Vocabulary practices

- Building of ontologies in the US seems to be mature example
- Take different viewpoints:
    - **Business use cases**
        - How to help communities to create vocabs
            - Importance of ease of use and interoperability. A huge vocab of everything not tractable.
        - ANDS services - 3 systems
            - RVA Portal - Discovery of vocabs for use, for easy access
            - SISSVoC - repository for visualisation
            - Pool Party - editing suite for vocabs
            - Low-level access for developers/geeks
            - Systems put together after approaches by several institutions
        - Supply a range of tools for users
            - E.g. easy creation of UI lists
        - Various levels of sophistication for governance
            - What is the right level? What is good governance?
        - Notification of changes of terms
            - 'Notify any changes to terms into collections'
        - Who is responsible for tracking the crosswalks?
        - Use of similarity measures rather than direct matches (don't over simplify)
        - Quantitative measure of relationship strength
        - Rely on the domain experts to resolve similarity of terms
            - 'Make your definitions born linked'
        - Cross-domain vocabs are hard
            - e.g. how to get the sub-set that you want to use
            - Or problem when for example - observed property not sitting in the right domain.
        - Authority; who signs off on vocabularies? Can be informal
    - **Example cases**
        - Subsetting terms
        - Observed properties, units of measure - how to make use of a range of vocabularies? E.g. link between NEII, ALA
        - Ease of use in finding vocabularies; need to create subsets.
        - How to build consensus on things? The RFC/internet example.
            - Reproducible process across many areas
            - RDA interest group vocabulary services specification (IETF)
            - Requirement for this group to identify needs for this RDA group?
            - Git pull request example
            - Two aspects of governance
                - Creation of the vocabularies (then subsetting and proxying)
                - Changes and ongoing management of vocabularies
    - **Key points**
        - We need to support some of these use cases with the infrastructure
            - No complete solutions yet
            - Subsetting of vocabularies
            - Observable properties space
        - Discovery and visualisation of vocabularies.
        - Community Ownership
        - This group to provide input into RDA interest group vocabulary services specification (bring together NCRIS capabilities in this space)
        - Work happening in OGC and GEO definitions space

CODATA recommendations

- **Doing an audit of existing vocabularies - create a list**
    - Create a star rating assessment
        - Community needs will be different, one star will be ok for some
    - How can different unions help each other to bring up/improve sophistication of vocabularies?
    - Involve RDA, OGC etc
    - Commonly don't look to re-use, this should be encouraged
- Multi-lingual support
- Is usage of vocabularies a metric?
    - E.g. my friends/community are already using this vocab
- Support content negotiation for different profiles
- In the marine space - international re-use is common
- ESIP is taking on MMI initiative
    - Any recommendations/experiences
- Sustainability and designing URIs that are flexible
    - E.g. if NASA gives up SWEET, what happens to sweet URIs?
    - For increased adoption and re-location - good to have generic name spaces
- Any working groups that are looking at transition plans for end of life of vocabs/URIs?
    - Responsible governance
    - Coordination of approaches
    - Agile infrastructure
    - **Define a best practice for succession**
- Use international groups who have resilience to changes in funding
- Guidelines to allow organisations to push back against bureaucracy
    - E.g. URI schemes with org names embedded

- How to avoid complexity?
  - Can be complexity behind the scenes, but risk is that if you make lots of simple solutions the aggregate becomes more complex