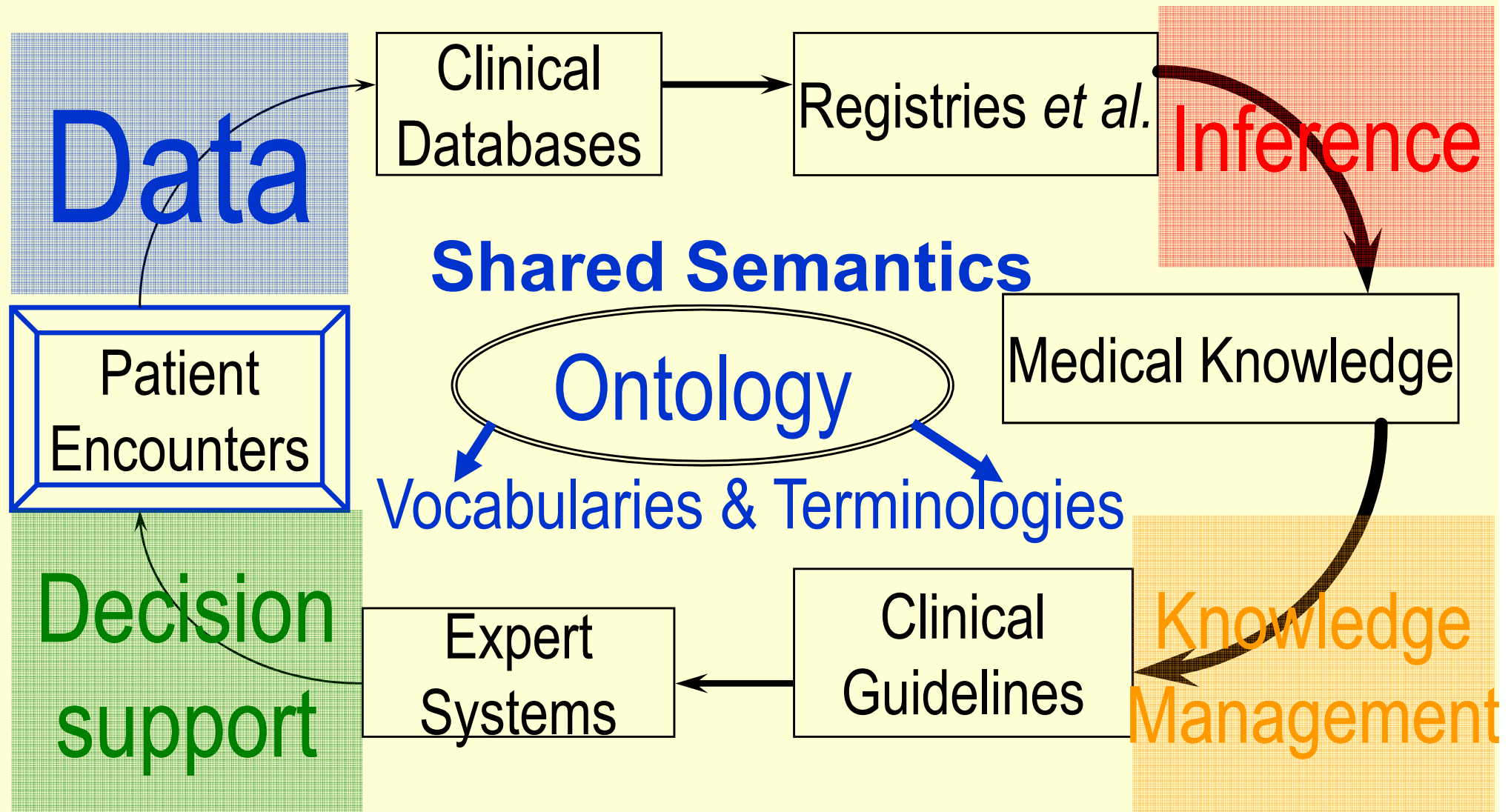


# **Terminologies and Classifications Health Data Interoperability and the ICD**

Christopher G. Chute, MD DrPH  
Professor, Biomedical Informatics  
Mayo Clinic College of Medicine  
Rochester, Minnesota

10<sup>th</sup> International HL7 Interoperability Conference  
Kyoto, May 2009

# From Practice-based Evidence to Evidence-based Practice



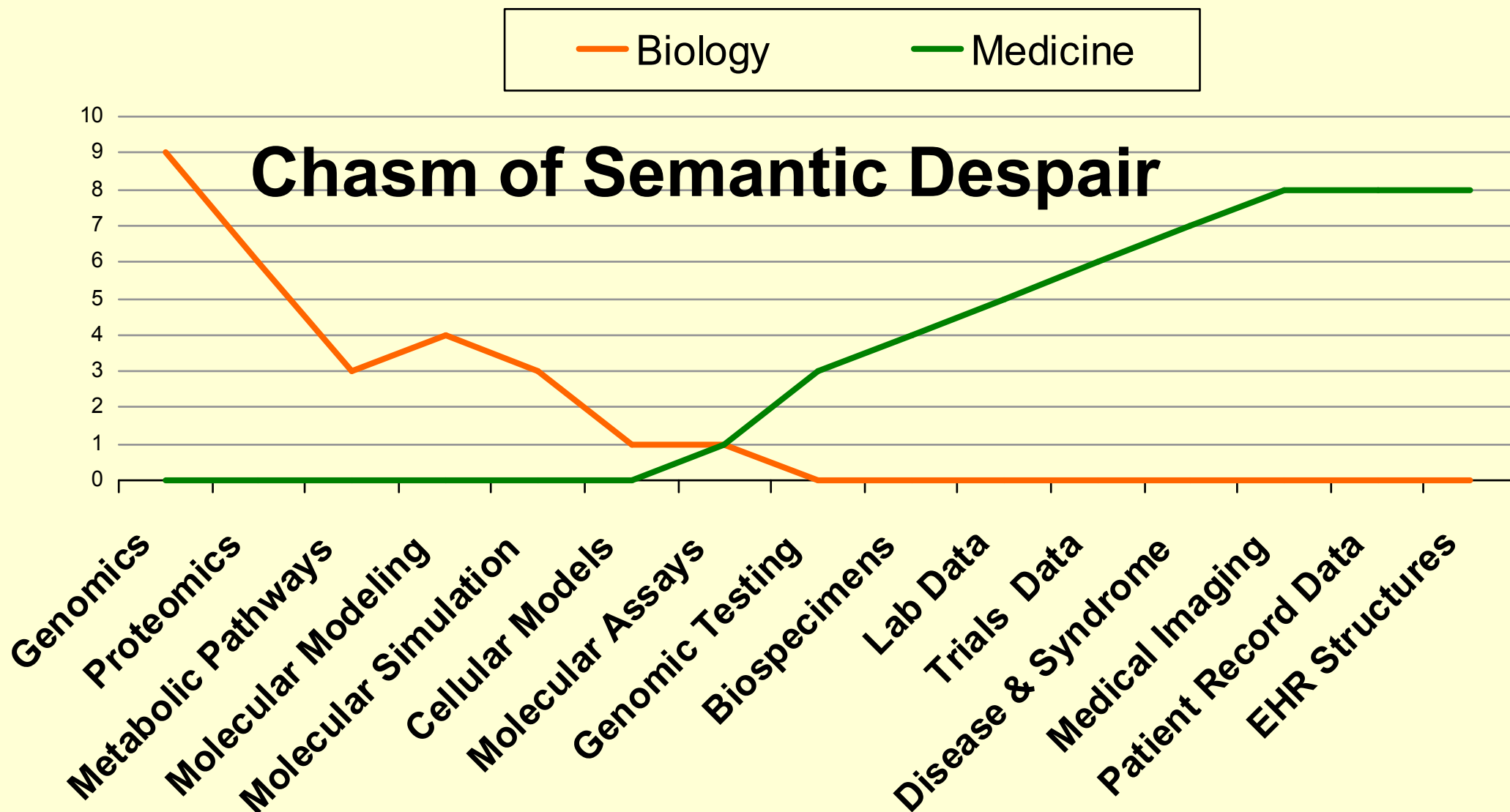
## **Blois, 1988**

### **Medicine and the nature of vertical reasoning**

- Molecular: receptors, enzymes, vitamins, drugs
- Genes, SNPs, gene regulation
- Physiologic pathways, regulatory changes
- Cellular metabolism, interaction, meiosis,...
- Tissue function, integrity
- Organ function, pathology
- Organism (Human), disease
- Sociology, environment, nutrition, mental health...

