

CALMIP: Tour d’horizon de la science ouverte.

March 5, 2021

Contents

1	Périmètre et organisation du groupe de travail	2
2	Open Data: Introduction et vue générale	3
3	Interfaces généralistes	4
3.0.1	European Open Science Cloud (EOSC)	4
3.1	data.gouv.fr	4
3.2	OpenAIRE2020	5
3.3	OPERAS	5
3.4	Eudat.eu	6
3.4.1	Rôle	6
3.4.2	Vue générale	6
3.4.3	Points d’intérêt	6
3.5	Zenodo	6
3.5.1	Rôle	6
3.5.2	Vue générale	6
3.5.3	Points d’intérêt	6
4	Formats de métadonnées et protocoles pour le dépôt en science ouverte	7
4.1	Dublin Core Metadata initiative (DCMI)	7
4.1.1	Rôle	7
4.1.2	Vue générale	7
4.1.3	Points d’intérêt	7
4.2	Open Archives Initiative (OAI)	7
4.2.1	Rôle	7
4.2.2	Vue générale	7
4.2.3	Points d’intérêt	7
4.3	Schema.org	8
4.3.1	Rôle	8
4.3.2	Vue générale	8
4.3.3	Points d’intérêt	9
4.4	dataCite	9

4.4.1	Rôle	9
4.5	Research Data Alliance	9
4.5.1	Rôle	9
4.5.2	Vue générale	9
4.5.3	Points d’intérêt	10
5	Interfaces disciplinaires	10
5.1	Astrophysique: IVOA	10
5.1.1	Description de l’IVOA	10
5.1.2	Rôle de l’IVOA	10
5.1.3	Points d’intérêt concernant l’IVOA	10
5.2	Astrophysique: VAMDC	10
5.3	Astrophysique: Europlanet	11
5.3.1	Rôle	11
5.3.2	Vue générale	11
5.4	Données géospatiales	11
5.4.1	Rôle de Open geo spatial Consortium (OGC)	11
5.4.2	Points d’intérêt concernant l’OGC	11
5.4.3	Exemple de service régional	11
5.5	Sciences de la vie	12
5.6	Sciences sociales	12
5.7	Santé	12
6	Les sources d’informations à étudier	12
7	Bilan	13
7.1	Généralités	13
7.2	Éléments pour orienter le travail de ce groupe	13

1 Contenu du document

Ce bref exposé aborde les points suivants:

1. Une cartographie des solutions de dépôt de données en science ouverte
2. L’organisation d’un accompagnement, pour les chercheurs qui le souhaitent, dans le choix d’une plateforme de dépôt adaptée à la nature (discipline, concepts, contenu) et à la destination principale (sujet de recherche, communauté concernée) des données à partager.

2 Open Data: Introduction et vue générale

Avec l’arrivée de grande masses de données scientifiques issues d’expériences, de modélisations ou d’observations différentes, vient la question de la réutilisation de ces données hétérogènes. Le coût humain et matériel de ces données impose d’en assurer la meilleure utilisation possible. Cette problématique n’est

pas nouvelle, et se pose dans tous les domaines scientifiques. Elle est étudiée simultanément par chacune des disciplines concernées, comme par les sciences de l'information qui l'envisagent de façon générale et le plus souvent indépendamment de ces cas particuliers d'utilisation. Elle comprend notamment des questions de référencement de données, d'interopérabilité et de description du contenu et du contexte des données. De multiples initiatives visant à résoudre tout ou partie de ces problèmes ont émergé et continuent de voir le jour.

Du point de vue du W3C, les initiatives liées à des disciplines spécifiques sont des "Vertical Applications"¹. Les vertical applications comme Open Geo Spatial Consortium (OGC²) et l'International Virtual Observatory Alliance (IVOA³), malgré tous leurs mérites (et ils sont grands), contribuent à créer des espaces isolés.

