

## ORCID Multiple Assertions Working Group Recommendations

### Working group members:

- Geoffrey Bilder - <http://orcid.org/0000-0003-1315-5960>
- Paolo Bouquet - <http://orcid.org/0000-0001-5305-3817>
- Paul Dlug - <http://orcid.org/0000-0003-3518-0710>
- Mark Doyle - <http://orcid.org/0000-0001-5919-8670>
- Janifer Gatenby - <http://orcid.org/0000-0002-4921-4101>
- Laure Haak - <http://orcid.org/0000-0001-5109-3700>
- Richard Ikeda - <http://orcid.org/0000-0003-4015-5528>
- Laura Paglione - <http://orcid.org/0000-0003-3188-6273>
- Chris Shillum - <http://orcid.org/0000-0002-1108-3660>
- Mike Taylor - <http://orcid.org/0000-0002-8534-5985>
- Vetle Torvik - <http://orcid.org/0000-0002-0035-1850>
- Simeon Warner - <http://orcid.org/0000-0002-7970-7855>
- Bruce Weinberg - <http://orcid.org/0000-0001-8856-1803>

### Document update history:

Date	Changes	Who
2013-10-08	Collected ideas from MAWG discussions	Simeon
2014-01-15	Fleshed out based on initial MAWG review	Simeon
2014-03-19	Minor revisions based on MAWG feedback	Janifer, Laure, Simeon (editor)

## I. Background, drivers and goals

The charge of the group is to develop recommendations for policies, business rules, and technology to manage multiple assertions (full charge in Appendix A). Key current drivers are:

- The primary driver for ORCID is participation from both individuals and organizations. Support for multiple assertions should take this into account.
- Currently the registry keeps and displays all information. This process results in information about the same item (works especially) being imported from several different sources. Users interpret this action as duplicates being added to their record and there is no good mechanism for them to handle it.
- Some organizations (particularly research institutions) are interested in making assertions about the people who have worked or studied at their location. The system currently supports only self-assertions or delegations of that authority, and not assertions by other organizations.

The following key goals for the support of multiple assertions motivate the recommendations of this group:

- Improve the quality of the data in ORCID by providing mechanisms to manage duplicate and conflicting data as the registry grows
- Improve the reliability of the data in ORCID by allowing authoritative sources to support or refute an individual ORCID Record holder's assertions
- Save user effort in creating and maintaining high-quality profiles

A number of other possible goals have been expressed but are less important:

- Provide general and/or detailed information about all types items that the Registry stores (people, works, grants, organizations, etc.)
- Provide a mechanism to resolve data differences in different external databases
- Provide a resource for identifying an "authoritative source" for the different types of information stored in the Registry
- Provide a mechanism to normalize the list of an individual's "associated objects" (works, associations, etc.) so that other databases will have similar lists

No distinction is made between the terms "claim" and "assertion", in the following the term assertion is used.

## 2. Use cases

Seven use cases were considered and are listed here. The first 4 are listed in approximate priority order, most important first. While there was not consensus, the working group was in general agreement about the importance of the first 4 use cases, and less so about ordering the others.

### 2.1 Use case: Indication of duplicate ORCID iDs

**Actors/goals:** Member organizations should be able indicate when they have identified two ORCID iDs that identify the same individual. The goal is to have this information available so that it might then either be acted upon by ORCID to initiate a combination process, or be flagged to the individual(s) to check and initiate a resolution.

**Assumptions:** We expect that a certain number of duplicates will be created either by individuals alone, by individuals and institutions, or by multiple institutions. Hopefully other mechanisms will limit the number of such duplicates created, but there will be some number of duplicates.

**Flow:** The key assertion is "ORCID1 is the same individual as ORCID2". It will be important to know who is making the assertion. It may also be useful to have other information accepted/stored to aid in deciding how to use the assertion (confidence or reason for duplicate identification).

