of Scoping Search (2)

Apart from RE+RI blogs, relevant cases may also be found in the general blogosphere, e.g. on blogs with a more general academic/research/university focus. The blogosphere is also continually evolving with new blogs appearing and old blogs disappearing, so the project will need to have a watching brief in relation to any new RE+RI blogs in order to identify them as they emerge.

Two approaches were used to search the general blogosphere: 1) using Google's blog specific search function (a sub-function in Google News search); and 2) using the search function of the largest blog provider WordPress (market share >25%).[1]

A number of different combinations of search terms were tested. Furthermore, the identified cases were compared to the cases reported in RE+RI blogs.

It was found that searching the general blogosphere does identify cases that are not reported in RE+RI blogs. It identifies additional cases that: either 1) do not involve formal retraction of papers but other RE+RI issues; or 2) have only reached local prominence but have not, or not yet, come to the attention of national media or the academic literature.

Almost any Google Search produces a very large number of results, and the optimal search strategy, therefore, necessarily involves finding an appropriate balance between exhaustiveness and specificity. When Google's blog search function is used the number of results found is not reported by the search engine and it is, therefore, not possible to provide a simple numerical estimate of the implications of different search strategies. In any case, this would be difficult since calculating an estimate of the specificity of a given search strategy requires separating the results found into relevant and non-relevant results and counting each group and this cannot be done automatically. It was determined by trial and error that two combinations of search terms similar to those to be used for the academic literature work well (i.e. strike a reasonable balance between exhaustiveness and specificity). These are, for research ethics cases, ((“research ethics”) AND (“violation” OR “unethical” OR “misconduct”)) and, for research integrity cases, (((“research integrity”) AND (“case” OR “instance” OR “example”)) OR ((“scientific misconduct”) AND (“case” OR “instance” OR “example”))).

A search using Google's blog search function seemed to identify all the blogs and cases identified by the more specific WordPress search, indicating that Google is able to efficiently search for and within WordPress blogs. The general blogosphere can therefore be searched by employing Google's blog search function as the only search engine.

[1] https://en.search.wordpress.com/

3.5.1 Summary

The resulting changes to the protocol can be summed up as follows:

1) A subset of Retractionwatch cases will be identified for tagging using the searchable database of Retractionwatch cases. The specific strategy is described above;

2) The general blogosphere will be searched at the beginning of the scaling up phase and then at regular four-monthly intervals in order to a) identify emerging RE+RI focused blogs, and b) identify additional cases of interest not picked up by the RE+RI blogs. The specific search strategy is described above.

3.6 Online Repositories

Details of the search methodology employed to search for RE+RI cases in online repositories are provided on pp. 56-64 of the protocol (D.5.1).

A scoping search was undertaken at the time of producing the original protocol (see *Table 8*). Using the methodology put forward in the protocol, we have been able to retrieve and register a large number of cases from different repositories listed in the protocol. All of these cases comply with the definition of the case mentioned in section 5 of the protocol.

Currently, three of the cases that have been tagged using the wiki data management template have been identified through the Office of Research Integrity (ORI), the Committee Of Publication Ethics (COPE) and the American Anthropological Association. Each repository offers different types of cases that have allowed us to test the flexibility of the data management template in terms of accommodating various cases. Cases provided by the ORI website include factual cases. The COPE website only provides factual-anonymised cases with questions and professional advice. Finally, the American Anthropological Association provides fictional cases with questions, comments and discussion points.

