

Introduction to Semantic Interoperability

| DIGITAL TRANSFORMATION TOOLKIT
KNOWLEDGE TOOLS

3

PAHO



Pan American
Health
Organization



World Health
Organization
REGIONAL OFFICE FOR THE
Americas

Why code health data?

Information systems play a key role in data management and health information. Data management is currently a basic requirement for evidence-based decision-making. Likewise, information systems are crucial for designing or adapting policies that enable more equitable, higher quality health care. Advances in information processing technology and new tools for data management open up endless possibilities in the public health field (1, 2). Having validated data makes it possible to advance towards the possibility of finally achieving effective exchange of information, consequently improving such aspects as health access, service prioritization, and adequate care, especially for populations in situations of vulnerability. This also has an impact on issues related to patient safety, medical error prevention, and resource optimization, as well as improving research, education, and management processes to reach the ultimate goal: providing better health care by reducing potential inequities.

The problem of data management in health is typical of this discipline. Health has its own language; one that is ambiguous, full of eponyms, homonyms, synonyms, acronyms, and abbreviations. Its nomenclature is huge and very rich in detail, highly context-dependent, riddled with jargon and localisms, and often lacks rigorous definitions. Unfortunately, computer systems have difficulty understanding natural, even if it is limited to a specific domain, such as that of the medical sciences. This is why information must be coded in advance, so that computer systems can process it. In recent decades, the understanding of natural language has become a very active field of study, with great progress on certain tasks. Therefore, we can expect that in the medium or long term, computer systems will be able to automatically and unrestrictedly process information expressed in natural language, whether written or spoken, which will greatly improve their use. In the meantime, we must continue to code data.

Given the above, it is necessary to design vocabulary control strategies so that the clinical information stored in health information systems can be used for multiple purposes. This is achieved through a process termed *health data representation* (1, 3, 4).

What is health data representation?

Health data representation is the process that aims to document professional thinking and perceptions, patient care, and the information generated during care. This information is stored as part of the care process. Thus, during a care encounter, the things that happen in the real world are represented as structured information so that they can be stored in a database and managed by computers.

Obtaining this data can be an obstacle when implementing information systems; therefore, it is necessary to find strategies that address user needs: documenting findings, processes, and care outcomes using free narrative text, in natural language, without restrictions. Added to this is the complexity of representing patient data when it comes from different sources, such as lists of diagnoses, symptoms, and signs; notes about clinical developments and procedures; lists of medications; results from laboratory analyses and complementary studies; social determinants and environmental health information; clinical decisions and treatments; genomics and proteomics; and patient preferences.

Semantic interoperability means that, once shared, data can be understood by health information systems.

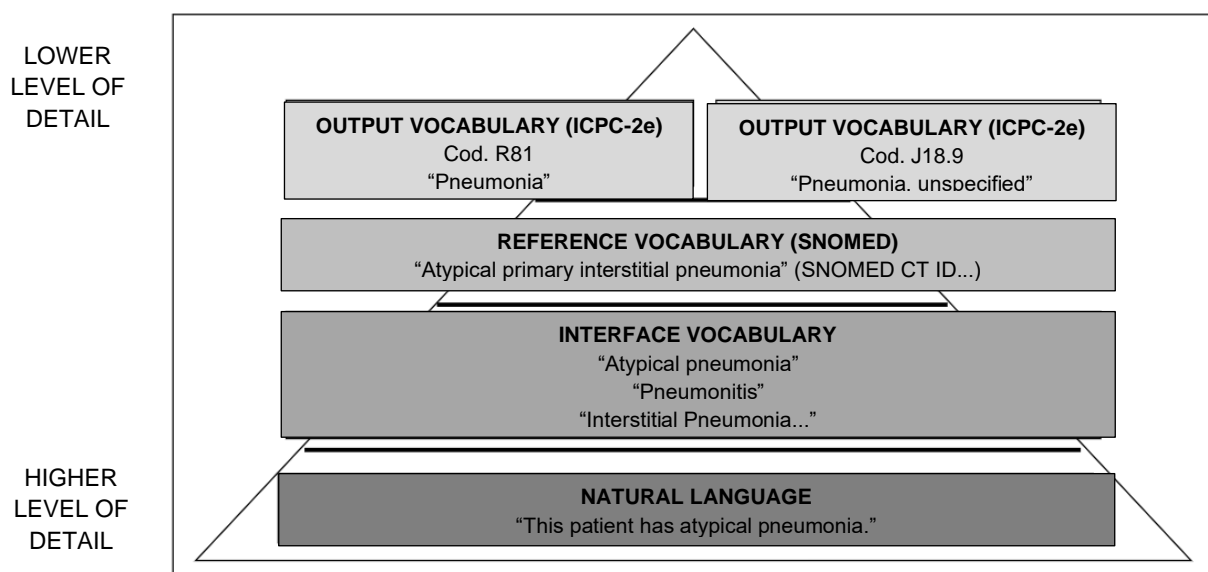
What is semantic interoperability?