**Excluded:** This use case does not include the processes that are used to resolve or combine ORCID identities that are indicated. The data may be used simply to provide advisory messages, or to initiate manual or automatic combination of ORCID iDs.

**Other considerations:** There may be some legitimate cases for individuals using multiple ORCID iDs for different personas, although this is not the normal expectation of ORCID. How might that be handled? It will be important to minimize the creation of duplicate ORCID iDs as well as having ways to gracefully handle duplicates. An alternate view is that this may not be such a big issue because researchers will find it in their best interest to ensure that they have just one ORCID iD.

## 2.2 Use case: Validated affiliations

**Actors/goals:** Member institutions should be able to validate affiliation self-claims in ORCID records. Desire for this ability has been expressed by member institutions and is a member benefit that may encourage institutional membership.

**Assumptions:** Individuals will add affiliation claims to their ORCID profiles and there will be some fraction of these that are not accurate, either because of error or intent to mislead. We assume that the validating institution knows the ORCID iDs of its members. There must be a mapping from member institution identities to affiliations that they can speak authoritatively about (in many cases there may be more than one affiliation depending on the granularity used to represent affiliations in ORCID records).

**Flow:** A member institution may make an assertion of the form "ORCID is/was affiliated with AFFILIATION" for an AFFILIATION they can speak authoritatively about where there is a corresponding AFFILIATION entry in the ORCID record. The ORCID record may have dates, closed or open-ended, associated with the affiliation information so there is a question about whether the member can/should/must provide dates with a validation assertion, or is simply validating whatever claim has been made in the ORCID record. The key implementation question is whether a validation assertion is stored as an addition to the ORCID record (some glorified flag indicating "validated"), or whether the validation assertion is stored as a separate assertion that is somehow matched with data in the ORCID record to result in an understanding that the affiliation in the ORCID is validated. The latter approach is more complicated but is stable in the case that the individual modifies their ORCID record -- say deleting and then re-adding a validated institution.

**Excluded:** Ability to refute an affiliation claim (separate use case with different privacy policy issues). Ability to correct suggested affiliations, correct minor errors, or complete information.

**Other considerations:** There are UI and policy questions about how such a validated affiliation would be indicated on the public profile (out of scope for MAWG). Validated affiliations are essentially claims about claims and thus this use case has architectural implications.

### 2.3 Use case: A single researcher has published works under multiple ORCID iDs

**Actor/goals:** Actors are researchers, publishers, and third parties such as CrossRef and others who may want to query the ORCID registry to retrieve or validate research publication lists. Goal is to ensure that ORCID can serve as a robust definitive source for information about researcher publications. This use case will focus on journal publications, although it can be generalized to other kinds of works with the roles of publisher and CrossRef replaced with other organizations.

**Assumptions:** Some portion of researchers will end up with multiple ORCID iDs, each with publications attached. ORCID will allow trusted member organizations to assert that a researcher with a particular ORCID iD is an author/contributor for a particular work. Researchers will be given the opportunity to add such works to their ORCID profile.

**Flow/Data:** The main data involved are DOIs or other unique identifiers for works and ORCID iDs. Typical flow would start with the publisher collecting an author's ORCID iD and associating with the author's profile (usually a one-time process). Author may have forgotten they have an ORCID and create a second one for another publisher or accept one created for them by, say, their new employer. Upon publication of the work, publisher deposits ORCID iD along with DOI into a service such as CrossRef which then makes this data available to ORCID. Alternatively, publisher may be a trusted organization that can directly make claims within ORCID. Second publisher does the same, but unknowingly uses a different ORCID iD for the author in their CrossRef deposit. For each work, the researcher is notified by ORCID that a new claim asserts that they are an author of a particular work. Author would typically acknowledge the claims and would like to add the works to their profile. Author eventually realizes there are multiple ORCID iDs with their publications and notifies ORCID. ORCID deprecates one of the duplicate ORCID iDs in favor of the other. ORCID would like to ensure that an API that returns works for the researcher includes all works from both the deprecated and non-deprecated ORCID iDs. ORCID may achieve this result either 1) by moving works from the deprecated ORCID iD to the non-deprecated iD or 2) by ensuring that queries traverse the relationship identifying the deprecated ORCID iD and the non-deprecated iD and return the total works as a union of the works associated with each iD.

