Gravitate Health

innovative medicines initiative

Gravitate-Health

WP1 – WP User needs, scenarios, KPI

D1.4 G-Lens specification - information models

Due date	31/10/21
Delivery date	29/10/21
Deliverable Type	R
Dissemination Level	PU

Lead contributor	16 KI Martin Ingvar / 32 Bayer Koen Nauwelaerts

Other contributors	UiO, PFIZER, HL7 Europe, DW, UPM, Empirica, EPF, I-HD, HIMSS Europe, NSE, Trinity, AHUS, ECHA, MINDVIEW, UCSC, SPMS, NoMA, The Synergist, Open Evidence, UCPH, SERMAS, CBG, Bayer, GRT, EliLilly, Viatris, NOVARTIS, ROCHE, UCB, Datapharm

The Gravitate-Health project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 945334. This joint undertaking receives support from the European Union's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations [EFPIA]. The total budget is 18.5M€ for a project duration of 60 months.

efpia Datapharm

Document History

Version	Date	Description
V0.1	June 2021	First draft
V0.4	July 2021	First full draft including addenda – review required
V0.7	September 2021	Review Giovanna Ferrari Project Lead and Anne Moen Project Coordinator
V0.8	29.09.2021	Final edits before formal review
V0.9	18.10.2021	Review by Petter Hurlen and Jürgen Hauck
V1.0	23.10.2021	Post review edits and resubmission for final approval
V1.1	27.20.2021	Final edits by Line Løw, edits and approval by Anne Moen, UiO and Giovanna Ferrari, Pfizer

Disclaimer 1: *The present deliverable reflects the authors' view and neither IMI nor the European Union, or EFPIA, are responsible for any use that may be made of the information contained herein.*

Disclaimer 2: *Some of the texts describing the different sources of information are derived from the source texts issued by the respective responsible issuer of the information.*

List of addenda:

D1.4 Addendum 1 Summary of the available information on the proposed scenarios from T1.1

D1.4 Addendum 2 Mapping the different scenarios on the information process model

D1.4 Addendum 3 A table analyzing the structure of the SmPC

TABLE OF CONTENTS

1.1 LIST OF FIGURES

1.2 LIST OF TABLES

1.3 Vocabulary

Here certain concepts are listed as to improve the accessibility of the text.

Focusing: Working concept definition in the Gravitate-Health project. Focusing entails the use of regulated information that is presented without changes except parts that are focused for presentation. All information is available and nothing is omitted from the original text.

Information model: An information model is a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. The advantage of using an information model is that it can provide sharable, stable, and organized structure of information requirements for the domain context[1].

Information process model: In informal model level that serves to unite the understanding of steps of information gathering, information processing and information presentation between different stakeholders in the project.

International Patient Summary: A Patient Summary is defined by ISO/TR 12773-1:2009 as a "Health record extract comprising a standardized collection of clinical and contextual information (retrospective, concurrent, prospective) that provides a snapshot in time of a subject of care's health information and healthcare."

Regulated information: A source that has a specified level adequate information, and a specified governance that includes a defined responsible issuer (Binary concept).

Structured data is comprised of clearly defined data types with patterns that make them easily searchable and possible to analyze. Structured data analytics is a mature process and technology.

Trusted information: A quality statement with no relation to regulatory bodies. Information of any sort may be used by health professionals based on their professional standing. Information to the public within the medical domain (disease prevention, diagnosis and treatment) must have a responsible body or HCP behind it. Verified ID patient generated data is considered authentic and therefore constitute trusted information.

Unstructured data – has a low or inconsistent internal structure and is therefore not easily used for computational purposes, such as focusing (or personalization/ filtering). Unstructured data is comprised of data in formats like free text, audio, and video. Unstructured data is of course human readable, but the informal hierarchical structure of natural language comes with low structure for computability.

[1] Y. Tina Lee (1999). "Information modeling from design to implementation" National Institute of Standards and Technology.

2 EXECUTIVE SUMMARY

Delivery 1.4 summarizes the findings from Task 1.3 in the Gravitate-Health project. The ambition of the task was to characterize the information landscape in which Gravitate-Health exist. We have focused on the ability to serve properly focused drug information to the end user which is the main focus of the Gravitate-Health project.

The analysis of information sources demonstrates that there are a number of official repositories with pan-EU coverage and full coverage of the drugs available. The overall characteristics of the current landscape are the de-facto result of an **exclusively manual regulatory procedure** with information produced and managed as text documents based on use of text-templates (often MS Word templates). The end result is overall information with low resolution and low structure. In crucial components some inconsistencies between document have been noted adding to the difficulties of amalgamating information from different sources.