Semantic interoperability is the ability of computer systems to exchange data, with unambiguous meaning. It is a requirement not only for health data be shared between different systems or applications, but for them to be understood. Semantic interoperability refers to the transmission of the meaning of data. This is achieved by linking each piece of data to a shared, controlled vocabulary. This shared, standardized vocabulary, in most cases agreed at the international level, is what provides the capacity for unambiguous interpretation of information. This vocabulary is called the *terminology standard* (5, 6).

What are terminology standards?

Terminology standards are mostly international agreements that unequivocally represent a health concept from a given domain. There are different types of standards for different areas of health, which differ according to their higher or lower level of detail, or according to their usefulness and their structure. Standards with a low level of detail are used, for example, to manage data grouped at regional levels. Those with a higher level of detail are more useful for managing data at the patient level, such as implementing support systems for decision-making. This is the case, for example, with the World Health Organization (WHO) International Classification of Diseases 11th Revision (ICD-11), which enables more accurate and detailed data recording and compiling, and greater clinical accuracy, than its previous versions. Usually, a terminology standard's level of detail is called its *granularity*.

Figure 1. Language structure and levels of detail (granularity)



The most commonly used terminology standards are detailed in table 1.

Table 1. Terminology standards

Standard	Description
ICD International Classification of Diseases (7)	This is the WHO classification family. Its version 11 (ICD-11) provides general improvements in coding that enable more accurate and detailed data collection and recording, and unprecedented clinical accuracy. It is the international standard for the systematic recording, reporting, analysis, interpretation, and comparison of mortality and morbidity data. Its previous versions (ICD-9, ICD-10) are by far the most used by countries worldwide for official data representation in health care.
ICPC International Classification of Primary Care (7)	ICPC is a taxonomy of the terms and expressions commonly used in general/family medicine. It encompasses reasons for consultation, health problems, and the care process.
SNOMED Systematized Nomenclature of Medicine (9)	SNOMED is a systematically organized computer-processable collection of medical terms providing codes, terms, synonyms, and definitions used in clinical documentation and reporting. In terms of data coverage, its scope is similar to ICD-11.
LOINC Logical Observation Identifiers Names and Codes (10)	LOINC is a database and a universal standard created to identify medical laboratory observations, although currently expanding its scope to other complementary studies and clinical documents.
NANDA-I Nursing Diagnosis Association (11)	NANDA-I is a standard system of representation of vocabularies related to the nursing care process. This coding system includes interventions (NIC - Nursing Interventions Classification) and results (NOC - Nursing Outcomes Classification). The combined use of these codes makes it possible to represent the entire nursing care process.
ATC Anatomical Therapeutic Chemical Classification (12)	ATC is a coding system based on codes assigned to a drug according to the organ or system on which it acts and how it works. This classification system is maintained by WHO.