**Excluded:** Cases in which a researcher denies a claim that they are an author/contributor on a work.

**Other considerations:** It is desirable that any identification of two ORCID iDs as belonging to a single researcher be reversible in case the merge needs to be undone. Resolution mechanism 1) is only reversible if additional claims are properly added to record the history of the pre-merged state. Resolution 2) is easily reversible by just undoing the association between the two ORCID iDs.

### 2.4 Use Case: Refuted affiliations

**Actors/goals:** Member institutions might wish to be able to refute affiliation self-claims in ORCID records. This would be an extension of the "Validated affiliations" use case.

**Assumptions:** Individuals will add affiliation claims to their ORCID profiles and there will be some fraction of these that are not accurate, either because of error or intent to mislead. We assume that refutations would come, or be allowed, only from institutions in the claim. There must be a mapping from member institution identities to affiliations that they can speak authoritatively about (in many cases there may be more than one affiliation depending on the granularity used to represent affiliations in ORCID records).

**Flow:** A member institution may make an assertion of the form "ORCID is/was not affiliated with AFFILIATION" for an AFFILIATION they can speak authoritatively about where there is a corresponding AFFILIATION entry in the ORCID record. A refinement might allow not the basic affiliation claim but the dates to be disputed.

**Excluded:** Ability to suggest correct affiliations or correct minor errors.

**Other considerations:** There are very significant policy questions about how the information would be used: it is likely against the ORCID policy to display on a user profile the fact that an affiliation has been refuted, but a refutation might trigger some dispute resolution mechanism. There may be issues created with possible refutation of affiliation in records that are duplicates and should instead be flagged as such. This is likely not a use case to implement in the immediate-term except via a trigger of ORCID's existing manual dispute procedures (<http://orcid.org/orcid-dispute-procedures>). It is perhaps better to have an environment where researchers strive to have their claims validated and where un-validated claims have low value, and so there would be less need for incorrect claims to be refuted.

## 2.5 Use Case: Academic Use and Contributions to ORCID

**Actors/goals:** The actors are academic researchers interested in the science of science and innovation. The goals are to (1) to improve the quality of ORCID, which will generate more complete and more accurate profiles, accelerating adoption and (2) use ORCID for academic research purposes.

**Assumptions:** There are 2 fundamental approaches to disambiguation – algorithmic and researcher controlled. Neither approach covers all documents with 100% accuracy. However, these two approaches are highly complementary: the ORCID approach has (potentially) higher accuracy but is limited by participation rates, while the algorithmic approach is easily scalable but limited in accuracy. Combining these two approaches will enable both static (e.g. correcting existing errors) and dynamic (e.g. refining best practices / algorithms) improvements in both approaches.

**Flow:** Researchers will access ORCID registry data, compare these data to existing algorithmic data (as well as data from other sources, potentially), identify likely errors in ORCID, and provide suggestions to ORCID for inclusion in the registry. Improved data would also be used for researchers studying the science of science.

**Excluded:** Publicly reporting data on individual ORCID members in research products.

**Other considerations:** One question is how suggestions for improvement can be included in the registry (e.g. do they have to be approved by users)? A second question is whether registry data that is not marked "public access" could be accessed by a trusted team working in a secure environment. A third issue would be to identify one or more funding agencies.

## 2.6 Use case: Declined claim of authorship

**Actor/goals:** Actors are researchers, publishers, and third parties such as CrossRef and others who may want to query the ORCID registry to retrieve or validate research publication lists. Goal is to ensure that ORCID can serve as a robust definitive source for information about researcher publications. This use case will focus on journal publications, although it can be generalized to other kinds of works with the roles of publisher and CrossRef replaced with other organizations.

