Open Science in Estonia

Open Science Expert Group of the Estonian Research Council
Principles and Recommendations for Developing National Policy

Contents

Preface

The objective of open science is to transform scholarly research into a process based on openness, creativity and cooperation by using the opportunities provided by modern information and communication technologies. In particular, the concept of open science denotes free access to research publications and research data, but it also includes open code (using of open source code soft and hardware) and open methodology (sharing of information about the methods, equipment and analysis techniques used in research), as well as open peer reviewing (quality is ensured by transparent and traceable process of pre-evaluation). The present document focusses on research data and research publications, which are collected and published by using financing from the public sector.

Many countries are in the process of developing, or have already adopted the strategies of open science, where open access is among the conditions for using financing from the public sector and among the criteria for evaluating the results of research. The European Union actively supports the principles of open science; according to the recommendations of the European Commission, all member states should formulate their national guidelines on ensuring open access to research information. The ambitious goal of Europe is to make all research articles and their basic data publicly accessible already starting from 2020.

In April 2015, the Estonian Research Council formed the Open Science Expert Group with the task of drawing up proposals for developing the Estonian national open science policy. The aim of the present document is to create a general framework, as well as to issue specific recommendations for developing both the national and institutional guidelines for this process. We also hope to increase the awareness of the Estonian scientific community and the public of the principles of open science and open access.

The present version of this document is a summary of previous discussions of the Group, of the open science seminar held in Tartu in October 2015 and of the feedback received in the course of a public consultation held in spring 2016.

June 2016, Tartu / Tallinn

Estonian Research Council Open Science Expert Group:
Olle Hints, Chairperson, Tallinn University of Technology
Reet Adamsoo, University of Tartu
Martin Eessalu, Ministry of Education and Research
Martin Hallik, University of Tartu
Jaak Järv, Estonian Academy of Sciences
Andres Kollist, Academic Library of Tallinn University
Urmas Kõljalg, Estonian Academy of Sciences
Marika Meltsas, Estonian Research Council
Viktor Muuli, Estonian Research Council
Raivo Ruusalepp, National Library of Estonia
Priit Tamm, Estonian Research Council
Kadri Vider, Center of Estonian Language Resources
Allar Viik, Ministry of Economic Affairs and Communications

1. Introduction

In the context of the present document, the concept of open science denotes free access to electronic research information, and foremost, to publications and research data which have been created and published with public funding. The open science ideology, already familiar for decades, has primarily been the researchers' initiative for supporting information exchange and cooperation. As an example, we should mention the whole human genome sequencing and sharing of the data among research groups in order to accelerate the completion of the work, or the repository of research publications in the fields of mathematics and physics arXiv (http://arxiv.org) which started operating already in 1991. The sharing of research results via social media and special environments created for this purpose, such as Academia.edu (https://www.academia.edu), ResearchGate (https://researchgate.net) and others has become increasingly popular. Open access as an ideology for increasing the accessibility of research publications was launched in the form of the Budapest Open Access Initiative in 2002, and carried on by subsequent declarations and international agreements.

Publicly accessible research information makes it possible to increase the effectiveness and speed of research by reducing the need for duplication and repeating of work; the same basic data would yield more results and the visibility and number of citations of the results are, as a rule, much larger. This fact has been confirmed by numerous studies. For example, the publications which are accessible via the arXiv accumulate several times more citations than the publications which have not been archived, regardless of the journals where they had been published[1]. From the researchers' viewpoint, the principles of open science help to increase the academic credit and citability of the authors.

As important it is the feedback of research to the state and society, as well as the wider possibilities for using research information in education, civil services and citizen science, where conventional research journals are mostly inaccessible. The ideology of open science has a remarkable effect on economics and innovation, outweighing the additional costs of infrastructure, etc. It has been estimated that in Australia, providing free access to research findings would result in economic profit of nine billion Australian dollars over 20 years[2]. Providing free access to three large research data repositories in the United Kingdom (*Economic and Social Data Service, Archaeology Data Service, British Atmospheric Data Centre*) would make it possible to increase the amount of funds invested in research and development activities by two to ten times over 30 years[3].