Si c'est efficace dans une optique où il est nécessaire d'avancer dans la résolution de problèmes spécifiques, en revanche cela ne favorise pas l'utilisations de données multidisciplinaires. Il faut en effet appréhender plusieurs ensembles de protocoles, de formats et de descriptions différents. Enfin, les problématiques d'ergonomie (en termes de découverte de données) ne sont souvent pas résolues (par exemple, l'utilisation des outils de l'IVOA peine à se diffuser dans la communauté astrophysique malgré tout le travail engagé, en raison du "coût d'entrée" pour utiliser l'architecture et les outils). Chaque source de données (instrument/expérience/logiciel...) a souvent son propre espace d'archivage, offrant un accès total ou partiel, plus ou moins confidentiel et plus ou moins ergonomique, aux données produites. On parle d'espaces isolés (silos) de données, mais on peut aussi parler d'espaces isolés de formats et de protocoles, puisque les formats et les protocoles utilisés pour décrire et accéder aux données sont souvent différents en fonction des disciplines, mais aussi fréquemment en fonction des sources elles-mêmes dans une discipline donnée.

Il existe donc d'un côté des initiatives "domaine-agnostiques" comme EUDAT et RDA, et d'un autre des initiatives spécifiques à certaines disciplines comme l'IVOA, ou l'OGC. Une famille d'outils transverses qui prendraient en compte les spécificités de chaque initiative disciplinaire mais capables de prendre en charge des données indépendamment de la discipline d'origine semble prometteuse. On pourrait imaginer que ce type d'actions "gère" jusqu'à un certain niveau de détail, puis "laisse la main" à des protocoles et formats plus spécialisés quand la compréhension des données devient trop fine et proche du domaine. Mais comment arriver à une convergence entre la vue scientifique experte d'un domaine particulier et la vue scientifique experte du traitement de l'information en général ? Une telle initiative (dans le domaine de l'information spatialisée) est conduite au sein du JWOC : *"The Spatial Data on the Web Interest Group is a W3C entity matched by a sub-committee of OGC's Technical Committee. Collectively, the two comprise the Joint W3C/OGC Organizing Committee, the JWOC"*⁴.

¹<https://www.w3.org/standards/semanticweb/applications>

²<http://www.opengeospatial.s/org/docas>

³<http://www.ivoa.net>

⁴<https://w3c.github.io/sdw/jwoc/#scope>

Pour couvrir le périmètre posé en point de départ de ce document, il est nécessaire de répondre aux questions :

1. Qu'est-ce qui est standardisé ? Par qui ? Pour quels besoins ? Avec quelles limites ?
2. Qui utilise quels standards ? Pour quelle discipline ?

3 Interfaces généralistes

3.0.1 European Open Science Cloud (EOSC)

EOSC⁵ fait partie des outils listés dans le plan national de la science ouverte, juillet 2018⁶.

Lancement officiel le 23 novembre 2018. La procédure pour rejoindre EOSC est décrite dans la FAQ⁷ au point "services" → 3. Il n'y a pas de détails techniques, mais les grandes lignes. EOSC vise à fédérer les initiatives dispersées comme OpenAIRE, EUDAT ou egi (<https://www.egi.eu/>) sous un portail commun. Le but étant d'éviter que les mêmes outils soient développés plusieurs fois, et également que les services offerts à la communauté soient rassemblés en un seul portail.

3.1 data.gouv.fr

data.gouv.fr⁸ est une plateforme mise en place par le gouvernement Français. Elle permet (après enregistrement) de publier et de récupérer des jeux de données. Il y a une API. Les détails techniques ne sont pas nombreux, il faudra voir si l'API permet de les préciser⁹.

Décliné par portails disciplinaires, par exemple pour des données géospatiales.

3.2 OpenAIRE2020

Open Access Infrastructure for Research in Europe 2020¹⁰. Les recommandations sont de faire un Data Management Plan (DMP), de trouver un annuaire (repository) si possible relevant de la discipline concernée, ou bien un annuaire général et de publier ses données dans cet annuaire en spécifiant quelle licence est utilisée.

Exemple d'exploration de l'annuaire OpenAire :

<https://explore.openaire.eu/search/find?keyword=meteo> Techniquement, utilise openArchive (et Datacite, et d'autres encore peut-être) pour les formats de métadonnées et OAI-PMH comme protocole d'accès.