**Assumptions:** There are a variety of reasons why a researcher might decline to accept a claim that they are an author of a particular work. Broadly, the claim may either be factually correct, but the researcher doesn't want to acknowledge authorship (research result is wrong, publication in a different field, etc.) or the claim may be factually incorrect. This use case is concerned with the factually correct case; it is assumed ORCID has a mechanism to report and correct factual errors.

**Flow/Data:** The main data involved are DOIs or other unique identifiers for works and ORCID iDs. One typical flow would start at the publisher collecting an author's ORCID iD and associating with the author's profile (usually a one-time process). Upon publication of the work, publisher deposits ORCID iD along with DOI into a service such as CrossRef which then makes available this data to ORCID. Alternatively, publisher may be a trusted organization that can directly make claims within ORCID. For each work, the researcher is notified by ORCID that a new claim asserts that they are an author of a particular work. Author would typically acknowledge the claims and add the works to their profile. However, in this use case, the researcher declines to add the work to their profile even though they are indeed an author of the work.

**Excluded:** Factually incorrect claims that a researcher is an author/contributor to a work.

**Other considerations:** The main consideration is what is the disposition of a claim that is in the ORCID system, but not accepted by the researcher as part of his or her profile? Does ORCID have a mechanism to track and store such claims? Can declined claims be exposed through the ORCID API with additional provenance information so that consumers of ORCID data can decide for themselves whether to accept the claim? Would this violate the ORCID registry privacy policy? Should ORCID maintain a secondary claims store independent of the ORCID registry itself that serve as a definitive source of factual claims independent of researchers' profiles?

## 2.7 Use case: ISNI Assertions pushed to ORCID

**Actors/goals:** ISNI and ORCID are using identifiers from the same scheme that are conformant with the standard ISO 27729. The goal is to align identities in the two separate databases using the same scheme and eventually to have a common identifier for an individual researcher. ORCID and ISNI are working in different but complementary ways to achieve unique identification of researchers.

**Assumptions:** ISNI will receive data for ingest into ISNI on a regular basis, e.g. monthly. This data will include all publically available data about all new ORCID assignments and also will include updated (enriched data). A full refresh may also be a possibility.

In the first place, the goal will be to align ORCID and ISNI identifiers with a longer term goal of being able to have just one identifier per individual. The process of achieving the latter will require coordinated development on both sides.

In most cases the email address of the individual will not be part of the public data. However, the back half of the email could be useful information. It would be useful to pursue the possibility of providing this information to selected trusted parties.

### **Flow/Data:**

- ISNI ingests data on a regular basis from ORCID.

- ISNI loads this data onto the ISNI database and produces matches with all its sources. (The data may match with existing assigned data but it may also make a new match and cause a record to be updated to assigned.)
- Via the ISNI notification system, data is sent back to ORCID, including assigned ISNIs as appropriate. The ISNI notifications are time and date stamped to prevent repetitive assertions being sent.
- The ISNI assertions are then loaded to a “pending review” place on the ORCID database and the relevant individual is notified by email.
- ISNI may also return “possible matches” where the match was between the “no match” threshold and the “definite match” threshold
- [The individual is able to express preference for the identifier to be used on both databases.]
- ORCID notifies ISNI of claimed assumptions in its next regular update to ISNI. (This may result in resolution of some possible matches on ISNI).

**Excluded:** A separate use case will be made for pushing ISNI assigned researchers that are not in ORCID to ORCID. A separate use case will be made for pushing unclaimed assertions in the ORCID database to ISNI.

**Other considerations:** The ISNI database already (2013-12) has data about 1.57 million researchers in ORCID’s scope and that will grow in 2013 with the addition of more theses data. This data is already available when a researcher registers for an ORCID via the SRU enquiry API, but this is an optional step. Furthermore, at the time an SRU search is made, it only has the name available and may produce result sets that are too big especially for common names. Matching fuller records post assignment by machine will therefore be fruitful in many cases.

