To Summarize…

Metadata is...

Metadata is Wonderful!

- If it is created
- If it is created and collected reliably
- If it is created and collected reliably by competently trained persons
- If it is created and collected reliably by competently trained persons who follow some sort of industry approved dictionary (RDF)

Metadata is…

Cheap & Easy!

Metadata is Cheap and Easy

- If it is generated by computers
- If it is generated by computers, captured, and stored reliably
- If it is generated by computers, captured, and stored reliably in industry-standard formats

Metadata is...

Metadata is Consistent

- If it conforms with industry-standard schemas
- If it conforms with industry-standard schemas that have been vetted and approved by all the stakeholders
- If it conforms with industry-standard schemas that have been vetted and approved by all the stakeholders who agree to use a dictionary of descriptors that represent the same thing

Metadata is…

Metadata is Sharable

- If it meticulously follows industry standards
- If it meticulously follows industry standards and is made available via formats (like XML) that are popularly utilized.

Metadata is...

Metadata is Wonderful!

- But too few information producers supply it
- But one person's metadata is another's data
- But many questions remain about the most appropriate metadata to record or create
- But the dream of the Semantic Web is still moribund after 20 years…

Metadata is...

Metadata is Cheap and Easy

- But low-hanging fruit, like computer-generated metadata is often neglected
- But human catalogers need a great deal of training and experience to be accurate
- But employing human catalogers is expensive
- But double-checking the metadata adds a layer of cost most are unwilling to pay

Metadata is…

Metadata is Consistent

- But many organizations prefer to use their own schemas and dictionaries
- But it takes years, if not decades, to get stakeholders to agree on schemas and dictionaries
- But a degree of customization is required by many organizations, introducing free radicals
- But getting people to fill in the values correctly, if at all, can be like pulling teeth

Metadata is…

Metadata is Sharable

- But many prefer to use their own schemas and dictionaries
- But many stakeholders, by nature, "do not play well with others"
- But there are often considerably difficult security and rights issues to consider
- Export and import issues remain with several file format options and sporadic precision problems
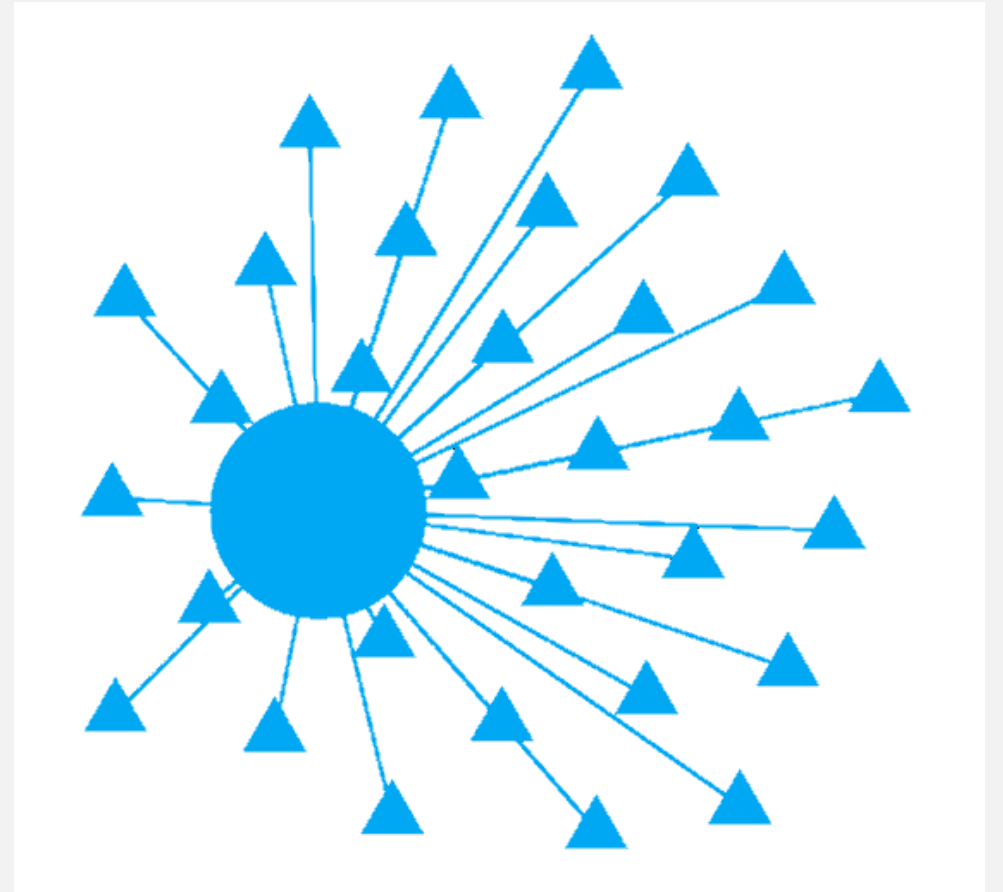
Uuuum… There are Humans Involved.

- As with most human enterprises, more stakeholders translate into more requirements

- It is damn hard (can take decades) to agree on standards

- Standards are always subject to change

- Users are not concerned about standards until they start getting bad results… long after the standards are set

- Designers (or users) have differing agendas, all of which are more important than others

- The only constant is change

Different Strokes for Different Folks

- How do we accommodate the world views and paradigms of different cultures and interest groups?

- Do the dominant cultures dominate?

- Can we record the alternative views? Then rank them?

- Who added the metadata and what is their background and training? Why are they adding metadata?

- Does the metadata help us to understand the legal and civil issues surrounding objects and their descriptors?

Computer-generated Concept Clouds

- A document is analyzed along with a thousand others bearing the same traits
- The document is then scored for similarities to the others
- Descriptors from its neighbors are applied to the document, with some measure of the similarity
- Descriptors added to the document are conversely applied to its neighbors
- Adds a third dimension:
entity – attribute – value - congruity

Things are Not as Simple as X = Y

- Recently, researchers around the world, including UNC and Duke, have noted that reducing blood volume, especially iron, has an effect on high blood sugars. (Leeches anyone?)

- How would you dial this up?

Bottom Lines

- Metadata is a wonderful idea that has yet to take full expression
- There's an enormous amount of work to be done in the field
- Still various organizations have made robust systems that meet a great deal of their needs, with a few small sacrifices.
- No two systems are exactly alike, but they share a lot of important commonalities.
- It takes a great deal of communication and interaction to develop a system that accommodates the widest range of stakeholders

Bottom Lines

As Winston Churchill would have said, if he weren't so drunk most of the time, *"Metadata is the worst form of organization, except for all the others."*

Bye Now!

- I wish you all the best!