# The Continuum Of Biomedical Informatics

## Bioinformatics meets Medical Informatics



# Whither Phenotype?

## Spans spectrum from enzymes to disease

- Pharmacogenomics – enzyme functionality
- Physiologist – cellular function
- Systems biologist – pathway circuit flow
- Sub-specialist – organ functioning
- Patient/Clinician – disease manifestation
- Public Health – population characteristics

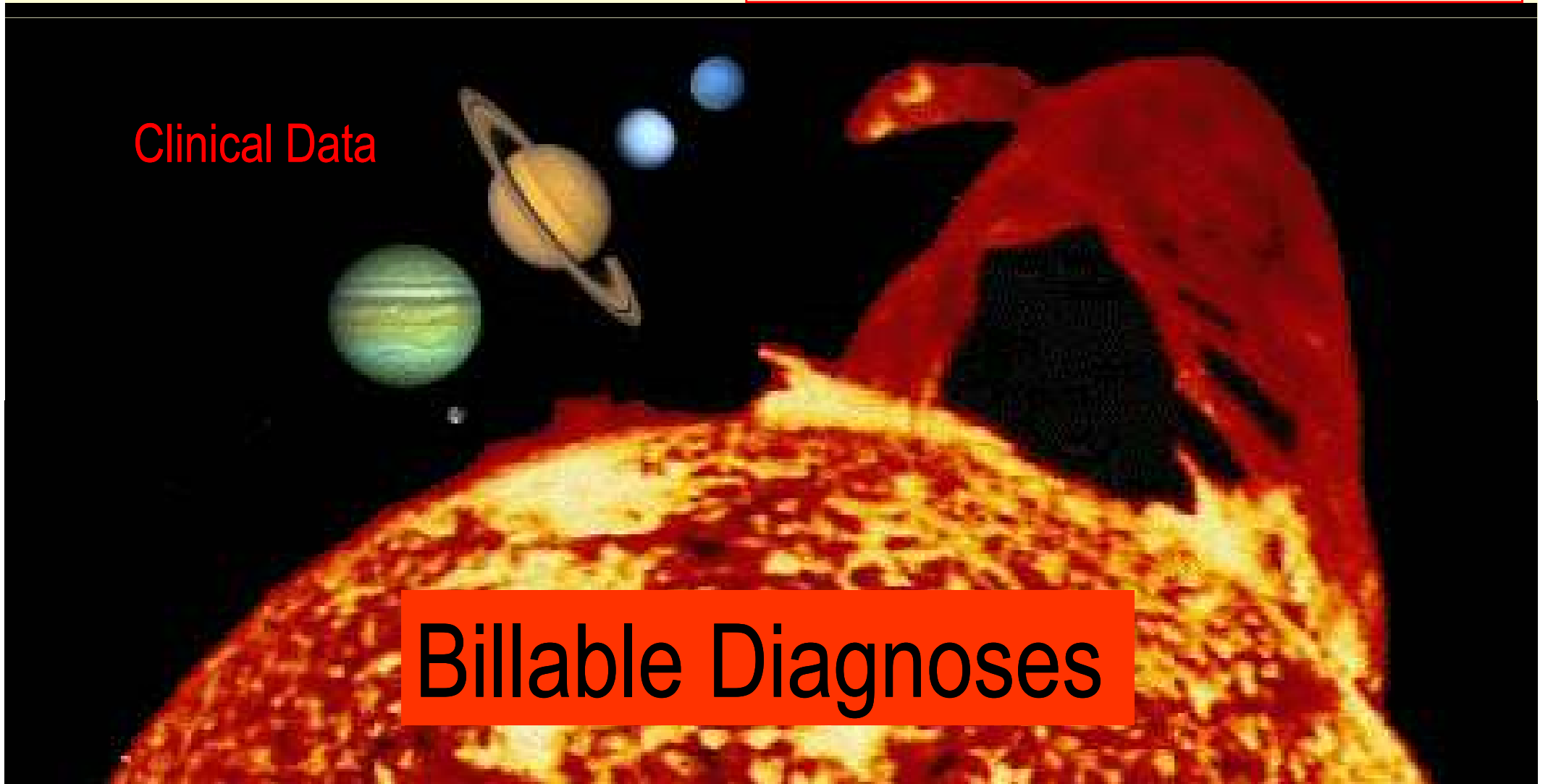
*Highly specific to use-case context*

# The Historical Center of the Health Data Universe

**Billable Diagnoses**

Clinical Data

**Billable Diagnoses**



# Copernican Healthcare

*(Niklas Koppernigk)*

Clinical Data

Billable Diagnoses

Clinical Guidelines

Medical Literature

Scientific Literature

Clinical Data

# First Formal Health Data Model

## London Bills of Mortality

- Commissioned 1542 (1598)
- Intended to Track Plague (Black Death)
  - ~60 disease categories
    - Various Defined
  - Collected by Parish Clerks (Chanters)
  - Irregularly Printed in Folio
  - Data Table Layout
    - 16<sup>th</sup> Century Spreadsheet



*Natural and Political*  
**OBSERVATIONS**

Mentioned in a following INDEX,  
and made upon the  
Bills of Mortality.

1910  
1665  
245.

---

BY  
Capt. **JOHN GRAUNT**,  
Fellow of the *Royal Society*.

---

With reference to the *Government, Religion, Trade, Growth, Air, Diseases*, and the  
several Changes of the said CITY.

— *Non, me ut miretur Turba, laboro,*  
*Contentus paucis Lectoribus.* —

---

The Third EDITION,  
much Enlarged.

---

LONDON,  
Printed by *John Martyn*, and *James Allestry*,  
Printers to the *Royal Society*, and are to be sold at the  
sign of the *Bell* in *St. Pauls Church-yard*.  
MDC LX V.



*The Table of CASUALTIES.*

[illegible]



# The Table of CASUALTIES.

The Years of our Lord	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1661	1662	1663	1664
Abortive and Stil-born	335	329	327	351	389	381	384	433	483	419	463	467	421	544	499	439	410	4
Aged	916	835	889	696	780	834	864	974	743	892	869	1176	909	1095	579	712	661	6
Ague and Fever	1260	884	751	970	1038	1212	282	1371	689	875	999	1800	2303	2148	956	1091	1115	11
Apoplex and Suddenly	68	74	64	74	106	111	118	86	92	102	113	138	91	67	22	36		
Bleach			1	3	7	2				1								
Blasted	4	1			6	6			4		5	5	3	8	13	8	10	
Bleeding	3	2	5	1	3	4	3	2	7	3	5	4	7	2	5	2	5	
Bloody Flux, Scouring and Flux	155	176	802	289	833	762	200	386	168	368	362	233	346	251	449	438	352	3
Burnt and Scalded	3	6	10	5	11	8	5	7	10	5	7	4	6	6	3	10	7	
Calenture	1			1		2	1	1			3							
Cancer, Gangrene and Fistula	26	29	31	19	31	53	36	37	73	31	24	35	63	52	20	14	23	
Wolf				8														
Canker, Sore-mouth and Thrush	66	28	54	42	68	51	53	72	44	81	19	27	73	68	6	4	4	
Child-bed	161	106	114	117	206	213	158	192	177	201	236	225	226	194	150	157	112	1
Chrisoms and Infants	1369	1254	1065	990	1237	1280	1050	1343	1089	1393	1162	1144	858	1123	2596	2378	2035	22
Colick and Wind	103	71	85	82	76	102	80	101	85	120	113	179	116	167	48	57		
Cold and Cough							41	36	21	58	30	31	33	24	10	58	51	
Consumption and Cough	2423	2200	2388	1988	2350	2410	2286	2868	2606	3184	2757	3610	2982	3414	1827	1910	1713	17
Convulsion	684	491	530	493	569	653	606	828	702	1027	807	841	742	1031	52	87	18	2
Cramp			1														1	
Cut of the Stone		2	1	3		1	1	2	4	1	3	5	6	4				
Dropfie and Tympany	185	434	421	508	444	556	617	704	660	706	631	911	646	872	235	252	279	2
Drowned	47	40	30	27	49	50	53	30	43	49	63	60	57	48	43	33	29	
Excessive drinking			2															
Executed	8	17	29	43	24	12	19	21	19	22	20	18	7	18	19	13	12	
Fainted in a Bath					1													
Falling-Sickness	3	2	2	3		3	4	1	4	3	1		4	5	3	10	7	
Flox and small Pox	139	400	1190	184	525	1279	139	812	1294	823	835	409	1523	354	72	40	58	5
Found dead in the Streets	6	6	9	8	7	9	14	4	3	4	9	11	2	6	18	33	20	
French-Pox	18	29	15	18	21	20	20	20	29	23	25	53	51	31	17	12	12	
Frighted	4	4	1		3		2		1	1				9	1			