This variety of standards, along with the complexity of clinical recording for health professionals (who in general are not trained to carry out these coding tasks), have made it necessary to create tools to facilitate this task, which is how *terminology servers* emerged.

What is a terminology server?

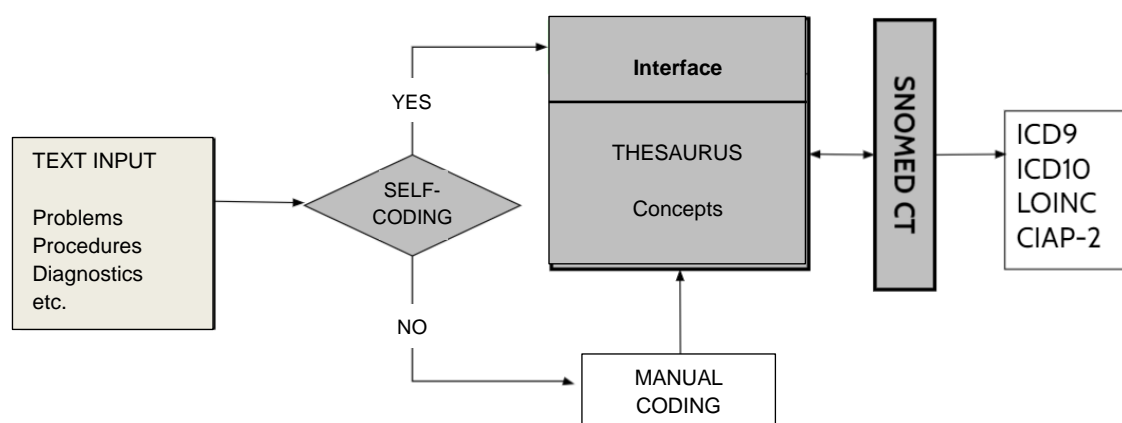
A terminology server is an application that, once integrated into any information system or EHR, makes it possible to represent health data in a flexible manner. In other words, the terminology server helps professionals by coding their data, so they do not have to know how to use the different standards; the only thing they have to do is to concentrate on documenting clinical actions as well as they can, and let the system do the coding for them (13, 14).

Every time someone enters a text, the system identifies it, enters it as part of the thesaurus of the institution, incorporating the term into what are called *interface vocabularies* and then,

depending on the need, it codes it using the different standards. If the server does not identify the text, it will go through a manual review process conducted by experienced coders before being incorporated into the thesaurus (15).

A **thesaurus** is a list of controlled words or terms used to represent concepts. It refers to a lexicon of ordered terms comprising the specialized vocabulary of an academic discipline or field of study, showing the logical and semantic relationships between the terms.

Figure 2. Flow of information on a terminology server



What are the differences between interface vocabularies and reference vocabularies (terminology standards)?

There is an added layer of complexity in the representation of health data. So far, we have seen that we could achieve semantic interoperability with the existing international standards, using terminology servers to assist health professionals in the arduous task of coding. Health vocabulary is not simple: the jargons of each specialty, localisms, or even unintended consequences of implementing information and communication technologies (ICTs) (e.g., typing errors, spelling errors) mean that servers do not always work properly (16).

Interface vocabularies can bridge this gap between natural language (how we speak and write) and terminology standards. These vocabularies are collections of terms, many of which are synonymous with a common medical concept, which comprise not only valid but also invalid forms, thereby increasing terminology servers' recognition capacity, and enabling the incorporation of local jargons into international standards (17–19).