The most prioritized source of information, and central to Gravitate-Health, is the regulatory approved product information in an electronic format; ePI –electronic product information – that holds the among other things the product leaflet (PL) and the summary of medicinal product characteristics (SmPC). There is an initiative currently in progress to establish an EU common standard for ePI.

Refinement of the information by focusing on relevant components will be done with structured information collected from the end-user. The end-user will be served by fitting the collected information on to archetypical dimensions for personalization. Also, other sources with the end user consent will be imported to the G-Lens, e.g., international patient summary and data from prescription/ dispensation registers.

An information process model and a low resolution information model according to this analysis is presented. The process model has been mapped on the proposed user scenarios. This analysis has shown that the presented information process is compatible with the proposed scenarios with the caveat that some of the scenarios are sparsely presented at this stage.

There are a number of considerations emanating in D1.4 regarding the ability for Gravitate-Health to make a sustainable difference:

1. Already the most relevant information for assistance in drug use is available online on most EU countries, in some cases via attractive and comprehensive solutions for local use, so there will need to be 'adding value' to the end-user to encourage uptake or support cross-border care.

2. The ability to present individually focused information with high precision is limited due to the quality of the information in the regulated sources at the present time.

3. The speed of implementation of the ePI may be too slow to provide the necessary basis for the Gravitate-Health project during the project period. This risk was highlighted at the outset of the project, with mitigations identified in the Description of Actions document. An important way of mitigation is to follow the evolving ePI landscape during the project.

3 Point of Departure

This is the report from task 1.3 that aims to specify a prototype for an information model that could provide a basis for an alignment across the Gravitate-Health Consortium on interoperability and support a rational development of software and data architecture. The quest to investigate the possibility of defining a common information model is very important given that the Gravitate-Health project is serving many countries, many cultures, has an ambition to serve subjects with limited digital maturity with an open-source platform that is created in a widely distributed setting.

The following steps have been identified in the task 1.3 that constitute this deliverable D1.4

A task objective is to take input from T1.1 (End user requirements) and T1.2 (User personas)

The task will identify, review and map other trusted information sources – such as the international patient summary (IPS) and health education material (well prepared / peer reviewed/endorsed by HCP + authorities), Product information, EPAR (European Public Assessment Report) and other trusted sources

A clear prioritization of the data sources and consider the access to structured data,

Align data sources with the goals of Gravitate-Health

Early creation of mock data with structured granularity that adheres to the requirements of the selected scenarios should be developed to benefit the progress of the different use cases.

As general mechanisms for individualization of the information from trusted sources will be adopted, the proposed scenarios should feed into and align with a common information model (T1.3).

Mapping to KPIs (T1.4)

A mapping of each proposed scenario against the information model (T1.3)

A strategy to consecutively include new developments in the information structure provided from trusted sources

These steps are accounted for below under separate headings.

Also, the task was prescribed to coordinate with other parts of the ongoing project.

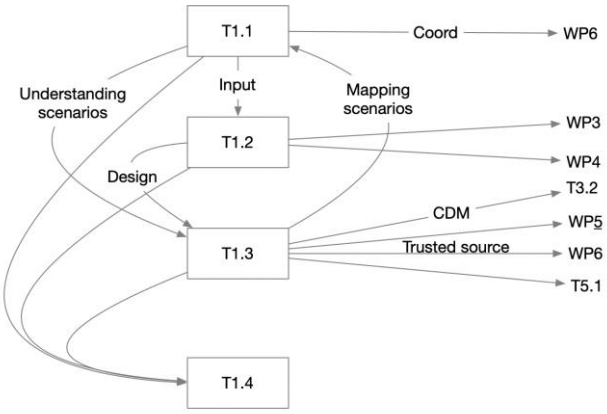

Figure 1. Task 1.3 received most input from other tasks in WP1.

Already at the outset of T1.3 potential hurdles to reach the goals of the project were identified and different ambition levels were identified in order to discuss the consequences of the findings in T1.3. Further mitigation strategies will be discussed and implemented going forward.