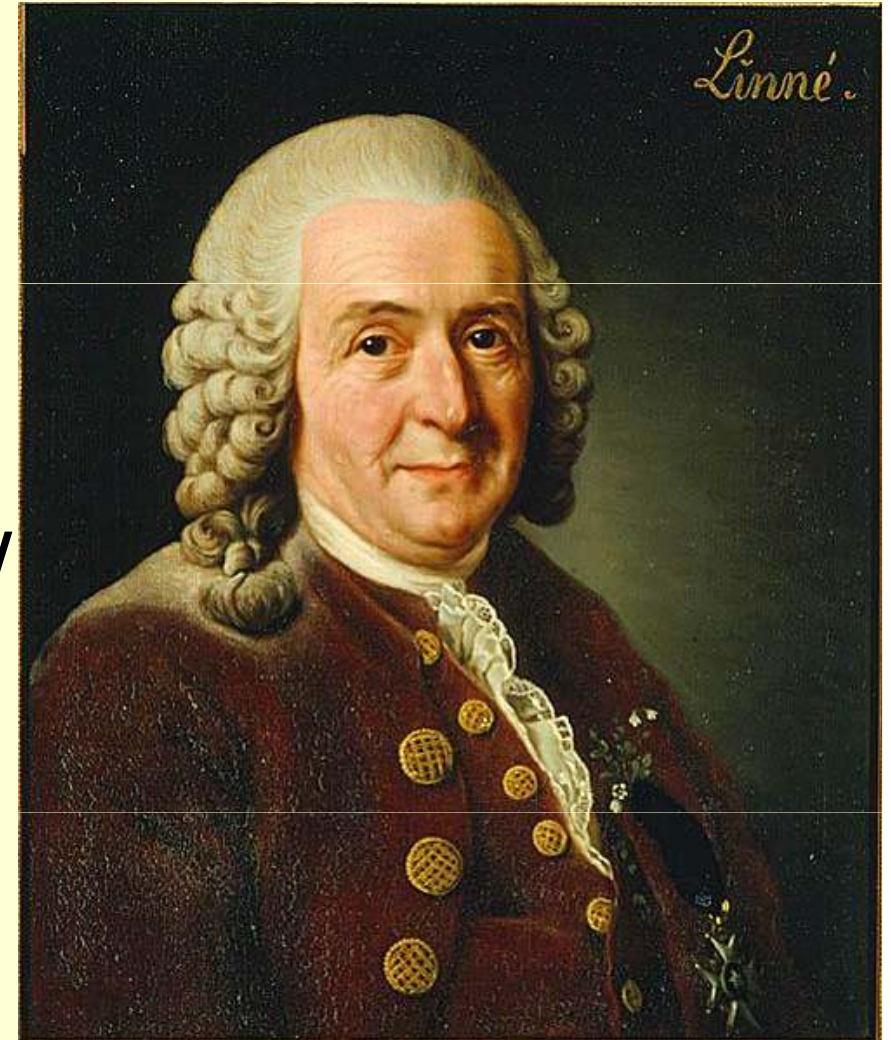


# Inferences from 16th Century Data Models

- John Graunt, First Epidemiologist - 1662
  - Estimate of Population
  - Age-specific Mortality
  - Disease Specific Mortality
  - Foci of Mortality
  - Recognize *Epidemic and Endemic Disease*
  - Notions of *Etiology and Causation*
- Modern Epidemiology and Demographics
- Work ignored for 200 years (Farr)

# Flawed Information Model

- Carolus Linnaeus  
Carl von Linné
  - *Genera Morborum* (1763)
- Underscored Content Difficulty
  - Pathophysiology vs  
Manifestation  
e.g. Rabies as psychiatric  
disease
- Lacked the Germ Theory of Disease
  - Was not incorporated into an information model



# Weights and Measures

“The nomenclature is of as much importance in this department of inquiry, as weights and measures in the physical sciences, and should be settled without delay.”

- William Farr, about Cullenian system
- First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England. London: 1839 p. 99.

# The Birth of a Formal Data Model for Cause of Death Classification

- First Statistical Congress, 1853 (Brussels)
  - William Farr (London)
  - Marc d'Espine (Geneva)
  - Achille Guillard (Paris) - Sponsor
- First International List (139 rubrics)
  - *Il y a lieu de former une nomenclature uniforme des causes de décès applicable a tous les pays.*
    - Sixteenth Annual Report of Registrar-General of England, 1853, Appendix, p. 73.

# Competing Concept Models about Cause of Death Classifications

- Farr – Anatomical Organization vs.
- d'Espine – Systemic Disease (Humors)
  - Necessary to Pick One
- Codes Linked to Disease Order
- Single Hierarchy Structure
- Vague Distinction between Term and Concept
  - Multi-Lingual (French Master)
- Revisions:
  - 1857, 1855, 1864 (model of Farr), 1874, 1880, 1886



# Jacque Bertillon

## ICD-1

- Introduced 1886 version in Chicago, 1893
  - International Statistical Institute (161 rubrics)
  - American Public Health Assoc. adopted 1898
  - Basis for decennial census
- Bertillon, Permanent Secretary General
  - International Commission (French Government)
  - Decennial revisions 1900, 1909, 1920 (Paris)

## ICD *not* Catalog of Disease

“The International List of Causes of Death makes no pretension of being a proper nomenclature of diseases or of including a scientific classification of diseases.”

Introduction ICD-2, 1909

# Post-Bertillon ICD

- French Government continued sponsorship
  - (1929, 1938) 5th revision: 200 rubrics
- WHO assumed sponsorship, 1948. (6th rev)
- Greatly expanded to include morbidity
  - 612 diseases and morbid conditions
  - 153 external causes of injury
  - 189 categories of injuries, by lesion, nature
  - decimal sub-divisions
  - 1955, 7th revision (minor)
- Pressure to Index Hospital Patients

# ICD-9-CM

## Resolution?

185 Malignant neoplasm of prostate

- 232,000 cases in the US each year
- Cannot distinguish stage or extent of disease

E845.9 Accident involving spacecraft injuring other person

- No person in history has ever died from this
- Not to be confused with death of astronaut or ground crew

# Combination Platter Coding in Diabetes Mellitus

- 250.0 DM without mention complications
- 250.1 Diabetes with ketoacidosis
- 250.2 Diabetes with hyperosmolarity
- 250.3 Diabetes with other coma
- 250.4 Diabetes with renal manifestations
- 250.5 Diabetes with ophthalmic manifestations
- 250.5 Diabetes with neurological manifestations
- 250.7 Diabetes with peripheral circulatory disorders
- 250.8 Diabetes with other specified manifestation
- 250.9 Diabetes with unspecified complications

