Building blocks of health information: Classifications, terminologies, standards

Bedirhan Ustün & Nenad Kostanjsek
WHO Geneva
Information challenges in health care practice

Physician’s diagnosis & interventions

Nurses’ diagnosis & interventions

Physiotherapist’s diagnosis & interventions

Social worker
SW Interventions

Occupational therapists diagnosis & interventions

Patient

Physician documentation

Nursing documentation

PT documentation

OT documentation

PT documentation

OT documentation

PT documentation

OT documentation
Health information challenges

- **DATA SOURCES**: Increasingly complex and multiple sources
  - Heterogeneous: *Variations in conceptual frameworks*
  - Incomplete: *not the whole population is covered*
  - Non-comparable: *no mechanism to assure that entities/properties are comparable*

- **DATA PROCESSING**: Aggregation
  - Which public health indicators? *Relevant concepts are not agreed upon*
  - How do you combine data? *Different sources of data & logic not fully specified - transparent*
  - Mechanisms not automated: *Meta data and other relevant structured formalism are not fully specified*
  - input – output specifications: *Norms and standards are not fully identified*

- **QUALITY**: Reliability and Quality Assurance
  - Possible error sources
  - Coverage & completeness
  - Possible duplication and delay
  - Audit trails

- **PARTICIPATION**: involvement and empowerment of citizens
  - Privacy Protection
  - Prevention of misuse
Health Information Systems
Requirements for Digitalization

- **Common terminology → ontology**
  - Universal descriptions for
    - genes, molecules, cells, diagnostic methods, signs, symptoms, interventions and other entities.

- **Common structure**
  - Information models
    - Disease, disability, risk factors, interventions

- **Common reporting methods**
  - casemix groupings
  - resource groupings
  - outcome measurement systems.
Constraints in the real world

- Human variability
  - People can’t always interoperate
    - Machines will never interoperate better than the people that use them

- Too many requirements … so priorities unclear
  - Intimately intertwined with EHRs, Public health, Decision support, Clinical care…

- Poor match of problem space & solution space
  - Poor definition of purpose
    - “What’s it for?”

- 20 years of intensive work in IT has not yet provided a solution

- Temptation to do more than is possible
WHO classification development in the 20th Century

Construction of ICD-10 & ICF:

- ICD: 8 Annual Revision Conferences (1982 - 89)

- ICD: 17 – 58 Countries participated
  - 1-5 person delegations
  - mainly Health Statisticians

- ICF: 61 Countries participated
  - 1-5 person delegations
  - Multi-disciplinary

- Manual curation
  - List exchange
  - Index was done later

- "Decibel"? Method of discussion

- ICF: Concept driven

- Output: Paper Copy

- Work in English only

- ICD: Limited testing in the field
  - ICF: drafts translated into / tested in 27 languages

- post-coordinated development of linkages to related classification, terminologies and assessment instruments
Placing WHO Classifications in HIS & IT of the 21st Century

- Population Health
  - Births
  - Deaths
  - Diseases
  - Disability
  - Risk factors

- Clinical
  - Decision Support
  - Integration of care
  - Outcome

- Administration
  - Scheduling
  - Resources
  - Billing

- Reporting
  - Cost
  - Needs
  - Outcome

KRs Mappings
Terminologies
Classifications

ICD
ICF
ICHI
ICPS
ICTM

e-Health Record Systems
World Health Organization
The desiderata for a WHO FIC in 21st Century

- Evolve a **multi-purpose** and **coherent** classification which is
  - **consistent** yet **adaptable** and **interoperable** across
    - different uses (public health, service management, research)
    - the spectrum of health care (Primary, Secondary, Tertiary)
    - in developing and developed countries
  - compatible with other **WHO classifications**

- Serve as an **international** and **multilingual** reference standard for scientific **comparability** and **communication** purposes

- Ensure that ICPS will function in an **electronic health records** environment.
  - **Link WHO FIC logically** to underpinning **terminologies** and **ontologies** (e.g. SNOMED, GO, …)
  - WHO FIC categories “**defined**” by "**logical operational rules**" on their associations and details
WHO FIC deliverables

1. **Print versions** "fit for purpose" in multiple languages

2. **Web Portal** that allows to
   - **access and browse** the classification with definitions, descriptive characteristics and semantic linkages with other WHO FIC classification and related terminologies
   - **maintain and update** classification using a Collaborative Authoring Tool with established workflows

3. **formalized language**: Logical and machine readable knowledge representation of WHO classification entities (concepts with attributes) and their relationships
Key workstreams & elements for developing WHO FIC

- Use cases
- Content model (parameter & value set)
- Population & peer review of content model
- Web based collaborative authoring tool (iCAT)
- Ontology development
Use Cases

- Who are the **users**: (Actors, settings, instances…)

- How do they **currently use classifications**?
  - Inputs
  - Process
  - Outputs

- What can be done to **improve the use**?

- Identify **requirements** for WHO FIC
What is a Content Model (CM)?

- Captures the **key parameters** for the definition of an classification entity
  - *What is a concept – entity in a classification?*
  - *How do you define it: basic attributes?*
  - *What different values it can take?*

- **in a standard & systematic way**
What is a CM Parameter?

- A group name for **Common Characteristics / Attributes**
- Which describes a **particular component or perspective**
  - kind, scope, size, location, origin, ...
- Useful to understand the entity mentally and define it in an **unambiguous** way
- Refer to the entities in a **systematic** way
  - Allow **sorting** into classes, **grouping**, indexing, searching...
  - Useful to teach computers *(also humans !!)*
Example Parameters

- Symptoms and Manifestations
  - signs, symptoms and findings

- Etiology
  - underlying explanatory mechanism(s)

- Course and Outcome
  - distinct pattern of development over time
What is a Value Set?

