CNRS ROADMAP
FOR OPEN SCIENCE

18 November 2019

TABLE OF CONTENTS

INTRODUCTION

The international movement towards Open Science started more than 30 years ago and has undergone unprecedented development since the web made it possible on a global scale with reasonable costs. The dissemination of scientific production on the Internet, its identification and archiving lift the barriers to permanent access without challenging the protection of personal data or intellectual property. Now is the time to make it "as open as possible, as closed as necessary".
Open Science is not only about promoting a transversal approach to the sharing of scientific results. By opening up data, processes, codes, methods or protocols, it also offers a new way of doing science.

Several scientific, civic and socio-economic reasons make the development of Open Science essential today:

- Sharing scientific knowledge makes research more effective, more visible, and less redundant. Open access to data and results is a sea change for the way research is done, and opens the way to the use of new tools. These tools result from recent research into artificial intelligence (text and data mining, automatic learning), facilitate research beyond boundaries and increase interdisciplinarity.
- Open Science changes the way research is part of society by helping to restore trust between citizens and scientists. Accessibility by many actors and disciplines empower responses to contemporary issues (e. g. multiple and/or cross-disciplinary challenges) and scientific integrity is strengthened.
- Researchers often pay to publish, pay to read articles and do the reviewing work for free, while subscription costs continue to rise. Scientists are regaining control over a system that had escaped them and has become financially unsustainable because of the constant increase in the costs of scientific dissemination through the currently accessible publication of scientific results.

Just over a year ago, France embarked on this vast transformation movement. Presented on 4 July 2018 by the Minister of Higher Education, Research and Innovation, the "National Plan for Open Science"[1] aims, in the words of Frédérique Vidal, to ensure that "the results of scientific research are open to all, researchers, companies and citizens, without hindrance, without delay, without payment".

As a result, an Open Science Committee (CoSO), chaired by the Director General for Research and Innovation (DGRI), was set up. It is divided into four colleges: publications, research data, skills and training, Europe and international. CNRS researchers and support staff, as well as members of the DIST (Scientific and Technical Information Department), are present in all the groups and jointly steer the operation of three of the four colleges. The National Open Science Plan is a commitment for research operating organisation to draft their Open Science policies.

As the work carried out within the joint research units by researchers and teacher-researchers is intricate by nature, the development of Open Science must be considered nationally and jointly with all research stakeholders, and in particular other research organisations and our university partners. This must happen in strong coherence with the action of the MESRI (Ministry of Higher Education, Research and Innovation) for Open Science.

The CNRS roadmap for Open Science will be implemented through an action plan regularly reviewed and placed under the responsibility of the Department of Scientific and Technical Information (DIST), which will put its three service units into action:

The Inist (Institute for Scientific and Technical Information), a CNRS service unit set up in Nancy in 1989, has just been reorganised into three departments linked to development directions in Open Science:

- Access to scientific information (a major and historical focus of Inist missions with the negotiations and the BibCNRS portal, but also the new ISTEX and PANIST platforms for publications);
- Research Data Development and Dissemination (a line that should expand in the future and which now takes advantage of the pioneering investment made by the Inist by developing tools around research data like DMP OPIDoR (Optimisation of Research Data Sharing and Interoperability) to make Data Management Plans, etc. and also to build data repositories for "small" data);
- Scientific Information Analysis and Mining (a forward-facing development line for Text and Data Mining tools).

The CCSD (Centre for Direct Scientific Communication), a joint service unit, the architect of the open archive HAL (Hyper Articles Online), and which subsequently developed Episciences, a platform for open access journal publications and other services;

PERSEE created in 2013 to digitise and develop and disseminate scientific heritage.

