

## Use Cases – 29 July 2015

Updated set following the first meetings of all three Advisory Groups on 23 July 2015.

Note: the user stories were initially collected from the project's first workshop. These were collated, merged and categorised to the set used for the Advisory Group meetings. Following this meeting, these will now become the more detailed use cases from which an initial set of requirements will be extracted. The terms user stories and use cases have been used interchangeably, although this is incorrect from an Agile perspective. From now on these will be referred to as use cases.

Whereas the original set were ordered numerically this version is ordered by category. They have been updated according to the Advisory Group actions and comments in the Google Doc. Some of the comments have been added to the Notes field.

### Categories

#### Admin

<b>Use Case ID</b>	UC19		
<b>Title</b>	Update platform software		
<b>Category</b>	Admin		
<b>Owner</b>	TBC		
<b>Description</b>	As a developer maintaining the registry, I want to be able to <b>quickly</b> and <b>easily</b> apply platform updates with minimal disruption and development effort so that the system can remain secure, bugs can be dealt with and new features taken advantage of.		
<b>Notes</b>	CKAN is an ongoing open source platform. ANDS would need to be forked, if selected for the registry, so would require significant effort to include any future updates to the ANDS code. Although a design factor this use case has been kept to ensure the ease of updating the service is included as an evaluation criteria.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Developer	Administrator	Technical User

#### Browse

<b>Use Case ID</b>	UC4		
<b>Title</b>	Make Connections with other data and researchers		
<b>Category</b>	Browse		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher, I want to discover related objects and resources (e.g. via a recommendation service)		
<b>Notes</b>	See UC2 This is not an Amazon-like recommendation service. It shows information related to the search criteria. This requirement will require further clarification during development. This would be useful from a visual research perspective, as much searching and browsing is serendipitous. The University of Creative Arts Library is investigating this for their Library catalogue.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>

Could	Researcher		
-------	------------	--	--

<b>Use Case ID</b>	UC18		
<b>Title</b>	Browse registry information		
<b>Category</b>	Browse		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher, I want to be able to browse through data so I can find records I was not aware of.		
<b>Notes</b>	<p>This should include filtering by data type (e.g. tabular data, sound, bitmap).</p> <p>Not all repositories make use of "data type": datasets, sound recordings, video, software, etc. Therefore it's difficult to find "data" as opposed to publications (especially for IRs that hold both). Need to agree to use a form of this metadata field, browsing could be enabled by these types of data.</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Researcher		

## Compliance

<b>Use Case ID</b>	UC15		
<b>Title</b>	National reporting to funders (eg RCUK, OpenAIRE)		
<b>Category</b>	Compliance		
<b>Owner</b>	John Kaye		
<b>Description</b>	As a national organisation, I want to be able to select datasets by characteristics (e.g. funder, publication) to report to international (EU) funding, e.g. provide a national OpenAIRE return for all UK institutions. These need funding and publication and compliance with OpenAIRE metadata standards (DC 3.0 or CERIF-XML (as published by OpenAire/euroCRIS))		
<b>Notes</b>	<p>Potential UC: As a sysadmin, I want to report broken links, etc</p> <p>Potential UC: Generate registry contents analysis. Details of the use case can define what filters can be applied by the user, and what can be reported upon. "Metric collection" use case (need to define what the metrics are e.g. number of views, visits, etc)</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Should	Project Manager	Funder	

<b>Use Case ID</b>	UC26		
<b>Title</b>	Certify system is compliant to the RDDS		
<b>Category</b>	Compliance		
<b>Owner</b>	TBD		
<b>Description</b>	As a system supplier/vendor/repository manager I want to certify my system is compliant to the RDDS (requires a set of standards, an indication of minimal/optimal mandatory/optional metadata elements and their effect on discovery).		
<b>Notes</b>	The solution can't do this (though it can support it). The use case to be satisfied is probably something like "access standards & operational documentation". (Which is more generic - lots of different users need		

	<a href="#">access to documentation/training etc. for various purposes</a>		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Could	System Supplier/Vendor	Repository Manager	