The interest and need of researchers for speedy and unconstrained access to research information on the one hand, and the understanding that everything created with public funding belongs to the public

---

[1] Gentil-Beccot, A., S. Mele and T. C. Brooks (2009), Citing and Reading Behaviors in High-Energy Physics: How a Community Stopped Worrying about Journals and Learned to Love Repositories. CERN-OPEN-2009-012. http://arxiv.org/abs/0906.5418

[2] Houghton, J. and P. Sheehan (2009), "Estimating the potential impacts of open access to research findings", *Economic Analysis and Policy*, Vol. 29, No. 1, pp. 127-42. http://ac.els-cdn.com/S0313592609500483/1-s2.0-S0313592609500483-main.pdf?_tid=9b5abd68-555c-11e5-bb0b-00000aab0f6c&acdnat=1441639373_f8eabbc13a850504379e43ecc0de8b73 (checked on 04.09.2015).

[3] JISC (2014), "The value and impact of data sharing and curation: A synthesis of three recent studies of UK research data centres", JISC, March, http://www.cni.org/news/jisc-report-value-impact-of-data-curation-and-sharing (checked on 04.09.2015).

on the other hand, as well as the significant development of technical solutions, are leading to open access to research information becoming a cross-disciplinary trend. However, open science is not an end in itself, but these policies form a means for improving the quality of research, promoting cooperation and increasing the cohesion of science and society, thus aiming at enlarging the social and economic impact of science and developing a scientific worldview.

The objective of the present document is to create a common framework for addressing the matters related to open science in Estonia and to contribute to the development of national and institutional guidelines. The Appendices to this document offer, for comparison, a list of selected references to open science policy basic documents from the EU and several countries, and the clarifications of more important terms and key words.

2. Trends in the European Union and its member states

In the couple of recent years, the European Commission has actively started to develop principles in the area of research and development and innovation for making knowledge better accessible for the public. The renewed conception of the European Research Area 2012 prescribes free access to knowledge, and its optimal circulation and transfer by using possibilities of the digital research area as one of the priorities. Proceeding from the principle that the publications and research data, which are the results of research done with public financing, have to be accessible to the public, the European Commission has published several policy documents providing practical guiding to its member states.

Accordingly, the European Commission has modernised its policy regarding research publications and research data, primarily, under the programme for research and innovation Horizon 2020. Differently from the previous European Union framework programmes, Horizon 2020 pays much more attention to the openness and accessibility to other interested parties of the research publications and data created under publicly financed projects. The initiative of the European Open Science Cloud encourages the member states to develop modern data infrastructures for archiving, sharing and re-using research data. In May 2016, the high-level advisory group Open Science Policy Platform was created at the European Commission for the application of the open science policy. The ambitious plan of the European Union stipulates the prevalence of open access and open research data by 2020.

Science Europe, the umbrella organisation for 47 research funding organisations from 27 countries, also supports the transition to open access in research results publishing and has formulated minimum requirements to the publishers who offer the gold open access business model: all publications must have the open licence (e.g., Creative Commons Attribution, CC BY, see Appendix 1: 'licences') and be machine readable.

More and more countries are adopting the principle of openness in research funding. They are drawing up national guidelines and recommendations with regard to providing open access to research publications and research data, their management and long-time preservation Requirements specifying that the applicants are obliged to submit data management plans are gradually introduced into national research funding instruments. Plans related to the preservinging and making accessible

the research data produced under research projects are becoming an inseparable part of these projects.

For a long time, the United Kingdom and the Netherlands have been the spokespersons for open science in Europe. Both countries have set themselves the aim of ensuring open access according to the conditions of the CC BY licence to the research publications created under public funding. They also draw attention to the fact that such results are attainable only when all the EU member states follow the same requirements in their national policies. The plan of the Swedish Research Council requires that starting from 2025, all publications issued with public funding must have the CC BY licence and provide immediate open access. A pilot project of additional funding is planned to facilitate the opening of research data. The Finnish roadmap of open science, published in 2014, sets as a subsequent objective that all research information should be openly and freely accessible in standardised data systems. A selection of references to open science strategies and guidelines of the EU and its member states can be found in Appendix 2.