The Open Science policy will grow in a differentiated way according to the different disciplines. It must be central to the CNRS institutes. In particular, these institutes operate the many research infrastructures that produce data. Some are thematic infrastructures for the management and sharing of data included in the national roadmap for research infrastructures. These resources play an important role in the national, European and international landscape, such as the Strasbourg Astronomical Data Centre, the Huma-Num platform for digital humanities, or the Data Terra Research Infrastructure (formerly the Earth System). The institutes have also developed publication initiatives, in particular the open access journal and book publishing platform "OpenEdition", the Mersenne centre developed for the open access publication of mathematical journals and currently in the process of opening up to any discipline in which articles are drafted in LaTeX format. The DIST is in permanent contact with the thematic institutes thanks to the Scientific and Technical Information correspondents (or STI correspondents). A network of "data correspondents" is being set up as part of the French node of the Research Data Alliance (RDA).

The implementation of the CNRS roadmap for Open Science aims to accelerate the process towards Open Science by relying on concrete actions around four main objectives:

1. Keep control of our scientific production and achieve 100% open access to CNRS publications during the term of this roadmap;

2. Develop a culture of data management/sharing among all stakeholders in the data life cycle: researchers, engineers, computer scientists, librarians, etc. based on the implementation of the FAIR principles (findability, accessibility, interoperability and reusability);

3. Develop and promote infrastructures and tools for fully independent mining and analysis activities on scientific contents;

4. Transform the individual assessment of researchers by making it compliant with the goals of Open Science together with taking their contributions to Open Science into account in assessments.

The whole process will be supported by training activities and an international strategy.

1. **bit.ly/nationalplanSite**

1 | PUBLICATIONS

Objective: Scholarly publications written and produced by CNRS researchers and financed mainly from public funds must be 100% accessible* and reusable during the term of this roadmap. Copyright must not be transferred.

The CNRS's strategy to reach 100% of scientific publications in open access is to promote "bibliodiversity". This means a diversity of research outputs as well as a diversity of publication methods, as there are several possible ways to publish in open access. Beside open archives, the models considered virtuous for publishing in open access, i.e. for non-profit purposes, are multiplying. Relationships with conventional publishers are still contemplated, but as part of negotiations aimed at reducing costs and promoting open access.

Action 1: Conduct a policy of support and development of the open HAL archive in conjunction with a policy of encouraging the deposit of scholarly publications

This involves the development of the CCSD, the Joint Service Unit that operates HAL. In 2019, the CNRS provided exceptional additional resources of €650k in addition to a position open to the public service recruitment competition. The sustainability of this increase in the resources needed for a new ambition will be covered in 2020 by a grant from the national fund for Open Science and further by user contributions based on a business model for an adapted development of HAL, in line with its many partnerships and which has been validated by the CoSO. This is consistent with the National Plan option to make HAL an open archive developed by the CNRS and its other supervisory authorities, INRIA, INRA and Lyon University, the central component of its policy. This infrastructure has been adopted for several years by many universities and research organisations with the development of 137 institutional portals.

All studies on open archives show that it is not enough to foster depositing in open archives but that a threshold of 80% of deposits can only be reached by a more incentive-based policy. Therefore, the CNRS requires that all scholarly publications resulting from research funded mainly by public funds, and which can be deposited in open archives on the basis of the Digital Republic Act should be accessible in HAL.

A first step was taken in 2019 by the CNRS, which requested that only publications available in HAL should be reported in the 2019 Annual Researcher's Report (or "CRAC"). A similar mandate for "RIBAC" reports (annual activity reports of humanities and social sciences researchers) is planned for 2020. This measure was announced to the unit directors by a letter from the Chief Scientific Officer (DGDS) dated 19 April 2019.

The increase in the deposits in HAL must be supported by an improvement of the HAL tool to meet the needs of the French research communities. The CNRS has undertaken to facilitate researchers' deposits in HAL. Interoperability with other international open archives such as ArXiv, PubMed Central and RePEc should be consolidated or developed. This effort also includes matters related to moderation, standardised referencing, harvesting to enhance the HAL Research Infrastructure.

