Goals for Persistent Identifiers

Goals

Note: Questionable goals are in *red*

For Identifiers (just the complete URI and local part, not about mechanisms)

1. Identifiers must be permanent
2. Identifiers must permanently resolve to a digital object, which may be surrogate for physical object
3. Simplicity (of description and implementation)
4. The identifier itself needs to support two parts: a local namespace prefix and a (fairly) arbitrary identifier part
5. Makes sense in context of VIVO, arXiv, OAIS (CUL), Voyager Catalog, and WorldCat
6. Local part should not be Cornell branded so that it is compatible with multi-institution collections where Cornell branding is inappropriate. Generally, these identifiers should be branded to Cornell through the use of a Cornell-based DNS name for the full URI/resolver. It should also be possible to use the local part with non-Cornell resolver for consortia and other non-Cornell-branded needs (cf. DOI, handle; not our purl)
7. Identifiers should be short (so use 26 letters plus numbers)
8. Identifiers should be easy to copy by hand - separate every 4 characters with a dash
9. It must be possible to support opaque identifiers
10. Must be resolvable through a web browser
11. Must be unique within our PID system without the DNS name portion of the URL
12. Recommended that newly created local identifier namespaces use identifiers that include a check "digit" (so that typing errors are not likely to result in a valid identifier).

For Resolver and System

1. Supports billions of identifiers with very fast resolution
2. Robust architecture and implementation - a highly available system
3. Need to avoid unbounded generation of surrogate persistent identifiers
4. There should be a separate service request using the identifier as a parameter that gives metadata about the identifier
5. Need a lightweight understanding of identifier equivalence
6. The identifiers and the associated content should be easily discoverable by Google
7. The overall system should integrate well with the "web architecture"
8. Should have a PID corresponding to every Cornell NetID, using 303 redirect to an appropriate surrogate.
9. Should give useful error when check digit is wrong.
10. Support for vitality checking

Out of scope / Not goals

1. Any extra characters for error correction (as an extension of any possible error detection)
2. Attempt to create a fixed-length of fixed-syntax form (that would aid recognition but cost in flexibility/length)
3. Support for "private" identifiers (e.g. for dark archive or internal digital objects) (this is about metadata stored with id and facilities provided depending on it)