PID discussion I:

Definition of PID (Persistent IDentifier) :

Lets start with Wikipedia:

...

"A persistent identifier (PI or PID) is a long-lasting reference to a document, file, web page, or other object. The term "persistent identifier" is usually used in the context of digital objects that are accessible over the Internet. Typically, such an identifier is not only persistent but actionable:^[1] you can plug it into a web browser and be taken to the identified source.

An important aspect of persistent identifiers is that "persistence is purely a matter of service". (...) No identifier can be inherently persistent, however many persistent identifiers are created within institutionally administered systems with the aim to maximise longevity."

And from there we already get many distinct approaches and services:

People and organisations:Uniform Resource Identifiers:Open Researcher and Contributor ID (ORCID)Digital Object Identifier (DOI)Research Organization Registry (ROR)Magnet link^[5] (decentralized, with BitTorrent)Publications:Archival Resource Keys (ARKs)Virtual International Authority File (VIAF)Uniform Resource Names (URNs)International Standard Name Identifier (ISNI)Extensible Resource Identifiers (XRIs)International Standard Book Number (ISBN)Persistent Uniform Resource Locators (PURLs)

PID discussion II:

Definition of PID (Persistent IDentifier) from the FDO people:

narrow the broad PID definition down to a specific system (based on their FDO paradigms): digital objects (DO) = bitsequences, DO Interface Protocol (DOIP), Repository System, Registry System (and items below)

(<u>https://fairdigitalobjectframework.org/</u>)

Identifier/Resolution Protocol (IRP): "*a rapid-resolution protocol for creating, updating, deleting, and resolving identifiers that are globally managed and allotted. Each identifier is associated with a record that clients can resolve to using this protocol.*".

Identifier/Resolution System: the system enables:

"allotment of unique identifiers to information in digital form structured as digital objects regardless of the location of such information or the technology used to serve such information;"

"the resolution of the identifiers to current state information about the corresponding digital object, e.g., its location(s), access & usage policies, timestamps, and/or public keys."

Problem:

- basically no scientific community can apply this 'ab ovo'
- it is (more or less) not dealing with many stages of the scientific production process
- abstracts away the specifica of the community processes (false generalisation of the interoperability)

PID discussion III:

PID (Persistent IDentifier) in scientific research processes:

1. assumption: a scientific research process comprises of at least (roughly) three stages:

- 'experimental phase': testing hypotheses without already a clear concept of how to organise results
- 'consolidation phase': processing data/ separating instrument and signal etc.,/ sharing and scrutinising results within collaboration
- 'publishing phase': publishing conclusions and data.
- 2. assumption: each stage will result in different data sets and identifiers for data
 - dependent on the community, its data structures and sets, the first two stages will be mostly using ideosyncratic identifiers and also metadata systems
 - dealing with these data also corresponds to different levels of scientifc competence in the specific field
 - interdisciplinarity at the first two stages is wasted effort (this does not preclude to learn about good solutions, but their application remains a task of the field)

3. assumption:need for systems of PID (and Metadata Schemata) to enable cross-disciplinarity for third level

- PUNCH approach:
 - use the already developed systems where possible
 - acknowledge the use of
 - the mix of externalized Metadata schemata (e.g. DOI) and community Metadata (e.g. UCD)
 - the implicit metadata built into well established software procedures in a community
 - encapsulating data/software/workflows wherever possible
 - try to combine this in a registry for the platform