Figure 3. Reference terminology (terminology standard) vs. interface terminology

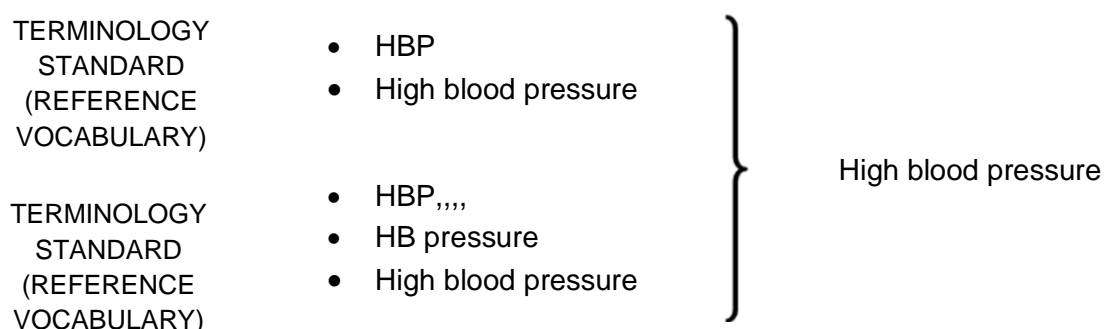


Table 2. Differences between terminology standards and interface vocabularies

CHARACTERISTICS	TERMINOLOGY STANDARDS	INTERFACE VOCABULARY
Domains (diagnostics, procedures, etc.)	Depends on each standard	Encompasses all domains
Acronyms and abbreviations	Few or none	All
Spelling errors	None	Many
New terms	According to the version	Online
INVALID terms	Never	Yes
Local jargon	Depends (whether the standard has local versions or updates by country)	Yes

Why is it important to have coded data, beyond semantic interoperability?

Going back to the beginning of this document, to achieve semantic interoperability we must have coded data; however, this is not the only advantage of the process of coding or representing health data (15).

Having coded data makes it possible to:

- Improve financial and administrative management
- Measure risk at the patient or population levels
- Implement clinical decision-making support systems
- Improve clinical reporting
- Promote research
- Integrate sources of information to provide evidence to professionals and patients
- Facilitate interaction with external software or different suppliers
- Improve training/education programs

What does ICD-11 offer as a new tool for term coding?

The International Classification of Diseases (ICD) is WHO's global terminology standard for health data, clinical documentation, and health statistics reporting. Its eleventh version, ICD-11, is completely and scientifically updated and designed for use in a digital world, with less training effort and easier implementation. Its development in multiple languages facilitates the processes of interoperability between countries, while its proposal platform enables stakeholders to participate in its continuous updating (7).

ICD-11 is the result of health, epidemiology, and statistical professionals working together, along with coding and IT experts from around the world. This latest edition is electronic and has more than 100,000 terms for medical diagnosis; its search logic can interpret more than 1.6 million terms, it and can be installed and used online or offline (7). ICD-11 represents a major improvement over previous revisions. It reflects critical advances in science and medicine and includes updates on the treatment and prevention of diseases, facilitating their use and improving coding accuracy.

The possibility of using extension codes enables the incorporation of relevant details for clinical documentation. Extension codes make it possible to adapt the degree of detail that can be reflected and adapted in the primary care environment, the surveillance of rare diseases, and the notification of adverse events as part of managing quality and patient safety.

ICD-11 is a more flexible, user-friendly system and has a greater capacity to represent health information, so it can be perfectly integrated into routine clinical documentation, mainly in computerized records systems (20).

How does this relate to the eight principles for the digital transformation of public health?

In mid-2020, the United Nations presented eight areas of collaboration based on recommendations from a high-level task force on technical cooperation in the age of digital interdependence. The Pan American Health Organization (PAHO) has adopted and adapted these areas into eight guiding principles to reflect the imperatives of the digital transformation of the health sector: 1) universal connectivity, 2) digital public goods, 3) inclusive digital health, 4) interoperability, 5) human rights, 6) artificial intelligence, 7) information security, and 8) public health architecture.