3.1 The task of creating an information model

Models represent a tool to convey a systems level understanding of the project and calibrate the understanding between participants that contribute to the different levels of the project (content, semantic understanding, software architecture, hardware architecture, legal and sustainability). The general description of the Gravitate-Health project sets the basis for a suggested information process model. The information process model defines, with low resolution, the different streams of information that need to be defined in order to reach the goal. This low-resolution model provides a basis for consensus on the general construct of the Gravitate-Health information landscape. The information process model defines parts of the information model which is a static model that is supposed to clarify how the matching of different information sources can be utilized together to create the desired end result. This means that all participants need an agreement on the information process model before the finalization of the information model. The model as such forms the basis for a consensus on how to design and pursue the development of the G-Lens in the distributed environment that will serve the different scenarios.

A common information model is the basis for the management of the development process and the goal of a proper service that is high in demand by the end users. A comprehensive information model is aiding in the final development of the user experience with the goals:

- You find first experience promising
- 2-3 clicks and you find your info

- The G-Lens talks to you as it has anticipated your needs
- Proper cross-referencing provides a solid basis for further search
- Similar information is presented the same way - you feel at home
- You easily decide what is trusted vs. non-trusted information
- The tool welcomes you!

The information model we construct has a three-tiered structure.

- The first tier of constructing the Information Model consists of the dimensions that identify how the information will be categorized and labeled for both internal and external use in Gravitate-Health. An important question is if the available information has computable properties (structure, vocabulary, dimensionality). In this step an information process model is used as a tool.
- The second tier assesses the information sources.
- The third tier provides structure for each information type, outlining the content units that authors use to build information types.

All through the work a clear vision is that the developed information model will assist all layers of the gradual development of the G-Lens (see fig 2).

Figure 2. To maintain interoperability in all layers the developed information model must continuously be challenged with these four perspectives.

3.2 The stakeholder requirements (T1.1 and D1.1)

The stakeholder requirements have been identified and analyzed in Task 1.1 and reported in Deliverable 1.1. An important conclusion is that while there was a wide span in the identified stakeholder requirements many of them were already incorporated in many of the proposed user scenarios. Some of the

identified requirements were identified as being directed towards general user friendliness.

Cross border services from the G-Lens were noted as a goal in the application. However, the finding that the information landscape varies considerably between countries means that while developed general services may provide substantial value in one country, in others more consideration needs to be given to ensure 'value add' given that national services may have a both wider and deeper scope above and beyond ePI, and already be deeply established in the local target groups. The cross-border services in the first round pertains to support of multiple languages separate from preferred language option.

The scenarios (as stakeholders) have in part already developed their strategies based on the local information landscape and several have reported to have access to structured drug information. The development speed of the ePI will be a critical factor for markets where structured information is not currently available.

The potential for two-way communication above and beyond the query situation/use of Real World Data was a recurring theme in the scenario feedback. This has issues of responsibility and legitimacy which will need to be appropriately considered within the project.

Support was found for the concept of focusing and enriching the drug information. The findings are in line with the literature[2]

Updates to product information may not be reviewed by the end users given that many only read the leaflet once and never revisited.

There is a strong need for putting any information that is given in the proper context.

If the G-Lens response guides the user to other external sources of information user attractiveness is lost.

3.3 MDR/GDPR considerations

While this is the task for others in the consortium some considerations are needed here because they influence the information process model and the information model. The full implication of the MDR/GDPR regulations will be considered elsewhere in the project. The regulation landscape of information and of the handling of information is not uncomplicated. A preliminary analysis of the examples of questions listed in the Gravitate-Health application demonstrates the reach of the regulatory measures.

Stated questions as examples in the Gravitate-Health application:	MDR implications[3]	GDPR sensitive
What are the symptoms?	yes	yes
What is the health problem?		yes

[2] Kusch MK, Haefeli WE, Seidling HM. How to meet patients' individual needs for drug information - a scoping review. Patient Prefer Adherence. 2018 Nov 6;12:2339-2355. doi: 10.2147/PPA.S173651. PMID: 30464421; PMCID: PMC6229142.
[3] Medical device regulation REGULATION (EU) 2017/745), took effect May 2021

What should I do?	yes	yes
Who should I see?	yes	yes
Where can I find out more?		yes
G-Lens can assist self-care	yes	yes
Facilitating access to the health system through services like eBooking and ePrescription,		yes
What does that mean?		yes
What are the outcomes?	yes	yes
What are my choices?	yes	yes
What are the risks?	yes	yes
What are the alternatives	yes	yes
How will it affect me?	yes	yes
How will it affect my family?		yes
What can I do?	yes	yes
How am I progressing?	yes	yes
How can I get help?	yes	yes
Anticipating therapy reactions and impact on their daily activities		yes
What can I do?	yes	yes
How can I help myself?	yes	yes
When do I need help?	yes	yes
What help can I get?		yes
How can I get help?	yes	yes