**Severity?**

# ICD-10

## Improved Resolution?

C61 Malignant neoplasm of prostate

- Alphanumeric format
- No substantial improvement in resolution
- Continues tabular (non-axial, unidimensional) format

# The ICDs

1853	Farr/d'Espine	139	
1893	Bertillon	161	
1900	ICD 1	179	
1909	ICD 2	189	
1920	ICD 3	205	
1929	ICD 4	214	
1938	ICD 5	200	
1948	ICD 6	954*	*Decimal subdivisions
1955	ICD 7	965*	
1968	ICD 8	1,040*	
1968	H-ICDA2	905	4,334
1975	ICD 9	1,164*	8,173
1979	ICD-9-CM	1,179	14,473
1993	ICD 10	1,967*	10,468
2007	ICD-10-CM		68,000

# ICD11 Use-Cases

- Scientific Consensus on Clinical Phenotype
- Mortality
- Morbidity
- Case Mix
- Primary Care
- Quality and Safety



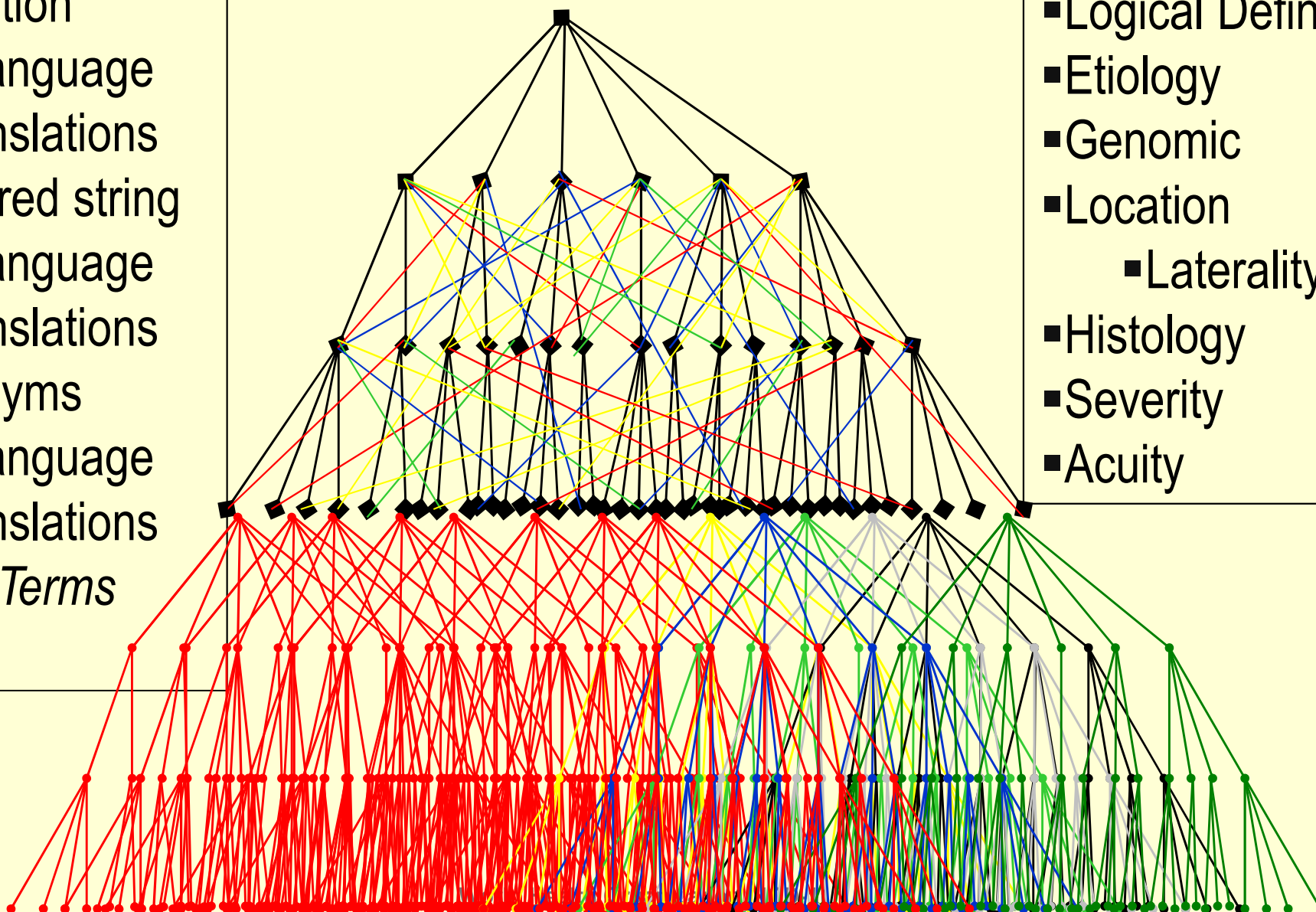
# ICD11 - An Enriched Information Model

## Concept name

- Definition
  - Language translations
- Preferred string
  - Language translations
- Synonyms
  - Language translations
- *Index Terms*

## Relationships

- Logical Definitions
- Etiology
- Genomic
- Location
  - Laterality
- Histology
- Severity
- Acuity



- **Manifestation Attributes**
  - Symptoms
  - Signs
  - Diagnostic results
  - Functional impact
- **Etiology**
  - causal agents
  - mechanism
  - Genomic characteristics
- **Temporal Relations**
  - chronicity (including acute)
  - episodicity
- **Severity and/or Extent**
- **Hierarchical relationships (parents and children)**

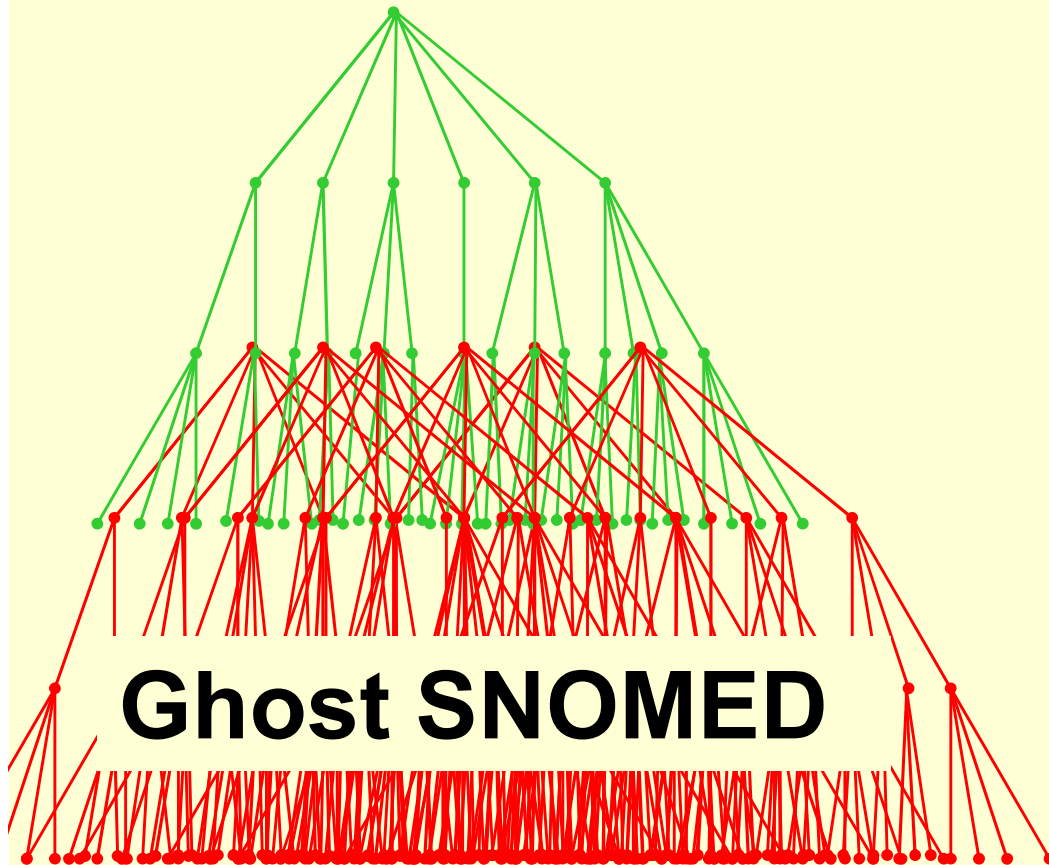
# Discussions with IHTSDO

## International Health Terminology (IHT)

- IHT (SNOMED) will require high-level nodes that aggregate more granular data
  - Use-cases include mutually exclusive, exhaustive,...
  - Sounds a lot like ICD
- ICD-11 will require lower level terminology for aggregation logic definitions
  - Detailed terminological underpinning
  - Sounds a lot like SNOMED

