

A Digital Curation Model focused on Semantic Enrichment

Hyewon Lee (Seoul Women's University)

Soyoung Yoon, Ziyoun Park (Hansung University)

Overview

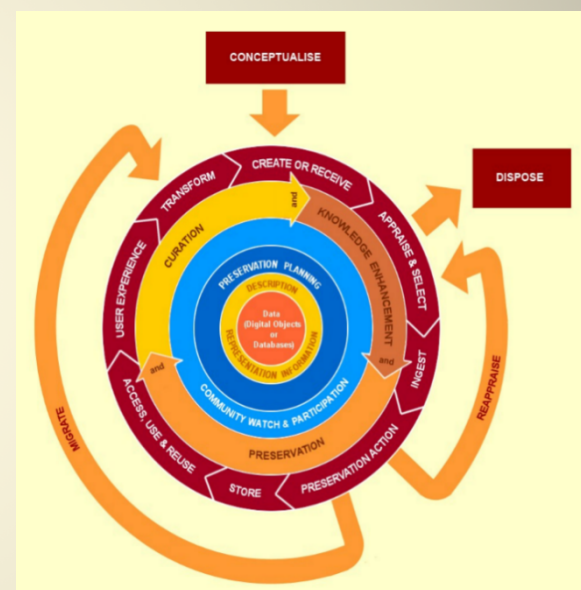
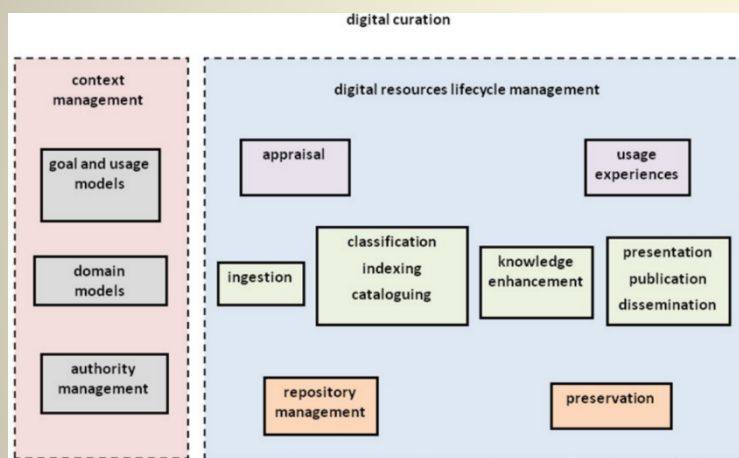
- Digital curation
 - to *maintain* current utilization and *add value* for future reuse.
 - interdisciplinary area
- Digital curation lifecycle model
 - the process similar to an organism's life cycle
 - UK Digital Curation Centre(DCC) Curation Lifecycle Model

Purposes

- Improving on the current DCC curation lifecycle model
 - focusing on the *description and representation* for digital objects(data and/or contents)
- Redesigned to
 - support the process of *adding value* to the core value of digital resources
 - consider the possibility of *accommodating external changes* to the agency's original functions
 - Move from digital objects-driven model to *user-centered model*

Description & Semantic Unit

- DCC&U Model: *Knowledge Enhancement* of DCC Model (Constantopoulos et al. 2009).



Knowledge enhancement

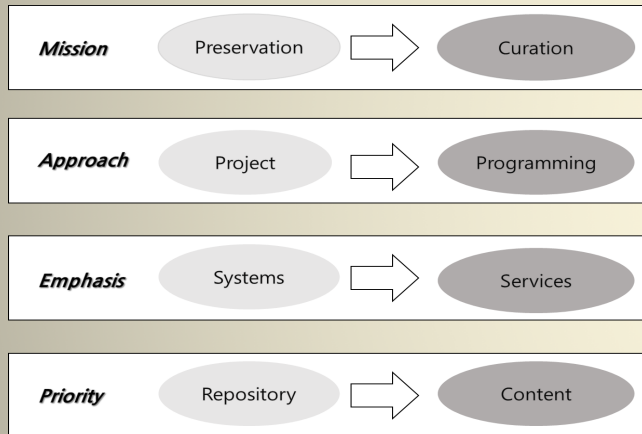
A process that refers to the real-world entities, situations, and events represented by digital resources, their wider context and domain, and the digital resources themselves

a modified lifecycle action

(Curate, Preserve, and Knowledge Enhancement) with information about the knowledge enhancement of data.