Table 1. A preliminary analysis of the questions listed in the application and their relation to GDPR and MDR

The ambition for the project is of course to adhere to regulations but also to enable development of a tool in such a way that unnecessary (costly) certifications and complicated interactions and on-boarding procedures for the end-user could be avoided. A bad initial user experiences are always a threat leading to an increase in churn. The discussion on how to avoid a heavy regulatory impact on the project led to a few conclusions:

1. Use focusing and not filtering of the information as serving full PIL will not be construed as personal medical advice.
2. When possible, formulate results as general information, not personal medical advice
3. Maintain the streams of regulated information separate from unregulated in order to allow and focusing to be localized peripherally in the Gravitate-Health eco-system.
4. Post G-Lens added enrichment information should be done as part of separate software that takes responsibility on MDR issues

3.4 The "PERSONA" information space (T1.2)

The departure for T1.2 was from the Patient healthcare engagement model (PHE) that strongly motivates the G-Lens project to really focus on end user interests and capabilities with a dimensionality of think, feel, and act. The tool that was used was PERSONA characterization. A persona is a fictional, yet realistic, description of a typical or target user of the G-Lens.

Personas have been described and defined in a high dimension model with some 25 characterizing dimensions, some of them binary whereas others are continuous parameters. Aspects of general health literacy, preferred modes of interaction, ongoing medications, personality and other traits are listed. It has been identified in T1.3 that this high dimensional space is an opportunity to create a space for focusing both for information content but also for aspects in the I/O domain served to the end user. A fit in the persona space can gradually be built by the interaction with the end-user and include background information from other sources such as the international patient summary, prescription registers and dispensing registers.

From an information point of view a "Persona fingerprint" could be created and expressed as a vector to guide the focusing process in order to serve the end-user with relevant information. Such a fingerprint could be set as a string where each dimension has a state (estimated/assumed) where each parameter has a defined outcomes space.

A preliminary analysis of the persona dimensions yielded a set of dimensions and outcomes (Table 2). These may be directly acquired in a formal onboarding procedure or gradually built as a result of communication between the G-Lens and the end user.

Dimension	Type	Example	Example	Example	Example
Age	Continuous	Low=0	High=120	Unit=year. If<6 months	(or category)
Social support/ family	Categorical	None	Low	Medium	High
Sex	Categorical	Female	Male	Other	Non-disclosed
WorkLife	Categorical	unemployed/retired	employed non manual	employed manual	
Smoking, ongoing	Categorical	No	1-5 per day	5-20 per day	>20 per day
Physical activity	Categorical	None	Low	Medium	High
Organized	Categorical	Low	Medium	High	
Extrovert/ introvert	Categorical	Low	Medium	High	

Emotional/ Rational	Categorical	Low	Medium	High	
# Diagnoses	Categorical	None	Low	Medium	High
Chronic affliction	Binary	No	Yes		
# Medicines	Categorical	single	some (2-3)	many	
Health advice received	Categorical	single	some	many	
Concern domains	Categorical	single	some	many	
Share info willingly	Categorical	Low	Medium	High	
Mood level	Categorical	Low	Medium	High	
Autonomy	Categorical	Low	Medium	High	
Health Literacy	Categorical	Low	Medium	High	
Digital literacy	Categorical	Low	Medium	High	
Tool support interest	Categorical	Low	Medium	High	

Table 2. A preliminary list of 20 dimensions that may be used to build a persona archetype

It is feasible to consecutively label all enrichment material in the same type of dimensions as to better select information resources that are considered relevant for the end user. For example, if an enrichment instructional video on how to avoid a drug interaction is too difficult to follow for a subject with estimated low health literacy a better option would be to gradually introduce the subject and pair it with queries for absorbed knowledge. Thus, such a selection could be standardized by labelling all elements in the library for enriched material with the level of health literacy that is needed to meaningfully present it to the end user. Essentially all persona dimensions (Table 2) could be used to label the elements in the enrichment material.