3. The situation of open science in Estonia

Open access to research information is not a novelty in Estonia. For example, the majority of our research journals have offered immediate open access starting from the time when electronic versions of articles became available more than ten years ago. Research libraries have actively promoted the ideology of open access by organising, among other events, the already traditional seminars and information days during the international Open Access Week. The Estonian official political position with respect to open science has been clearly defined by the government decision concerning the availability and preservation of research information, formulated in answer to the European Commission communication and recommendation on 17.7.2012 no. C(2012) 4890. This decision states, among other issues, that: (1) research publications should be openly accessible, (2) in case of research financed by public funding, open access should be included in financing conditions, (3) Estonia supports the harmonisation of open access policies at the European Union level, (4) it is necessary to explain the possibilities of open access to both the researchers and the public, (5) Estonia supports the broadening of free access to research data, at the same time insisting on the careful consideration of large-scale investments.

The Estonian Research and Development and Innovation Strategy 2014-2020 "Knowledge-based Estonia" refers to open access as one of the measures for increasing the social and economic benefits of research and development: "To encourage open access to public-financed research results and research data." So far, the strategies and development plans are lacking more specific action plans and separate funding for broadening open access to research information.

Starting from 2013, the competition-based funding instruments (institutional and personal research funding) include a requirement for open access: "The publications that result from the implementation of the research theme shall be freely available to the public in the Estonian Research Information System (ETIS), unless set forth otherwise in the conditions for publication, and for the protection of copyright or intellectual property rights." Covering of the open access publication processing charges from the institutional and personal research funding is allowed, but it is still a relatively rare practice.

This is understandable, considering the rather limited capacity and the fact that publication costs should often be covered by other expenditure.

An extract from the Estonian Research Information System of October 2015 shows that of all the publications which resulted from the registered competition-based research projects (institutional and personal research funding, the Estonian Research Council research funding, targeted financing of research topics), more than 40% were openly accessible. Of publications related to institutional and personal research funding, published in 2013-1014, open access articles represent already more than 50%. During the recent decade, Estonian researchers have published articles in more than 4200 different journals, 355 titles of which (8%) are gold open access journals. By the end of 2015, Estonian scientific publishers issued 46 peer-reviewed journals; about 3/4 of them have the gold open access. Of the 11 Estonian esearch journals indexed in Web of Science, only nine are openly accessible. Not all Estonian open access journals have a clearly defined copyright ownership and licensing conditions, and not all of them are correctly included into the DOAJ list and the SHERPA/ROMEO registry.

The attitude of Estonian researchers toward the matters of open access has not been systematically studied, but the initial feedback shows rather scarce knowledge about the matter. General knowledge about copyright, different versions of open access and licensing conditions (incl. Creative Commons) is poor.

The legal status of research data would need a more precise regulation and so far, citation of research data has not been reflected in any of Estonian funding instruments. A national policy document Open Data Green Book was compiled on the initiative of the State Information Systems Department of the Ministry of Economic Affairs and Communication, with the objective of explaining the principles and more important policy options within the area, facilitating cooperation in Estonia and developing an action plan. This is a broad-based document, but matters related to research data have been left to the background.

Although research data-related activities have not yet been coordinated, some significant steps have been taken in recent years, and technical solutions have been created for depositing research data and making them available for researchers and the general public. Among important landmarks are several objects of the national research infrastructure roadmap, such as NATARC, the Center of Estonian Language Resources, Estonian Centre for Genomics, and the Estonian integrated portal e-varamu. On the initiative of the research infrastructure roadmaps, Estonia joined the international DataCite Consortium in 2014, making it possible to launch data repositories and to start allocating DOI identifiers to research data. The first data sets were registered at the DataCite Estonia Centre in 2014. In early 2015, Estonian universities cooperated in order to create the DataCite Estonia Consortium for the further coordination and continuation of these activities. DataCite helps to ensure better traceability and retrievability of research resources created by the member research institutions. Several speciality-based interfaces have already been developed and DataCite Estonia has by now allocated more than 500 000 DOI identifiers, mostly via the biodiversity database PlutoF and the research infrastructure roadmap object NATARC. The use of international data repositories by Estonian researchers has not been analysed, but the traditions and intensity of use differ much by specialities. Regarding the international coordinating of research data, it is important that the Estonian state and research institutions cooperated with several different pan-European research infrastructures and data networks, among others with BBMRI, EATRIS, ELIXIR, CLARIN, European Social Survey, GBIF and others.

