

Establishing semantic interoperability of biomedical metadata registries using extended semantic relationships

Yu Rang Park, Young Jo Yoon, **Hye Hyeon Kim**, Ju Han Kim

Seoul National University Biomedical Informatics (SNUBI),
Division of Biomedical Informatics,
Seoul 110-799, Republic of Korea



AGENDA

- INTRODUCTION
 - ✓ Metadata and the ISO/IEC 11179 standard?
 - ✓ Limitation of ISO/IEC 11179
- METHODS
 - ✓ Semantic relationship
- RESULTS
 - ✓ Three types of semantic relationships
 - ✓ Evaluation
- DISCUSSION & CONCLUSION



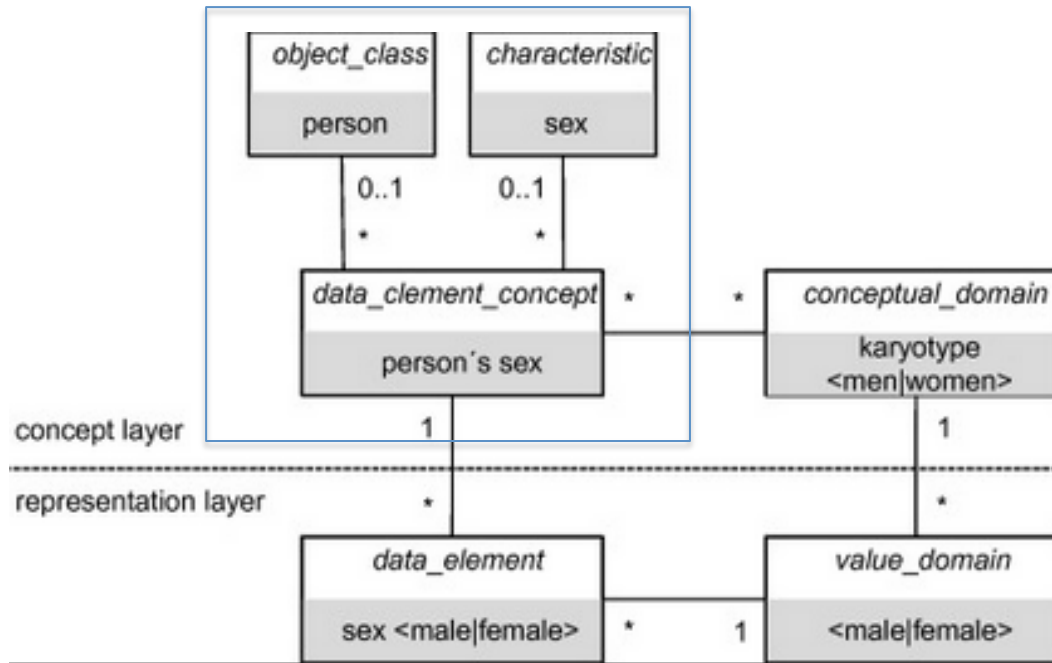
- Achieving semantic interoperability is fundamental and critical for sharing biomedical information.
- Metadata based on the ISO/IEC 11179 metadata standard (MDR) is considered one of the ways to solve term interoperability.
 - ✓ Representative MDR implementation is NCI's caDSR.
- However, several studies have been shown semantic and structural limitation of ISO/IEC 11179
 - ✓ Low term reusability: Most of new data elements, newly defined, are not from the existing data elements. It causes the high increase of data elements.
 - ✓ No representation of inter-relating concepts: Hard to represent the relationships among data elements.



LIMITATION OF ISO/IEC 11179

Fundamental reason for this issues

“Single semantic perspective” of ISO/IEC 11179



Definition of relationship is existed in ISO/IEC 11179 model

✓ User-specified relationships



ANSI/NISO Z39.19-2005

ISBN: 1-880124-65-3

Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies

Abstract: This Standard presents guidelines and conventions for the contents, display, construction, testing, maintenance, and management of monolingual controlled vocabularies. This Standard focuses on controlled vocabularies that are used for the representation of content objects in knowledge organization systems including lists, synonym rings, taxonomies, and thesauri. This Standard should be regarded as a set of recommendations based on preferred techniques and procedures. Optional procedures are, however, sometimes described, e.g., for the display of terms in a controlled vocabulary. The primary purpose of vocabulary control is to achieve consistency in the description of content objects and to facilitate retrieval. Vocabulary control is accomplished by three principal methods: defining the scope, or meaning, of terms; using the equivalence relationship to link synonymous and nearly synonymous terms; and distinguishing among homographs.

