Org-ID Use Cases

Introduction

This document summarizes the target use cases for the Org-ID study. Use Case owners (and others on the working group) are asked to first complete a narrative description of their use case below, being as specific as they can on the exact process involved. The CASRAI team will then parse these narrative descriptions for follow-on tasks.

CASRAI use cases are not the traditional, fully detailed, use cases for software development. They are simpler narratives more akin to user stories and they focus specifically on an interchange or interoperability step in a business process.

As defined in Wikipedia a user story is "*one or more sentences in the everyday or business language of the end user or user of a system that captures what a user does or needs to do as part of his or her job function. It captures the 'who', 'what' and 'why' of a requirement in a simple, concise way, often limited in detail by what can be hand-written on a small paper notecard.*"

A typical template for such user stories is:

"As a <role>, I want <goal/desire> so that <benefit>"

Some examples (from Wikipedia):

- **Search for customers**: As a user, I want to search for my customers by their first and last names.
- **Modify schedules**: As a non-administrative user, I want to modify my own schedules but not the schedules of other users.
- **Run tests**: As a mobile application tester, I want to test my test cases and report results to my management.

Use Case Summaries

OrgID-UC1 - Researcher - applying for funding

As a researcher applying for funding, I need to list multiple organisations related to my proposal in order to enable the target funder to uniquely identify previous employers and other funders, collaborators or industry partners and beneficiaries.

WG Owner: Working group

Component Parts: For application form (presume online):– identify previous employers, projects and funders; identify affiliation of previous collaborators/co-authors; identify specific intended beneficiary organisations. OrgIDs should make it easier for the researcher as producer of this information by reducing time spent in collecting, collating and then manually entering data. They should make it easier for the funders as consumers of this information to gather and analyse this data with minimal manual intervention.

Example 1.1: I am a post doc researcher, have been employed on research projects in my department before but now looking at being a PI in my own right for the first time. I need to look at the requirements of the funder - do they require a European or international partner and do X and Y organisations satisfy the criteria? Having identified a funder, during the application process I need to uniquely identify my past employers, my doctorate awarding institution, current and past affiliations of my international collaborators and my possible referees' affiliations. My research is in the area of clinical practice and I want to use OrgIDs to identify the particular health organisations where I expect this work will have impact. This will require international coverage for the OrgIDs.

Example 1.2: I am a member of staff at a funder. I need to analyse the data submitted by the researcher during the application process to access past record of all applicants' organisations in managing and making use of previous funding. I need to check intended beneficiaries against the stated aims of this funding call and our wider strategic aims. I need to ensure that all the collaborating organisations qualify to be members of this consortium and whether they are eligible for funding. If the application process requires (or encourages) the researcher to submit OrgIDs then the analysis can be automated to some extent.

OrgID-UC2 - Funder - minimising conflicts of interest

As a funder preparing to find referees or reviewers, I need to be able to identify suitable people in order to minimize conflicts of interest (through potential co-location at host institution).

WG Owner: Kevin Dolby

Component Parts: Funder staff need to identify suitable people to review funding applications. As well as ensuring that reviewers have an appropriate level of expertise in the

relevant subject area, it is important to ensure that there are no potential conflicts of interests. Part of this process will be ensuring that reviewers are not based at the same institution as the applicant, or at least in a completely separate part of the institution. Correctly identifying the institution, down to an appropriate level of faculty/department, is a key requirement.
Example 2.1: I am a member of staff at a funder. In finding national and international reviewers and referees, I will use internal and external systems and informal advice networks and colleagues. I need to be able to check the current and past affiliations of such referees and check the relationships between those organisations and the organisations mentioned as current and past affiliations by the principal applicants in the funding request. Where there is a match, it would be very useful to be able to quickly recognise how, for example, two parts of a large organisation are related or sit within the "tree" of the large organisation.

OrgID-UC3 - Funder - tracking published outputs

As a Funder, collating outputs in end-of-research reports, I need to be able to track published outputs in order to understand our contribution & successful collaborations.

WG Owner: Muriel Mewissen

Component Parts: Identify affiliations of all authors; Identify funder(s) – who funded the research; Identify funder(s) – who paid for any APC; Identify data steward (often an organisation rather than a person).

Example 3.1
I am a member staff at a funder. I want to acquire information on which publications have resulted from work partly or wholly funded by us and which funding calls/streams produced this work. Where work has been funded from more than one source, I wish to assess which funding collaborations have been the most productive. If authors submit funder and employer information on the form of OrgIDs this will allow me to gather this information accurately and with minimal manual intervention. I will be consuming the data produced by the author in 3.2 and provided by the publisher in 3.3.
Example 3.2: I am an author, part of an international team, working on a funded project. I wish to publish in a journal and clearly identify the funding bodies, my institution and the institutions of my collaborators. I want to ensure that the publisher recognises the institution so that any bulk deals on APC charges are honoured and any special deals regarding publication embargoes are recognised. I want to ensure the publisher recognises the funder(s) so that any Open Access (Green or Gold) conditions imposed by the funder are honoured. I want to clearly identify within the article the organisation hosting the datasets so that potential users can easily access both the data and their conditions of use. I will benefit from using OrgIDs by avoiding repetitious multiple manual input of the same data, by receiving correct credit from funders for my research outputs and by avoiding errors in inputting key organisation data (e.g. mis-spelling Wellcome Trust).