3.5 Classifying sources of information

The sources of information may be characterized in many dimensions that summarize all of the describing traits that are important when creating an information model. Relevant medicinal product information has many sources with different jurisdictional status, accessibility, and regulatory status. A unique feature of medicinal product information is that the general right to freedom of speech is constrained for the manufacturer by regulatory measures along the way of information to the end user. Thus, a text may be regulated and approved for general use and certified to adequately describe drug properties. Such a text may also loose the regulated status by instigation of minor changes, as changes and revisions to product information must be managed by defined regulatory processes.

Another dimension of sources may be that they may carry different levels of trust in the eyes of different groups (end users, regulatory bodies, health care professionals etc.). Examples of sources with non-regulated status but rated as carrying high trust include WEBMD.com and health-oriented websites from caregivers with strong reputation such as NHS Health A-Z[4]. A third distinct dimension is whether the information is structured properly. A fourth pertains to an estimation of the granularity. A fifth dimension regards how accessible for the non-professional user. This is especially prominent in healthcare with its strong tradition of professionally specific language combined with the issue that the knowledge mass underlying the information often is very complex. The explanation of a complex knowledge mass in accessible terms is a challenge. A sixth dimension is the level of coverage of the field i.e. whether e.g. all drugs are covered, nationally available drugs, drugs within a certain diagnosis group or chemical class. Hence, a source may be reliable, trustworthy, highly structured, cover the whole knowledge field but fail completely on accessibility. The geographic reach is also a factor to weigh in.

Dimension	Comment and example of sources	Rating scale
1. Regulated status	To be classified as regulated an issuing public body must be identified. Example ePI, SmPC	Binary yes/no. Once switched to no, a reconfirmation is necessary before reaching regulatory status
2. Trustworthiness	A measure of the external general perception of the source regarding trustworthiness. No real measure exist and the trustworthiness is influenced by many factors such as organization trademark (e.g. ASK_MAYO_CLINIC.com), high presence in the public domain (Google searches) and the fact that the issuer is a public body (IPS, 1177.se)	Scale Very high, High, medium, low, very low, none
3. Structured consistently	Any source of information published on the net has a structure and if it is consistent across the source it is good otherwise it is a faulty structure.	Estimate of consistency (good /bad).
4. Granularity of information	This is independent from the structure as granularity identifies the hierarchical taxonomy and allow the correct identification of elements all the way to the highest granularity. ePI does not have a consistent structure and has a low level of granularity.	Granularity can only be determined following the study of the taxonomic structure of the source (High/low)
5. Accessibility for the non-professional	This is a simple estimate of the content accessibility based on the language, choices of concepts, level of abstraction. In the ePI the PIL is supposed to be accessible for the general public whereas the SmPC is more directed towards the health care professionals.	Estimate of accessibility (High/low)

[4] https://www.nhs.uk/conditions/

6. Coverage of the knowledge field	Highly specialized information sources may be high on all other dimensions but low	Described individually for each source
7. Geographic reach	A source may reach across EU or be limited regionally e.g., as a national source	EU wide-national- Targeted select group-other

Table 3. Classification of information sources – important dimensions to consider in G-Lens development

As the main focus of the Gravitate-Health project is to serve end users with reliable and accessible information on medication, we can also classify the sources as either being part of the information base that will be conveyed or sources that may be used to modify the presentation in what we have conceptually named focusing.

When assessing the different information sources the structure and granularity of the information in these sources is of central interest. Low granularity and poor structure (low consistency) creates difficulties in manipulating the information mass and adjust it to the needs and preferences of the end user. Therefore, one of the primary goals of the Gravitate-Health project is threatened if the projected information sources have poor quality. Different remedies must be sought if this is the case, including the use of tools like natural language processing and manual post-annotation. Structure-at-source is always superior to post hoc remedies against low structure as the cost rise very quickly as remedies may be labor intensive or the precision may decrease drastically.

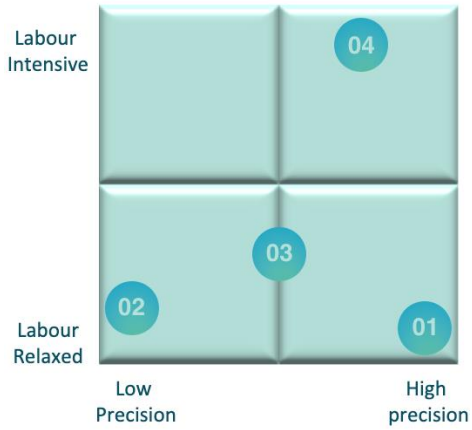

Figure 3. Structure-at-source (1) is the ideal. Natural Language Processing of source text (NLP) (2) yields low precision as the amount of available information is limited. Adding partial structure manually improves the NLP results a little (3). Extensive post-annotation (4) becomes prohibitively labor intensive and is most likely not an option for Gravitate-Health.