For the 1.57 million researchers, ISNI has already made 9 million links among its sources and is diffusing the data to VIAF and other sources as linked data. These links are both inter-domain and cross domain.

The ISNI database is constructed by loading from existing databases. IDs are assigned to a researcher because at least two data sources in ISNI share information about that researcher and his or her associated works. Via ISNI, ORCIDs will thus be more widely diffused to researcher databases that are not ORCID selection databases. Is there a limit to the number of selection databases to propose to a researcher?

### 3. Types of assertion to support (and not)

This section is an attempt to abstract some of the ideas and the classes of assertion that are involved in the use cases. It is hoped that these classes will provide guidance beyond the initial implementation of high priority use cases.

**Stand-alone assertions** - It must be possible to make assertions outside of an individual’s profile to support our use cases. A number of examples are given below but these might include "University (member) asserts individual is affiliated with them". Duplicate assertions from multiple sources indicate agreement and thus more reliable information. Storage of duplicate assertions is more stable under changes (such as one of a duplicate pair being removed) than storage of an assertion that another assertion is correct.

**Counter assertions and refutation** - It must be possible to either make a counter assertion or to refute a assertion. In both of these cases it is not sufficient to simply make a different assertion. For a counter assertion

it must be known which assertion is countered so that the counter assertion does not appear simply as a separate assertion (if identifying information doesn't match up at all for example). For a refutation without a counter assertion there is no alternate information, just the assertion that the original assertion is wrong. This implies a mechanism for referring to other assertions stored by ORCID.

**Example assertions that ORCID should support** – The following list is a set of example assertions that should be supported. In all cases "individual" should be read as an "individual identified by an ORCID iD":

- All currently supported assertions by individual (contributed to work, has bio details, has affiliation, etc.)
- An individual asserts that another individual is a co-author of a work
- A university asserts that an individual is currently affiliated with them
- A university asserts that an individual, where the individual's public profile asserts an affiliation, is NOT affiliated with them
- A publisher asserts that an individual is an author of a work (outside of the current mechanisms for update of an ORCID profile via a trusted relationship)
- A granting body asserts that an individual is a recipient of a grant (outside of the current mechanisms for update of an ORCID profile via a trusted relationship)
- A professional association asserts that an individual is a member (type of affiliation)
- A professional association asserts that an individual made a presentation (cf. publisher says individual authored a work)
- A professional association asserts that an individual serves as a reviewer
- Individual agreement or disagreement with an assertion about the individual (agree with or discount an assertion that someone else has made, i.e., person A says they did something, and I agree/disagree)
- Counter assertions (express disagreement and supply corrected information, i.e., person A says that this happened, but it actually was that)

**Example assertions that ORCID should not support** - In some of the following cases it may not be easy to determine whether or not the restrictions are met and thus be hard to implement a rule excluding them, however in principle we should not support them:

- "Like" (i.e. ORCID should aim to support truth, not popularity)
- Information harvested from friend-of-a-friend or similar pages
- Assertions of a malicious intent - hard to determine and control
- Should not allow individuals to take advantage of credit for works they didn't contribute to by individual with similar name
- Things that may be inferred from other assertions, or assertions that are implicit in other contexts (i.e. can't assert that you are a co-author when that can be implied from both being authors of the same work). e.g I am a co-author with Y (in absence of the work). Unnecessary if work record present because implied, hard to verify in absence of work.

#### 4. Features, requirements and design considerations

This section describes some features, requirements and design considerations generalized from the use cases to help inform the design and architecture of ORCID's support for multiple assertions.

#### **4.1 Provenance**

It is essential that the provenance of assertions be recorded and maintained. Identical assertions from different sources should be retained intact with provenance information. This allows understanding of the provenance, and is robust in the case that the assertion from one source is later updated or deleted.

