# Questions from Webinar Participants and Panel Responses

Responses are from:

- CF: Claudia Fabian, co-founder International Standard Manuscript Identifier
- LM: Lorna Mitchell, Head of Library, Archives & Publications, Royal Botanic Gardens Edinburgh
- RB: Rebecca Bailey, Programme Director, Towards a National Collection
- RK: Rachael Kotarski, British Library, Principal Investigator, PIDs as IRO Infrastructure

- What about the way papers are linked to research data in PID/knowledge graphs (see scholix for example)? Isn't that a feasible model to link PIDs  for IROs' collections too, regardless the objects? Ta
  - Answer: Absolutely. It would be a big benefit for heritage organisations to be able to make use of approaches like scholix to demonstrate how the items they collect, curate and preserve have important research value. It's true that this should work no matter what the object is, as long as it's citable as evidence in research. Having PIDs that can be used by researchers for citation in papers and datasets would allow us to start to do this for heritage objects (RK).

- Everyone knows what "persistent" means and what "identifier" means - but you are defining PID as something more. Should you not use a term like 'persistent hyperlinked identifier ' PHIs?
  - Answer: I think in the context of this project, a clear definition is enough – even without a new name. But it might be interesting to investigate how such not yet hyperlinked identifiers can be made hyperlinked. I think that is a technical issue, but there is also the need to relate them to an existing datapool (e.g. a VIAF number to the VIAF system, a GND-number to GND (Gemeinsame Normdatei) etc. (CF)

- What PIDs can anyone suggest for use with art and design research outputs, e.g. exhibitions or artefacts?
  - Answer: In general, Digital Object Identifiers provided by DataCite would be suitable for any items that are the outputs of research, including exhibitions and artefacts. Both or these have reasonable item types within the metadata ('Event' and 'PhysicalObject' respectively, see https://schema.datacite.org) (RK).

  - I would look into the Getty Art and Architecture Thesaurus (but are these identifiers hyperlinked?). Many exhibitions can be found in authority files, like the German GND, and there they have an authority number which can be used as a

permanent identifier. The issue here may be: how to bring entities not yet present within this context into the authority file. (CF)

- Cannot DOIs be used for specimens too? That would ensure they are harvested and fed into interlinking PIDs graphs without the need to radically change the used metadata schemas.
  - Answer: It absolutely would be possible to use DOIs for specimens. It is technically possible now with DataCite, as it has room in the schema (resourceType 'PhysicalObject', see https://schema.datacite.org), but the question is whether the whole metadata schema and services offered by DataCite are relevant and appropriate for the sector to use at scale. Whether the whole sector wanted to use DataCite or set up a heritage sector registration agency for DOIs is an open question, and one we'd like to answer more fully (RK).

- How do we draw the distinction between a PID that refers to the physical object and a PID that refers to the museums documentation about the object (i.e. the object record)
  - Answer: We answered this a little in the session. A PID should have relevant metadata that makes it clear whether the PID is for the object or it's records. It would be sensible to allow for both in many circumstances, and to make clear links between the PID for an object and PID for its record (RK).

- Approach of minimal metadata + PID certainly pragmatic and useful but in many cases identification is a (scholarly) issue in itself. Modular approach is key.
  - Answer: I'd agree that accurately identifying items in our collections is a scholarly issue but the first step in that process is bringing together the items and the relevant scholars. The creation of a basic catalogue record (i.e. metadata) and a PID would enable that first step and allow the confirmed identification to be permanently linked to the specific item either by adding to the existing metadata or by linking the item's PID to the relevant scholarly publication. (LM)