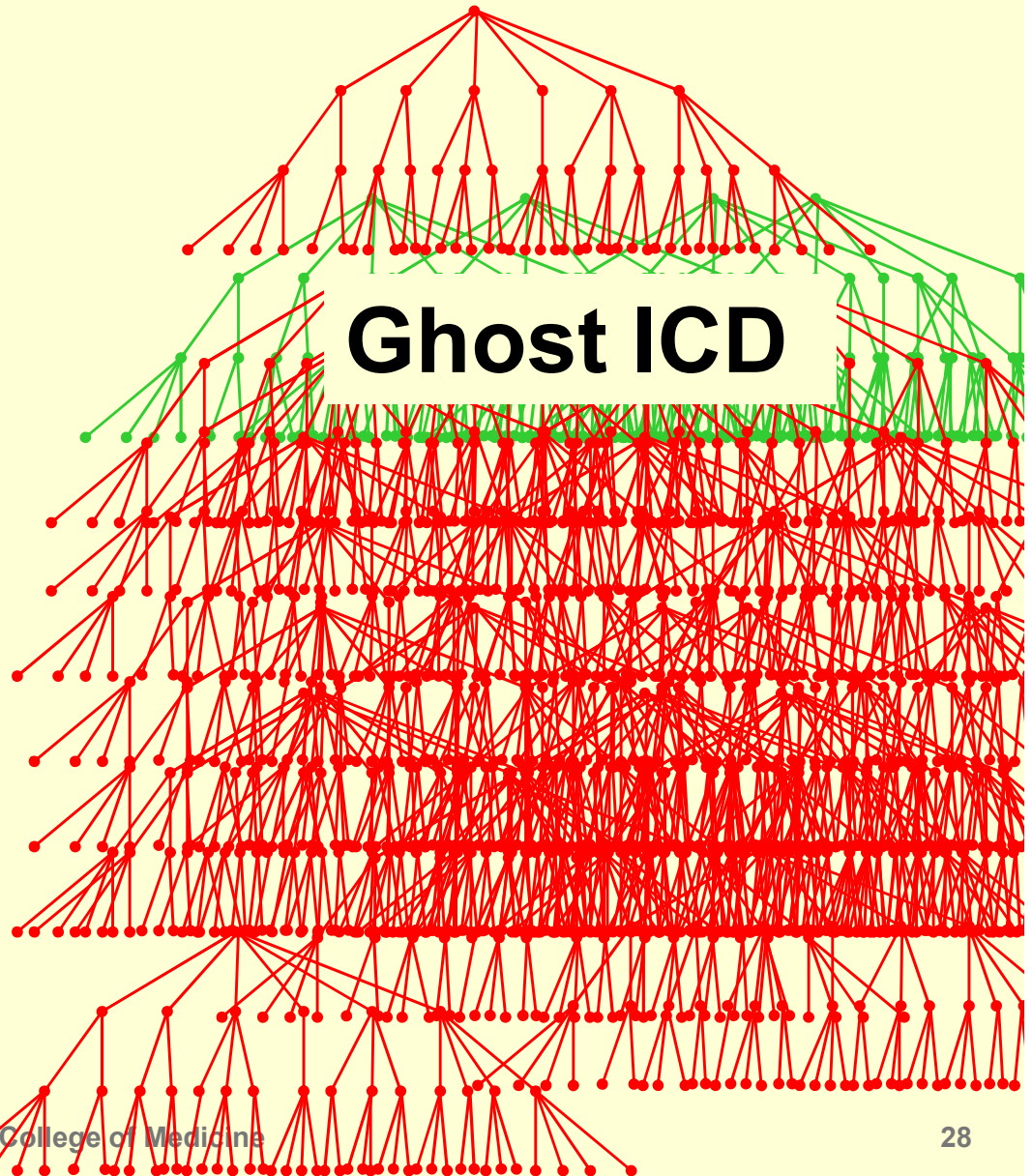
# Potential Future States

ICD-11



**Ghost SNOMED**

SNOMED



**Ghost ICD**

# Effort

# SNO

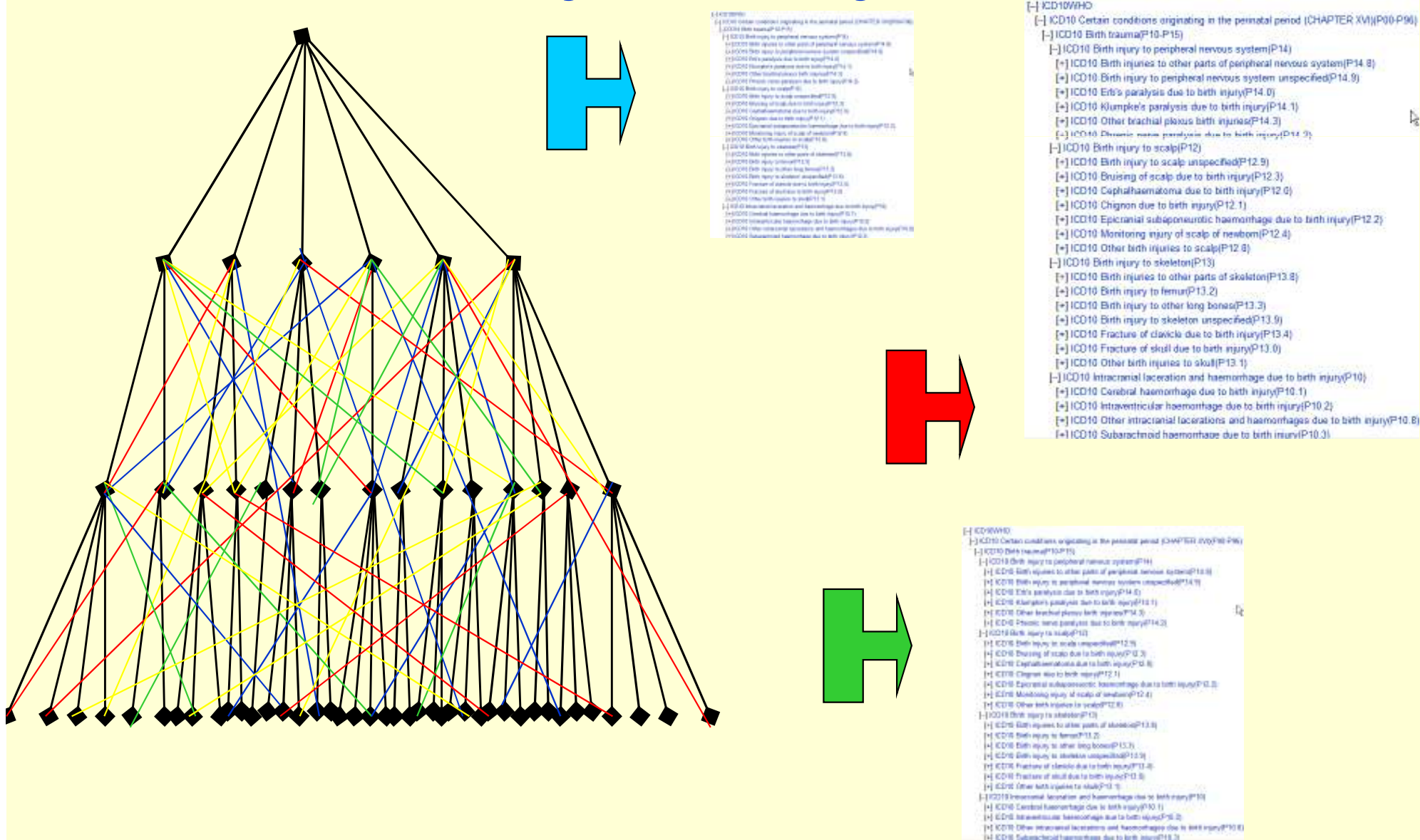