4. Open access to research publications

Estonia's strategic objective 2020

The academic community is familiar with and accepts the principles of open access. Research articles published with the support of public funding are freely accessible to all one year after their original publishing date at the latest; at least half of them are immediately accessible in the final form and their long-time maintenance is guaranteed. All research journals, published in Estonia with public funding, follow the open access principles and apply the conditions of free content licensing.

Principles and recommendations

- Estonia must support the establishment of a common European Union open science policy and actively participate in pan-European organisations which stand up for broadening open access and reducing the costs of publishing and subscribing for research literature.
- In case of publicly funded research, open access must be included into the funding conditions. In case of competitive-based funding instruments (institutional and personal research funding), it is necessary to establish requirements for ensuring open access to publications; the criteria for supporting/preferring open access must also be added to the regulations concerning other instruments of public funding of research.
- Ideally, all research publications, published with the support of public funding, should have immediate open access to the final version (gold open access; Clarifications of terminology, see Appendix 1). This objective is today attainable only with additional special-purpose financing. Ensuring the green open access to research articles can be attained without substantial additional funding. The maximum allowable embargo period with green open access is one year.
- Researchers and authors must maintain their freedom to decide and choose the publishers and publications where to publish and the repositories where to preserve their research results, following the good practice of scientific publication.
- Open access must be applied in accordance with the European Union and Estonian legislation, including copyrights. Of different content licences, it is preferable to use for research publications the Creative Commons Attribution 4.0 International licence (CC BY 4.0) or other licences that meet the definition of free cultural works.
- The Estonian Research Information System (ETIS) is developing into the Estonian research information database, including the central repository of research publications. It is necessary to ensure its continuous development, and compatibility with the OAI-PMH protocol (Open Archives Initiative Protocol for Metadata Harvesting) and other open standards (e.g., ORCID).
- The principle must be introduced that in order to confirm a research publication in ETIS it is necessary to add its full text (pdf) and indicate whether the publication has closed access, or it is self-archived with an embargo period (green open access) or has immediate open access (gold open access). It must be possible to include the licensing conditions into the record of the research publication. In addition to other functions, ETIS must facilitate comfortable sharing of closed access publications with colleagues.
- Research journals, published in Estonia with the support of public funding, must have

immediate open access; they must be entered into the register of open access research journals DOAJ, must comply with the common principles of open access, use the CC BY or some other licence, and guarantee long-time preservation of the publications and access to the archive of the articles.

- It is necessary to consider the costs and benefits of the state membership in the activities of open access publishers where the membership fee reduces the rate of publication fees (i.e. BioMed Central/Springer Open Shared Support Membership).
- In case of public funding of research, minimum requirements must be established for open access research journals, such as their inclusion in the Directory of Open Access Journals list (DOAJ, http://doaj.org), or in the Web of Science, SCOPUS, PubMed and EBSCO databases. Publication fees of the journals which are not listed in these databases are not eligible.
- Considering the fact that the number of journals with chargeable open access is increasing but their scientific quality and credibility cannot always be verified, it is important to make sure that formal requirements should not have a negative effect on the quality of research and publications.
- It is also necessary to ensure that the application of the principles of open science should not worsen the access of Estonian researchers to subscription research publications and other research information.
- It is important to increase the target groups' awareness of open access, copyright and different licensing conditions (incl. different versions of Creative Commons licences).

5. Open access to research data

Estonia's strategic objective 2020

The academic community is familiar with and accepts the principles of open science and open data. Research data collected with the support of public funding are freely accessible and re-usable to all, if not restricted by legal requirements. Research data are preserved in open repositories which meet certain standards and are made available at the earliest opportunity.

Principles and recommendations

- The applications for public funding for research projects must include data management plans, which contain information about publishing and licensing of the data.
- Research data must be securely preserved, in compliance with the requirements and customs of the research domain. In preserving the data, it is preferable to use the international disciplinary or universal repositories which meet the certain quality requirements. Research data concerning Estonian state, culture, population and territory are (also) preserved in Estonian repositories. The repositories used must meet certain standards[4].
- Research data are preserved in as original form as possible without reducing their quality and

---

[4] E.g., OpenAIRE Guidelines for Literature Repositories (https://guidelines.openaire.eu/en/latest/literature/index.html)
OpenAIRE Guidelines for Data Archives (https://guidelines.openaire.eu/en/latest/data/index.html)
Data Seal of Approval (http://www.datasealofapproval.org/en/)

level of detail.