### Evaluate/use results

<b>Use Case ID</b>	UC2		
<b>Title</b>	Understand metadata quality (promote data sharing quality and practices)		
<b>Category</b>	Evaluate/use results		
<b>Owner</b>	Veerle Van den Eynden		
<b>Description</b>	As a researcher (or other user), I want to understand from the RDDS record the quality of the metadata associated with the dataset and encourage researchers to document data to a high standard and deposit it for reuse, <del>to improve impact and REF rankings</del> (remove this text as incorporated into UC28).		
<b>Notes</b>	see also UC28 <a href="#">Data quality aspects are now in UC44.</a> <a href="#">UC5 has been merged into this UC.</a> <a href="#">Provide feedback to data centres re quality of each harvested record?</a> <a href="#">What standards to employ?</a> <a href="#">Are different quality standards applied to different resource types (data vs other research outputs?)</a> <a href="#">How do you represent "not assessed?"</a> <a href="#">Or take an ISO-19115 style approach (cite the assessment, the standard assessed, and the outcome)</a>		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Should	Researcher	Research Manager	Repository Manager

**Commented [CB1]:** Not convinced need a category to include what might be done with the search results. These could all be included within the Search category.

<b>Use Case ID</b>	UC44		
<b>Title</b>	Understand data quality		
<b>Category</b>	Evaluate/use results		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher (or other user), I want to understand from the RDDS record the quality of the dataset, how the dataset has been packaged, whether it is in a form that will enable me to reuse it, and whether there is published literature that describes the dataset and draws conclusions from it.		
<b>Notes</b>	see also UC28 <a href="#">Need to consider what information the user needs in order to successfully (and efficiently) follow quoted URLs to the various resources.</a>		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Should	Researcher		

<b>Use Case ID</b>	UC36		
<b>Title</b>	Data and metadata provenance		
<b>Category</b>	Evaluate/use results		

<b>Owner</b>	TBD		
<b>Description</b>	As a researcher I want to realize the origin of data and metadata published		
<b>Notes</b>	<p>Added by Vasily Bunakov</p> <p>"INSPIRE (for example) has the concept of "lineage" which is typically implemented as just a free text field, but can be used in a more sophisticated way to hold the chain of derivation in terms of original sources, intermediate states and the process steps that were followed in creating the resource as it is now. How much detail is needed? I wonder how many providers actually have this sort of information."</p> <p>Difficult to show provenance as not part of metadata.</p> <p>Out of scope.</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Won't	Researcher	Repository Manager	

<b>Use Case ID</b>	UC38		
<b>Title</b>	Network characteristics		
<b>Category</b>	Evaluate/use results		
<b>Owner</b>	TBD		
<b>Description</b>	As a networking researcher I want to find datasets that contain mobility or movement of particular nodes in a network. This can be people, or animals, or nodes in a graph. So I need to look for coordinate information within the data themselves.		
<b>Notes</b>	<p>See also UC29</p> <p>Added by Vasily Bunakov</p> <p>Very specific search criteria. More about mining the underlying data than what might be in the metadata. Out of scope for the registry to set to Won't</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Won't	Researcher		

### Impact/Analytics

<b>Use Case ID</b>	UC9		
<b>Title</b>	Data centre impact		
<b>Category</b>	Impact/Analytics		
<b>Owner</b>	Jim Halliday (STFC/UKERC)		
<b>Description</b>	As a Data Centre Manager, I want to promote the existence of datasets I hold and data links I know about, for the benefit of the UK research community – particularly for cross-disciplinary research.		
<b>Notes</b>	<p>Similar to UC28 but from a specific Data Centre perspective</p> <p>Merging with UC1, UC8 and UC12 was an option but keeping separate as each one relates to the needs of specific stakeholders.</p> <p>Need a clear definition of what is meant by "promote". Does this include sharing results or highlighting them within the registry?</p> <p>"External Data Centre Manager" is a manager from other data centre who wants to cooperatively publish their own datasets, or compare practices or performance of different data centres</p> <p>"Information Scientist" could be someone who evaluates the research impact using specific techniques like bibliometrics, i.e. for whom the</p>		

	techniques (not the impact per se) are of a professional interest		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Data Centre Manager	External Data Centre Manager	Information Scientist

<b>Use Case ID</b>	UC12		
<b>Title</b>	Funder return on investment		
<b>Category</b>	Impact/Analytics		
<b>Owner</b>	Christopher Brown / Veerle Van den Eynden		
<b>Description</b>	As a funder / programme manager, I want to discover all datasets (and relevant statistics) that I have funded so that I can demonstrate / evaluate return on investment.		
<b>Notes</b>	Can we show ROI? Should it be "find resources by funder identity"? Similar to UC15		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Funder		

