Imagining 2029 webinar series:
Moving Towards European Health Data Space(s)

Towards European data spaces for medicines
3rd EHTEL/ELO Network factsheet

This factsheet is supported by UNICOM project

EHTEL Association, a.i.s.b.l.
Avenue de Tervuren, 168, box 2 B-1150 Brussels Belgium Tel: +32 2 230 15 34 Fax: + 32 2 230 84 40 info@ehtel.eu
www.ehtel.eu

This factsheet is mainly based on the information collected during the third EHTEL/ELO virtual meeting dedicated to the theme "Towards European data spaces for medicines". After a brief summary of the content of the previous two webinars, this factsheet explains why the concept of European Health Data Spaces is particularly relevant for **medicinal products** with a focus on **pharmacovigilance**. It explains how the **implementation of the identification of medicinal products** (IDMP) standards, supported by the UNICOM project, could be **a real game-changer**. It also analyses what has already been **done by Member State**s in this field, with **Belgium** as an example, and considers the necessary **next steps**. The webinar was hosted by EHTEL on Monday, 21st September 2020, with the support of the **EHTEL ELO Network** and its Co-Chair **Andreas Grode** (Gematik, Berlin - Germany).

Table of contents

Introduction to data spaces and medicinal products

Health data lakes exist everywhere in Europe and have been created for specific purposes.[1] Some have a name and are duly referenced; others are totally unknown and exist in total isolation. Connecting existing lakes to create data space(s) is not a trivial issue. It usually requires the **intervention of intermediaries** whose job is to make the data exploitable for another purpose while keeping in mind the original context in which the data were created.

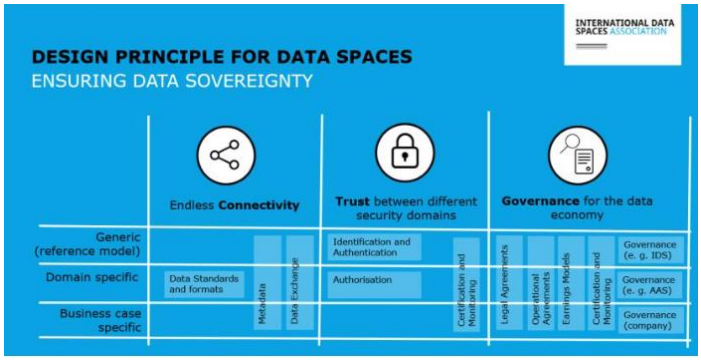
Correctly identifying existing data lakes to create an embryo of **a National Data Space** is already an important challenge which requires a dedicated action plan. However, the national level is often too limited to support many crucial objectives related to e.g. patient safety, research, knowledge management, innovation support, or artificial intelligence (AI). It is thus important that the **key building blocks** which enable the creation of these data spaces are capable of supporting a cross-border or a cross-domain perspective.

The vision proposed by the **European Commission** goes beyond linking together existing data lakes but aims, from the very start, at **enabling data to become part of new, open data spaces**. Aside from the investment in data-based digital services and strong leadership and cooperation, **a radical cultural change** is needed to allow ethical data sharing.

The ELO team has already outlined the key aspects and concepts set out by the European Data Strategy of the European Commission and explored several **reference architectural propositions**, in particular, one developed by the International Data Spaces Association (IDSA) that has the ambition to become a '**cross domain reference architecture'**. The reference architecture differentiates between what is domain specific and what is not. The IDSA model was presented by the Open Dei project during an

[1] One of the most well-known health data lakes is the Patient Shared Electronic Health Record which combines both structured data and free text.

EHTEL webinar held on 20 May 2020: European Strategy for Data: Pathways for moving towards (Health) Data Space.



The IDSA model relies on **multiple layers** such as business, function, process, information and, system. Between, and common to, all these layers are **transversal functionalities** that foster security, certification, and governance. In the **business case**, the model specifies the roles for actors that would govern the data flows between different domains or data spaces. Key participants (i.e., actors in the system) would be the Data Owner, Data Provider, Data Consumer, and Data User or Broker Service.