Figure 4. Eight principles for the digital transformation of public health

1	Universal connectivity 	Ensure universal connectivity in the health sector by 2030
2	Digital public goods 	Co-create digital public health goods for a more equitable world
3	Inclusive digital health 	Accelerate towards inclusive digital health with an emphasis on the most vulnerable
4	Interoperability 	Implement interoperable, open, and sustainable digital health and information systems
5	Human rights 	Mainstream human rights in all areas of digital transformation in health
6	Artificial intelligence 	Participate in global cooperation on artificial intelligence and any emerging technology
7	Information security 	Establish mechanisms for trust and information security in the digital environment of public health
8	Public health architecture 	Design public healthcare architecture in the era of digital interdependence

Principle 4: Interoperability

Implement open, sustainable, and interoperable information systems for health


Health information systems—timely and open access to properly disaggregated data, integrating national and local systems, digital health, and ICT—facilitate the effective identification, reporting, and analysis of cases and contacts; the early case tracing and detection; and the definition and monitoring of the population at risk in a manner as secure, interoperable and personalized as possible.

Coded data (which makes semantic interoperability feasible) contributes to the incorporation of the dimensions contained in this principle: management and governance of information systems; management of data and IT; information and knowledge management and innovation; and digital integration and convergence—that is, the possibility of accessing the same content from different devices.

It also contributes to the aim of achieving interoperability between databases and applications, facilitating access to reliable data and knowledge at the right time, in the right place, and in the right format.

Semantic interoperability is crucial to designing governance schemes for data generated by the interoperability of health systems.

8 Guiding Principles for Digital Transformation of the Health Sector



Interoperability

4

Implement interoperable, open, and sustainable digital health and information systems

Information systems for health—with timely and open access to properly disaggregated data, integration of national and local systems, digital health, and ICT—facilitate effective identification, notification, and analysis of cases and contacts, early case search and detection, and definition and monitoring of the at-risk population in a manner that is safe, interoperable, and as personalized as possible.

Call to action

- Include the following four dimensions in national plans and public policies on information systems and digital health: 1) management and governance of information systems; 2) data management and information technology; 3) information and knowledge management and innovation; and 4) integration and digital convergence, meaning the ability to access the same content from different devices.
- Ensure the implementation of information systems and digital health strategies under governance that ensures the convergence of investments and actions, as well as the interconnection and interoperability of databases and applications to facilitate timely access to reliable data and knowledge in the right place and the right format.
- Consolidate infrastructure for the exchange of open data and critical information focused on ethical and cybersecurity criteria in information flows.
- Adopt a digital literacy program based on detected needs and considering different contexts, in order to reduce inequalities.
- Provide the conditions and necessary support to strengthen existing initiatives and build a “multi-stakeholder network that promotes comprehensive and inclusive approaches to building digital capacity for sustainable development”.
- Define governance schemes for the data generated by interoperable health systems in order to promote secondary use of information, generating data for tactical and operational decision-making.
- Develop secure mechanisms that allow the exchange of clinical documentation (syntactic interoperability) based on existing standards.

“ There is a need to stop seeing ‘interoperability’ as an isolated IT-related term. It is a key concept for having timely open access to properly disaggregated data and the integration of national and local systems.”

PAHO / WHO Statement during the High-Level Conference on Information Systems for Health, February 2021.

Principle 8: Public health architecture.

Design a renewed public health architecture for the age of digital interdependence

In the age of digital interdependence, public health architecture must be designed within the framework of a digital governance agenda. It must be cross-cutting, bringing together the different aspects of governance and optimizing strategic planning and resource management. It must be based on the use of rules and procedures in many different areas, not just in the field of health; this is the case of connectivity and bandwidth, which have an influence on health, education, and every other sector.

This point will be crucial to achieving one of the objectives of this principle, which is to have the necessary technologies to enable the exchange of standardized data, for national use and benefit. It will be important to design robust mechanisms to exchange information that can be understood, and ultimately used for improving public health.



Public health architecture

8

Designing public health architecture in the age of digital interdependence

Public health architecture in the era of digital interdependence must be framed within each government's digital agenda. It must be transversal in order to include the different aspects of governance and optimize strategic planning and resource management. It should be based on the use of standards and procedures that favor multiple areas, not only the health field. This is the case of connectivity and bandwidth, which have an impact on health, education, and all other sectors.