Table 8: Details of Scoping Search

Repository	Quantity and other features of cases
American anthropological association	Association's aim is to educate members on ethical issues within anthropology. As part of this mission, the Committee on Ethics provides case studies on how to approach ethical dilemmas in the discipline. ➢ On 15/08/17 contained two fictional cases including questions for discussions and comments. ➢ Cases are neither tagged nor categorised.
American medical association – Journal of ethics	The ethics journal of the largest association of physicians and medical students in the US. Based on the monthly theme case and commentary are gathered from experienced physicians and

	other experts in the field who can help readers think productively about that topic. ➢ On 15/08/17 contained 1570 cases. Mostly based on historical scenarios but rewritten for publication. Cases are not accompanied with questions but include experts' commentary. ➢ Cases are categorised under several headings. Lots of duplicates and rather chaotic. Cases are neither tagged nor searchable.
American physics society	Society provides a series of case studies on ethical issues that can arise in the course of doing physics research, and intended to be an educational resource for researchers, mentors, and students. ➢ On 14/08/17, contained 39 fictional cases that were specifically designed for doing physics research. Cases come in two versions, one for students and another including possible assignments and discussion guides for teachers. ➢ Cases are classified under 9 topics in a report. Cases are not tagged.
American sociological association	The association has 21000 active members and aims to articulate policy and implement programs to impact sociology on a national and international level. ➢ On 14/08/2017, contained 103 fictional cases on ethical standards that should be followed in research. Cases are followed by questions and discussions. ➢ Cases are neither categorised nor tagged.
Austrian Agency for Research Integrity	The agency is responsible for investigating alleged cases of scientific misconduct in Austria, evaluating the severity of each violation and proposing consequential measures with the first two tasks being assigned to the Commission for Research Integrity. A selection of cases and inquiries from 2009 until 2015 is made available in English. ➢ On 18/08/2017, contained 38 anonymised case summaries incorporated in the annual reports. ➢ Cases are neither categorised nor tagged.
Bioethics outreach - Iowa state university	Cases are developed by the office of biotechnology as a method for improving bioethics outreach. ➢ On 14/08/2017, contained 45 cases with questions, moral issues and some discussion points highlighted. ➢ In addition to the English cases, there are also two cases in Spanish and two in Portuguese. ➢ Cases are not tagged and are not searchable.

Centre for ethics in professions – University of Puerto Rico	Centre provides cases for different fields of applied ethics aiming to emphasise the ethical dimensions involved in the science based professions. ➢ On 14/08/17, contained 12 cases on research ethics and plagiarism, including discussion questions. ➢ Cases are not tagged and are not searchable.
Center for the Study of Ethics in the Professions - Illinois Institute of Technology	The centre provides a large library of various resources including case studies and codes of conducts. The 'Research ethics case studies' collection provides cases from several disciplines. ➢ On 14/08/2017, contained links to 110 fictional cases that are available in several other online sources. ➢ Cases are tagged and can be searched and filtered based on subject and discipline
Committee on Publication ethics (COPE)	COPE provides advice to editors and publishers on all aspects of publication ethics and, in particular, how to handle cases of research and publication misconduct. The case repository stores anonymised cases of real scenarios that were discussed in COPE forums. ➢ On 18/08/2017, contained 570 cases together with the advice given by COPE. More recent cases include follow-up information and outcome too. ➢ Based on the type of violations or issue at hand, cases are classified under 10 categories. Cases are also tagged and searchable.
CTSPEDIA (A knowledge base for clinical and translational researchers)	CTSPEDIA is a knowledgebase that provides useful resources for clinical and translational researchers. Fictional cases relevant to clinical research are submitted by field experts. ➢ On 14/08/2017, contained 112 short cases, some with and some without questions for discussion. ➢ Although some other resources are tagged, but cases are neither categorised nor tagged, and therefore, are not searchable.
Danish Committee on Research Misconduct	The Committee handles Danish cases of research misconduct. As a general rule a case is raised via a complaint handed in at the research institution where the research in question was carried out. All anonymised case summaries are available in Danish, including a few in English. ➢ On 18/08/2017, contained 8 case summaries in English. Cases are anonymised and differ in the provided context ➢ Cases are neither categorised nor tagged.
Economic and social research council of the UK	A non-departmental public body for funding research on economic and social issues in the UK. The repository aims to raise