- The preserved research data must be machine readable and interoperable; scanned texts must be alphanumerically identifiable and allow data mining.
- Research data must be available as early as possible, preferably, simultaneously with the appearance of publications. If the data is not accompanied by a publication, it must be available by the completion of the related research project at the latest.
- Research data must be accompanied by standardised metadata and unique digital persisent identifiers (e.g., DOIs).
- Relationships between research data and publications must be reflected in the metadata, which forms the basis for mutual linking and referencing.
- For licensing research data it is preferable to use, similarly to publications, the open content licences. The licence must allow accessing, preservation, copying, forwarding and re-using of the data.
- Similar principles must be applied in preservation and making available the research collections, which have been the basis for the research, or have been formed in the course of the research, and other supporting materials.
- It must be possible to register the published datasets at the metadata level in ETIS similarly to publications; it must be possible to use them as an additional criterion (depending on the speciality) in evaluating researchers and research groups.
- Research data must be available for reuse if it is not limited by restrictions related to national security, personal data, ethics, business interests, intellectual property (copyright or industrial property) or to other legal issues.

6. Implementation of principles

The effect of the external environment, especially, the success of the European Union and its member states in adhering to and applying the principles of open science, is an important factor in the implementation of Estonia's principles of open access. For example, it is unlikely that Estonia alone could achieve significant results in influencing international research publishers, but cooperation with other countries could open up possibilities for fundamental changes. General trends in Europe support the implementation of the principles and recommendations set out in this document.

In order to implement these principles, it is first necessary to impose additional requirements on the research project applications and reporting. It must be taken into account that such requirements would add to the researchers' workload and administrative burden and, most probably, also to the costs. Therefore, it is sensible to apply the requirements of open access only gradually. The economic aspects of open science would need a special analysis. The researcher's freedom to choose where and how to publish their work must be maintained. Without additional funding it is not possible to impose the strict requirement for open access or the proportion of open access on researchers.

One of the greatest challenges in implementing these principles is the necessity to demonstrate to different stakeholders how open access to research information creates new possibilities, helps to increase the efficiency and quality of research, supports the academic career of researchers and

authors and contributes to the wider development of society. At the same time it is necessary to avoid the situation where the requirement for open access would have a negative effect on the level of research and publications or lessens the amount of publications and other research information accessible to Estonian scientists at present.

Key stakeholders and their roles in the implementation of the principles of open science

Ministry of Education and Research

- Is the contact point for open science and open access in Estonia, ensures the communication with the European Commission and with national research and development institutions.
- Cooperates with other funders and Ministries in the application of similar principles under the authority of the Ministry of the Environment, Ministry of Rural Affairs, Ministry of Culture and others.
- Finances, and seeks opportunities to find and channel additional funding into the widening of open access to research information and speedier application of the principles of open science.
- Ensures the ability of ETIS to act as a central Estonian repository of research publications by ordering technical solutions necessary for fulfilling this task.
- Ensures the sustainability of the e-infrastructure of Estonian research and its potential for inter-operability with the European research infrastructures.
- Ensures the strengthening of the position of Estonian research journals, which meet the international criteria, in the field of research publishing.

Estonian Research Council, Archimedes Foundation, Enterprise Estonia, and other research funding organisations

- Establish the principles of institutional open science.
- Establish clear rules to funding measures which require open access to published publications and research data, define the allowed and preferable licensing conditions and apply other measures for facilitating open access.
- Apply the information about the publication and citations of research data as one of the criteria for research performance evaluation.
- In cooperation with research and development institutions, develop basic forms of data management plans.
- In cooperation with universities and research libraries, coordinate the awareness and knowledge-raising in the matters of open access and open science, and the developing of common terminology.
- Estonian Research Council monitors and analyses, by using ETIS, the trends in the changes in the proportion of open access articles, in the scientific level of open access articles and journals and the statistics of research data publishing.

Estonian Academy of Sciences

- Supports and promotes the principles of open science and the unrestricted availability of research information.
- Organises the publishing of research journals, ensuring the immediate open access to research works and the openness of the journal archives, and uses as free as possible content licences (preferably, CC BY).