Figure 1: IDSA design principle for Data Spaces

The first factsheet explored briefly **models and initiatives developed by Member States** to create their own National Data Space.

The second factsheet focused on the key question of the **trust** between European Member States, companies, platforms and, more generally, data ecosystems[2] which is an absolute pre-requisite for the creation of data spaces. It examined, in particular, the questions of governance, transparency and consent. It streamlined the **key proposals made by Member States to support governance**, and explored – with the InteropEHRate project[3] – how a solution which fully empowers the patient in term of data control can contribute to this change in paradigm. It discussed how **privacy** and **ethics compliance "by default"** could be implemented in a seamless way at European level.

In this **third factsheet**, the spotlight is turned on another crucial condition for the effective deployment of data spaces: **the wide availability of quality and interoperable data**. The literature abounds with examples that show the tremendous importance of the availability of adequate **high-quality data** to support advanced processes such as AI. It is widely recognised that *ex post* data cleaning is often ineffective and costly, especially in the absence of proper metadata and domain knowledge sources.

To illustrate the importance of high-quality data, a specific health domain has been selected: that of **medicinal products**. It shows how the use of adequate interoperability standards along the whole value chain – inclusive of all actors potentially impacted – would make an important difference for the creation of exploitable data spaces.

Medicinal products-related data are the most widespread of health data. They are usually the priority first type of data to be entered into clinical systems as they have direct links with third parties e.g., reimbursement schemes. These data are also of a very complex nature. A lot of effort has already been devoted to the **development of standards** which allow their detailed and complete description in support of multiple use cases. Due to the multiplicity of actors involved and client systems, quite a few Member States have dedicated **important resources** to provide an integrated response to the needs related to medicinal products information for all the clients in their national eHealth ecosystem(s).

[2] See also EHTEL factsheets dedicated to data ecosystems.

[3] https://www.interopehrate.eu

The lack of readability, availability and accuracy of prescribed and/or dispensed drugs is one of the key factors in explaining **medical errors**. In 2016, the World Health Organization (WHO)[4] estimated average error rates of 3% at the dispensing stage while other studies[5] have reported that in-patient hospital medication error rates are between 4.8% to 5.3%. These errors have a very important impact on patient safety, quality of life, and public finances.

The current **digitalisation of the health sector** may contribute substantially to reduce **the threats to patient safety**. However, without the adoption and concrete implementation of common conceptual and semantic standards to identify such drugs, the expected added-value cannot be delivered in full (e.g., if data are unable to travel between systems, regions, countries or use cases). Although currently still of limited scope, **ePrescription at European level** experiences important limitations due to the difficulty of correctly identifying and substituting a defined product by another.

Pharmacovigilance: An example of a data space in the medicinal product field

Looking at patient safety from the angle of **pharmacovigilance**[6], the reference space is truly European.

The COVID-19 crisis is providing multiple examples of the **urgent need to correctly identify medical products** due to the need to rely on quality information in order to take adequate decisions on pharmacovigilance, drugs shortages, and dynamic clinical trials. A global solution to this issue would vastly improve challenges linked to **polypharmacy** and **decision support**, including decisions related to **pricing**, extending far beyond national or regional borders.

Miriam Sturkeboom (i~HD, Medical University of Utrecht) provided the webinar with several enlightening examples of how a **better identification and description of medicinal products and other critical data** in Europe, and beyond, can support pharmacovigilance objectives.