⁵<https://www.eosc-portal.eu/>

⁶http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.p

⁷<https://www.eosc-portal.eu/about/faqs>

⁸<https://doc.data.gouv.fr/faq/developpeur>

⁹<https://doc.data.gouv.fr/faq/producteur/>

¹⁰<https://www.openaire.eu/>



Figure 1: Les étapes pour être "Open Science Compatible" selon OpenAire

Quels formats de métadonnées sont servis?

http://api.openaire.eu/oai_pmh?verb=ListMetadataFormats

Quels sont les datasets en métadonnées oaf?

http://api.openaire.eu/oai_pmh?metadataPrefix=oaf&verb=ListRecords

3.3 OPERAS

Operas¹¹ est une autre infrastructure Européenne. Elle paraît moins avancée que EOSC et data.gouv.fr, elle semble en cours de construction (<https://operas.hypotheses.org/design-study/ds-technical-mapping>).

3.4 Eudat.eu

3.4.1 Rôle

"The establishment of the EUDAT CDI is timely with the imminent realization of the European Open Science Cloud which aims to offer open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines".¹²

3.4.2 Vue générale

Rassemblement de fournisseurs de données. Pour faire partie de l’infrastructure en tant que fournisseur de données, il faut signer un accord de collaboration (qui à autorité dans un projet pour signer un tel accord?). Il y a deux niveaux d’intégration, interoperable (on partage et on est référencés) et intégré. Intégré

¹¹<https://operas.hypotheses.org/>

¹²<https://www.eudat.eu/eudat-cdi>

→ on entre dans leur infrastructure collaborative y compris au niveau des droits d'accès, on doit fournir une interface commune avec leurs autres services et on doit utiliser leur framework de services.

3.4.3 Points d'intérêt

B2Find, outil de recherche de datasets

Exemple de recherche tomographie: <http://b2find.eudat.eu/dataset?q=tomography>

L'infrastructure utilise le standard OAI/PHM (<http://www.openarchives.org/>)

3.5 Zenodo

3.5.1 Rôle

Zenodo est une interface de partage de données en science ouverte ou à accès restreint.

3.5.2 Vue générale

Zenodo est hébergée au CERN et opérée par le CERN et OpenAire. L'utilisation de Zenodo demande une authentification pour le dépôt de données ou la consultation de données en accès restreint.

Zenodo propose une interrogation et un mode de dépôt de données par l'utilisation d'un API REST comme alternative à l'utilisation de l'interface Web.

3.5.3 Points d'intérêt

L'interface est accessible en ligne¹³.

4 Formats de métadonnées et protocoles pour le dépôt en science ouverte

4.1 Dublin Core Metadata initiative (DCMI)

4.1.1 Rôle

Propose un format de métadonnées.

4.1.2 Vue générale

DCMI ¹⁴ est une initiative déjà ancienne (1995), qui a bien essaimé. Le format de métadonnées Dublin Core est notamment le seul format obligatoire pour pouvoir diffuser des données dans le protocole OAI-PHM (voir ci-après), très utilisé.

¹³<https://zenodo.org/>

¹⁴<http://dublincore.org/>

4.1.3 Points d'intérêt

La description du format: <http://dublincore.org/documents/2008/09/01/dc-ds-xml/>

4.2 Open Archives Initiative (OAI)

4.2.1 Rôle

OAI¹⁵ a pour rôle d'élaborer et proposer des standards d'interopérabilité pour le partage de données.

4.2.2 Vue générale

Les protocoles "Protocol for Metadata Harvesting" (PMH) et "Object Reuse and Exchange" (ORE) de OAI sont utilisés par plusieurs initiatives de partage de données (Eudat,RDA...).

4.2.3 Points d'intérêt

Liste des data providers enregistrés dans open archive initiative :

<http://www.openarchives.org/Register/BrowseSites>

Version XML :

<http://www.openarchives.org/pmh/registry/ListFriends>

OAI-ORE ORE¹⁶ est un des protocoles de OAI. Pour mettre en œuvre la norme¹⁷, il faut monter un endpoint (URL of the OAI provider server on data provider site).

"Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of Web resources. Support for Open Archives Initiative activities comes from the Digital Library Federation, Coalition for Networked Information, and National Science Foundation Grant No. IIS-9817416 (Project Prism)".

OAI-PMH Protocol for Metadata Harvesting (PMH)¹⁸ est un protocole simple, avec seulement 6 opérations.