Action 2: Recommend the use of preprint servers which host manuscripts submitted to journals, to provide for rapid open access distribution channels through non-profit platforms

Many communities have developed or are beginning to develop preprint servers like ArXiv. This is a pioneer in physics and mathematics, for the rapid dissemination of research results in parallel with the certification processes provided by journals. These include ChemRxiv in chemistry, BioRxiv in life sciences and SocArXiv in human and social sciences (HSS).

These platforms make it possible for each manuscript not yet reviewed to guarantee the attribution to authors, dating, permanent identification, hosting and the attribution of a license defining the conditions for reuse. Some of them provide for the direct submission of manuscripts to a series of journals and direct linking to an article's final version as and when it is published.

(*) this does not affect the protection of personal data and intellectual property or any other necessary protection. The goal is to be "as open as possible, as closed as necessary."

In an initial phase, the CNRS will select a certain number of platforms to support to facilitate these new methods of open access scientific communication according to their importance for the communities concerned, in coordination with the national Open Science fund. It will also seek to advance interoperability between the different deposit platforms. We will mobilise resources so that the CCSD can propose solutions jointly with those in charge of these platforms and develop tools that will facilitate single deposits for researchers.

Action 3: Support electronic publishing platforms that host open access texts and provide publication and certification venues

The ultimate goal is that every research community, especially disciplinary, can find alternatives to channels its publications towards open access publishing systems, with economic models institutions can support so that it is not the author who pays, either to publish or to read.

Open access publication platforms offer different types of material: journals, books, recommendations, referees' evaluation reports, readers' comments, annotations or other forms of discussion around publications. In fact, some platforms are hosts, as they host and promote scientific content produced by other organisations (journals, publishers, organised communities, open archives) while others select and produce their own content directly. In all cases, these platforms allow for more or less open peer reviewing (open reports, open identities). The vast majority of them publish documents under a creative commons license, with intellectual property being retained by the authors.

Even if some of them aim to cover all disciplines, they are shaped by a history of their own and more restricted options. For example, Copernicus is mainly specialised in geosciences, Scipost in physics, the Mersenne centre in mathematics, OpenEdition in HSS, E-life and F1000 in Life Sciences, Episciences in mathematics, computing and HSS and PCI (Peer community In) in environmental sciences. These choices can be related to the authors' writing format options, in particular the use of Latex or other word processing software. This is the case, for example, of the Mersenne centre and OpenEdition respectively. On the publication output itself, different software is used to produce documents in HTML, PDF, Epub or XML (e.g. Lodel, OJS).

The CNRS will support the scientific communities, discipline by discipline, in the development of alternative publication solutions with a publishing and peer review context comparable to traditional publishers.

Action 4: Support a documentary strategy that aims to reduce the burden and cost of journal subscriptions and increase the accessibility of manuscripts in the published version.

Each year, the CNRS pays more than €12 million excluding VAT to some 50 publishers for its laboratories to access to scientific documents kept behind paywalls (through BibCNRS portal implemented by the Inist). Some access for the CNRS joint units is financed and made available through partners. Access can be provided through university portals, organisations. Some subject libraries finance their own complementary subscriptions (e.g. as part of the mathematics network). As a result, researchers spend more and more time accessing scientific production. Negotiations under Couperin for contracts with publishers who corner the market must result in lower journal costs and an open publication policy, and especially compliance with the French Digital Republic Act. The payment of publication charges to make an article freely accessible in an otherwise subscription-based journal is an approach that should be strongly discouraged, as it is not in line with the national plan for Open Science, and causes additional costs that are difficult for an institution to monitor and control. These are the so-called hybrid journals which are not apparently virtuous since they are in fact subject to double payment. Publication in such journals, without paying publication charges, together with a deposit in HAL preserves the freedom of a researcher to choose his/her publication channel.

Action 5: Request that all publications resulting from research work funded by a CNRS call for projects be made available for open access

This objective is rather a minimal, tangible and immediate action to be implemented, consistent with the national Open Science plan and the requests of Europe and the ANR in their own calls for projects. Open access to a publication will be possible according to the diversity of possible avenues outlined above: open archives, open access journal publishing platforms. The open archive approach is a virtuous alternative to the requirement of APCs (article processing charges), which are very high for journals deemed "prestigious".