- Defining semantic relationship between concepts is a well-established topic in the development of controlled vocabularies.
- Three types of semantic relationships
 - ✓ Equivalency
 - ✓ Hierarchy
 - ✓ Associative
 - 1) Dependent
 - 2) Composite
 - 3) Variable



MATERIALS

College of American Pathologists

www.cap.org/apps/cap.portal?_nfpb=true&cntvwrPtl_t_actio

Commitment

cap

Home Advocacy Reference Resources and Publications Learning

Log In/Register

JOIN THE CAP

RENEW MEMBERSHIP

CAP Events

CAP '13 - THE Pathologists' Meeting™
October 13-16, 2013

House of Delegates
Residents Forum Meeting
October 12, 2013

View more events >>

Cancer Protocols

Updated May 28, 2013

CAP contributes to the International Collaboration on Cancer Reporting for the development of global evidence-based core data sets.

The International Collaboration on Cancer Reporting which includes Pathology Colleges and Associations Australia (representing a population of approximately 20 million people) was established in 2011 to examine the practicability of agreed and evidence-based cancer pathology data.

Visit the ICCR website

The College of American Pathologists
- The International Collaboration on Cancer Reporting

Copyright, Disclaimer and Authorized Use Licensing Information

Protocols	Download			
Breast				
DCIS – Breast UPDATED	PDF (357 KB)	Word (464 KB)	2009 Version (PDF, 215 KB)	
Posted: June 15, 2012				
Invasive Breast UPDATED	PDF (1.2 MB)	Word (2.3 MB)	2009 Version (PDF, 547 KB)	Work Aid (Word, 137 KB)
Posted: June 15, 2012				
Central Nervous System				
Brain/Spinal Cord UPDATED	PDF (217 KB)	Word (187 KB)	2012 Version (PDF, 189 KB)	
Posted: January 30, 2013				
Endocrine				
Adrenal Gland	PDF (815 KB)	Word (632 KB)	2011 Version (PDF, 818 KB)	
Posted: June 15, 2012				
Appendix NET UPDATED	PDF (513 KB)	Word (301 KB)	2011 Version (PDF, 149 KB)	
Posted: June 15, 2012				
Colon NET UPDATED	PDF (512 KB)	Word (305 KB)	2011 Version (PDF, 154 KB)	
Posted: June 15, 2012				
Pancreas (Endocrine) UPDATED	PDF (1.4 MB)	Word (1.8 MB)	2011 Version (PDF, 1.5 MB)	
Posted: June 15, 2012				
Small Intestine NET UPDATED	PDF (522 KB)	Word (304 KB)	2011 Version (PDF, 113 KB)	
Posted: June 15, 2012				
Stomach NET UPDATED	PDF (81 KB)	Word (100 KB)	2011 Version (PDF, 81 KB)	
Posted: June 15, 2012				
Thyroid UPDATED	PDF (2.4 MB)	Word (1.2 MB)	2011 Version (PDF, 272 KB)	Work Aid (Word, 165 KB)
Posted: June 15, 2012				
Gastrointestinal				
Ampulla of Vater UPDATED	PDF (722 KB)	Word (952 KB)	2011 Version (PDF, 374 KB)	
Posted: June 15, 2012				
Anus UPDATED	PDF (766 KB)	Word (1 MB)	2011 Version (PDF, 406 KB)	
Posted: June 15, 2012				