Pour la mise en place d'un serveur conforme à OAI-PMH, le protocole est décrit ici: <http://www.openarchives.org/OAI/openarchivesprotocol.html#ProtocolFeatures>
Une vue détaillée de OAI-PMH est disponible dans un article publié en 2010[?], qui illustre le fonctionnement du protocole, son lien avec les métadonnées et avec le Linked Open Data (LOD).

¹⁵<http://www.openarchives.org/organization/>

¹⁶<http://www.openarchives.org/OAI/openarchivesprotocol.html>

¹⁷<http://www.openarchives.org/OAI/2.0/guidelines.htm>

¹⁸<https://www.openarchives.org/pmh/>

Des outils pour OAI-PMH BASE¹⁹ est un moteur de recherches d’annuaires de données compatibles OAI-PMH. Il propose une API, une intégration simple de recherches dans une page Web et d’autres facilitateurs pour utiliser le format.

Il existe un validateur de services OAI-PMH²⁰ qui permet de vérifier la conformité d’un annuaire avec le standard en utilisant les opérations prévues par ce dernier. Un autre validateur est disponible sous BASE (<http://oval.base-search.net/>), un peu moins complet.

DSpace²¹ et EPrints²² sont deux serveurs OAI-PMH ”clés en mains”. D’après les retours sur le blog de BASE(<https://www.base-search.net/blog/baseoi/>), DSpace paraît être le plus utilisé.

DataVerse²³ est une autre plateforme de partage de données, plus orientée vers les données scientifiques.

4.3 Schema.org

Format proposé par Google

4.3.1 Rôle

Décrire des jeux de données ou le contenu de pages Web.

4.3.2 Vue générale

Schema.org n’est pas une initiative comme les précédentes. Il s’agit d’un schéma de description de jeux de données ou de contenu de pages Web. Il est discuté au sein du W3C (dans un groupe de discussion), mais il ne s’agit pas d’un standard W3C. Les participants els plus actifs sont Google, Microsoft, Pinterest, Yandex.

4.3.3 Points d’intérêt

En octobre 2018, Google sort la bêta de son moteur de recherches de données, ”Google dataset search engine”²⁴ basé sur la description en schema.org des datasets indexés. On connaît la force de frappe du géant, et le potentiel d’outils annexes qu’il amène avec lui. Schema.org est un format simple à générer, et l’indexation dans de dataset search engine sera automatique à partir du moment où un dataset aura été décrit, et une page Web dédiée comportant cette description sera créée.

C’est certainement une première étape simple à passer, et qui peut déjà apporter un service concret aux utilisateurs pour indexer leurs données dans le

¹⁹<https://www.base-search.net/about/en/index.php>

²⁰<http://validator.oaipmh.com/>

²¹<https://wiki.duraspace.org/display/DSDOC6x/Installing+DSpace>

²²<http://files.eprints.org/>

²³<https://dataverse.org/>

²⁴<https://toolbox.google.com/datasetsearch>

search engine et les rendre disponibles immédiatement (ou presque, le temps que l'indexation automatique se fasse).

4.4 dataCite

4.4.1 Rôle

DataCite²⁵ propose une description généraliste pour un dataset quel que soit son domaine. Un moteur de recherche est disponible sur leur site, par exemple une requête sur les datasets de météo:<https://search.datacite.org/works?query=meteo>
Le schéma lui-même → <http://schema.datacite.org/meta/kernel-4.1/>. Un DOI est obligatoire, l'information fait partie des mandatory fields.

Sur le blog de DataCite, Google dataset Search Engine est accueilli avec enthousiasme (Figure ??)... Même si on peut noter que, bien que des datasets avec DOI peuvent effectivement être retrouvés dans le Gogle datasets search engine, ces DOI *ne sont pas* obligatoires.

4.5 Research Data Alliance

4.5.1 Rôle

RDA²⁶ définit son rôle ainsi: *"The RDA Mission: RDA builds the social and technical bridges that enable open sharing and re-use of data"*²⁷.

4.5.2 Vue générale

Initiative internationale, avec un noeud Français³ récent, depuis mi-2018. On peut rejoindre comme partenaire personnel ou institutionnel. Rejoindre comme partenaire institutionnel, ce n'est pas gratuit. Individual membership gratuit, "mais" attend une vraie implication. A voir selon ce que l'on peut tirer d'une telle collaboration. Peut-être s'intégrer dans le Working Group "Research Data Repository Interoperability WG"?