2 | RESEARCH DATA

Objective: Data (raw data, texts and documents, source codes and software) produced by CNRS researchers or using resources implemented by the CNRS must, as far as possible, be made accessible and reusable. This is according to the FAIR principles for the consolidation of knowledge essential for the development of a more efficient research. The data should be "as open as possible, as closed as necessary".

Action 1: Develop a culture of data management/sharing among all stakeholders in the data life cycle: researchers, support engineers, computer scientists, librarians, etc.

With the current massive emergence of digital data produced during research, it is essential to develop and share best practices (FAIR) between research actors. Data in the broadest sense refers to raw or processed data, texts and documents, source codes and software. The OPIDoR services developed by the Inist help with the assignment of DOI (Digital object identifier) identifiers to data sets (via DataCite) and support for the implementation of Data Management Plans (DMP). The European Commission and more recently the French ANR research (funding) agency have requested DMPs when writing research project submission. In addition, the notion of "Data stewardship" reflects the prospect of identifying resource personnel. The Mission for the management and relations with regional delegations and institutes (MPR) has developed a methodological data management tool that can help in the implementation of a data sharing and management policy. We will set up a network of resource persons who will be responsible for data in research structures and projects and will receive training in the use of specific services.

The principles will be implemented in different ways in different disciplinary contexts, in an undogmatic approach and in accordance with established practices. The very definition of data varies greatly from one scientific community to another. This involves in-depth discussion within communities on the data life cycle, its backup, access, archiving, and reuse. All CNRS institutes and their communities have not reached the same level of progress in developing a culture of data management and sharing. The challenge is to raise awareness among all research stakeholders in order to achieve the implementation of the FAIR principles.

Action 2: Develop data papers, joint deposits of publication and underlying data and support researchers in the use of data management tools

This objective is a good means to engage concretely in the issue of data sharing. The data on which publications are based must be made accessible and reusable as soon as possible. This requirement is already implemented within the European H2020 programme and must be extended to any data produced with public funds. It will therefore be a matter of facilitating the deposit of the underlying datasets of a publication in a repository simultaneously to the deposit of the publication in an open archive (which does not imply that the data should be deposited in that open archive). The underlying datasets of publications will be deposited in an appropriate data repository, if possible a thematic repository best matching the quality assurance and dissemination of the data. More generally, the CNRS encourages the dissemination of all structured data by a deposit in thematic or generalist data reservoirs, possibly supplemented by the publication of data papers. For this, the CNRS Institutes in conjunction with the DIST services may identify resource structures around them, such as University libraries, human sciences research resorts (Maison des Sciences et de l'Homme), as well as resource staff able to support scientists in these new practices. This data curation floating staff foreshadows the actual recognition of data stewardship. The Institutes which already have a journal support policy have additional leverage and could expand it with a publications data component in the future.

Action 3: Support and enable data producing research infrastructures in setting up and implementing data policies

The CNRS is widely involved with its partners in national and international Research Infrastructures (RI), which are the venues where research data are created and analysed, e.g. analytical instruments, computing infrastructures, data infrastructures, observatories, etc. The CNRS will publish an infrastructure charter to broaden the application of FAIR principles to all disciplines and as a commitment for infrastructures to adopt FAIR practices and quality standards and to release data policies drafted in concertation with the scientific communities using them. Some infrastructures (such as Progedo and Huma-Num at the Institute for Humanities and Social Sciences (INSHS)) have moved quite forward in this process, while others like the Chemistry research infrastructures are at support stage. The SOLEIL synchrotron has also launched a data management policy. There are many examples and they tend to be generalised. These developments must be correlated with certifications (CoreTrustSeal type) when the infrastructures also disseminate their data.