Universities and other research and development institutions

- Draw up institutional principles and the action plan for open science.
- Promote open access to research information in teaching and learning as a component of open and knowledge-based society.
- In publishing, prefer open access and as free as possible content licences (preferably CC BY).
- Cooperate with DataCite Estonia in the development, operation and promotion of a central and speciality-based research data repository.
- Ensure the reasonable use of copyright issues in the employment contracts of academic staff, giving the authors the right to submit and publish their articles in open access journals without the employer's additional permission.
- When evaluating the performance of researchers and teaching staff, consider besides publications also the publishing of data and citations of the data.
- Take measures to prevent publishing of research results in paid open access journals, whose quality and reliability cannot be verified (including a thorough scrutiny before the confirmation of the publications in ETIS by the institution).
- Open and develop curricula on research data (data scientist).
- Provide the researchers of their institutions with necessary tools for data management.
- Research libraries advise and train Estonian research community in open access and data management issues and, in cooperation with research funders, develop the basic forms for data management plans.

Researchers

- Recognise the principles of open science and consider the public accessibility of research results and sharing of data and other basic materials as a norm.
- When choosing publishers, make informed decisions by proceeding from the good practice of research publishing, knowing the possibilities for open access and self-archiving offered by the chosen journals, and when all other conditions are equal, preferring publishing with immediate open access.
- Are familiar with the main discipline-specific research data repositories in their field, can use them for finding comparison data as well as for depositing their own data.
- Use the EU and other available opportunities to apply for additional funding for publishing their research results in open access journals and data repositories
- Do not make concessions on the quality of their research for keeping up with the open access principles.

Implementation schedule

2016–2017

- The Expert Group presents the working version of the document "Estonian open science principles and recommendations" to different stakeholders and to the public (2016).
- Implementation of the principle that in order to confirm publications in ETIS it is required to submit also the machine readable full text (pdf) or DOI together with an indication about the status and conditions of open access.
- Estonian Research Council starts/continues to monitor and analyse the open access-related situation and trends on the basis of the data available in ETIS.
- Cost-benefit analyses will be carried out (2016-2017) in order to get the whole picture of the relevant costs for Estonia, incl.:
  - The open access journals and publishers, used by Estonian researchers for open access publishing will be mapped and these costs will be analysed.
  - Access costs for traditional journals and research databases will be mapped.
  - Analysis of the research data preservation costs, based on real-life practice (research and development institutions, the existing research infrastructure objects, such as NATARC, ETAIS, Center of Estonian Language Resources, etc.) will be carried out.
  - Policy recommendations will be developed based on the analyses mentioned above.
- Estonian Research Council has established the institutional and the Ministry of Education and Research the national open access principles, which are followed by other institutions within their scope: the principles contain precisely formulated objectives and indicators for monitoring the achievement of results.
- Requirements for broadening the percentage of open access research publications are implemented gradually: at least 50% of the published results of competitive projects must have open access.
- Applicants for personal research funding are required to submit the data management plan; implementation of the plan is evaluated in the mid-term and final reports (starting from 2019). This requirement will be applied at all the following funding application rounds.

2018–2020

- Research and development institutions have agreed upon/confirmed the institutional open access rules and guidelines.
- The proportion of open access publications and research data will rise in project applications and reports to at least 75%.
- In the appearance of additional resources, additional funding is allowed for ensuring gold open access; the proportion of gold open access can in this case be applied as a separate criterion (e.g. 50%).

Appendix 1. Clarifications of terminology

Article processing charge (APC) – costs related to publishing an open access publication, incl. content and language editing, proofreading and formatting, adding of links, etc. The APC for high-quality journals is usually 1000-3000 EUR. See Gold Open Access.

Data Management Plan – a plan, associated with research, outlining which data will be collected/produced in the course of research, how they are managed and archived and which part of them will have open access. The aim of the data management plan is to describe the whole data management process during the project and after its completion. The data management plan is submitted together with the research project application.

Dataset – a collection of data elements. Most often a dataset corresponds to one table in a database or one data matrix.