  - I fully agree with all these comments. The "entity" problem is that a real world object can only be presented, never present on the internet. Even a digital copy is not the thing itself, and there may be different digital copies around. The identifier should allow us to bring all kinds of information about this object together … but he should not himself get (too much) involved in issues of describing this object. I am aware about a number of "grey zones" in this approach. E.g. an object has a life in time (which might be taken into account in the object identifier) – when do we need a new identifier because it is no longer the same object? Two scholarly descriptions of a mss may be so different that we think it is two objects … but it is only one.
    Nevertheless: better an IMPERFECT (which for me would be a "broad") identifier than no identifier. Better a solution for 95% of "easy" cases than make a small portion of difficult issues prevent any definition of an identifier. (CF)

- We have a DOI for a 3d digital model of one of our marine reptiles - which also has a MDA style registration number
  - Answer: Thanks to whoever supplied this comment. If that's you, please do get in touch (you can email rachael.kotarski@bl.uk) to share the example with us, and we'll share the link. For more information on MDA codes, please see: https://collectionstrust.org.uk/resource/mda-codes/

- How do we handle versioning of PIDs, objects change over time, museum documentation is updated, how do we let people refer to a PID at a point in time ?
  - Answer: This is a question that needs to be approached based on the needs of various users of the PID. If we're looking at supporting researchers, the object needs a new PID whenever it is changed in some way to support the provenance and trust in their work. The organisation will have its own collection management needs, and all this should come together to identify a versioning policy for the identifiers (RK).

- How will TANC link to DiSSCo?
  - Answer: Discussions have taken place between the two initiatives. There is not a formal link between them, although the Natural History Museum London are a partner in the Heritage PIDs work as well as DiSSCo.

- https://collectionstrust.org.uk/ might be a good organisation to help push the need of PIDs
  - Answer: Absolutely, I (Rachael) have had an initial conversation with them and look forward to more as the project goes on (RK).

- The irony is that Collections Trust" already run the "Persistent" "Identifiers" (MDA Code system) within Museums
  - Answer: My initial discussion with Collections Trust was on exactly this topic, and I'm hoping we can work together looking at the actionability part to make MDA codes a big part of the solution (RK).

- Does it matter how we create PIDs? Do PIDs need to be subject specific?
  - Answer: PIDs do not need to be subject specific. What is important is their ability to meet the needs of the organisation assigning them to collections and their users. PID technologies themselves (e.g. DOIs, ARKs, Handles) are already subject-agnostic, but the services built on top of these technologies (specifically those from Crossref or DataCite which are built on DOIs) are what may be more tailored, and it's what those services provide are where more tailored support is given in some areas (RK).

- There's been reference to institutions holding collections in perpetuity. Sharing that Project Omega at TNA is looking towards a linked data catalogue and PIDs.
  - Answer: Many thanks for this recommendation. There are some blog posts about this project here https://medium.com/@catalogueprojects

- Would a manuscript PID identify the material object or the digital surrogate?
  - Answer: It depends which mss. PID … The ISMI should be for the REAL object … but we already have numbers of URNs for digital copies, standard identifiers for catalogue records … (CF)

- For Rebecca: given that foundation projects go on beyond deadline for discovery project applications, how can you ensure that emerging learning from foundation projects is considered by applicants?
  - Answer: Applicants are encouraged to make contact with the Foundations Projects relevant to their proposals and will potentially be able to take part in project webinars. Each Foundation Project will issue an interim report in time to inform the full development of shortlisted Discovery Project proposals (RB).

- "There should not be more than one PIDs". How do we police people creating PIDs for items which already have a PID? Especially if PIDs are difficult to find.
  - Answer: There are cases where more than one PID for an item has been created. This primarily causes problems with tracking and metrics, since you'd have to aggregate information across all PIDs for the item to understand its use.

    It's difficult to 'police' who creates PIDs for items they might not have control over, but my feeling is if organisations have visible and usable PIDs in place for their own content already, it's less likely that someone would try and create them separately (RK).