Action 4: Support and follow-up of Data Infrastructures - Implementation of a coordinated service with the institutes to promote the deposit of data for all staff of CNRS units

Thematic data infrastructures play a national or international role. Some are included in the national roadmap for research infrastructures. This is in line with the structuring measure of the National Open Science Plan, which recommends "developing thematic and disciplinary data centres". The CNRS will continue to support these infrastructures, and will support the development of new thematic data reservoirs and services. This support will be conditional on an assessment of their impact, their adequacy to scientific needs, and the quality of their management. CoreTrustSeal certification will be sought.

Data sharing is already a widely developed practice in some disciplines; most often international disciplinary repositories have developed over the last several decades, particularly in astronomy. This has not materialised yet in other disciplines for which the international structuring of research data is less straightforward and/or established, and because the very nature of the data produced (size, heterogeneity, complexity, etc.) requires prior consideration of the very nature of the data. Thus, many units can create a large number of "small data" (in the sense of data storage, as opposed to "big data"). This is called "long-tail data". The CNRS will study the opportunity to create a general data repository for the long-tail data for which there are no identified thematic repositories. This reflection will be part of a national review about the management of long-tail data and the assessment of the National Plan requirement "to develop a generic service for the reception and dissemination of simple data". Since the reuse of the data sets thus collected can only be effective within a recognised and trusted framework, a CoreTrustSeal certification mechanism will be implemented.

Action 5: Create and display a directory of repositories and data services for which the CNRS is responsible and in which it participates

This directory would make it possible to display an important element of the CNRS's contribution to Open Science. It would include the data services of the above-mentioned research infrastructures set up under Action 3, the thematic data centres and services (Action 4) and the general data repository possibly created under Action 4.

3 | TEXT AND DATA MINING AND ANALYSIS

Objective: Facilitate text and data mining by developing infrastructures, tools and skills to enable fully independent analysis capacities on scientific content.

In recent decades, the volume of information produced by research has increased. In a given field, a researcher is in no position to "ingest" all this information produced, when she/he can afford to access to it. The new support services for researchers in the field of scientific and technical information (STI) must move towards content analysis (content mining) in addition to their offering in services like abstracting and indexing (bibliographic databases, catalogues) and access (portals) to scientific production. These services rely on proven technologies (indexing based on authority files, search engines, identifiers, etc.) which are correctly mastered by the STI arena. Meanwhile, content analysis requires the development and implementation of technologies that are not yet well-mastered by all scientists (ontologies, semantic description, knowledge representation languages, etc.). Thus, content mining consists in exploring texts and data with computer programs to extract information from them. This process, repeatedly performed on separate corpora, is likely to produce new scientific knowledge. The search for texts and data is thus seen by researchers as an extension of their "right to read".

Action 1: Support and develop infrastructures for content analysis

It is important to invest resources in content mining activities. Recently, deep learning methods created new research opportunities to process massive and large data. The CNRS steered the ISTEX platform in this direction, with the idea that the aggregate collection of data from various sources enabled the development of tools and services for content exploration in response to researchers' demand. The Inist was the chosen operator because of its combination of technical infrastructures and terminological know-how and its scientific knowledge.

The VisaTM platform developed at the Inist and supported by MESRI has linked the Istex and the European OpenMinted platforms. It provides an overview of the text mining tools available. The VisaTM project shows the need for an infrastructure that provides a researcher with the technical and scientific environment necessary to carry out his/her research using content mining technologies.

Action 2: Legislative framework: Support, transpose and inform

In October 2016, France recognised and took into account the need to legally open any lawfully accessible content for research and teaching purposes in its Digital Republic Act. However, this intention was not followed by an implementation decree, pending the European Copyright Directive. Europe recently passed the text of this Copyright Directive(March 2019). It creates an exception for research allowing searches into any lawfully accessible content for academic purposes.
Its transposition into French law requires to remain alert to:
1. The constraints possibly imposed on the means of access authorised for text mining;
2. The implementation by platforms of systems to preserve the integrity of networks and servers;
3. All usage data (queries and downloads) collected, stored and potentially circulated by the platforms;
4. The licenses negotiated with publishers and how their clauses take into account the TDM legislative framework.