#### **4.2 Supporting updates**

It is essential that updates from the same source be possible without creating duplicate assertions. ORCID should retain information about the source plus any identifiers for the item or assertion. Thus information about the same item or assertion from the same source can be understood to be either identical (no change) or an update (changed information), without creating duplicates. Also, if an assertion has been removed/deleted by an action at ORCID, some record sufficient to avoid re-import should be retained.

#### **4.3 Logical Separation of Assertions from the ORCID Record**

Assertions should be stored and accessed with logical separation from the Record in the ORCID Registry. This distinction will allow separation of the notions of assertions in an ORCID Record (where the privacy policy says the individual has absolute control) from assertions made by other parties that have either 1) not (yet) been approved by the individual, 2) are disputed by the individual, or 3) perhaps not even associated with an ORCID record. Even with such separation it will be essential to control user perception of how possibly inaccurate or unwanted data is made available to data consumers.

#### **4.4 Assertion granularity**

Assertions should be stored at the level of a work/affiliation/grant/etc. (such as “ORCID is an author of a work with title X, DOI Y, publisher Z,...”) rather than at a more atomic level (such as “this work has title X”). This might mean that an edit of data from another source would be a copy and then an edit of parts of the data.

#### **4.5 Who can make assertions?**

Only member organizations and the individuals concerned in assertions should be allowed to make assertions. All member organizations will have agreed to appropriate terms as part of their membership agreement. Individuals may generate assertions as a result of editing the information in their profiles or rejecting suggestions of information to be included in their profiles.

#### **4.6 Communication with ORCID iD owners**

A result of accepting assertions will be that possibly new information will be available to enhance an ORCID record, or that a user might want to refute. Reduced effort in maintaining their profile will be a potential benefit to users but we must provide a balance between alienating them with too much email, and having them surprised by what is added to their ORCID Records. Any alert emails should be infrequent (e.g. 6 months or yearly) by default, and this frequency should be controllable. It may be useful to have notions of a class or classes of trusted parties so that a user may then opt-in to trusting an entire group of parties rather having to set up and manage trust relationships one-by-one.

#### **4.7 Negative assertions and assertions about assertions**

Correcting inaccurate information is in everyone's interest. ORCID's privacy policy gives ORCID iD holders control over what is displayed on or shared from their Records. It is therefore impractical to imagine direct

negative or disputed assertions in an individual's record because most individuals would not allow them. Instead we should emphasize validated claims through conspicuous display of the validation, and ease of making such a claim.

In the case of inaccurate information, instead of a display of a negative assertion, we should provide a mechanism to allow individuals and organizations to inform ORCID iD holders of errors in their records, perhaps providing corrected information, or other details that would enable the member organization to validate the claim. In cases of unresolved disagreement or fraud, ORCID's established dispute procedures (<http://orcid.org/orcid-dispute-procedures>) would handle misrepresented information.

While architecturally it must be possible to make some forms of assertion about assertions to support validation and (perhaps) refutation, there are currently no use cases suggesting the need for more complex relationships involving arbitrary assertions about assertions, or chains of assertions about assertions.

#### 4.8 Queries supported

The design of multiple assertions will need to support several types of queries and reports. While it is difficult to anticipate all possible queries, there are several that should be supported early on including:

- What information in this Record was verified, when, and by whom?
- Who claims to be affiliated with my organization, and who have I not yet validated?
- Who else has claimed to be a contributor to one of my works?
- Which ORCID iD holders have had five or more refused claim validations?
- What percentage of validatable claims have been validated? On this Record? In the Registry?
- What information from my organization has been rejected by ORCID iD holders?

Some of these queries may be available to the general public, while others might be part of a value-added benefit of membership for those financially supporting ORCID's mission. In addition to finding ways to support scholarly research access to data, it might also be useful to provide research access to some query facilities.