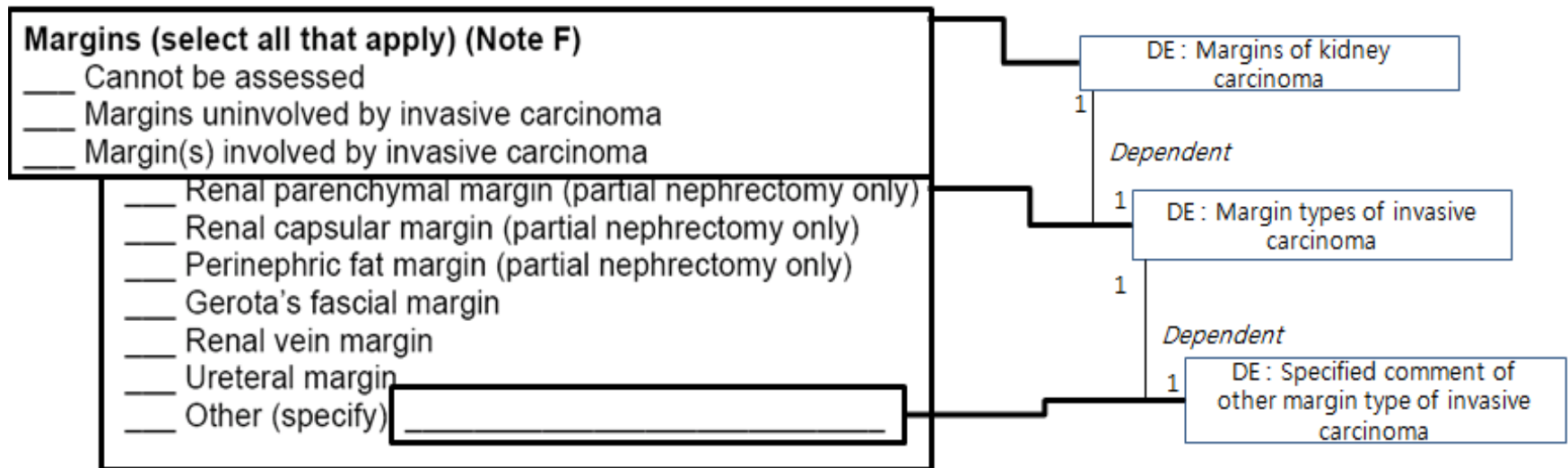


RESULTS (1/3)

- Three types of semantic relationships
 - Dependency relationship

CAP Approved

Genitourinary System • Kidney



- Composite relationship
- Variable relationship



RESULTS (2/3)

CAP Approved

Head and Neck • Larynx

■ T

Histologic Type (select all that apply) (Note B)

☐ Squamous cell carcinoma, conventional

Variants of Squamous Cell Carcinoma

☐ Acantholytic squamous cell carcinoma

☐ Adenosquamous carcinoma

☐ Basaloid squamous cell carcinoma

☐ Papillary squamous cell carcinoma

☐ Spindle cell squamous cell carcinoma

☐ Verrucous carcinoma

☐ Giant cell carcinoma

☐ Lymphoepithelial carcinoma (non-nasopharyngeal)

Neuroendocrine Carcinoma

☐ Typical carcinoid tumor (well differentiated neuroendocrine carcinoma)

☐ Atypical carcinoid tumor (moderately differentiated neuroendocrine carcinoma)

☐ Small cell carcinoma, neuroendocrine type (poorly differentiated neuroendocrine carcinoma)

☐ Combined (or composite) small cell carcinoma, neuroendocrine type

☐ Mucosal malignant melanoma

Carcinomas of Minor Salivary Glands

☐ Adenoid cystic carcinoma

☐ Mucoepidermoid carcinoma

☐ Low grade

☐ Intermediate grade

☐ High grade

☐ Other (specify): _____

☐ Other carcinoma (specify): _____

☐ Carcinoma, type cannot be determined

DE: Histology types of Larynx carcinoma

1

Composed with

1

DE: Histology type's variants of Squamous Cell Carcinoma

Composed with

1

DE: Histology types of Neuroendocrine carcinoma

Composed with

1

DE: Histology types of Minor Salivary Glands Carcinoma

CAP Approved

Histologic Type
☐ Squamous cell carcinoma, conventional

Variants of Squamous Cell Carcinoma
☐ Acantholytic squamous cell carcinoma
☐ Adenosquamous carcinoma
☐ Basaloid squamous cell carcinoma
☐ Papillary squamous cell carcinoma
☐ Spindle cell squamous cell carcinoma
☐ Verrucous carcinoma

Giant cell carcinoma
☐ Lymphoepithelial carcinoma (non-nasopharyngeal)