<b>Use Case ID</b>	UC22		
<b>Title</b>	Library return on investment		
<b>Category</b>	Impact/Analytics		
<b>Owner</b>	Christopher Brown / Veerle Van den Eynden		
<b>Description</b>	As a librarian, I want to use analytics that show the value of my dataset. (Help with impact. Aggregation level).		
<b>Notes</b>	Similar to UC15 but for different stakeholder. See UC15 notes on ROI. Could be part of the general reporting use case (UC12 / UC15). More of a business benefit to be demonstrated once the system is operational.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Librarian		

<b>Use Case ID</b>	UC28		
<b>Title</b>	Research impact		
<b>Category</b>	Impact/Analytics		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher, I want to list the datasets/papers derived from my dataset so that I can demonstrate impact for REF and grant applications (via DOIs)		
<b>Notes</b>	Combined UC13+14 "Research Impact" may make more sense than "Researcher Impact", taking into account that research is often cooperative (multi-authored) and that a researcher conducts multiple experiments/investigations. So this is in fact "many-to-many" relationship, and what can be (potentially) directly measured is impact of a research (a paper, a dataset); researcher impact is always "synthetic"/derivative - obtained through certain aggregations or averages of research impact(s). What does "derived" mean? Does this only mean papers that use data or new datasets that are transformations of the data? Or does it include papers that cite data (e.g., as examples of data in a field), or data that have been collected using the same methodology as a paper (i.e.,		

	reproducibility vs repeatability). May need another Actor (e.g. a Thomson Reuters or similar web service to query DOI impact). You probably wouldn't want to scope in-house the development of solutions that T-R already spend a very large amount of money on....!		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Researcher	Information Scientist (see UC9)	

## Policy

<b>Use Case ID</b>	UC5		
<b>Title</b>	Promote data sharing quality and practices		
<b>Category</b>	Policy		
<b>Owner</b>	Veerle Van den Eynden		
<b>Description</b>	Encourage researchers to document data to a high standard and deposit it for reuse, to improve impact and REF rankings		
<b>Notes</b>			
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Should	Research Manager	Repository Manager	

Commented [CB2]: Incorporated into UC2.

<b>Use Case ID</b>	UC6		
<b>Title</b>	Ease of data harvesting		
<b>Category</b>	Policy		
<b>Owner</b>	David Wilson		
<b>Description</b>	To make it easy for a repository / data centre to provide data to encourage contributions		
<b>Notes</b>	Action: Replace with UC33, but retain 'ease' aspect in evaluation criteria. Usability is different for different people. Need to engage with a range of users. If ease of harvesting is critical, need to set quality criteria to assess against, and design a testing/ongoing monitoring regime to prove that criteria are being met.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Repository Manager	HEI	Data Centre

Commented [CB3]: Incorporated into UC33

<b>Use Case ID</b>	UC30		
<b>Title</b>	Metadata rights/ownership		
<b>Category</b>	Policy		
<b>Owner</b>	Veerle Van den Eynden		
<b>Description</b>	As a data centre / repository I want to be sure that publishing metadata to RDDS and how such metadata are licensed respects the rights / licensing of my metadata, in particular when they are provided by/copyright of data depositors. Investigate how metadata can be optimally shared to the RDDS.		
<b>Notes</b>	This use case is similar to the one added by VADS - UC41. Keeping separate as this use case is specific to the licensing issues around the metadata. UC41 is about ensuring there is rights information about the underlying data held in the metadata.		

	Although encourage metadata to be OA, there is the use case at VADS where there are licensing issues with the metadata so this use case will look at how this can be resolved in the registry.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Data Centre		

Use Case ID	UC41		
Title	Copyright and licensing, including third party materials		
Category	Policy		
Owner	Leigh Garrett/Amy Robinson (VADS)		
Description	As a visual researcher I want to know the copyright and licensing arrangements for data, including third party visual materials; as a data provider I want to ensure that rights information is clear in order to easily enable re-use where possible and to protect materials from unauthorized uses.		
Notes	Relates to UC30 – perhaps could be merged Added by VADS. Keeping separate. See UC30 notes.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Researcher	Data Centre	

## Search

Use Case ID	UC1		
Title	Find Research Outputs from my institution		
Category	Search		
Owner	Laura Molloy		
Description	Complete list of all data from my institution		
Notes	<p>Could merge with UC8, UC9, UC12?</p> <p>Need a clearer definition of what is meant by "data" in this context. Clarify role of institution in this use case: as data host, as affiliation...</p> <p>Would this include data from repositories such as NBCI, EMBL etc? Only data from participating HEIs/DCs at this stage.</p> <p>Part of a more generic search/browse? What do we mean by "from"? data.gov.uk has (still has) trouble with the differences between "harvest source", "publisher of the metadata", and "publisher of the resource". In a federated set up these things can be very significant.</p> <p>In Datacite terms need to be able to search for institution as Publisher (of the data), institution as affiliation of Creator or Contributor, institution as source of metadata. (Unlike for Datacite, I would like to see affiliation of Creator and Contributor as mandatory)</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Should	Research Manager	Researcher	