# 11

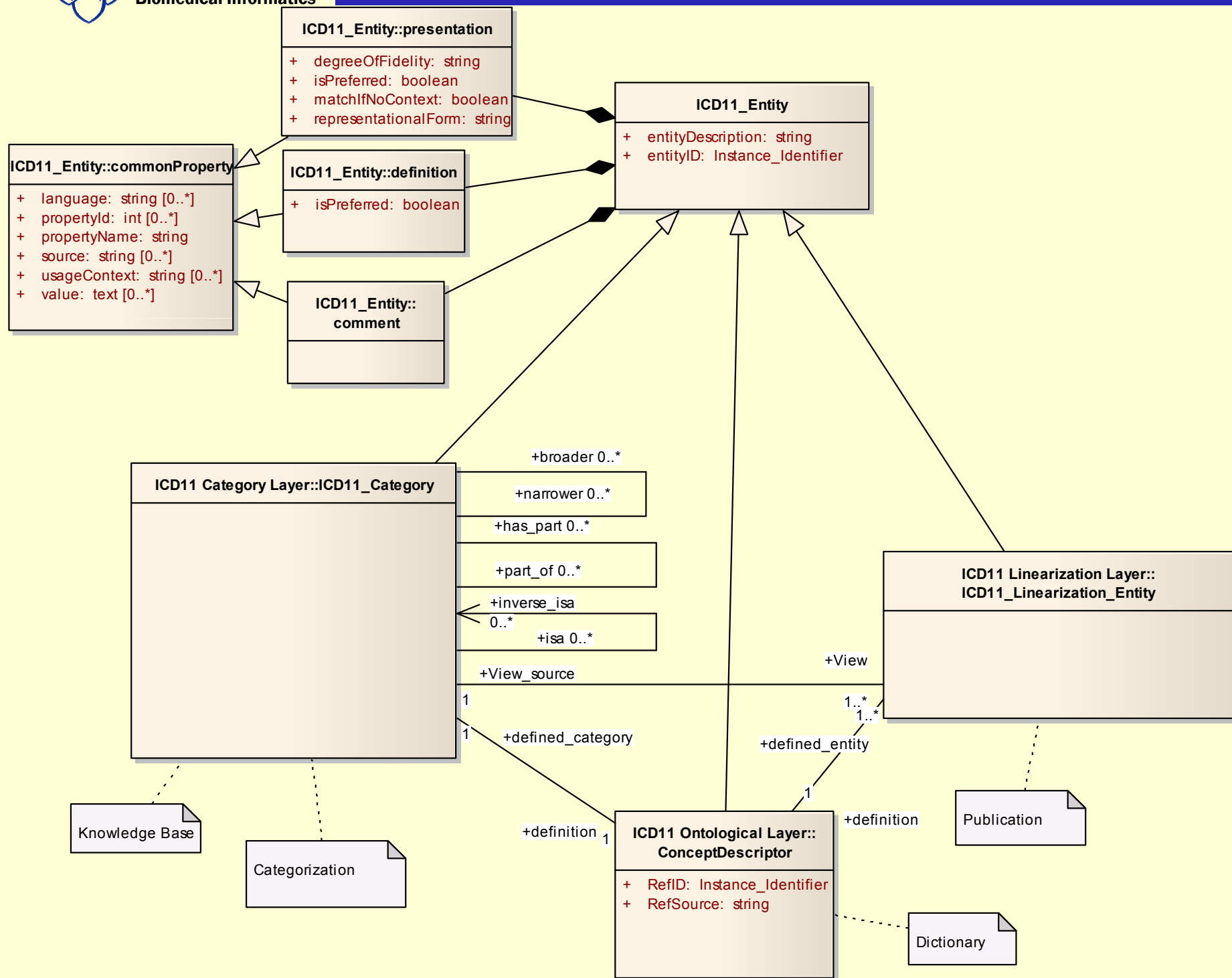




# Linear views may serve multiple use-cases

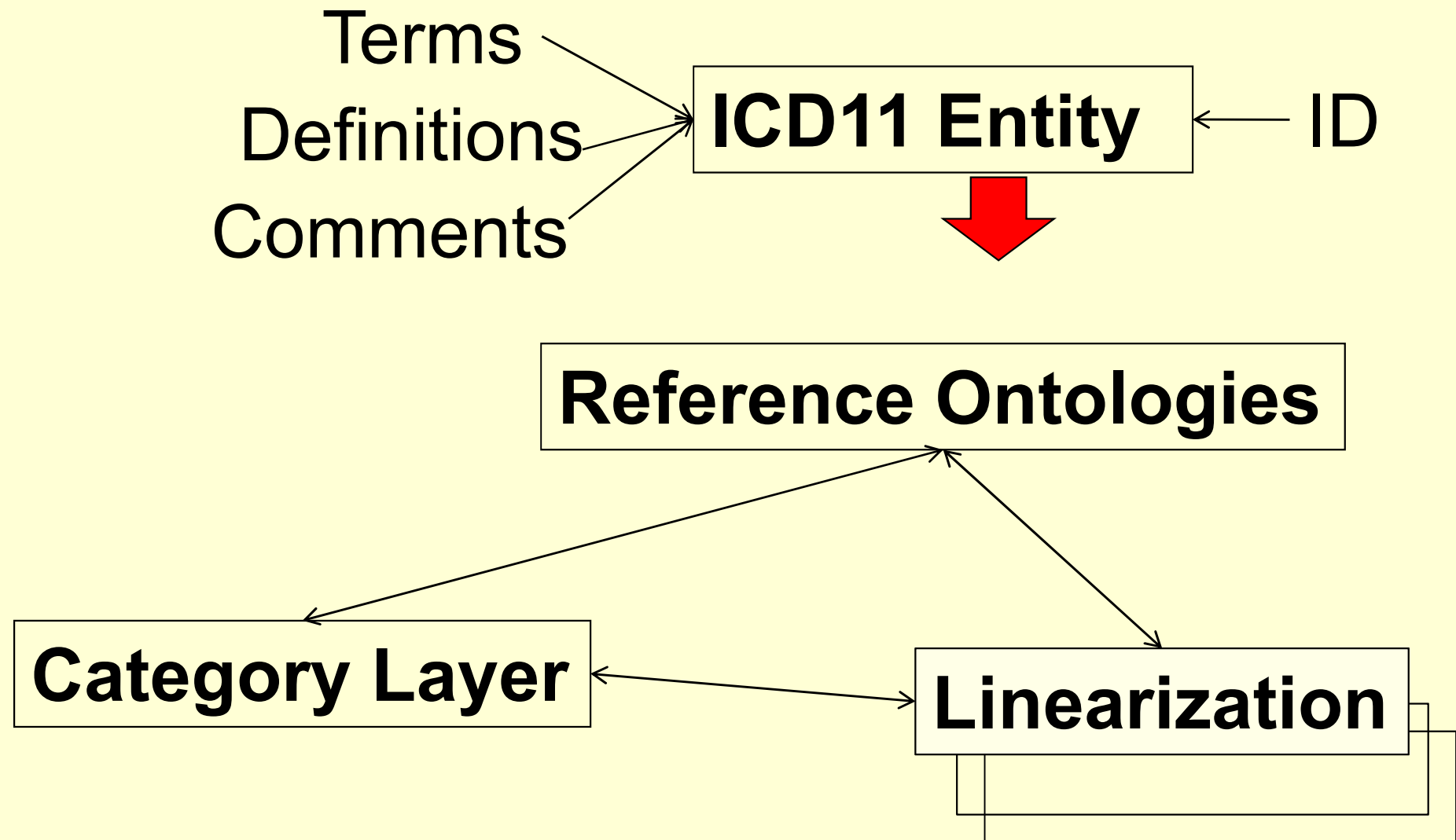
## Morbidity, Mortality, Quality, ...