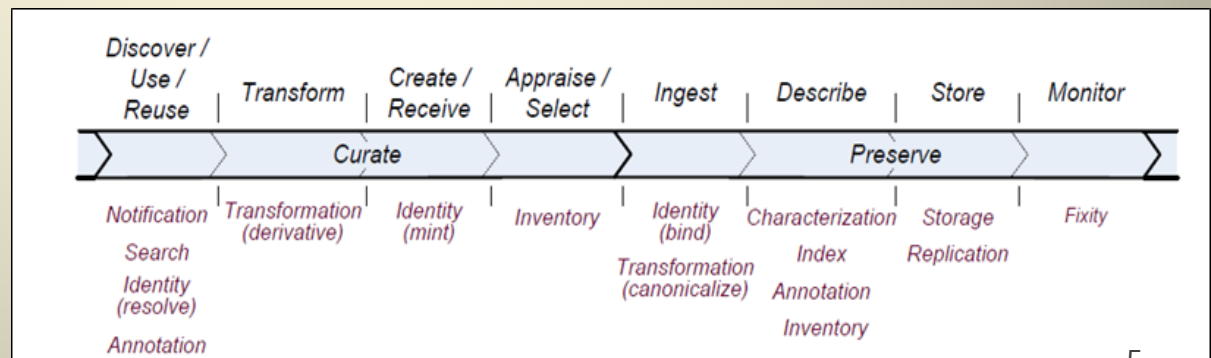
Description & Semantic Unit

- University of California Curation Center (UC3)
Micro Service (CDL 2010)



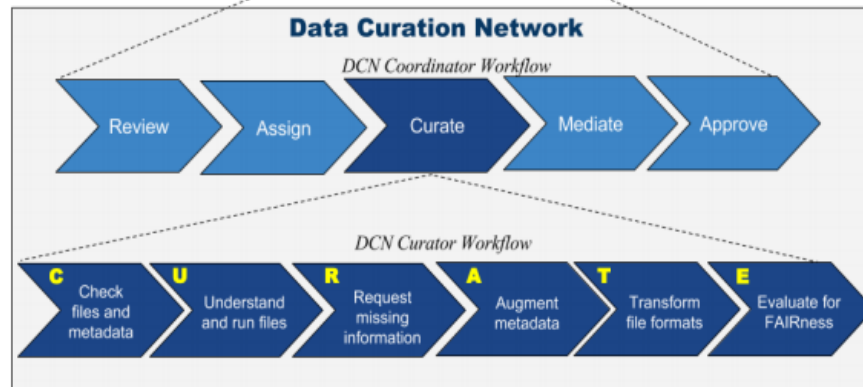
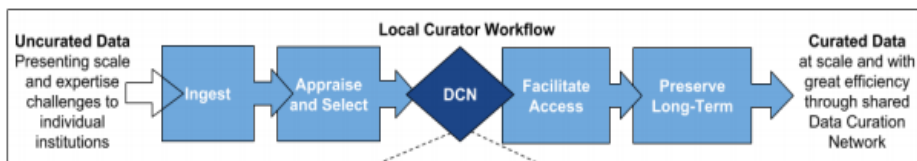
a process of reinvention involving ***significant transformations*** of the outlook, effort, and infrastructure

the semantic aspect through value, Service, and context in micro ***service*** stage



Description & Semantic Unit

- The Data Curation Network(DCN) (Johnston et al. 2017)



- ✓ **"Augment Metadata"** step is also represented semantic augmentation of the data.
- ✓ **DCN curators** take standardized and file type-specific actions when reviewing the data for fitness for reuse using their expert skills and domain-specific knowledge.