- A question for Lorna: Did you suggest more focus on digitising analogue collection items and less focus on metadata would help with the adoption of PIDs? How? / Why?
  - Answer: My suggestion was that adopting a "digitise with minimal metadata" approach might help us to tackle the digitisation challenge that so many collections are struggling with, i.e. how do we make our collections accessible digitally when we don't know what we have in them. Including a PID as part of the minimal metadata that's created when an item is digitised would still enable the item to be uniquely identified but without the need for time and effort to be spent creating detailed catalogue records / metadata. [I should point out that I wouldn't generally advocate a minimal metadata approach - I'm a big fan of very detailed catalogue records, especially for the collections that I'm responsible for, but, at the rate we're going that is going to take a really, really long time and so a more pragmatic approach may be required.] (See also my response to the question above about identification as a scholarly issue. LM)

- Would the panel agree that there could be a place for Wikidata as being a brokerage system for PIDs since there will generally never be a single PID solution for everyone?
  - Answer: Yes, if we can come up with a way to ensure confidence in that system (LM).

  - I am convinced that the wikidata structure is extremely helpful in this context. But today it is fully open for everybody to contribute and edit. For a controlled system this might be a problem. (CF)

- Claudia is talking about centrally-maintained PIDs. What are the trade-offs between a 'top-down' strategy (from the centre) and a bottom-up one (where individual institutions assign their own PIDs)?
  - Answer: There is value in centrally maintained PIDs, especially for authority control that can be used across organisations and communities. It's important to note that centrally maintained systems can still be directed by the community as a whole, and largely still enable local applications where institutions assign PIDs as they see fit (within community guidelines). For instance, with the DOI registration agencies, they enable publishers, institutions, and other holders of content, to assign DOIs to their own content as they see fit. This is different again from institutions not just assigning their own PIDs, but building their own PID system, which is a truly bottom-up strategy (RK).

  - I am not sure that I talk about centrally maintained PIDs – this is a more technical issue. I think there must be a (central) system where I can go to look for the number – in order to avoid duplicates. And it must be made sure that the same number is only allocated once … But then, cooperating systems could allocate this number and contribute the information to a central system … e.g. in Germany, if GND has a number for each mss., it could attribute an ISMI as well … but this is more on an organisational level …
  E.g. for ISIL (international standard identifier for libraries) – would it not be easy to have a central system to look it up? I haven't identified one yet. But there is one for Germany, one for UK … (CF)

- Do any of the TANC plans include investigating Provenance mechanisms for creation of PIDs?
  - Answer: Not currently, but could feature in one or more Discovery Project bids (RB).

- Do we need a new book PID maybe? ISBNs used by book industry for stock management which makes them unsuitable as a PID (e.g. they differ for different manifestations of the same edition of a work)
  - Answer: I'd be quite keen to look at the feasibility of using ISBNs as PIDs purely because we'll already have ISBNs in our collection management systems and so

if we can re-use them it might be easier than starting from scratch. In part it depends on whether we want an identifier for the content (in which case a new PID might be required) or the item that contains that content. (LM)

- Re making links between objects and exhibitions, etc, see Spectrum Use of Collections procedure. Re distinguishing between analogue originals and digital facsimiles see Spectrum Reproduction procedure
  - Answer: Many thanks for this recommendation, please see links: https://collectionstrust.org.uk/resource/use-of-collections-the-spectrum-standard/

  - https://collectionstrust.org.uk/resource/reproduction-the-spectrum-standard/ (RK)

- A question for Claudia: is there a webpage for the research/project group she is involved with? Thanks!
  - Answer: There is not yet a webpage, but two links to where we are: http://digitalia.sbn.it/article/view/2486 https://www.manuscript-cultures.uni-hamburg.de/files/mss_cataloguing_2018/Cassin_pres.pdf (CF)


- "In relation to Rebecca's comment about who do we need to influence. Many institutions are tied to commercial suppliers with limited potential for development. How might we approach that?
  - Answer: Commercial companies are eligible to be part of TaNC Discovery Project bids, perhaps including them there could be useful? (RB)

  - There must also be scope for consumer power here, i.e. if organisations require suppliers to support their adoption of PIDs then it might start to happen. (LM)