Neuroendocrine Carcinoma
☐ Typical carcinoid tumor (well differentiated neuroendocrine carcinoma)
☐ Atypical carcinoid tumor (moderately differentiated neuroendocrine carcinoma)
☐ Small cell carcinoma, neuroendocrine type (poorly differentiated neuroendocrine carcinoma)
☐ Combined (or composite) small cell carcinoma, neuroendocrine type

Mucosal malignant melanoma

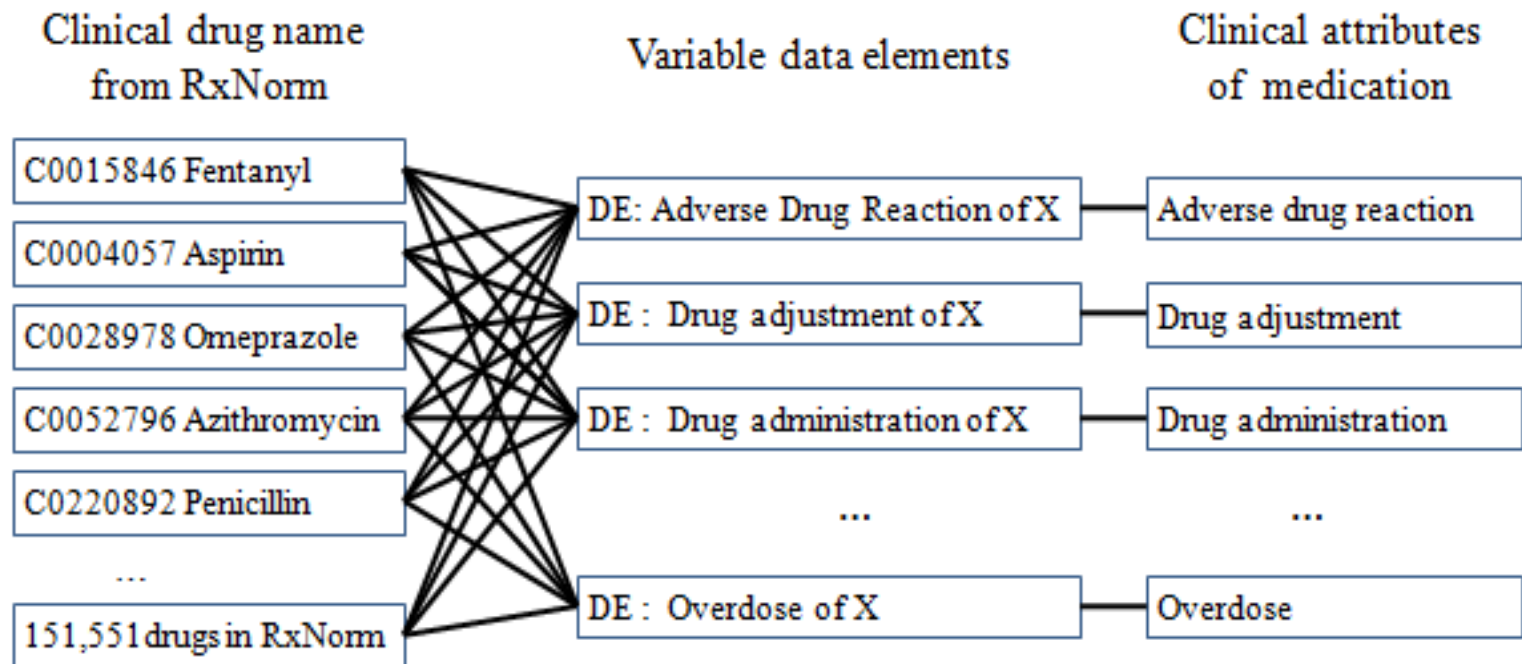
Carcinomas of Minor Salivary Glands
☐ Adenoid cystic carcinoma
☐ Mucoepidermoid carcinoma
☐ Low grade
☐ Intermediate grade
☐ High grade
☐ Other (specify): _____