One example was related to the need to rely on **big data** and thus to access large data spaces to evaluate medicines. Studies have shown that, if massive access to the data contained in electronic health records had been granted, **major side effects related to new drugs** – which could only otherwise be documented after years – could have been identified in just a few months. In the **United States of America (USA)**, the Federal Drugs Administration (FDA) thus concluded that it had to totally change the way it was evaluating medicines safety and that a "system of systems" had to be created, giving access to a 100 million medical records to ensure safety analysis[7]. This commitment was translated into a FDA law amendment. In **Europe**, several initiatives[8] have been taken to link together different data lakes focused mainly on distributed analytics and a Common Data Model. The European Medicines Agency (EMA) has initiated a number of European studies to evaluate the effects of certain drugs classes. New pandemics, such as H1N1 in 2009, have provided supplementary **motivations to federate data lakes** and **develop common data models**.

---

[4] WHO, Medication Errors: Technical Series on Safer Primary Care, ISBN 978-92-4-151164-3

[5] See for example: Belén Jiménez Muñoz A, Muiño Miguez A, Paz Rodriguez Pérez M, Dolores Vigil Escribano M, Esther Durán Garcia M, Sanjurjo Saez M. Medication error prevalence. International journal of health care quality assurance (2010)

[6] Pharmacovigilance is drug safety.

[7] https://www.fda.gov/safety/fdas-sentinel-initiative

[8] Examples include the EU-ADR, SOS and IMI-Advance projects.

In pharmacovigilance, the harmonisation of quality data has become ever more important. Efforts at harmonisation were initially mainly related to the coding of events (i.e., diagnosis). However, even if solutions have now been found to map the different codes and terms used in different countries, the process still requires an important investment of time and resources on the part of the data producers to comply with the agreed requirements. In the example of a study related to an **anti-inflammatory drug**, once the mapping had been completed, access to multiple data lakes allowed both to confirm the impact of the drug on people who had a heart attack or experienced heart failure and to develop much better overall estimates. The **lack of harmonisation of attributes attached to drugs, such as dose and strength**, is a serious limitation to the generalisation of research results.

Other examples of the need for data harmonisation include **vaccine safety** and **medicine safety during pregnancy**. In the example of H1N1 vaccine safety monitoring, the semantic resources used to describe and code the vaccines through accessing many data lakes were unable to provide the level of detail required. (This will probably also be the case when upcoming COVID-19 vaccines are monitored.) In a more recent project[9] launched in 2019 aimed at monitoring and communicating medication safety in pregnancy and breastfeeding, the previous approach has been modified to reach improved results. While the syntax of all the data analysed still needs to be taken care of by the local data producers' systems, the **semantic harmonisation** of all the data is now done centrally.

Coding systems used in regular dictionaries do not allow for proper identification and datasources cannot even code to ATC

Property category	SNOMED-CT	Read-2	MeSH	ATC	BNF	AHD
Pathogen	✓	✓	✓	✓	✓	✓
Disorder	✓	✓	✓	✓	✓	✓
Vaccine strategy	✓	✓	✓	✓		✓
Ingredient		✓	✓		✓	✓
Route		✓	✓	✓		✓
Valence		✓	✓	✓		

Figure 2: Status of vaccines' description in dictionaries used in the IMI-Advance project

Today, it is obvious that – beyond the access to a wide range of data lakes – creating **a Data Space for Pharmacovigilance** would still require **substantial levels of investment** in order to deal with the heterogeneity of pharmaceutical data and, in particular, medicinal products-related data.

Adapted user interfaces, which reduce the need of intermediaries for data coding and global reliance on standards such as IDMP, can play a very important role in improving this situation. Adapted user interfaces will help not only to improve pharmacovigilance, but also to make sure that the **available data lakes can be used for a wide range of purposes which have positive impacts on society and the economy**.

As this examination of the field of pharmacovigilance has shown, without **advanced semantic interoperability** the benefits emerging from the creation of European Data Spaces will remain limited.

[9] The Conception project is funded by the Innovative Medicines Initiative (IMI).

The need for semantic interoperability

Various **barriers to semantic interoperability** remain. The most important barrier to semantic interoperability is probably related to the long-term existence of **silos** at all levels. Every outcome, every domain, and even every pathology, has defined its own universe while taking into consideration the needs and roles of a somewhat closed system. Many silos have achieved results for the very specific purpose they were pursuing and are now reluctant to leave their comfort zone or to see the added-value of a multi-purpose approach. Today, **there are not yet any 'fit-for purpose', globally agreed standards (of concepts, data models, or resources), coding systems, or implementation guidelines which can ensure high-quality data at all levels of use**. As a result, data do not travel well.