4.5.3 Points d'intérêt

Switchboard: Connecter de multiples annuaires de datasets → <http://www.rd-switchboard.org/> RD switchboard propose un portail Web pour chercher des données, mais aussi un service RESTful.

Exemple d'utilisation OAI/PMH pour un registry de RDA:

<https://www.ands.org.au/online-services/research-data-australia/rda-registry/oai-pmh-data-provider/oai-pmh-metadata-harvesting>

Liste de standards de métadonnées établie par RDA :

<http://rd-alliance.github.io/metadata-directory/standards/>

Climatologie, météo :

²⁵<http://schema.datacite.org/>

²⁶<https://rd-alliance.org>

²⁷<https://www.rd-alliance.org/about-rda>

<http://rd-alliance.github.io/metadata-directory/standards/cf-climate-and-forecast-metadata-conventions.html>

5 Interfaces disciplinaires

5.1 Astrophysique: IVOA

5.1.1 Description de l'IVOA

L'IVOA²⁸ est le mastodonte de l'astrophysique. Grosse infrastructure. Toutefois, il y a peu de pénétration de ces outils dans la communauté astrophysique (par rapport aux développements engagés). Une des raisons est la difficulté d'accès (ergonomie, au sens large) et le sentiment de "perdre du temps" par rapport à des habitudes déjà installées. On voit facilement l'effort à faire, moins le gain que l'on en attend.

5.1.2 Rôle de l'IVOA

Fournir des formats de description, des protocoles d'accès et des outils de traitement utilisant ces formats et ces protocoles pour les données astrophysiques.

5.1.3 Points d'intérêt concernant l'IVOA

Liste de formats et des docs techniques²⁹.

5.2 Astrophysique: VAMDC

VAMDC³⁰ utilise surtout TAP (protocole IVOA) pour une fédération de bases de données avec portail unique pour l'astrochimie.

5.3 Astrophysique: Europlanet

5.3.1 Rôle

Fournir des protocoles d'accès aux données, une bibliothèque de logiciels³¹

5.3.2 Vue générale

Europlanet-2020³² est un portail destiné à la communauté de planétologie. Il propose des outils et des protocoles d'accès³³ Il utilise par exemple, comme l'IVOA le protocole Table Access Protocol (TAP).

²⁸<http://www.ivoa.net>

²⁹<http://www.ivoa.net/documents/>

³⁰https://portal.vamdc.eu/vamdc_portal/home.seam

³¹<http://planetaryspaceweather-europlanet.irap.omp.eu/tools.html>

³²<http://www.europlanet-2020-ri.eu/research-infrastructure>

³³<http://www.europlanet-vespa.eu/standards.shtml>

5.4 Données géospatiales

5.4.1 Rôle de Open geo spatial Consortium (OGC)

OGC³⁴ est le pendant de l'IVOA pour les données spatialisées. Geoserver³⁵ est un produit issu de l'OGC, qui constitue l'implémentation de référence des recommandations issues de l'OGC. Il permet de récupérer et de partager des données, comme de les utiliser avec des services externes comme Yahoo Maps ou Google Maps.

5.4.2 Points d'intérêt concernant l'OGC

GeoSPARQL permet des requêtes vers des schémas RDF, peut-être intéressant pour des aspects ERV amarrant des raisonnements impliquant des ontologies? *"Catalog Service for the Web (CSW) is a standard for exposing a catalogue of geospatial records in XML on the Internet (over HTTP)"* Les catalogues de données de l'OGC peuvent être référencés en utilisant le "Catalog Service" de l'OGC, qui joue le rôle de OAI-PMH (schématiquement). Pour voir le standard: <http://www.opengeospatial.org/standards/cat>

OS Geo Open Source Geodata foundation GeoNetwork
(<https://www.geonetwork-opensource.org/>) API pour client en JAVA qui utilise Dublin Core et d'autres formats spécifiques géo.

5.4.3 Exemple de service régional

<https://www.crige-paca.org/> pour la région PACA.

5.5 Sciences de la vie

Elixir³⁶ pour les sciences de la vie.
On peut également citer Health Care and Life Sciences (HCLS)³⁷, qui vise à décrire un dataset des sciences du vivant mais supposément "broadly applicable".