Use Case ID	UC8		
Title	Research data promotion		
Category	Search		
Owner	Jim Halliday (STFC/UKERC)		
Description	As a researcher, I want to be able to publicise the existence (and		

	location) of datasets I've created.		
<b>Notes</b>	<p>Find included in UC29</p> <p>Could merge with UC1, UC9, UC12?</p> <p>"External Researcher" is a researcher who did not create the datasets but is interested in them.</p> <p>What does "publicise" mean? How would you assess whether or not the promotion use case is being satisfied? (related to UC7 and UC28).</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Researcher	External Researcher	

<b>Use Case ID</b>	UC20		
<b>Title</b>	Find research outputs by geospatial location		
<b>Category</b>	Search		
<b>Owner</b>	TBC		
<b>Description</b>	As a multidisciplinary researcher, I want to be able to specify a geographical area, and find all the available data about that area, so I can look for correlations between phenomena investigated by different disciplines.		
<b>Notes</b>	<p>Andy Turner – "I am interested in the implementation detail of this. Piping or joining queries together should be possible to narrow down a search and perhaps not just spatially. Perhaps as well as parametrising a spatial location, some temporal bounds could also be set as well as some way of narrowing down by other attributes. For example, can I search for all recorded interview data for interviews that took place in the UK since 2012 and that contain the phrase "referendum"? This could be related to UC3."</p> <p>Peter Vodden - "Could be part of a wider search/browse use case (UC18). Agree with Andy - spatial and temporal searches merit special consideration. A "multi-disciplinary" researcher may or may not be a special type of Actor - I suspect not.</p> <p>EDINA's UNLOCK (middleware) service can be used do this. An example is ShareGeo, in which a user can draw a bounding box around a map both for metadata input and for searching - crosswalked intelligently to placenames, geo areas, etc.</p> <p>Review priority – not all datasets will have geospatial metadata available.</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must/Should	Researcher		

<b>Use Case ID</b>	UC25		
<b>Title</b>	Find datasets I can learn from		
<b>Category</b>	Search		
<b>Owner</b>	TBD		
<b>Description</b>	As a researcher, I want to find exemplar datasets as inspiration on how to do my research		
<b>Notes</b>	<p>See UC29</p> <p>Note: a) find exemplar dataset as inspiration on how to do your research b) metadata includes links to datasets such as dissertations and theses (not filtered out).</p>		

	Does not include inclusion of records for student-created datasets. If this means making it possible to find datasets that are a good example of how to go about doing something, this should not be categorised student / learner - any researcher might be interested in or need to find this information. Could be part of a wider search/browse use case (UC18) - unless there are specific search requirements above and beyond the general case.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
	Researcher		

<b>Use Case ID</b>	UC29		
<b>Title</b>	Find research outputs by targeted search - existing/non-existing data		
<b>Category</b>	Search		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher, I want to search for data in my research area and only my area, or specific data related research, or data linked to a specific researcher, so that I can quickly find relevant results or confirm that there isn't already data out there, or at least in the RDDS system, that would suit my needs. As a funder, I want to ensure that data an applicant plans to collect doesn't already exist. (Policy and guidance available).		
<b>Notes</b>	Combined UC8 +UC11+UC16+UC21+UC23+UC24 Difficult to prove non-existence of data. This use case allows users to search for specific research so they can show they've made the relevant search and not found any results. This proof is useful in grant applications. Semantic-enabled solution needed if "my research area" search is to be effective - how will you specify your research area? Is there an agreed ontology? Maybe the metadata publisher used different concepts, different words, or synonyms?		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Researcher	Funder	

<b>Use Case ID</b>	UC37		
<b>Title</b>	Data anonymisation		
<b>Category</b>	Search		
<b>Owner</b>	TBD		
<b>Description</b>	As a security researcher I want to find datasets that have been anonymised or sanitised in a particular fashion so that I can attack or improve these anonymisation schemes.		
<b>Notes</b>	See also UC29 Added by Vasily Bunakov This is a search for specific information in the research data and would, therefore, require data mining. This is not a function of the registry, which can only provide searches on the information contained in the metadata. Mary Donaldson - "You could argue that is (a function of the registry). Security research is a valid academic field. Researchers in that field should be able to search for datasets to use in their research. Whether		

	this needs to be a distinct use case is another consideration.” Out of scope.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Won't	Researcher		