The universe of the **correct identification of medicinal products** is also of a very complex domain which requires the **investment and collaboration of multiple actors**. In many of its aspects, it could be compared to a "nested doll". No country or region can deal with this complexity alone. This challenge requires a high level of **awareness and motivation** to face the many problems created by disruptive change and, of course, a **capacity to mobilise the subsequent necessary resources.** Many Standards Development Organisations need to cooperate actively and structurally in order to deliver **implementable results**.

Most resource-rich countries maintain at least basic **national (electronic) repositories and databases of medicinal products** which have gone through the stipulated national regulatory process to be marketed in that specific national healthcare system. Unfortunately, these **uncoordinated national regulatory procedures** have resulted in a host of unintended consequences and impacts that endanger patient safety and hinder better healthcare service delivery, particularly in international contexts. Across health systems, exactly the same medicinal product may have different names. Across countries, the same name may identify a different product with a different active substance. The number and kind of medicinal products authorised for national marketing may also differ very considerably.

In **cross-border ePrescription services** this situation thus necessitates substitution in many, if not the majority of, instances – when a particular medicinal product is specified in a prescription. Similar challenges apply to the **electronic recording of medicinal products in other healthcare contexts**, e.g. in electronic patient summaries, health records, clinical decision and ordering systems, or in ePrescribing software.

Although **global medicinal product semantic interoperability** might simplify lots of the processes that the pharmaceutical industry needs to manage, it could somewhat hinder their marketing strategies which are very much built at national level.

> **Karl Stroetmann** (empirica, coordinator of the UNICOM project) noted that: "while the EMA has registered around 600,000 medicinal products, only an average of 20,000 are recorded for use at national level".

Quite a number of these medicinal product data are also quite old and would need to be updated, thus representing an important administrative burden to resolve for the pharmaceutical industry.

Aside from their impact due to the quantity of information produced, medicinal products-related data are also used by an impressive number of actors and systems in all segments of the value chain. The most important expected benefits from the use of IDMP is to **enable the seamless exchange and sharing of health data** related to medicines across all actors and stakeholders involved in handling or consuming such data, thus making it possible to **re-use the data** for multiple purposes. Other benefits include the improvement of the **quality of the clinical documentation** used by healthcare professionals and

facilitation of the use of **personalised decision support systems**. Another advantage is the direct impact on communication to the patient who will have **access to clearer and actionable information**. It will also create **important efficiency gains** in both public and private sectors as a lot of expensive mapping and duplication work will be avoided. Semantic interoperability related to medicinal products will deliver **individual, social, societal, and economic benefits**.

The urgent need for interoperable and qualitative data is applicable to all data produced in the healthcare domain.