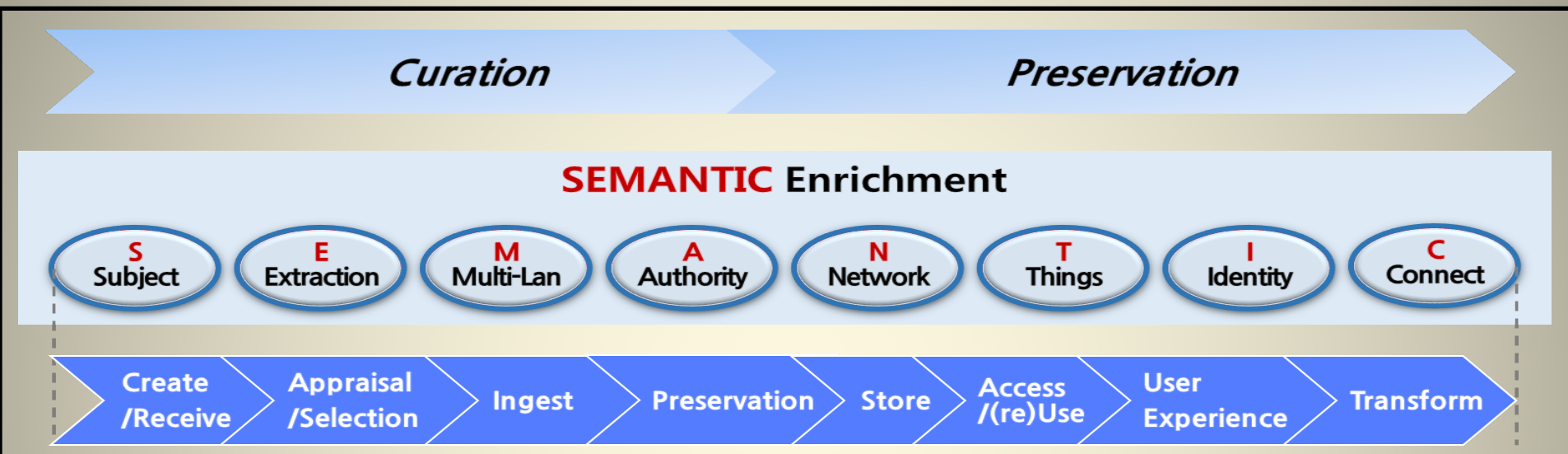
Semantic Enrichment of Digital Curation Model

- Analyzing the *description and representation information* action of the full life cycle action of the DCC model
- The purpose of *Semantic Enhancement*
 - Emphasizing a more **elaborate description** of the typical curation model
 - Supporting **user-centered services** as a way to utilize digital objects that are produced, managed, and preserved through their life cycle
 - Suggesting **considerations** that should be included in digital curation models of research institutes and universities
 - Assisting in the development of model that considers **links with external** institutions

Characteristics of Semantic Enrichment

- *affects all Sequential Actions* ;
 - Semantic Enrichment action in the Full Life Cycle
- carries the word "SEMANTIC," taken from *collecting the first letters* of the eight terms that express each property of description and representation information
- observes no meaning in the order of the spellings of SEMANTIC and is only *a single word for the concept*, highlighting the properties that description and representation should have

The Semantic Enrichment Action



- ✓ **Standardized metadata** can be applied to describe and control data and may be necessary to understand and process digital objects and their metadata.
- ✓ *The identifiers* have important implications in a digital curation model.
- ✓ The action manages *the name authorities and the subject authorities* that can be processed as properties of the digital objects.

SEMANTIC Enrichment

Subject

Build subject authority data

- Control alternative form and hierarchy description
- Integrate into thesaurus or ontology model

Extraction

Extract important attributes of content and data

- Bring out information that fulfills user needs
- Extract and express according to standards

Multi-Language

Construct Multi-lingual dictionary and thesaurus

- Express language notation for concept and present differentiation

Authority

Build authority data

- Gather and control the name of person and organization
- Control alternative form and hierarchy description
- Differentiate homonym, especially Korean people

SEMANTIC Enrichment

Network

Structuralize data connection

- Link between contents(ex. research reports and academic papers, academic papers and tables), link between data(ex. author and author, author and subject), link between contents and data (ex. academic papers and author)
- Use identifier

Thing

Include data and its description

- "Data" is the original information object created by the producer (author, contributor, etc.).
- "Description" is considered as metadata added in digital curation process.