E-infrastructures – unique, information technology-based infrastructures that offer their users easy and safe web-based access to resources, tools, methods and means needed for research and thus contribute to the transforming of research even more complex, global and multidisciplinary. E-infrastructure can be, e.g., high-capacity and high-performance data networks (e.g. GÉANT), dispersed computer infrastructures (networks and clouds), supercomputing infrastructures, infrastructures that collect and preserve research data, and the services accompanying these activities.

Embargo period – a period, required by the publisher, during which it is not allowed to make the full text of the publication publicly accessible. Most of the open access regulations establish that open access to articles published in peer-reviewed research journals should be ensured immediately after publishing, but not later than 12 months after the first publishing in the humanities and social sciences, and maximally six months later in other fields of science. The standard agreements, offered to the authors by leading scientific publishers, often do not grant the rights for archiving research in repositories; usually they require a longer delay in allowing open access. See *Green Open Access.*

Free cultural work – a work, including a research publication, whose licensing conditions allow unrestricted using and sharing on all kinds of purposes. The definition of free cultural work is met by CC BY and CC BY-SA licences, but not by, e.g., CC BY-NC, which does not allow using the work for commercial purpose (see https://creativecommons.org/freeworks).

Machine readable – a file format structured in such a way that software applications can easily recognise, verify and read specific data, including single facts, and their internal structure.

Open Access – electronic unrestricted access to research information. The activities of international commercial research publishers have up to now mostly been grounded on the order-based business model (paid by the reader, the library or the state) and the result, i.e. research publications, is not distributed free of charge. Due to the pressure exerted by open access policies, but having also seen new income opportunities, the leading scholarly publishers have created new scientific publications which use the gold open access business model (as a rule, the author pays the publication fee). The business model of core publications has usually not been changed, but an opportunity is offered to

the authors of the so-called hybrid publications to "redeem" their articles. However, in case of hybrid publications there may be a danger that taxpayers' money will be spent twice – for subscribing for such publications by libraries and for paying the publication fees from the grant funding. An overview of licensing conditions of scientific publications and self-archiving rules can be found in the SHERPA/ROMEO register (http://sherpa.ac.uk/romeo). Open access is divided into:

- **Gold Open Access** – the publication is immediately and in its final form accessible for all on the publisher's web page. Publication costs may be the responsibility of the author (see Article processing charge), or the publication of the journal may be supported by some institution or organisation, e.g., university, speciality association or academy of sciences. Most of Estonian scientific journals use the latter model and the authors do not need to pay the publication fee. Precise conditions for using the gold open access journals are specified by licences. Most of the open access journals use the so-called Creative Commons licences. Such publications can also be self-archived in their final form in an institutional (e.g. the digital archive of the university), the state (e.g. ETIS) or international repositories (e.g. ArXiv, PubMed Central).
- **Green Open Access** – the author deposits a published research article or a peer-reviewed manuscript into an institutional (e.g. the digital archive of the university), the state (e.g. ETIS) or international repositories (e.g. ArXiv, PubMed Central), where it is accessible for all. Self-archiving conditions may differ by publishers and journals. Usually, there is an embargo period set by the publisher, meaning that self-archiving is allowed not earlier than 6-12 months after the first publication of the article. Generally, self-archiving of the final version (see Publisher's PDF) is not allowed. In case of green open access, the authors do not pay the publishing fee. Operational costs of the publishers are covered by the journal and article subscriptions and additional services, which are paid by libraries and research institutions.

Open licences – in general, publications and (research) data are protected by copyright and open access has to be established by licences (Creative Commons: CC-Zero, CC BY, and others; more information http://www.hitsa.ee/teenused/autorioigused, http://creativecommons.org).

Open methodology – public and free access to methods, means and analyses used for carrying out research.

Open peer review – ensuring of scientific quality by transparent and monitored process of pre-evaluation.

Open source – free public access to the source code of software products.

Open source software – source code and related software are publicly accessible for all users both for consultation and rewriting.

Peer review – independent evaluation preceding the publication by at least one appropriate expert.

Persistent identifier – identifiers used for searching, identifying and linking of publications and datasets, e.g., ISBN (monographs), Handle ID (repositories), DOI (publications and datasets).

Post-print – peer-reviewed draft of an article.

Pre-print – draft of an article which has not yet been peer-reviewed.