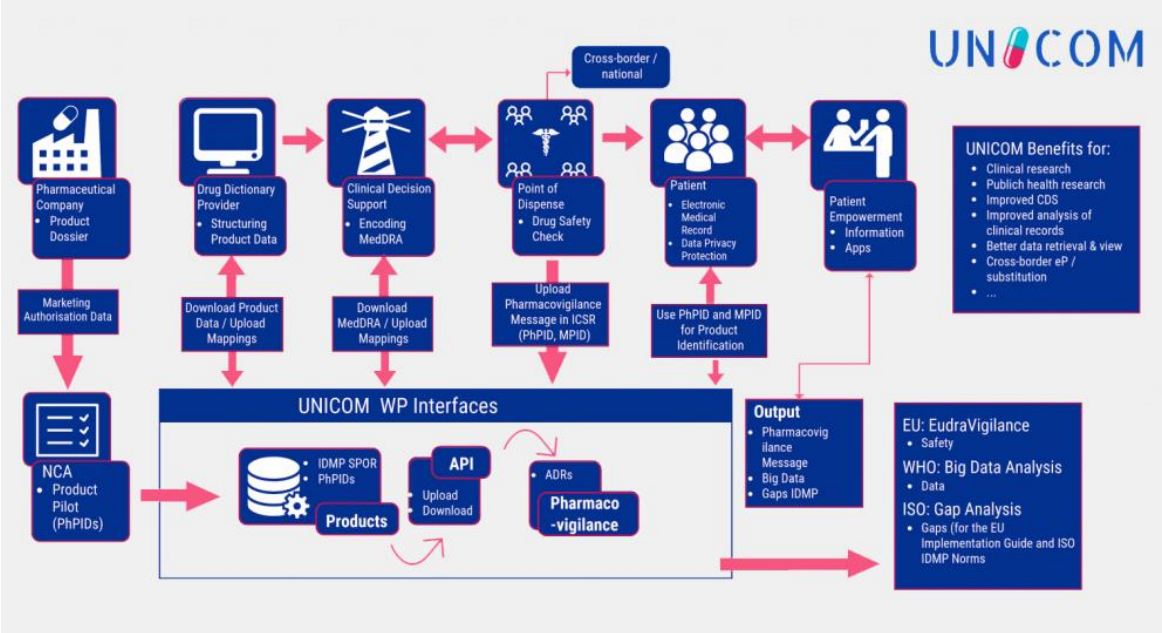


Figure 3: The UNICOM IDMP-related value chain is in essence multi-purpose

The example of semantic interoperability related to medicinal products is a powerful example, but much **more action** is of course needed. The need for real semantic interoperability is present for **all categories of health data** (e.g., diagnostics, procedures, observations, parameter risks, allergies, laboratory results, and images). It is also needed to ensure the conditions for safe sharing of such data, such as the availability of **validated and federated identifiers** for patients and healthcare professionals. The level of global harmonisation of data will thus have a direct impact on the real **usability and efficiency of the data spaces** created.

This harmonisation work on semantics in the health sector was initiated some decades ago. Even if progress has been made, **much remains to be done** to create semantic interoperability at the point of care. The use of intermediaries – whose job is to format or prepare the data for a specific purpose – is thus still likely to last for a while. Loosely or partially harmonised open data spaces will continue to exist alongside highly structured and controlled data lakes.

National Health Data Space perspectives in the medicinal product field: The Belgian example

Many Member States recognised early on **the importance of harmonising medicinal products** at least at national level. The willingness to implement a safe and legal ePrescription has been a strong driver, but other factors have also played a role.

In Belgium, the concept of "Authentic Validated Source" is a key component of the national e-Gov and e-Health strategies. An **authentic source** is the gold standard in any national IT organisation for obtaining specific data. It offers specific guarantees in terms of the **accuracy, completeness and availability** of the data. This source is the only legally binding reference for the data it contains. It has thus been applied to the domain of medicinal products.

In order to create that authentic source, it was necessary to collect the requirements of all the data producers and the data users. This meant mainly bringing together administrative and clinical perspectives of data and linking them altogether in a single conceptual model. The administrative perspectives included market authorisation, pharmacovigilance, and reimbursement. The clinical aspects involved scientific information, documentation, and modalities of prescription. Belgium, like quite a number of other nations, adapted a model developed earlier in the United Kingdom (the NHS **dictionary of medicines and devices [dm+d]**[10]) and adapted it. No fewer than six Belgian public and non-profit entities, which used to work in silos, had thus to agree to adapt their data model and work together in an integrated infrastructure.

> "Creating a virtuous national infrastructure for the identification of medicinal products is also about creating interoperability between all the public bodies involved", said **Luc Nicolas** (EHTEL).