Call to action

- Update digital agendas as a public policy with the necessary regulatory framework for digital applications in health.
- Capitalize on international efforts to develop projects to deploy health information technology.
- Consider establishing technologies to allow data exchange; for example, an interoperability bus with services for consulting standardized databases of national benefit and use.
- Define processes to evaluate emerging technologies related to big data, machine learning, artificial intelligence, and "omics" sciences, among others.
- Create change management teams that accompany digital transformation, specifically in the health environment, attending to the needs of different actors (patients, health professionals, administrative staff, leaders, coordinators, etc.).
- Incorporate new underlying frames of reference to predict the adoption and use of technologies in the health sector, both from the perspective of providers and users.
- Develop mechanisms to search, promote, and exchange information on good practices; and create effective mechanisms for the exchange of knowledge to avoid 'adatomism' in different sectors and locations.

"Public health architecture in the age of digital interdependence should be designed within the framework of a digital governance agenda. This architecture should be cross-cutting, permitting proper coordination of the different areas of governance and achieving optimization of strategic planning and management of the resources allocated to it"

PAHO / WHO Statement during the High-Level Conference on Information Systems for Health, February 2021.

How does this relate to the PAHO IS4H initiative?

Regarding the PAHO maturity model—specifically in the area of data management and information technology—there is a need for quality interoperability standards to achieve a national health information architecture. The model expects that at the maximum level of maturity, health information systems will be interoperable, thanks to a national infrastructure that uses up-to-date standards, technologies, and architectures, which will be fully implemented at the national level. This model places special emphasis on the implementation of a minimum set of standards, which includes:

- WHO Family of International Classifications (e.g., ICD-9/10/11, CIAP)
- Clinical Procedures Terminology (CPT)
- Standards for laboratory data (e.g., LOINC)
- Standards for pharmaceutical data (e.g., national drug codes)
- Rules relating to basic national data (data dictionary)



Where can I find more information?

Revisión de estándares de interoperabilidad para la eSalud en Latinoamérica y el Caribe [Revision of Interoperability Standards for eHealth in Latin America and the Caribbean]. Available at: <https://iris.paho.org/handle/10665.2/28188>.

Electronic Health Records (EHR) and Interoperability: Understanding two key concepts for better public health response. Available at: <https://iris.paho.org/handle/10665.2/52004>.

Interoperability in Public Health. Available at: <https://www3.paho.org/ish/images/toolkit/IS4H-KCI-EN.pdf>.

Contacts:

- Sebastián García Saiso, Director of the Department of Evidence and Intelligence for Action in Health (EIH)
- Myrna C. Marti, EIH Advisor for Information Systems and Digital Health
- Marcelo D'Agostino, EIH Senior Advisor for Information Systems and Digital Health

Acknowledgements. PAHO/WHO thanks the Spanish Agency for International Development Cooperation (AECID), the United States Agency for International Development (USAID), and the Government of Canada for their support.

Special thanks. This capsule was prepared under the leadership of Dr. Carlos Otero, PAHO's International Consultant for the IS4H Initiative, and was developed in collaboration with the Division of Social Protection and Health of the Inter-American Development Bank (IDB), the Department of Health Information of the Italian Hospital of Buenos Aires (PAHO/WHO Collaborating Center for Knowledge Management), the Open University of Catalonia (PAHO/WHO Collaborating Center for e-Health), the Center for Health Informatics of the University of Illinois (PAHO/WHO Collaborating Center for Health Information Systems), the Central American Network of Health Informatics (RECAINSA), and the PAHO Network of Experts in Health Information Systems (IS4H).