Publisher's PDF – the last, peer-reviewed pre-press version of a publication where all details are finalised.

Repository – the environment and infrastructure for long-time preservation of publications and data.

Research data – all the data that have been created or collected in the course of research; data that have been created for some other purpose, but can also be used in research, cannot be considered as research data (e.g., environmental monitoring data).

- **Open data, open research data** – electronically publicly available data, created and used in research by researchers, research groups, research institutions or the public sector.

- **Metadata** – standardised amount of information used for describing an object (publication, data, etc.). Authorship, title, institution, time of publication, and other data are generally required.

Research publication – a peer-reviewed research article or scientific monograph, accessible in the form of full text (printed or electronically available), or indexed in subject databases and reference journals.

Self-archiving – see *Green Open Access.*

Text and data mining, TDM – an automated process for finding useful information and patterns of events in large datasets. It consists of pre-processing, mining and post-processing of data.

Appendix 2. A selection of strategic documents concerning open research data from the EU, OECD and different countries
(read and cited on 24.05.2016)

OECD (2015), "Making Open Science a Reality", OECD Science, Technology and Industry Policy Papers, No. 25. http://dx.doi.org/10.1787/5jrs2f963zs1-en

OECD (2007), "OECD Principles and Guidelines for Access to Research Data from Public Funding". www.oecd.org/sti/sci-tech/38500813.pdf

"A Reinforced European Research Area Partnership for Excellence and Growth" (COM(2012) 392). http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-partnership-excellence-growth_en.pdf

"Towards better access to scientific information: Boosting the benefits of public investments in research" (COM(2012) 40). http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf

"Commission Recommendation on access to and preservation of scientific information" (COM(2012) 4890). http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf

"Guidelines on Data Management in Horizon 2020" (2016). http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

"Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020" (2016). http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

"Science Europe Principles on Open Access to Research Publications": D/2015/13.324/. http://www.scienceeurope.org/uploads/PublicDocumentsAndSpeeches/WGs_docs/SE_POA_Pos_Statement_WEB_FINAL_20150617.pdf

"European Open Science Cloud" (2016)
https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

Draft European Open Science Agenda (2016)
https://ec.europa.eu/research/openscience/pdf/draft_european_open_science_agenda.pdf

"The transition towards an Open Science system - Council conclusions" (2016)
http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/

Finland: "Open Science and Research Handbook" (2014). http://openscience.fi/handbook

Sweden: "Proposal for National Guidelines for Open Access to Scientific Information" (2015). https://publikationer.vr.se/en/product/proposal-for-national-guidelines-for-open-access-to- scientific-information/

Norway: "The Research Council's Policy on Open Access to Research Data" (2014). http://www.forskningsradet.no/en/Newsarticle/Research_data_must_be_shared/1254000848864/p1177315753918

Ireland: "National Principles on Open Access Policy Statement" (2012). http://www.dri.ie/sites/default/files/files/National%20Principles%20on%20Open%20Access%20Policy%20Statement%20%28FINAL%2023%20Oct%202012%20%29.pdf

Slovenia: "National Strategy of Open Acess to scientific publications and research data in Slovenia 2015-2020" (2015). http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National_strategy_for_open_access.pdf

Estonia: Open data green book (2014). https://opendata.riik.ee/roheline-raamat

Appendix 3. A selection of national and international repositories of research data and publications

(read and cited 24.05.2016)

- **Estonian Research Information system ETIS**, contains a repository of Estonian research publications: http://etis.ee
- **DataCite Estonia**, a network of research data repositories through which Estonia participates in the work of the DataCite Consortium; it also assigns DOI persistent identifiers to research data: http://datacite.ut.ee
- **arXiv**, a global repository of publications in the fields of physics, mathematic, computer science, etc.: http://arxiv.org
- **PubMed Central**, a global repository of publications in the field of life sciences: http://www.ncbi.nlm.nih.gov/pmc
- **Pangaea**, a global repository of data in the fields of geo and environmental sciences: http://pangaea.de

Additional useful links:

(read and cited 24.05.2016)

- SHERPA/ROMEO (http://www.sherpa.ac.uk/romeo): Information about the conditions of open access and self-archiving of research journals;
- DOAJ, Directory of Open Access Journals (http://doaj.org): a global register of open access research journalsr.