	awareness about some of the ethical issues that can arise in research. ➢ On 14/08/17 contained 21 research ethics cases. Case studies are inspired by real issues that happen during the process of research, and were rewritten for educational purposes, and highlight the ethical process and lessons to be learned. ➢ Cases are subsumed under several categories but are neither tagged nor searchable.
Hale chair in applied ethics	The Hale Chair in Applied Ethics is devoted to research and teaching in ethics within and about the professional disciplines. ➢ On 15/08/17 contained 23 fictional cases developed for pedagogical purposes, including questions for discussions. ➢ Cases are neither tagged nor categorised.
Markkula Center for Applied Ethics -at Santa Clara University	The repository includes resources of different types on applied ethics across a variety of fields. ➢ On 14/08/2017, contained real fictional cases on educational integrity (25), bioethics (21) and 191 cases from other contexts (e.g. business ethics, journalism ethics, etc.). Cases also include questions for discussions. ➢ Within the categories of educational integrity and bioethics, cases are subsumed under several other categories, but they are not tagged, nor searchable.
National academy of engineering (online ethics centre)	The online repository of the US national academy of engineering, that is regularly updated and well maintained. In addition to engineering cases, the platform provides hundreds of cases on life and environmental science, social and behavioural sciences, computer, maths and physical sciences. ➢ On 14/08/17, contained 565 cases including 117 historical cases, 216 fictional cases, 183 open-ended scenarios, 37 mini-cases, and 12 role-plays. ➢ Cases are tagged and categorised by type of resource, topic, field, author and other publication information. Easy to search and filter based on various criteria.
National Center for Case Study Teaching in Science - State University of New York in Buffalo	A peer-reviewed collection containing cases in all areas of science. Under the 'ethics' topical area, a collection of cases that are all hosted on the platform can be found. ➢ On 14/08/2017 contained 135 curated cases. Most cases include questions for discussions and relevant illustrations/images. ➢ Cases are tagged, filtered and searchable based on several attributes.

	➢ More resources including teaching notes and answer keys are under paywall.
National institute of health (NIH)	The sourcebook of NIH provides case studies to use by the medical community to help staff involved in research define or refine their own standards. Each year a theme is specified and, fictional cases about that topic are developed and presented (except for 2005 & 2006). ➢ On 14/08/2017, contained 65 cases on 15 topics. Cases are designed for medical settings and include questions. ➢ Cases are subsumed under each year's theme, but they are not tagged.
National center for professional and research ethics – Ethics CORE project	The repository is the final deliverable of a project that was founded by the US National Science Foundation (NSF) to help the scientific community share ethics resources and work together. ➢ On 14/08/2017, contained 195 fictional cases with a short abstract. Some are hosted on the platform and are easy to download as a PDF file; some others are linked to other online sources. ➢ Cases are tagged and the repository can be easily searched.
National science foundation	The US National Science foundation's office of inspector general provides independent oversight and aims at preventing and detecting fraud, waste, and abuse. ➢ On 18/08/2017 contained 3245 records of investigated cases from 1989 until present. Cases are strictly anonymised. ➢ Records are neither categorised nor tagged.
Office for Human research protection (OHRP)	The OHRP provides leadership in the protection of the rights, welfare, and wellbeing of subjects involved in research conducted or supported by the U.S. Department of Health and Human Services. ➢ On 18/08/2017, contained 34 fictional cases explaining a situation, without further questions or discussions. ➢ Cases are organised based on host entities (the context) and issues, but cases are not searchable.
Office of Research Integrity (ORI)	The ORI oversees and directs Public Health Service (PHS) research integrity activities on behalf of the Secretary of Health and Human Services (HHS) with the exception of the regulatory research integrity activities of the Food and Drug Administration. In addition to fictional case studies for educational purposes, ORI stores information about publicly known cases of misconduct that were adjudicated by HHS since 2008.