Action 3: Develop the use of data processing and visualisation tools and techniques

Disciplines communicate through texts and data, whether underlying texts or not. These information elements are important and numerous and difficult to understand as a whole by human beings alone. The possibilities of information extraction are multiplied by the use of these tools.
The units affiliated to the CNRS have already taken up these issues. For example, the Inist-CNRS has developed techniques and tools for corpus processing. Thus, the Inist offers content extraction services. Huma-Num, the very large research infrastructures of the Human and Social Sciences, offers a list of tools to extract information from texts. The ISC-PIF (Institut des systèmes complexes de Paris Ile de France, or Paris institute for complex systems) provides a platform for content extraction and visualisation, among other things. Researchers and support staff in laboratories must be familiar with these technologies and tools and so these should be made easier to use. Any type of questioning can be addressed by these tools: scientific issues by assembling scientific corpora matching specific inquiries or steering and decision support by studying a set of reports.

4 | INDIVIDUAL EVALUATION OF RESEARCHERS AND OPEN SCIENCE

Objective: Reconsider the individual evaluation of researchers by using an approach that is compatible with the objectives of Open Science and by taking into account in his/her evaluation the contribution of a researcher to Open Science.

Many studies and publications show that the current system of research evaluation is the main barrier to a transition to Open Science. This is a double-faceted question:

1. Today, evaluations are largely based on bibliographies. Assessing the quality of productions based on the prestige of the journal or publisher means entrusting responsibility for evaluation to the "referees" hired by the publishers. In some disciplines, this prestige is based on "closed and unverifiable" indicators like the journal impact factor. In this case, responsibility for the evaluation is proxied to "proprietary" algorithms based on "closed" data;
2. The energy and means put by researchers into making their production as openly available as possible are not in any way valued by the assessment procedures.

A number of field initiatives have been taken. These include the Jussieu Call[1], the San Francisco Declaration on Research Assessment (DORA)[2], drafted in 2012 by a group of editors and publishers of scholarly journals, signed by many organisations and individuals.

The CNRS signed DORA on July 14, 2018. This is a commitment to avoid the use of bibliometrics and to prefer a more qualitative assessment, as well as to take into account the full variety of research production types.

The sections and commissions of the French National Committee for Scientific Research (CoNRS) must determine and publish the criteria they use to evaluate researchers. However, the CNRS wants certain common criteria to be complied with by all. Some of these criteria match with the implementation of Open Science policies.

In this context, each section and each commission should include the following four principles in its evaluation criteria:

1. It is the results themselves that must be evaluated, not the fact that they may have been published in a prestigious journal or other well-known media: CoNRS members must take responsibility for their judgment and not rely on anonymous publisher evaluations or algorithms. This should be reflected in assessment reports.
2. For each production cited in the evaluation files, a researcher must explain the scope, impact, and individual share of his/her contribution. Providing a full and complete list of productions is unnecessary.
3. All types of production must be accepted as part of the evaluation[3]: In particular, in all cases where it makes sense, the data underlying the publication and the source code necessary to produce the results must be provided. Pre-prints and other working papers are acceptable production for evaluation. The same applies to "data papers".
4. All productions cited in the evaluation files must be accessible in HAL or possibly in another open archive[4]: these are understood to be the actual outputs themselves and not their bibliographic reference. It should not be necessary to provide them in the file, as an active link to the archive may be sufficient.

1. jussieucall.org
2. sfdora.org/read

3. See HCERES research output guides: www.hceres.fr/en

4. Three exceptions to this rule are admissible:
• Results that are too recent may be under embargo. In this case they must have been filed in HAL though, under an embargo period not exceeding those provided for by law (6 months in STM, and 12 months for HSS). They are to be provided as a private link to HAL (or included in the evaluation folder).
• This rule cannot be an absolute recruitment rule for candidates working abroad in foreign or international institutions, or private institutions.
• The production type may not be accepted in HAL.