<b>Use Case ID</b>	UC39		
<b>Title</b>	Data collection methodologies		
<b>Category</b>	Search		
<b>Owner</b>	TBD		
<b>Description</b>	As a researcher I want to find data that were collected using a particular methodology, or had particular problems during the collection of a dataset.		
<b>Notes</b>	See also UC29 Added by Vasily Bunakov. Like UC37 this is searching for specific information held in the research data and unlikely to be in the metadata. Out of scope in first version of registry.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Won't	Researcher		

<b>Use Case ID</b>	UC43		
<b>Title</b>	Visual searching		
<b>Category</b>	Search		
<b>Owner</b>	Leigh Garrett (VADS)		
<b>Description</b>	As a creative arts researcher I want to find data using a visual rather than text-based search.		
<b>Notes</b>	Added by VADS.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Researcher		

### Syndicate (or Machine access)

<b>Use Case ID</b>	UC3		
<b>Title</b>	Linked Open Data (Semantic Web)		
<b>Category</b>	Syndicate (or Machine access)		
<b>Owner</b>	TBC		
<b>Description</b>	As a researcher, I can interrogate and find data using semantic web technologies (SPARQL/RDF) This relates to making the records (metadata) available in RDF form.		
<b>Notes</b>	Andy Turner – “I am interested in this. Could I find all verbatim transcribed interviews that have audio recordings? I think that in some ways this could be linked with a geospatial search using GeoSPARQL ( <a href="http://geosparql.org/">http://geosparql.org/</a> ). So this could be related to UC20.” Peter Vodden – “Using semantic web technologies - OK, but the technologies won't deliver benefit unless a critical mass of resources have been semantically marked up/enriched in a consistent manner using standard ontologies/vocabulary keywords. Who is going to do the markup? Or are we expecting that some form of advanced data-mining will "generate" the semantics (UC35)? Maybe need to better define the		

	initial aspiration? (To manage expectations)”		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Could	Researcher	Machine	Technical User

Use Case ID	UC7		
Title	Indexing data to external services		
Category	Syndicate (or Machine access)		
Owner	TBC		
Description	As a repository/data holder, I want to have metadata indexed by Google, etc. so I don't have to work to get the records indexed myself.		
Notes	<p>Pauline Ward – “esp Google Scholar, and any other popular current awareness services, and for as many metadata fields as possible to be picked up by those services. Or else if good integration is not possible, for the registry to have its own alerting service.”</p> <p>Peter Vodden – “Some clarity required - are we simply saying that the records are to be discoverable by web spiders (fairly trivial to establish and prove - records need to be exposed as web URLs - a design requirement), or are we specifying search engine optimization (SEO) criteria (e.g. page rankings) that must be met for certain types of searches? (complicated and potentially expensive).”</p> <p>This is not about Search Engine Optimization (SEO), but simply making the records discoverable by web spiders.</p>		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Must	Research Manager	Repository Manager	

Use Case ID	UC34		
Title	Harvestable Registry		
Category	Syndicate (or Machine access)		
Owner	TBD		
Description	As a developer I want to harvest metadata from the registry so that it can be reused or aggregated.		
Notes	What formats will be offered? HTML, RDF, JSON, CERIF, etc. Optional schemas? (e.g. Dublin Core?)		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3
Must	Machine	Developer	

Use Case ID	UC35		
Title	Text and data mining		
Category	Syndicate (or Machine access)		
Owner	TBD		
Description	As a developer I want to perform text mining over metadata or data mining over data assets (that metadata designate and point to). A specific case of data mining could be automated dataset characterization (defining metadata by a machine agent looking into data)		
Notes	<p>Added by Vasily Bunakov.</p> <p>In this use case, the RDDs would be open to text mining but not provide any text mining functionality itself.</p>		
MoSCoW	Stakeholder1	Stakeholder2	Stakeholder3

Commented [4]: This use case was not in the Christopher's list; I've added it for consideration

Could	Machine	Developer	
-------	---------	-----------	--

### Technical

<b>Use Case ID</b>	UC31		
<b>Title</b>	Show data services		
<b>Category</b>	Technical		
<b>Owner</b>	Veerle Van den Eynden / Peter Vodden		
<b>Description</b>	I want the RDSS to also harvest data services related to datasets		
<b>Notes</b>	<p>Peter Vodden – “Search/browse for Services is an extension of the general search/browse use case.</p> <p>There is a special use case which is "find services available for this dataset" (or vice versa - what datasets does this service operate on?) Which could be a specific application of a more general requirement to relate various resources together e.g:</p> <p>Resource A supersedes (is superseded by) resource B  Resource C has child resources "parts" D, E, F.  etc., etc..”</p> <p>Requires clarification.</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Data Centre		

**Commented [CB5]:** “Technical” category kept for now. If removed need to move use cases to more relevant categories or create new ones.