4 Sources of information on medicinal products

4.1 SmPC

SmPC stands for Summary of Product Characteristics. The SmPC is meant for healthcare professionals, such as doctors, nurses and pharmacists, and explains

how to use and prescribe a medicine. SmPCs are written and updated by pharmaceutical companies and are based on their research and product knowledge. The SmPC is then checked and approved by the European Medicines Agency (EMA) or alternatively a national competent authority (HMA), according to the way in which the medicine has been licensed. SmPCs have to contain certain numbered headings and information. Given the importance of the SmPC information for the Gravitate-Health project a brief summary of the headers is given below in table 4.

Section	Sub-section	Description
Section 1. Name of the medicinal product		What is the brand or trade name of the medicine?
Section 2. Qualitative and quantitative composition		What are the names of the active ingredients in the medicine, that make the medicine work? How much active ingredient does it contain? For example, paracetamol 500mg.
Section 3. Pharmaceutical form		What is the physical form of the medicine, for example a tablet, injection, ointment or syrup.
Section 4. Clinical particulars		This section explains how the medicine should be used or taken
	Section 4.1 Therapeutic indications	What diseases or medical conditions is the medicine approved to treat? Sometimes a healthcare professional might decide to use a medicine to treat conditions that are not listed on the SPC. If you are not sure why you have been given a medicine, please talk to your doctor or pharmacist.
	Section 4.2 Posology and method of administration	Posology means dose. What dose, or dose range, is used?
	Section 4.3 Contraindications	Contraindications are situations where a medicine should not be used. This section tells the prescriber when a medicine shouldn't be used or taken
	Section 4.4 Special warnings and precautions for use	Medicines always need to be taken or used carefully. This section tells the prescriber when to be extra careful when prescribing a medicine for some people.
	Section 4.5 Interactions with other medicinal products and other forms of interaction	Is your medicine known to react or interfere with any other medicines, herbal or dietary supplements?
	Section 4.6 Pregnancy and lactation	Information about taking or using a medicine if you are pregnant,

		thinking of becoming pregnant or are breast-feeding a child.
	Section 4.7 Effects on ability to drive and use machines	Will the medicine affect your ability to drive or use machines?
	Section 4.8 Undesirable effects	This section tells you about the side effects that people can get when they take or use the medicine. It tells you how often the side effect happens, how severe it might be, how long it might last for and what you should do.
	Section 4.9 Overdose	What could happen to you if you take or use too much of the medicine?
Section 5. Pharmacological properties		How does the medicine affect your body and what does your body do to the medicine?
	Section 5.1 Pharmacodynamic properties	How does the medicine have its effect on the body?
	Section 5.2 Pharmacokinetic properties	How the medicine gets into your body, gets to the part of the body where it needs to act, how the body changes the medicine and then removes it.
	Section 5.3 Preclinical safety data	Information about the tests that were carried out in a laboratory or on animals, before the medicine was used in humans. It includes the test results which are relevant to prescribers
Section 6. Pharmaceutical properties		This section gives information about the ingredients in a medicine, the packaging and how it should be stored.
	Section 6.1 List of excipients	What other 'ingredients' are in the medicine, apart from the active ingredient?
	Section 6.2 Incompatibilities	This section, along with section 4.5, tells you if there are any other medicines that should not be mixed or taken with this medicine.
	Section 6.3 Shelf life	What is the maximum amount of time the medicine can be stored for?
	Section 6.4 Special precautions for storage	How and where to store your medicine.

	Section 6.5 Nature and contents of container	Information about the medicine's packaging.
	Section 6.6 Special precautions for disposal and other handling	How to make-up or give the medicine and how to get rid of any left-over medicine.
Section 7. Marketing authorization holder		The Marketing Authorization Holder is the name of the pharmaceutical company who own the license to sell the medicine. Sometimes the Marketing Authorization Holder allows a different pharmaceutical company to sell their medicine.
Section 8. Marketing authorization number(s)		When the regulatory authority approves a medicine, they give it a number - the marketing authorization number.
Section 9. Date of first authorization/ renewal of the authorization		The date of first authorization is the date the regulatory authorities first approved the medicine. If the marketing authorization has been suspended and then granted again, there may also be a renewal of the authorization date.
Section 10. Date of revision of the text		If an SmPC changes, this is the date the pharmaceutical company sent the changes to the regulatory authority for their approval.
Section 11. Dosimetry		This section is only on SmPCs for radiopharmaceutical products. It tells you how much radiation you are exposed to.
Section 12. Instructions for preparation of radiopharmaceuticals		This section is only on SmPCs for radiopharmaceutical products. It tells you how to get rid of any unused or waste products safely.
Legal category		Every medicine has a legal category. POM means prescription only medicine - you can only get the medicine on prescription. P means pharmacy - you can only get the medicine from a retail pharmacy. GSL means general sales list - you can buy the medicine without a pharmacist, for example in a supermarket. CE Mark is used on devices.