	➢ On 18/08/2017 contained 35 historical cases of misconduct including names, affiliations and the decision of the court. ➢ Additionally, ORI's casebook contains 24 fictional cases including 8 role-plays. These cases are categorised in 8 categories based on the involved issue, and include questions and discussions questions for facilitators.
Research involving children	Financed by the UNISEF office of research and few academic institutions, the International Research Involving Children (ERIC) project provides a repository of information and resources to assist researchers, and guide and improve research involving children. ➢ On 14/08/2017, contained 27 cases inspired by historical events and rewritten by experts. Ethical issues are highlighted and motivations for the decisions are explained. ➢ Cases are categorised under four categories (Harms and Benefits, Informed Consent, Privacy and Confidentiality, Payment and Compensation) but they are not tagged.
The center for bioethics and human dignity studies, Trinity international university	A repository to analyse biomedical cases from a Christian ethics perspective. ➢ On 14/08/17 contained 16 case studies based on historical cases including recommendations and comments from several experts and an audio podcast. ➢ Cases are tagged based on the topic. ➢ Four of the published cases are also available in Chinese.
The global health network - WHO	The online platform made to present cases that originally appeared in a casebook published by the World Health Organisation (WHO) in 2009. ➢ On 14/08/2017, contained 64 fictional cases about ethical issues raised in health research. Cases include study questions. ➢ Cases are subsumed under eight categories but are not tagged.
The graduate school of Michigan state university	A collection of research integrity resources for graduate students. ➢ On 14/08/2017, contained 17 cases and 8 video case studies, all of which present fictional cases. PDF files also include questions for discussion. ➢ Cases are categorised under seven topics but are not tagged.
The MC Graw Hill Bioethics case studies	An educational source in the context of general human biology. ➢ On 14/08/17 contained 23 fictional cases including questions for discussions.

	➢ Cases are neither categorised nor tagged.
The Netherlands Board on Research Integrity (LOWI)	LOWI is an independent advisory body, established in 2003. Anonymised cases are stored on the website from 2007 until 2017 (in Dutch). ➢ On 18/08/2017, contained 24 anonymised case summaries in English including an explanation of relevant considerations. ➢ Cases are neither categorised nor tagged.
UK clinical ethics network - UKCEN	The network aims at providing up to date and reliable information on ethical issues that are commonly presented to clinical ethics committees or arise in clinical practice. ➢ On 15/08/2017 contained 3 fictional cases based on common scenarios in clinical practice. Cases are followed by questions for discussions and highlight several principles for moral deliberation. ➢ Cases are neither tagged nor categorised.
UK research integrity office	The UK Research Integrity Office (UKRIO) is an independent charity, offering support to further good practice in academic, scientific and medical research. Sample case study packs are provided for education and training purposes. ➢ On 18/08/2017, contained one free case study pack including 4 Fictional scenarios based on real life situations. Some suggested points for discussion accompany each case study. ➢ Cases are neither categorised nor tagged.
Washington university of saint Louise – centre for clinical and research ethics	A library of Research Ethics Case Studies for use in small group case discussions, and train-the-trainer support for units. ➢ On 14/08/2017, contained 106 case studies mainly on ethical issues raised in clinical research. Cases also include questions for group discussions. ➢ Cases are neither categorised nor tagged.
Yale Interdisciplinary Centre for Bioethics	The centre provides a series of cases rooted in the work of hospital IRBs and drafted by their members. ➢ On 14/08/2017, contained 6 fictional cases. They are very elaborate and highly contextualised cases in research ethics. Each case comes with several expert comments that shed light on various aspects of the case. ➢ Cases are neither categorised nor tagged.

3.6.1 Summary

On the basis that the scoping search has retrieved a large (yet manageable) number of pertinent RE+RI cases utilising the online repositories listed in the protocol, no changes have been made to this section of the protocol.

It should be noted, however, that, in order to avoid duplication of work and duplication of cases on the online platform, where sources have been retrieved through the Grey Literature search, the same sources will not be reported in the online repositories searches.