<b>Use Case ID</b>	UC32		
<b>Title</b>	Duplicate records		
<b>Category</b>	Technical		
<b>Owner</b>	TBC		
<b>Description</b>	I want to ensure that no duplicate records show in the RDSS, should they be harvested from different sources (compare unique identifiers) and different versions of a dataset are uniquely identified. Two parts - no duplicates, show multiple versions.		
<b>Notes</b>	What reports should the system produce to show any duplication or versions of the metadata?		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	System Manager		

<b>Use Case ID</b>	UC33		
<b>Title</b>	Harvest datasets		
<b>Category</b>	Technical		
<b>Owner</b>	TBD		
<b>Description</b>	As a System Manager, I want my metadata records harvested into a registry so that they can be discoverable.		
<b>Notes</b>	<p>Incorporates UC6 (Ease of data harvesting)</p> <p>Need to ensure that the system is designed to be easy to use. Usability is different for different people. Need to engage with a range of users in HEIs and Data Centres.</p> <p>Needs to include the RDSS reporting back to the source provider if there are errors/issues with the harvest.</p>		

	Rowena Rouse – “We use Equella as our (Oxford Brookes) repository software and I'd like to make our metadata harvestable by UKRRDS and then discoverable by Resource Discovery engines so happy to own UC33 if can focus on Equella. Can we learn from current data archives?”		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	System Manager	Administrator	Repository Manager

<b>Use Case ID</b>	UC40		
<b>Title</b>	Presentation of visual materials		
<b>Category</b>	Technical		
<b>Owner</b>	Leigh Garrett/Amy Robinson (VADS)		
<b>Description</b>	As a visual researcher I want to browse search results visually (e.g. thumbnails) rather than just text-based search results.		
<b>Notes</b>	<p>Added by VADS.</p> <p>Amy Robinson – “For the last 4 use cases from VADS (UC40-43): requirements from visual researchers to search/browse/access/re-use data can be very different to other researchers - some of these we are aware might be a 'nice to have' for this project like the visual search (UC43) - but the UC40 and UC 41 are more essential from our perspective. We have labelled accordingly.”</p>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	Researcher		

<b>Use Case ID</b>	UC42		
<b>Title</b>	Visual tools		
<b>Category</b>	Technical		
<b>Owner</b>	Leigh Garrett/Amy Robinson (VADS)		
<b>Description</b>	As a visual researcher I want appropriate tools for my needs (e.g. a lightbox to save particular visual materials, ability to compare images side by side)		
<b>Notes</b>	Added by VADS.		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Could	Researcher		

### Update

<b>Use Case ID</b>	UC17		
<b>Title</b>	Force refresh of registry information		
<b>Category</b>	Update		
<b>Owner</b>	Veerle Van den Eynden / TBC		
<b>Description</b>	<p>As a data centre, I want to be able to trigger re-indexing/harvesting of my data holdings or a subset of my data holdings so that the RDDS is correct and up to date; in particular deletion of datasets at source repository must be actioned rapidly in the RDDS.</p> <p>As a system manager, I've published information I shouldn't have and need it removed.</p>		
<b>Notes</b>	Peter Vodden – “Data centres may expect to have both scheduled harvest, and re-harvest on request.”		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>

**Commented [CB6]:** Edit and Update categories should be merged

Must	Data Centre	Repository Manager/HEI	System Manager
------	-------------	------------------------	----------------

**Edit**

<b>Use Case ID</b>	UC27		
<b>Title</b>	Remove information from the registry		
<b>Category</b>	Edit		
<b>Owner</b>			
<b>Description</b>	As a system manager, I've published information I shouldn't have and need it removed.		
<b>Notes</b>	<a href="#">Merged into UC17</a>		
<b>MoSCoW</b>	<b>Stakeholder1</b>	<b>Stakeholder2</b>	<b>Stakeholder3</b>
Must	System Manager		

**Commented [CB7]:** No use cases under this category so remove. Covered by "Update".