Other carcinoma (specify): _____
Carcinoma, type cannot be determined

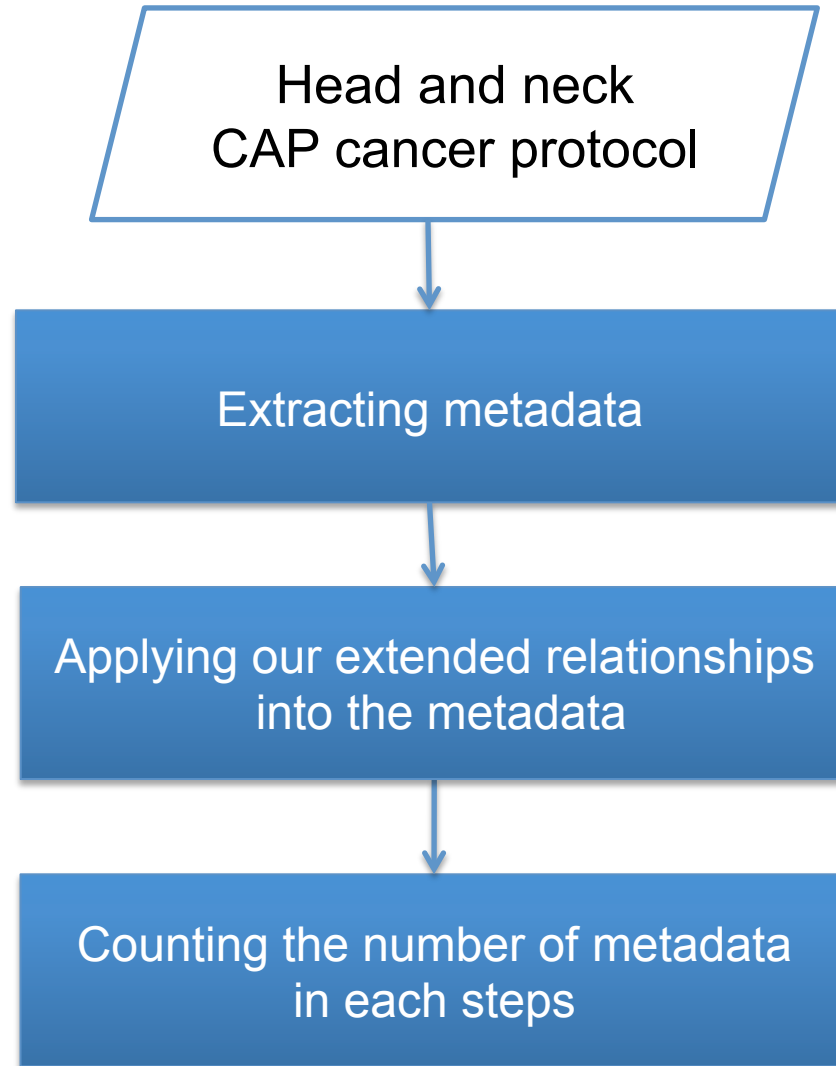


RESULTS (3/3)

- Three types of semantic relationships
 - Dependency relationship
 - Composite relationship
 - Variable relationship



EVALUATION (1/2)



*Classified 6 sub categories
55 cancer protocols*

EVALUATION (2/2)

CAP Cancer protocol	Larynx	Lip and Oral Cavity	Major Salivary Glands	Nasal Cavity and Paranasal Sinuses	Pharynx	Thyroid	Total
Metadata Extraction	79	85	59	81	89	91	484
Metadata integration	67	73	50	71	75	57	357
Application of semantic relationships	54	53	45	57	66	52	167

27%

66%

DISCUSSION & CONCLUSION (1/2)

- The ISO/IEC 11179 MDR standard has been recognized as one of the most powerful solutions for achieving semantic interoperability in biomedical domains.
- Several researches, however, demonstrated the semantic and structural limitations of the MDR standard.
 - ✓ Especially, MDR does not define or register the associative relationships between data elements.
- Our study introduced semantic extensions for MDR by defining new semantic relationships: 1) dependency, 2) composite, 3) variable relationships.
- To evaluate the semantic relationships, we brought up the use of CAP cancer protocols and compared the number of raw metadata, integrated metadata, and metadata after the application of the semantic relationships.



DISCUSSION & CONCLUSION (2/2)

- We assumed a decrease in the number of metadata can be translated to improved efficiency as showing high reusability
- As a future work, we will 1) bring up more direct method of evaluation to demonstrate improved efficiency and utility, and 2) consider the trade-off between the comprehensiveness and complexity



THANK YOU!

Hye Hyeon Kim: hyehyeon2@snu.ac.kr