Example 3.3: I work at a publisher. I also consume the author's data produced in 3.2 but have an intermediary role producing the data which funders and others will consume. One of my most time consuming tasks is ensuring that data supplied by authors on affiliation (often multiple authors and multiple affiliations) and funding (sometimes multiple and internationally distributed funders) is complete and accurate and compliant with funders' requirements. Widespread use of OrgIDs will save time and increase accuracy.

[Org-ID- UC4 deleted, numbering preserved to prevent confusion]

OrgID-UC5 - Researcher or research manager - reporting academic impacts to funders

As a research producer, I need to report academic impacts to different funders with different requirements.

WG Owner: Huw Charles

Component Parts: Identify co-author affiliation; Identify destinations of staff; Identify impact beneficiaries.

Example 5.1: I am a senior researcher responsible for reporting progress, success and impact of past funding. One key measure is the career progress of junior researchers previously employed on my funded projects. Using a combination of ORCID and OrgID, I need to accurately identify primary and secondary destinations of previous staff and doctoral students.
I also need to report further uses of the dataset(s) we have created by other UK and international research organisations. I need to include other "translational" research bodies, eg NGOs, NHS, government advisory bodies, European bodies, even start ups who are impact beneficiaries. I need to report to the REF the impact of the research environment etc. Increasingly there are (or will be) evidential requirements. Career paths and destinations will require international coverage.
Example 5.2: I am a member of staff at a funder. It is important for me to monitor the impact of past funding. Our funding can have impact nationally and internationally and I need to identify beneficiary organisations on and the impact on research careers and destinations of researchers. I am "the consumer" of the information submitted by the researcher in example **5.1**

OrgID-UC6 - Researcher - tracking organisations across time

As a researcher I need to preserve the historical integrity of organisational names at the time of data creation, collection or deposit (and other, specified times); it is similarly important, however, to record and retain the links between these differing names, so that any user can see which data came from which organisation, even if the organisation name has changed.

WG Owner: Lucy Bell

Component parts: capture and preserve the historically accurate organisation name associated with a data collection or, for example, a patent application, at specified times in the lifecycle of the data; link multiple organisation names with the single entity; map mergers, splits, takeovers and reincarnations of organisations; linking disparate names of the same organisation may require transparent and simple alias creation or, as with authority lists or controlled vocabularies, the association of one concept/organisation with multiple alternative labels (as in SKOS).

Example 6.1: I am a data archivist / librarian. I need to capture and preserve the identity of an organisation at the time of data collection and I also need to be able to link to up to date information on that organisation's changes of identity in the future without polluting or changing the original information.

Example 6.2: I am data user and wish to search for, use and cite a data collection, using the historically correct affiliation information. These historical names are crucial in data citation and reference lists.

Example 6.3: I am creating a search application for data resource discovery. I need to set up systems to link the multiple incarnations of organisations' names, so that a user searching for one incarnation of an organisation will also see results associated with one of its aliases.

Example 6.4: I am a data creator and wish to link all my correctly and accurately cited data collections. These data collections will sometimes be associated with organisations I used to work in, whose names have now changed, or indeed with my current organisation, whose name has changed. I need the citations to remain historically accurate while the organisations' provenance is clear.

OrgID-UC7 - Repository manager - populating repositories, managing automation

As a repository manager I need to be able to uniquely identify my repository, whether or not its location or URL changes; this will enable me to control semi-automated population of repository records.

WG Owner: Pablo de Castro & Muriel Mewissen

Component Parts: Creating a persistent identifier for a repository that is suitably associated with (but not the same as) the unique ID of the organisation responsible for the repository.

Example 7: The Jisc Publications Router (JPR) http://broker.edina.ac.uk/ is a tool for automatically distributing a set of inputs (usually journal articles and ideally delivered by publishers in a metadata + digital object form) onto the repository network. Its main conceptual aim could be to offer content providers (publishers) the possibility of helping their authors to comply with funder mandates by automatically depositing the final manuscripts into their institutional repositories. EDINA have developed the Organisation and Repository Identification Service (ORI), http://ori.edina.ac.uk/, but there are no identifiers for institutional repositories. This is mainly due to the fact that repositories are volatile entities which will often change their name and even their home URL, which is a strong reason for providing them with persistent identifiers whose data may be updated and maintained by the repository manager. More information can be found at: http://edina.ac.uk/projects/rjb-root/index.html

OrgID-UC8 - Developer - Directory services

As a developer for research funders, I need to link an OrgID within my application to a directory service. This will allow an end user or a machine to verify identity and contact details.

WG Owner: Richard Moore (dmoore@mooreanswers.co.uk)

Component Parts: The directory service use case allows either a machine or a user to enter an OrgID and retrieve a record or put in part of a record and retrieve any record(s) that match the record fragment, together with their OrgID(s). A more refined use case allows queries to be made against a directory service that extract all matching records.

There are two main variants of this use case:

1. Enter an OrgID - the directory services retrieves and returns the organisation record or reports that no such record exists.

2. Enter fragment of organisation record - the directory services retrieves record set(s) for all records that match the fragment. The records would include the OrgID and would be delivered as a series of record sets to accommodate large volumes of records matching the search fragment.

An application programming interface (API) is used to connect services programmatically. This enables 3rd party rich services to be constructed that can access ORGID's . A user interface sat 'on-top of the same API allows users to query the service.

Example 8: I am working on a project for a research funding agency building an online application form. I want to present a user with an interface which allows her, when stating institutional affiliation, to enter a few letters or words and then choose from a drop down box a list of matching institutions (and my form will record the orgID in the background) [variant 2]. Alternatively the applicant may enter an orgID and the interface will present the details of an institution to confirm that he has the right ID [variant 1].