5.6 Sciences sociales

CLARIN³⁸ pour les sciences sociales sont des projets liés à Eudat. Ils visent à remplir le rôle d'application verticale dans leur domaine.
Digital Curation Centre (DCC)³⁹ est une organisation Britannique déjà ancienne (2004). Elle a le même point de vue qu'OpenAIRE: Faire un DMP, exprimer ses métadonnées dans un standard et les déposer dans un annuaire struc-

³⁴<http://www.opengeospatial.org/docas>

³⁵<http://geoserver.org/>

³⁶<https://www.elixir-europe.org/about-us/what-we-do>

³⁷<https://www.w3.org/TR/2015/NOTE-hcls-dataset-20150514/#s4%E2%82%82>

³⁸<https://www.clarin.eu/content/clarin-in-a-nutshell>

³⁹<http://www.dcc.ac.uk/>

turé. Pour les sciences de la Terre, un des formats utilisés pour les métadonnées est CIM⁴⁰ CIM - Common Information Model → Décrit métadonnées, outils et méthodes.

5.7 Santé

Health Data Hub⁴¹ est un portail destiné au partage de données de santé. Pour y déposer des données, il est nécessaire de fournir un protocole scientifique, une autorisation de la CNIL, une déclaration d'intérêt public et d'autres documents destinés à assurer la sécurité et l'intérêt des données déposées.

6 Les sources d'informations à étudier

Une équipe d'experts missionnée par le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI) a produit deux documents d'intérêt pour notre groupe de travail. le premier⁴² est une étude comparative des éléments suivants:

- Australian Research Data Cloud (ARDC), Australie
- Dataverse Norway (DNO), Norvège,
- Dutch Archiving And Networking Structure (DANS), Pays-Bas
- Dépôt Fédéré des Données de Recherche (DFDR), Canada
- JISC Open Research Hub (JORH) Royaume-Uni
- Research Data Repository (RADAR), Allemagne
- UK Data Service (UKDS), Royaume-Uni

Le second document⁴³ se focalise sur les besoins des usagers des interfaces de science ouverte.

La mission gouvernementale "Bothorel" a fourni un rapport, qui n'est pas spécifique aux données scientifiques. Ce rapport considère non seulement la question et les enjeux (notamment politiques) de l'ouverture des données, mais également celle de l'ouverture des codes sources de logiciels.

⁴⁰<http://www.dcc.ac.uk/resources/metadata-standards/cim-common-information-model>

⁴¹<https://www.health-data-hub.fr/qui-sommes-nous>

⁴²<https://www.ouvri.lascience.fr/wp-content/uploads/2021/01/Etude-comparative-des-services-nationaux-de-donnees-de-recherche-HC.pdf>

⁴³<https://www.ouvri.lascience.fr/etude-comparative-des-services-nationaux-de-donnees-de-recherche-facteurs-de-reussite/>

7 Bilan

7.1 Généralités

Les protocoles d'accès et les formats de description de métadonnées sont définis par des initiatives transverses indépendantes des domaines d'application (solutions générales), et par des applications verticales qui sont liées à un domaine d'application qu'elles ont vocation à organiser.

Des plateformes "Open Data", "Open science"... se donnent le rôle de fédérer les usages en mettant en commun les pratiques d'où qu'elles proviennent dans le but d'harmoniser les formats et les protocoles, et de proposer des points d'entrée les plus généraux possibles vers les données. Les sections précédentes n'ont pas pour ambition de fournir un état des lieux exhaustif des formats et protocoles existants. Simplement de proposer une vue générale et synthétique. Une étude plus exhaustive a été menée par Jenn Riley⁴⁴.

7.2 Éléments pour orienter le travail de ce groupe

Il apparaît important, lors du recensement des possibilités de partage de données locales à l'université de Toulouse, nationales ou internationales, de les replacer dans le contexte général brossé dans ce document.

Il sera également nécessaire de placer les possibilités par rapport aux besoins et aux contraintes des utilisateurs, telles qu'exprimées notamment dans le second document du MESRI.

⁴⁴<https://datasupport.researchdata.nl/fileadmin/content/Essentials4DataSupport/Documents/seeingstandards.pdf>