5 | RECASTING SCIENTIFIC AND TECHNICAL INFORMATION FOR OPEN SCIENCE

Scientific and technological information has long been based on private mechanisms and systems, owned by for-profit companies that charge a high price for their services. In addition to its economic cost, this configuration limited control over data and the ability of public operators to enrich, merge and reuse them.
Open Science, as examplified in the Open citation initiative, should make it possible to change this situation and make the research carried out at the CNRS more visible, while streamlining* the information acquisition processes needed to facilitate researchers' and research support staff's work (limiting the number of surveys and multiple requests from units) and helping to manage these operations.

Action 1: Promote researchers' Orcid ID registration

The CNRS joined Orcid in May 2019 as a member institution and hopes that its research staff will widely adopt and use this unique identifier. In addition, the CNRS will take action to encourage its scientific staff to obtain an ORCID number and create their profile to gain international visibility. In return, it undertakes to clarify the information exchange processes between Orcid and the CNRS Information Systems "in order to simplify their lives".

Action 2: Work on new bibliometric indicators

The market for databases on publications, patents or other scientific products is undergoing a major transformation, with the emergence of solutions competing with well-established players (Web of science, Scopus) and the willingness of research funders, research institutions and publishers to make the metadata of these objects more accessible and usable, in addition to opening access to the products themselves. Moreover, these new actors, rather than presenting a selection of research outputs on the basis of the "prestige" of these certification sites, aim to give the most exhaustive possible views of the state of production, whatever the type of object considered (articles, books, conference proceedings, etc.) and the language used. In this context, the CNRS will regularly reassess its needs for tools and promote their adoption and distribution in line with its general Open Science policy (as a signer of DORA and the Jussieu call for Open Science and Bibliodiversity).

(*) Entered by the information holders, no double entry…

6 | TRAINING AND SKILLS

Objective: Support and train all research personnel, in particular doctoral students, and also to offer training sessions for research support professionals adjusted to the Professional Activity Branches and professional networks concerned.

The objective of adopting and appropriating the principles of Open Science and new practices for research will be achieved if this process provides the scientific community with adequate support and guidance. Insights into new uses and ultimately guidance to new practices must be based on an understanding of new challenges in open scholarly research.

The tools, services and infrastructures, in particular those set up by the CNRS, are crucial Open Science drivers. However, the simple presence of these elements does not mean that Open Science will become standard practice. To achieve this goal researchers and support staff must gain the knowledge and skills to embrace new scientific practices, to know and use tools, services and infrastructure, and to assist their colleagues during the transition phase.

Open Science Skills Acquisition[1]

Action 1: Develop the necessary skills and expertise for open access publication

- Platforms;
- Open archives (HAL and thematic archives);
- Licensing.

Action 2: Develop research data management skills

- Support services relating to the production, use/reuse of data, their dissemination according to the "as open as possible, as closed as necessary" principle;
- Compliance with FAIR principles;
- Help in creating a data management plan;
- Work on metadata, interoperability, etc.

Technical skills, particularly in data engineering, data science and data management.

Action 3: Develop scholarly skills to conduct open research, including skills in research integrity, ethics and law

Develop the skills to act within and beyond one's own scientific and disciplinary community.
A significant range of training courses that tend to develop skills in Open Science already exists. The plan for future training should focus on improving the quality and relevance of skills. To facilitate this task, it will be necessary to offer and promote traditional and/or online training courses and courses for scientists adapted to their career level.

Action 4: Develop support result mining and analysis skills in laboratories

The CNRS is involved in the development of the skills necessary to use tools to analyse texts and data, and visualise analysis results. The aim is to raise awareness of new information practices and strengthen user communities around these tools and technologies.

1. "Providing researchers with the skills and abilities they need to practice Open Science"- European Commission Open Science skills Working Group – July 2017

Objective: Take a place among international organisations in order to shape and share strategies and good practices for Open Science (both for scholarly publications and research data).