3.7 WP2's Focus Group Sessions

Details of the search methodology employed to retrieve RE+RI cases from WP2's stakeholder consultation are provided on pp. 65-68 of the protocol (D.5.1).

The results of face-to-face and online focus groups have, at the time of writing, not been supplied to WP5. An addendum to this report will follow once WP5 receives the results of the stakeholder consultation.

4. Changes to the System of Categories for Tagging Cases

Details of the proposed system of categories for tagging cases are provided on pp. 77-85 of the protocol.

At the time of producing the protocol, it was noted that the proposed system for tagging cases presents an idealised model, one that is, ultimately, dependent on the plans for the online platform to be determined by the EnTIRE project's executive committee, consultants, stakeholders and work packages in collaboration with the end-users.

The discussions regarding the tagging and presentation of data on the EnTIRE platform are now taking place across work packages. Furthermore, the IT structures are still in development and, ultimately, contingent on these cross-consortium discussions.

WP5 made it clear in its protocol that the proposed tagging 'system' should remain open to adaptation and innovation. The system for both tagging and presenting data is in constant development. Not only is it being affected by inter-work-package discussions and decisions, it will continue to be revised as the project develops in response to stakeholder consultation and user engagement.

It is worthwhile mentioning some of the changes that have been made to the protocol's tagging 'system' following the initial stages of cross-consortium collaboration. However, as already mentioned, these changes are contingent and likely to be revised.

4.2 Data Management Template

Our general approach to presenting information about cases has changed. Previously, we aimed to register case data regardless of the source of that data. However, registering case data without distinguishing between the data provided by different sources can lead to confusion, especially if two sources provide different interpretations of the same case. As a result, we distinguish the

information that is provided by each source from the purely factual elements of the case. As a result, we are registering case features as they are provided in individual sources, which, when combined with other sources, builds a better picture of the dimensions of the case. Depending on the case type (fictional, factual anonymized or factual), the factual elements of the case are currently restricted to:

Fictional	**Factual anonymized**	**Factual**
Case Title	Case Title	Case Title
	Year/Date(s)	Year/Date(s)
		Geographical Setting(s)
		Research Area(s)
		Affiliation(s)

Secondary features, such as URL, book and journal details, description, outcome, principles, values and virtues, types of violations and discussion points are currently treated as source specific.

4.2 Tagging Categories

Research Area(s) is Complemented by a Taxonomy

In the protocol, it was suggested that 'branch of science' and 'research discipline' should not be predefined. However, after the data management template was reviewed by other members of the EnTIRE project, it was concluded that given the importance of the branch and discipline of the case in question, allowing users to link disciplines to branches on an arbitrary basis would create more work for those who monitor the platform, confuse the tagging system and compromise the platform's search capabilities. On the basis of advice received from another work package leader, users can now select the research areas associated with a particular case from a European Research Council (ERC) taxonomy.

New Category: Outcome(s)

This 'outcome(s)' category was not mentioned in WP5's protocol. For training and education purposes, it is deemed to be a useful category for inclusion. It also allows users to produce more elaborate and specific content for a particular case, thereby, making the platform more user-friendly and less like a database.

Removed Category: Implicit Dimensions of the Normative Framework

The following was raised in the D.5.1 protocol:

"On the basis of the definition of a case in section 5, and if the cases are violations of the normative frameworks governing RE+RI, we will begin tagging RE+RI cases according to whether they are violations of the explicit dimension of RE+RI in terms of rules, regulations and legislation and/or the implicit dimension in terms of those norms developed in practice and through processes of deliberation regarding research misconduct".

After consulting with other work package leaders, it was decided that meaningful reference could only be made to the explicit dimensions of research ethics and research integrity, for example, policy documents, codes of ethics, regulations and legislation.