On the users' side, in the past, Belgian pharmacists, clinicians, researchers, and hospitals used to rely on their own **medicinal products dictionary** or on the one provided by the software that they purchased (which would then each make use of their official available data). The creation of the authentic source has completely changed this situation: all systems now have a full and direct access to all the data they need to provide the related services to their clients. The medicinal products' authentic source is also capable of supporting advanced functionalities, such as international non-proprietary name (INN) prescribing[11] or conditional reimbursement.

The way forward

The creation of a reference national medicinal products database is **a huge investment** for a country. A considerable amount of medicinal products' data is being created in a European context through a European registration process with EMA. Member States should ideally concentrate on the attributes necessary for the functioning of their health system and the associated legal requirements, while the main clinical-related data should be the same for all.

> "We need to invest in a more clever way," said **Luc Nicolas** (EHTEL).

This is the promise offered by **the wide implementation of the IDMP suite of standards** that UNICOM is aiming at. **Linking national drugs databases to the IDMP standard** will already provide large benefits, although the full harmonisation of national drugs database(s) will certainly take more time. The domain is a very complex one. People cannot wait to have solutions in place to all the pending issues to progress. Any progress is useful that will already allow people to improve cross-border data migration (such as prescription or knowledge databases); ensure comparison between countries (such as prices, national therapeutics arsenals, drugs utilisation, quality of prescribing and dispensing); and foster cross

[10] https://datadictionary.nhs.uk/supporting_information/nhs_dictionary_of_medicines_and_devices.html

[11] E. Van Bever, et al., Operational rules for the implementation of INN prescribing, Int. J. Med. Inform.(2013),http://dx.doi.org/10.1016/j.ijmedinf.2013.09.004

national cooperation (such as the creation of a **European Data Space for multi-centres pharmaco-epidemiology**).

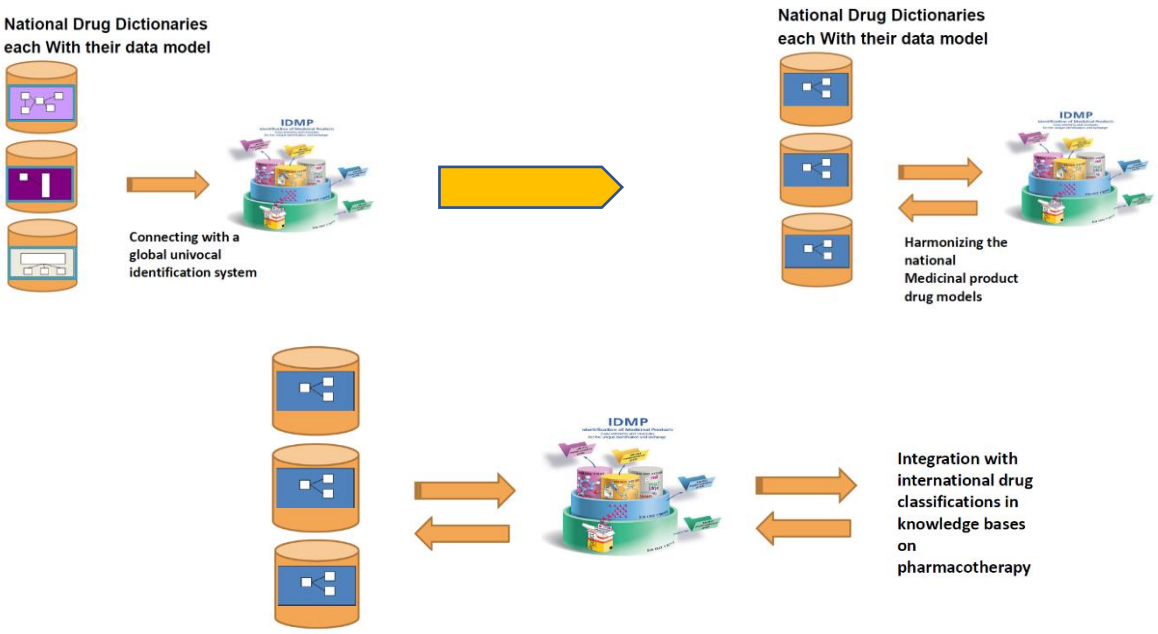


Figure 4: The journey towards IDMP harmonisation of drugs' databases

> "If, in four years' time – the UNICOM time frame – we are ready with simple medicinal products, this will already be a big achievement. The complex aspects need to be addressed but we need to give us the time to do it right. Let's thus all pick up on the achievements and put them into practice", said **Robert Vander Stichele** (i~HD).