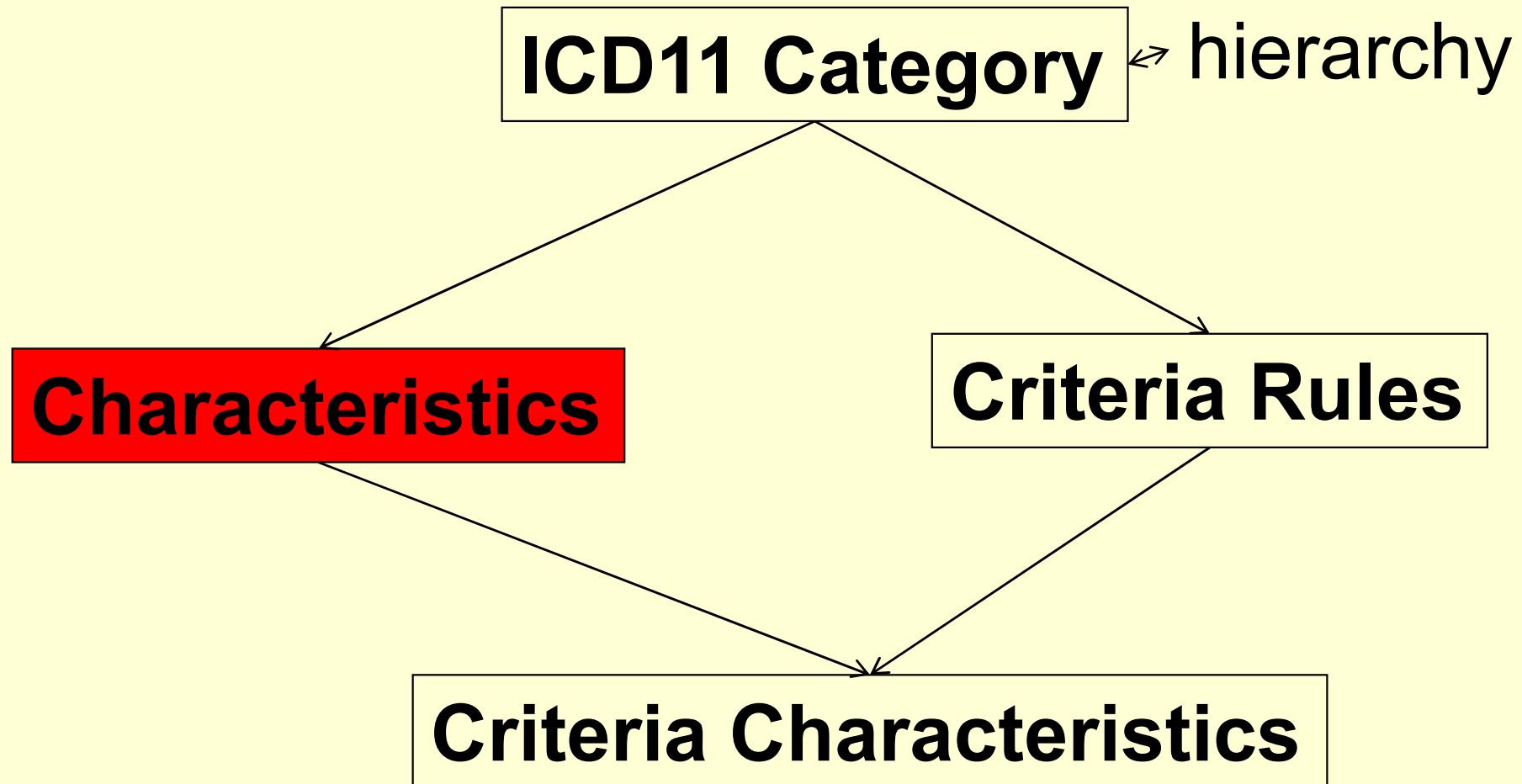
# Core ICD11 Model

# High Level Structure – Core Model

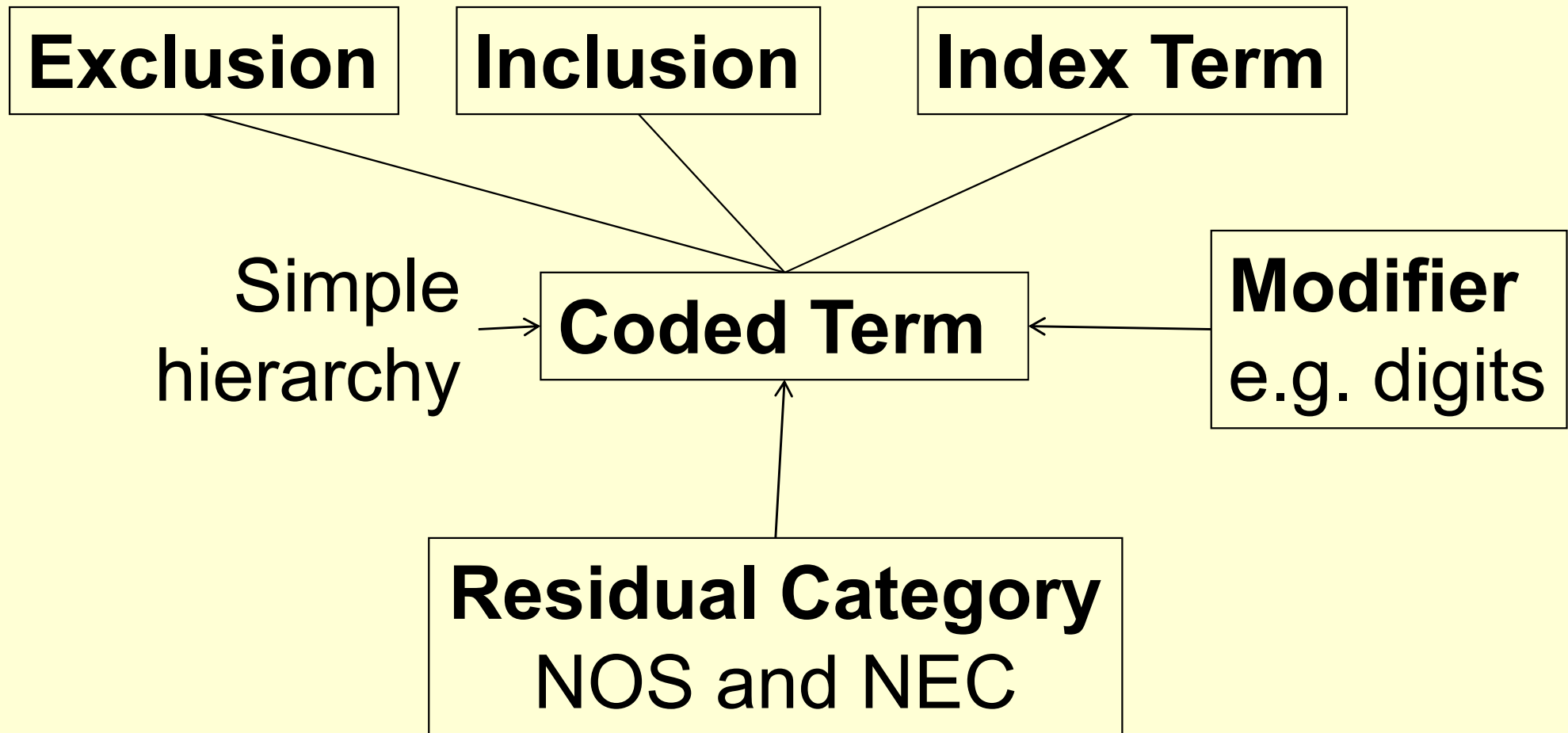




# Category (Clinical) Layer



# ICD Linearization



# Wiki

## Basis for *Light-weight editor*

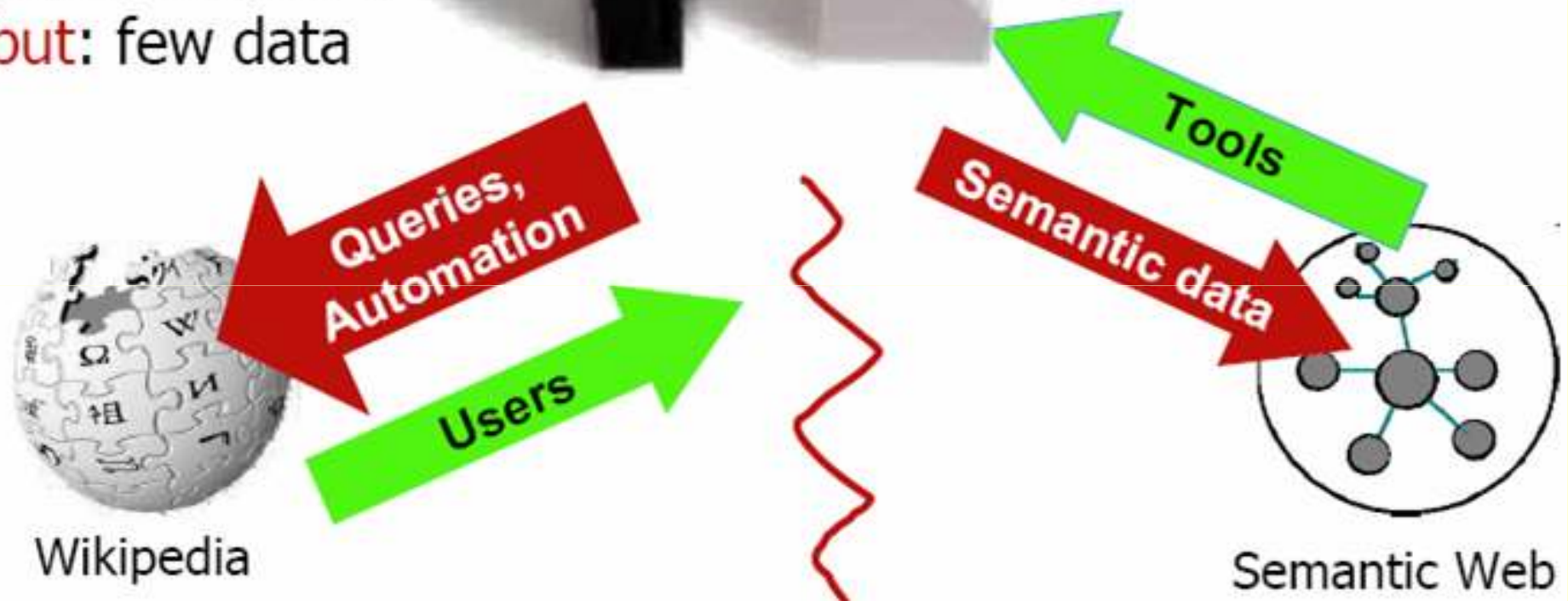
- The largest and most successful “social computing” project in history
- Tools and resources to permit huge communities to asynchronously author and edit complex resources
- Computer engineering of impressive capacity
  - Wikipedia manages >12,000 hits/second
  - Includes transactional edits

# W3C Semantic Web

- Explosion of methods, standards, tools
- Transform the practicality of complex concept management
  - XML – simple, interoperable syntax
  - RDF – simple data structure for semantic content
  - OWL – ontology authoring and interchange

## Goal: Marrying Wikipedia and the Semantic Web

- Wikipedia has many users and content, **but:** many manual processes and text-based search
- Semantic Web has sophisticated information processing, **but:** few data



From:



search ?

navigation

- [Main Page](#)
- [Task List](#)
- [Recent changes](#)
- [Feedback](#)

documentation

- [Help](#)
- [Video Tutorial](#)
- [Beginner's Guide](#)
- [MeetingNotes](#)

classifications

- [ICD11 Draft](#)
- [ICD10WHO](#)
- [ICD10CM](#)
- [ICECI](#)
- [Orphanet](#)

terminologies

- [SNOMED CT](#)

special

Add a proposal: Proposal for ICD11 Acute transmural myocardial infarction of anterior wall(I21.0) by Mayo at 20081028074507

Save this proposal

Meta

Name/Definition

Type/Anatomy/Pathophysiology

Manifestation

Etiology

Temporal Relation/Severity

### Symptom(Manifestation Attribute)

**Symptom:** SCT Squeezing chest pain(371030007)

**Defining Strength:** Pathognomonic

**Source:** Harrison's Principles of Internal Medicine

**Reference URL (optional):**

**Reference PMID (optional):**

**Reference ISBN (optional):** 978-0070202917

Remove this Symptom(Manifestation Attribute)

Add another Symptom(Manifestation Attribute)

# Adding Genomic Characteristics

<b>Genomic Characteristic:</b>	HUGO proline/serine rich coiled coil 1(PSRC1)
<b>Defining Strength:</b>	Associated Occasionally ▼
<b>Source (optional):</b>	J Mol Biol
<b>Reference URL (optional):</b>	
<b>Reference PMID (optional):</b>	18649068