The CNRS policy is part of an international environment. The DIST is involved in many international actions and collaborations, either as such or through units or individuals. Within the CoSO, the Europe & International College is commissioned to coordinate French actors on all actions, monitoring and foresight across the board with other colleges. The CNRS is very active in this process.

Action 1: Align CNRS positions with the European and international Open Science framework on the theme of data

Open Science is a very broad international reflection undertaking. This is carried out top-down both by science policy bodies (DG Connect, DG Research of the European Union) and by funders (Science Europe, etc.), and from the bottom up by the scientific communities themselves (GDR, GO FAIR, Codata, etc.). The Research Data Alliance (RDA) consists of nearly 9,000 members from 133 countries who are organised into 88 working groups or interest groups. In 2019, CNRS experts created the French RDA national node.

The national implementation of these efforts is organised around the MESRI's CoSO, SPSO (Permanent Secretariat for Open Science) and its various colleges in conjunction with the department for the management of research infrastructures. The CNRS is already taking its full place in these mechanisms to bring a multidisciplinary point of view and will strengthen the commitment of its institutes that undertake initiatives to open up the data produced by their scientific communities.

Being part of the implementation of the EOSC and establish data services that it is to provide is a key challenge for the CNRS.

Action 2: Support initiatives that work towards identifying the elements of data FAIRisation

Data FAIRisation requires work at the disciplinary level to work out data formats and their description. More general work is also needed on the technological and sociological aspects of data sharing. This task has to be carried out at the international level. As put forward in the National Open Science Plan, RDA describes good practices in data sharing. It can accommodate disciplinary discussions in the absence of an international relevant setting. The CNRS will support the participation of its teams in RDA and its disciplinary initiatives.

Action 3: Communicate with our European and international partners on open access scholarly publication strategies

The exchange of scientific information and the sharing of research results through scientific publications only makes sense when international. Open Science can only advance if it is correlated internationally. The CNRS takes its rightful place in this global environment.

Knowledge-Exchange is an original collaboration between six partner organisations from six European countries that promote open scholarship and have an important - albeit different according to the partners - role in building, supporting or financing digital infrastructure and services for research and higher education. With its five partners (JISC [UK], SURF [NL], DFG [GER], CSC[FIN] and DAFSHE [DK]), the CNRS, through the DIST, can thus co-produce studies and recommendations, exchange practices and even coordinate itself, by involving CNRS experts or other institutions. For example, a report on the use of preprints has just been produced by Knowledge-Exchange.

The CNRS, through the DIST and the CCSD, is an active member of the Confederation of Open Archive Repositories, aka COAR, a world confederation of open archives and a key forum for coordinating initiatives to ensure that research maintains control over the information it produces. Interoperability standards for the publication reservoirs of the future are being developed, allowing new services to thrive. The 2019 General Assembly was organised by the CCSD in Lyon.

The CNRS regularly participates in the Berlin conferences on the alignment of strategies to promote open access, which in December 2018 brought together participants from 37 nations and 5 continents (China, South Africa, California, Germany, etc.). A document was produced for the convergence of the Jussieu call for Open Science and Bibliodiversity and the German OA2020 plan. The challenge is to align the criteria for supporting open access journals, publishing platforms and infrastructures, and to coordinate investment efforts at a strategic level. It should be noted that Plan S modified its recommendations to account for the various open access publication channels, particularly under French pressure, and by organising a broad international consultation.

References:

White Paper - Open Science in a Digital Republic, DIST, CNRS (2016):
bit.ly/WhitePaperDigitalRepublic

National Plan for Open Science (4 July 2018):
bit.ly/plannational

Turning FAIR into reality, Final report and action plan from the European Commission
expert group on FAIR data, European Commission (2018):
bit.ly/turningFAIR

The San Francisco DORA Declaration:
sfdora.org/read

Accelerating scholarly communication - The transformative role of preprints
(KE, Sept 2019):
bit.ly/KnowledgeExchangeAccelerating