Table 4. Headings and sub-headings in the SmPC

A detailed scrutiny of the regulating document for production of the SmPC, i.e., the QRD template (quality review of documents, shows that there is a substantial freedom regarding the content under the determined headers and also cross

references between them are allowed. The cross referencing to information is not regulated in a detailed fashion which lowers the structural consistency across all SmPCs. Our analysis showed that there are >200 classifiable recommended items of information in the document but only some recommendations have strict formulations (=structure) whereas others are loosely formulated. While the QRD regulates the content of both the SmPC and the PIL (see below) there are differences between the two in how headings are applied. Also, our analysis shows that more than half of required information is not attributed under separate headings but is required as free text within a heading. The mapping of this is given in addendum 3.

Comment from T1.3 on the SmPC: The texts in the SmPC is not primarily directed towards the general public. The structure is given in the QRD template which thereby guarantees a certain consistency across all SmPCs. The granularity is limited to headers/text and there is no controlled vocabulary. The computability for the G-Lens focusing process is by that limited for the SmPC.

Summary of SmPC characteristics:

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	++	0-1	+	+++	pan EU

4.2 PIL or PL

The package leaflet must be prepared in accordance with the SmPC and be written and designed in such a way as to be clear and understandable. Package leaflet (PL) and Patient information leaflet (PIL) are used as synonyms. There is a great overlap but not full congruence in the information content between the PIL[5] and the SmPC[6]. An example of a package leaflet is given here [7]. The European Union authorization of a medicinal product includes the text of the package leaflet, which is the same throughout the Union. However, for products registered through national routes, there may be some differences in the text of the package leaflet on a country by country basis. Accessibility of the general public is granted in that the package leaflet reflects the results of consultations with target patient groups to ensure that it is legible, clear and easy to use. Article 63(2), 1st sub-paragraph of the Directive provides that "*the package leaflet must be written and designed to be clear and understandable, enabling the users to act appropriately, when necessary, with the help of health professionals. [...]*". Multilingual PILs are allowed but the official languages of a country must be included. The directives also include "*the marketing authorization holder shall ensure that the package information leaflet is made available on request from patients' organizations in formats appropriate for the*

[5] https://ec.europa.eu/health/sites/default/files/files/eudralex/vol-2/2018_packaging_guidelines_en.pdf CL2001L0083EN0110010.0001.3bi_cp 1..1 (europa.eu)
[6] An example: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-core-smpc-package-leaflet-68ge/68ga-generator-first-version_en.pdf
[7] https://www.medicines.org.uk/emc/files/pil.9542.pdf

blind and partially-sighted". All changes in a PIL not coming from a change in the SmPC must be re-confirmed with the regulatory body.

4.2.1 The headers of the PIL[8]

Below the structure of the PIL is given.

Section	Sub-section
Start information	Package leaflet: Information for the <patient> <user> {(Invented) name strength pharmaceutical form} {active substance(s)} <Read all of this leaflet carefully before you start <taking> <using> this medicine because it contains important information for you. - Keep this leaflet. You may need to read it again. - If you have any further questions, ask your <doctor> <,> <or> <pharmacist> <or nurse>. <- This medicine has been prescribed for you only. Do not pass it on to others. It may harm them, even if their signs of illness are the same as yours.> - If you get any side effects, talk to your <doctor> <,> <or> <pharmacist> <or nurse>. This includes any possible side effects not listed in this leaflet. See section 4.> Read all of this leaflet carefully before you start <taking> <using> this medicine because it contains important information for you. Always <take> <use> this medicine exactly as described in this leaflet or as your <doctor> <,> <or> <pharmacist> <or nurse> <has> <have> told you.
1. What X is and what it is used for	[(Invented) name, active substance(s) and pharmacotherapeutic group] [Therapeutic indications] [Information on the benefits of using this medicine]
2. What you need to know before you <take> <use> X	[Contraindications][Allergies][Warnings and Precautions][Children and Adolescents][Interactions][Tell you doctor of other medicines][Interactions with food and drink] [Use by pregnant or breast-feeding women, information on fertility] [Effects on the ability to drive or to use machines] [Excipients warnings]
3. How to <take> <use> X	[Dose (SmPC section 4.2)] [Route(s) and/or method of administration (SmPC section 4.2)] [Duration of treatment (SmPC section 4.2)] Including texts on deviances
4. Possible side effects	[Description of side effects in order of seriousness and separately other side effects] [**Additional side effects in children <and adolescents >]**
5. How to store X	[Expiry date] [Storage conditions] [Where applicable, shelf life after reconstitution, dilution or after first opening the container] [Where appropriate, warnings against certain visible signs of deterioration]