#### 4.9 Display of assertions

All claims and other information about a single association should be grouped together in some way (for example, claims/information about the relationship about a work or an affiliation to an ORCID iD). This information may contain varied accounts about the metadata used to describe the object, as well as multiple data sources, and multiple entities claiming the validity of claim. There needs to be an easily understandable way to enable iD holder control over which information is displayed and shared from the metadata. The interface also needs to be clear about the data sources for the information, and about who supports/agrees with/validates the claim. If data has been changed since supporting claim(s), there should be some indication of the changes made since.

## Appendix A. – Meeting the charge, an annotated version of the Multiple Assertions Working Group Charter from 2013-06-10

*Annotations shown in italics.*

### ORCID Background

ORCID has a two-fold mission. We provide an open Registry of unique researcher identifiers, with associated activities, affiliations, and works, primarily managed by researchers themselves. We also work collaboratively with the research community to enable linkages between these identifiers and other identifier schemes and systems, and allow assertions about these records to be made by organizations other than the individual researcher.

To date, ORCID has taken a cautious approach to enabling both self- and organization-derived assertions to co-exist in the registry: organization-derived assertions require either explicit acknowledgement by an individual researcher of assertions made about them, or the organization making the assertions must have an employment or similar relationship with the individual. While this approach attempts to minimize conflicts, they are inevitable as individuals change organizational affiliation. In addition, the current process limits the ability of less strongly affiliated organizations to make assertions about researchers.

*In essence, all data in the current system is asserted by the individual owning the record, or by a proxy of theirs. The data source is recorded but the assertion source is implied.*

### MAWG Charter

This working group will review options for managing multiple assertions, and develop recommendations for policies, business rules, and technology. Specifically, the group is charged with answering the following questions. The group will meet at least monthly and will produce a report listing recommendations by December 2013.

#### I. Making Assertions

- What kinds of assertions and by whom should be permitted by ORCID?

*Some examples are given in the “Use cases” section, and attempts to extract rules in “Types of assertion to support (and not)”.*

- How should the data in the ORCID registry be structured to manage assertions made by multiple parties about a referenced data object?

*Assertions made by different parties should be stored separately with appropriate provenance information. This includes assertions about the same data object and even identical assertions from different sources.*

- Should the ORCID Registry distinguish assertions made when no researcher participated in the assertion from those one was? If so, how?

*Source of assertion should be recorded and so an assertion by a researcher will be distinguished from an assertion by a third party.*

- What technologies should be explored to support multiple assertions about the same data object?

*The MAWG has stopped short of recommending specific technologies. These choices will be strongly influenced by existing technologies used and by staff and infrastructure. This report describes features and requirements.*

## 2. Addressing Conflicting Information

- How should ORCID handle conflicting assertions about a relationship between two data objects (e.g., a researcher and a specific work)?

*Conflicting assertions must be stored (logically) separately from researcher records/profiles. Suggestions to ORCID iD holders might show conflict and ask for input (support/refute) which would then be stored as new assertions.*

- In what circumstances would a conflicting assertion indicate a duplicate data item or record?

*If assertions from multiple sources are stored separately then a true duplicate occurs only when the same assertion appears from the same source. In other cases, a “duplicate” is to be seen as a supporting assertion. Resolution of near duplicates – such as different descriptions of the same work – will be facilitated by the ability to store user assertions that two items are indeed the same, and this information will be robust under future updates from the original sources.*

## 3. ORCID's Role

- What is ORCID's role in disambiguating non-name data?

*ORCID should provide facilities that allow individual ORCID record holders to see and understand where there is ambiguity, and then facilitate data correction by the appropriate parties.*

- What is ORCID's responsibility/obligation for providing “true”, “unambiguous”, “correct” or “up-to-date” data?

*ORCID cannot support extensive manual processes to ensure “correct”/“true”/“unambiguous” data. Support for multiple assertions can facilitate processes whereby the relevant actors help improve data quality, and ORCID can indicate where information has been validated by an authoritative party.*

- What role should ORCID play in establishing data or identity authority, or facilitating data normalization or correction?

*ORCID should support validation, and perhaps refutation, of assertions using data from authoritative sources (such as affiliation data from member institutions).*