Conclusions: The relevance of the medicinal domain for European Health Data Spaces

To ensure the benefits of data-driven health services, **a broader range of health data should be available** for reuse and to be shared in European Data Spaces. Beyond the ability to share data in a transparent, open and safe way, the main obstacles to the effective use of the emerging spaces remain **insufficient data compatibility and interoperability**. The more **structured and coded data** is to be found in clinical records, the more emerging data spaces will be able to deliver the expected benefits.

In comparison with other sectors which are key targets for the emergence of EHDS, such as the environment, energy, mobility, agriculture and manufacturing sectors, **the health sector needs to address a very high level of complexity**. Beyond the data itself, the **context** of creation of the data is of major importance and needs to be captured in order to guarantee correct interpretation and use. Medicine is indeed not pure science.

The **medicinal products domain** offers a unique opportunity to work on the major key components necessary to the creation of EHDS: documented use cases, established networks, trusted architectures and, last but not least, a global standard for data harmonisation.

Pharmaceuticals and vaccines have become a hard topic to bypass in the COVID-19 era, and this fact will without any doubt provide a major supplementary impetus. **Medicinal products-related use cases**, in particular in the pharmacovigilance domain, will thus most probably be one of the key testing grounds for European Health Data Spaces. For all the actors involved in medicinal products' data processing,

taking an active part in **implementing the IDMP standards** will make those spaces not only real but also truly operational.

To conclude, **the medicinal products-related European Health Data Space is one of the spaces that will offer the highest of opportunities to enhance and improve Europe's health systems and services, health of the people, and offer solid opportunities to European businesses and research centres**.

These spaces will act as a major **game-changer.** Their development will deeply affect a number of firmly established parameters and rules associated with the "private" access and use of data, and modify the rhythm and cycle of innovations.

Useful additional information

Harmonisation of clinical events in the pharmacovigilance domain: EU-ADR project

Vaccine collaboration monitoring for Europe: Vac4EU

IMI projects working with the European Medicines Agency (EMA) on COVID-19: EHDEN ConcePTION and ADVANCE/VAC4EU

European Medicines Agency (EMA): IDMPs standards overview

Belgian Medicinal Products "Authentic source": SAM

The Joint Initiative Council (nine Standard Development Organizations) white paper that offers practical solutions to real-world problems

Future EHTEL symposium and webinars will go on exploring the benefits and challenges of **European Health Data Spaces**. One session of the December 2020 EHTEL Thought Leaders Symposium will be specifically dedicated to users' acceptance of and engagement with this concept. Other Imagining 2029 webinars and factsheets focusing on **digital therapeutics** and **integrated care** might also be of interest for the readers of this factsheet.
Check out the EHTEL website for further details.

Acknowledgements

This report content is drawn from presentations made by EHTEL members and friends and, in particular, **Miriam Sturkeboom** (i~HD and Utrecht UMC, Department of Epidemiology, Belgium/Netherlands), **Karl Stroetmann** (empirica, Germany), **Jos Devlies** (Eurorec, Belgium), **Robert Vander Stichele** (i~HD, Belgium) and **Luc Nicolas** (EHTEL). EHTEL is especially grateful to the European Union-supported project UNICOM for its assistance in facilitating the webinar on which this factsheet is based.

.

For more information about EHTEL's ELO Network and its work on European data spaces:
Contact the EHTEL Secretariat - communication@ehtel.eu

The UNICOM project is funded by the European Union (EU) Horizon 2020 programme under Grant number No. 875299.