References

1. Márquez Fosser S, Gaiera A, Otero C, Benitez S, Luna D, Quiroz F. Automatic loading of problems using a comorbidities subset: One step to organize and maintain the patient's problem list. *Stud Health Technol Inform.* 2017;245:1358.
2. Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS. Toward a medical-concept representation language: The Canon Group. *J Am Med Inform Assoc.* 1 May 1994;1(3):207-17.
3. Vreeman DJ, Richoz C. Possibilities and implications of using the ICF and other vocabulary standards in electronic health records. *Physiother Res Int.* December 2015;20(4):210-9.
4. Pan American Health Organization. Electronic Health Records and Interoperability: Understanding two key concepts for a better public health response. COVID-19 Factsheet No. 2, 13 April 2020 [accessed 2 July 2021]. Available at: <https://iris.paho.org/handle/10665.2/52003>.
5. Dolin RH, Alschuler L. Approaching semantic interoperability in Health Level Seven. *J Am Med Inform Assoc.* January 2011;18(1):99-103.
6. Moreno Conde A, Moner D, Dimas da Cruz W, Santos MR, Maldonado JA, Robles M, Kalra D. Clinical information modeling processes for semantic interoperability of electronic health records: systematic review and inductive analysis. *J Am Med Inform Assoc.* 2015 Jul;22(4):925-34.
7. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 11th revision (ICD-11) website [accessed 13 July 2021]. Available at: <https://www.who.int/standards/classifications/classification-of-diseases>.
8. Global Family Doctor–WONCA. WONCA website [accessed 13 July 2021]. Available at: <https://www.globalfamilydoctor.com/groups/WorkingParties/wicc.aspx>.
9. SNOMED International. SNOMED website [accessed 13 July 2021]. Available at: <https://www.snomed.org/>.
10. LOINC. Logical Observation Identifiers Names and Codes website [accessed 13 July 2021]. Available at: <https://loinc.org/>.
11. NANDA International Inc. NANDA website [accessed 13 July 2021]. Available at: <https://nanda.org>.

12. World Health Organization. Anatomical Therapeutic Chemical (ATC) Classification website [accessed 13 July 2021]. Available at: <https://www.who.int/tools/atc-ddd-toolkit/atc-classification>.
13. Luna D, Lopez G, Otero C, Mauro A, Casanelli CT, González Bernaldo de Quirós F. Implementation of interinstitutional and transnational remote terminology services. AMIA Annu Symp Proc AMIA Symp. 13 November 2010;2010:482-6.
14. González Bernaldo de Quirós F, Otero C, Luna D. Terminology services: Standard terminologies to control health vocabulary: Experience at the Hospital Italiano de Buenos Aires. Yearb Med Inform. August 2018;27(01):227-33.
15. Campos F, Otero C, Kaminker D. Principios de interoperabilidad en salud y estándares (Principles and standards of interoperability in health). E-book. Italian Hospital of Buenos Aires. 2018.
16. Bakhshi-Raiez F, Cornet R, de Keizer NF. Development and application of a framework for maintenance of medical terminological systems. J Am Med Inform Assoc. 1 September 2008;15(5):687-700.
17. Kanter AS, Wang AY, Masarie FE, Naeymi-Rad F, Safran C. Interface terminologies: Bridging the gap between theory and reality for Africa. Stud Health Technol. Inform. 2008;136:27-32.
18. Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: Facilitating direct entry of clinical data into electronic health record systems. J Am Med Inform Assoc. 1 May 2006;13(3):277-88.
19. Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. A model for evaluating interface terminologies. J Am Med Inform Assoc. 1 January 2008;15(1):65-76.
20. Fung KW, Xu Julia, Bodenreider O. The new International Classification of Diseases 11th edition: A comparative analysis with ICD-10 and ICD-10-CM. J Am Med Inform Assoc. 1 May 2020;27(5):738-46.

PAHO/EIH/IS/21-023

© **Pan American Health Organization, 2021.** Some rights reserved. This work is available under license [CC BY-NC-SA 3.0 IGO](https://creativecommons.org/licenses/by-nc-sa/3.0/).