<b>Reference ISBN (optional):</b>	

Remove this Genomic Characteristic(Etiology)

<b>Genomic Characteristic:</b>	HUGO symbol withdrawn see CELSR2(EGFL2~withdrawn)
<b>Defining Strength:</b>	Associated Occasionally ▼
<b>Source (optional):</b>	J Mol Biol
<b>Reference URL (optional):</b>	
<b>Reference PMID (optional):</b>	18649068



# Partial I21.0 Anterior Transmural MI Defined

**Symptom(Manifestation Attribute):** Category:SCT Squeezing chest pain(371030007) (Pathognomonic)

(Source:Harrison's Textbook of Medicine)

3. *Harrison's principles of internal medicine*. New York: McGraw-Hill, Health Professions Division, 1998. isbn:978-0070202917. [isbn978-0070202917]

**Sign(Manifestation Attribute):** Category:Not Specified(Source:not specified)

**Diagnostic Finding(Manifestation Attribute):** Category:Not Specified(Source:not specified)

**Functional Impact(Manifestation Attribute):** Category:Not Specified(Source:not specified)

**Causal Agent(Etiology):** Category:Not Specified(Source:not specified)

**Mechanism(Etiology):** Category:Not Specified(Source:not specified)

**Genomic Characteristic(Etiology):** Category:HUGO proline/serine rich coiled coil 1(PSRC1) (Associated Occasionally)  
(Source:J Mol Biol)

4. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nitschke P, Hajat C, Mangino M, Hengstenberg C, Stark K, Ziegler A, Caulfield M, Burton PR, Schunkert H, and Tobin MD. *The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol*. J Mol Med 2008 Nov; 86(11) 1233-41. doi:10.1007/s00109-008-0387-2 pmid:18649068.  
[PubMed](#) [pmid18649068]

**Genomic Characteristic(Etiology):** Category:HUGO symbol withdrawn see CELSR2(EGFL2~withdrawn) (Associated Occasionally)(Source:J Mol Biol)

4. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nitschke P, Hajat C, Mangino M, Hengstenberg C, Stark K, Ziegler A, Caulfield M, Burton PR, Schunkert H, and Tobin MD. *The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol*. J Mol Med 2008 Nov; 86(11) 1233-41. doi:10.1007/s00109-008-0387-2 pmid:18649068.  
[PubMed](#) [pmid18649068]



# Conclusion

- Comparable and consistent health data has five centuries of history
- Modern ontology principles are being leveraged to create the next-generation ICD
- Social computing infrastructure (wiki) can greatly extend WHO's capacity in this effort