- This **list of possible terms** or options within each parameter
  - Every word, phrase, or statement that can be used by a computer and a human to describe a diagnostic entity or intervention
Example Value Sets – Body Systems

- Cardiovascular System
- Digestive System
- Endocrine System
- Integumentary System
- Musculoskeletal System
- Neurological System
- Reproductive System
- Respiratory System
- Urinary System
Why do we need a Content Model?

- To organize knowledge in a consistent, structured way
- To facilitate efficient and productive drafting
- To prepare for computerized terminologies and ontologies
  - Electronic Health Records
- To ensure the most useful document for end-users
ICD 11 is no longer just lists...it is based on a content model

| A000  Cholera due to Vibrio cholerae 01, biovar cholerae |
| A001  Cholera due to Vibrio cholerae 01, biovar eltor |
| A009  Cholera, unspecified |
| A0100 Typhoid fever, unspecified |
| A0101 Typhoid meningitis |
| A0102 Typhoid fever with heart involvement |
| A0103 Typhoid pneumonia |
| A0104 Typhoid arthritis |
| A0105 Typhoid osteomyelitis |
| A0109 Typhoid fever with other complications |
| A011  Paratyphoid fever A |
| A012  Paratyphoid fever B |
| A013  Paratyphoid fever C |
| A014  Paratyphoid fever, unspecified |
| A020  Salmonella enteritis |
| A021  Salmonella sepsis |
| A0220 Localized salmonella infection, unspecified |
| A0221 Salmonella meningitis |
| A0222 Salmonella pneumonia |
| A0223 Salmonella arthritis |
| A0224 Salmonella osteomyelitis |
| A0225 Salmonella pyelonephritis |
| A0229 Salmonella with other localized infection |
| A028  Other specified salmonella infections |
| A029  Salmonella infection, unspecified |
| A030  Shigellosis due to Shigella dysenteriae |
| A031  Shigellosis due to Shigella flexneri |
| A032  Shigellosis due to Shigella boydii |
| A033  Shigellosis due to Shigella sonnei |
| A038  Other shigellosis |
| A039  Shigellosis, unspecified |
| A040  Enteropathogenic Escherichia coli infection |
| A041  Enterotoxigenic Escherichia coli infection |
ICD 11 Foundation Component and Linearizations

ICD-11 content model parameters
- Definitions, synonyms
- Clinical descriptions
- Manifestation Attributes
- Causal mechanism & risk factors
- Functional impact

Value Set

SNOMED-CT, International Classification of Functioning, Disability and Health (ICF), International Classification of External Causes of Injury (ICECI)…
Populating and reviewing the content model

Managing Editor
Postmaster

Peer Review
Min 3 reviews – 2 approvals

TAG Review

Filter

O.K.

Other TAGs

RSG
Web based collaborative authoring tool (iCAT)

- **display & browse** taxonomy with its content model rubrics
- allow user to **comment** on the content
- allow users **editing** the content and facilitate the use of value sets derived from other classifications and terminologies
- allow user **restructuring** the classification
- Incorporates **multiple level** of **user access**
- supports **multilingual** representation
- **ontology** tooling interface with description logic technology
An Ontology for WHO FIC

● ENTITIES:
  - categorization of information
    • using a standardized set of concepts
    • with agreed definitions
    • Uniquely identified terms
  - based on explicit attributes/values

● RELATIONS:
  - relationships between the key concepts
ICPS Incident Concept
as proposed in ICPS categorial model

- Hazard
  - has cause
- Contributing Factors
  - Mitigating Factors
  - has circumstance
- Incident
  - has type
  - has people involved
  - has detection
  - has location
  - has consequence
  - Action
  - Harm
- Incident Type
- Person
- Detection
- Care Setting
What difference ontologies can make?
Coding cycling accident in ICD-10

- V12.24 Pedal cyclist injured in collision with two- or three-wheeled motor vehicle, unspecified pedal cyclist, non-traffic accident, while resting, sleeping, eating or engaging in other vital activities
The history of cycling codes

- 1972 ICD-9 (E826): 8 codes
- READ-2 (T30..): 81 codes
- READ-3: 87: 82 codes
- 1999 ICD-10: 587 codes
Defusing the exploding bicycle: 500 codes in pieces

- 10 things to hit…
  - Pedestrian / cycle / motorbike / car / HGV / train / unpowered vehicle / a tree / other

- 5 roles for the injured…
  - Driving / passenger / cyclist / getting in / other

- 5 activities when injured…
  - resting / at work / sporting / at leisure / other

- 2 contexts…
  - In traffic / not in traffic

V12.24 Pedal cyclist injured in collision with two- or three-wheeled motor vehicle, unspecified pedal cyclist, nontraffic accident, while resting, sleeping, eating or engaging in other vital activities
Conceptual Lego… it could be...

Goodbye to picking lists...

Structured Data Entry

Cycling Accident

What you hit

Your Role

Activity

Location
WHO classification development in the 21st Century

- Internet-based permanent platform
  - All year round
  - Open to all people in a structured way
  - Content experts & users are empowered

- Digital curation
  - Wiki enabled collaboration
  - Ontology based

- Enhanced discussion & peer review

- Electronic copy → print version

- Work in multiple languages

- field tests
  - Based on Use Cases
What is the answer? ... what is the question?

Computers are useless. They can only give you answers.

Pablo Picasso (1881-1973)