Identity

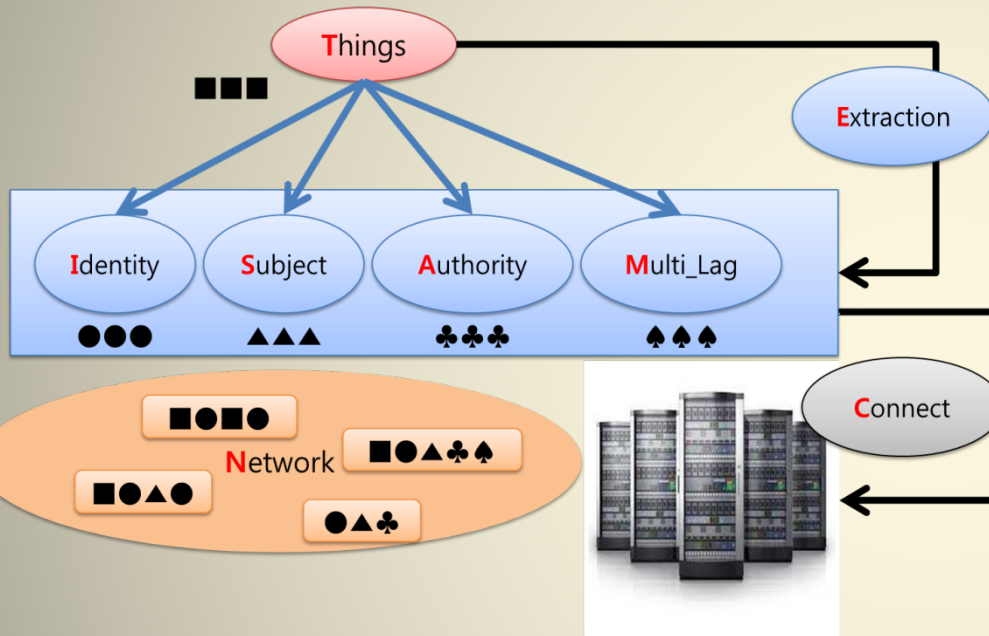
Identify unique character and symbol

Connect

Use data as a resource for various information systems

- Data is not dependent on any particular system or application and implies independent value.

SEMANTIC



- ✓ Individual digital objects(data and/or content) is considered a "**Thing**," and its attributes are **extracted** from the **identifier**, **subject**, **authority**, and variants(**multi-Language**).
- ✓ These attributes are then utilized in new services through a "**Network**."
- ✓ These new services are **connect**ed to an integrated system or other institutions' services or systems.

Conclusion

Showing how *external changes are accepted* with the unique values and functions of digital objects maintained.

Accommodating the needs of diverse users and *refining* the description and representation of digital objects accordingly

Presenting the *considerations* for the creation and management of the description and representation of digital objects in specific research areas and institutions

References

- CDL. UC3 Curation Foundations (2010). Retrieved from <<https://confluence.ucop.edu/download/attachments/13860983/UC3-Foundations-latest.pdf>>
- Constantopoulos, P., Dallas, C., Androutsopoulos, I., Angelis, S., Deligiannakis, A., Gavrilis, D., Kotidis, Y., & Papatheodorou, C. (2009). DCC&U: An Extended Digital Curation Lifecycle Model. *International Journal of Digital Curation*, 4(1), 34–45. <<https://doi.org/10.2218/ijdc.v4i1.76>>
- DCC Homepage A. Curation Lifecycle Model (n.d.). Retrieved from <<http://www.dcc.ac.uk/resources/curation-lifecycle-model>>
- Johnston, L. R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozlowski, W., Olendorf, R., & Stewart, C. (2017). Data Curation Network: A Cross-Institutional Staffing Model for Curating Research Data. Retrieved from <<https://conservancy.umn.edu/handle/11299/188654>>

THANK YOU

Hyewon Lee (*hwlee@swu.ac.kr*)

@HyewonMomoLee