Removed Category: Form

In the protocol, we stated that we would wish to tag a case according to whether it was a 'case study', a 'case report' or a 'case review'. When discussing this aspect of the data template with other work package leaders, it became apparent that such terms are of no use to the typical user. Furthermore, it is extremely difficult to ascertain what distinguishes a case study from a case report and a case report from a case review. As a result, this category was removed.

What these examples demonstrate, and what we, as a work package, have become increasingly aware of, is that when it comes to tagging and data presentation:

1) The users' perspective is of primary importance;
2) If there is a debate regarding the inclusion/exclusion of category or tag, then it is clear that the question demands collaborative engagement and cannot be arbitrarily passed over;
3) There is a need for cross-consortium, user-focussed agreement regarding the content and 'look-and-feel' of any metadata collection and presentation;
4) Databases are not user-friendly; users should be allowed the opportunity to elaborate on a case and to go into detail, rather than just 'tick the box';
5) Metadata presentation and tagging needs a trade-off, sometimes an intuitive trade-off, between predefining categories for the purpose of searching for data and leaving the data open to interpretation and open to new and novel forms of tagging and categorising.

These points affirm the approach to data presentation and case tagging that was adopted in the protocol (D.5.1). On that basis, we, as a work package, expect our revised tagging system to be continuously updated.

5. Summary of Changes to the Protocol

1) The protocol keywords have been refined for academic literature sources. For research ethics cases, the search term will be (("research ethics") AND (violation OR unethical OR misconduct)). For research integrity cases, the search term will be (("research integrity") AND (violation OR unethical OR misconduct));

2) The reference and citation databases for academic literature searches have been limited to six of the most available and commonly-used databases: Pubmed, OVID, Science Direct, Web of Science, JSTOR and SCOPUS;

3) The process for retrieving cases from RECs, RIOs and Regulatory Bodies has been significantly revised, with the new methodologies detailed in Tables 3 and 4 of this report. The most important revisions include:

    - Initially, WP5 will undertake an internet-based search for publically-available, online RE+RI cases using the list of those RECs, RIOs and Regulatory Bodies detailed in Appendices 1, 2 and 3 of the protocol. Where there is duplication of sources/cases already retrieved though Debrecen's Grey Literature search, the duplicated sources/cases will be reported in the Grey Literature systematic review and NOT in the RECs, RIOs and Regulatory Bodies searches;
    - In the case of RECs and RIOs, the electronic request for cases will be disseminated by EUREC and ENRIO, respectively. When it comes to regulatory bodies, the organisations will be contacted directly by WP5 as detailed in the original protocol;
    - Our requests for cases are restricted to those cases that are already publicly available

- Cases provided by RECs, RIOs and Regulatory Bodies must be electronically available via a URL link;
- Our requests for cases can only be made in English;
- All requests will be made in electronic format.

4) The protocol keywords have been refined for grey literature sources. For research ethics cases, the search term will be ((“research ethics”) AND (violation OR unethical OR misconduct)). For research integrity cases, the search term will be ((“research integrity”) AND (violation OR unethical OR misconduct));

5) The reference and citation databases for grey literature searches will be Biosis, Grey Literature Report, Greynet/OpenGrey, NTIS/NTRL and PolicyFile/Dissertation and Thesis Abstracts/ProQuest;

6) Where sources have been retrieved through the Grey Literature search, the same sources will not be reported in the media outlets, blogosphere, RECs/RIOs and online repositories searches.

7) A subset of Retractionwatch cases will be identified for tagging using the searchable database of Retractionwatch cases;

8) The general blogosphere will be searched at the beginning of the scaling-up phase and then at regular four-monthly intervals in order to a) identify emerging RE+RI focused blogs, and b) identify additional cases of interest not picked up by the RE+RI blogs;

9) The system for both tagging and presenting data is in constant development. Not only is it being affected by inter-work-package discussions and decisions, it will continue to be revised as the project develops in response to stakeholder consultation and user engagement.