[8] Annex 3 https://www.ema.europa.eu/en/documents/template-form/qrd-product-information-annotated-template-english-version-102-rev1_en.pdf

6. Contents of the pack and other information	[Full statement of the active substance(s) and excipient(s)] [Pharmaceutical form, nature and contents of container in weight, volume or units of dose] [What X looks like and contents of the pack][Name and address of the MAH and of the manufacturer responsible for batch release, if different] [Marketing Authorization Holder and Manufacturer]

Table 5. The headers of the PL

4.2.2 PIL information structure

The PIL structure is directed from the QRD definition. An example of the internal structure in the PIL demonstrates the lack of semantic and vocabulary structure:

[Route(s) of administration according to "Standard Terms" published by the Council of Europe and an additional patient-friendly explanation may be given if necessary.
Method of administration: directions for a proper use of the medicine, e.g. "Do not swallow", 'Do not chew", "Shake well before use" (user testing experience has shown it is useful to state the reasons for the inclusion of such a statement, e.g. "Do not break or crush the tablet(s). If you do, there is a danger you could overdose because this medicine will be absorbed into your body too quickly").

When applicable, there should be descriptions (if useful with illustrations) of opening techniques for child-resistant containers and other containers to be opened in an unusual way.
Where relevant, guidance should always be included to clarify if the medicine must be taken with food, during/before meals, or clearly state if food/meals have no influence, etc.]

<The score line is only there to help you break the tablet if you have difficulty swallowing it whole.> <The tablet can be divided into equal doses.>
<The score line is not intended for breaking the tablet.>

[If appropriate, especially for medicines available without prescription, precise statements should be included on:

· the usual duration of the therapy;
· the maximum duration of the therapy;
· the intervals with no treatment;
· the cases in which the duration of treatment should be limited.]

[For some medicines it may be necessary to include some additional information in this section although this need not be covered in all cases. The following headings can be used as a guide:] **<If you <take> <use> more X than you should>**

[Describe how to recognise symptoms if someone has taken an overdose and what to do as per SmPC section 4.9.]

<If you forget to <take> <use> X>

[Make clear to patients what they should do after irregular use of a medicine, e.g.: if information is available, try to include information on the maximum interval the missed dose can be caught up as per SmPC section 4.2.]

<Do not take a double dose to make up for a forgotten <tablet> <dose> <...>.>

<If you stop <taking> <using> X>

[Indicate withdrawal effects and how to minimise them as per SmPC section(s) 4.2 and/or 4.4**.**
A statement on the potential consequences of stopping the treatment before finishing the course of treatment and the need for a prior discussion with the treating physician, pharmacist or nurse should be included as appropriate.]
[Close this section with:]

<If you have any further questions on the use of this medicine, ask your <doctor> <,> <or> <pharmacist> <or nurse>.>

T1.3 comment on the PIL: The PIL is explicitly prepared to reach the public and to be accessible. It has overlapping information content with several other documents (most notably the SmPC) but not a corresponding structure in the headers. There is no defined vocabulary. It has a certain match to the SmPC but is not congruent in headers nor in the cross-reference structure. The computability for highlighting is limited and co-computability between SmPC and PIL is a challenge given the partial overlap but lack of common headers.

Summary of PL characteristics:

Regulated status	Trustwothyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	++	0-1	+++	+++	pan EU

4.3 A comparison of content: PIL vs SmPC

As the PIL and SmPC both are defined in the QRD template with the aim to provide corresponding information we did a test of the congruence between the SmPC and the PIL. This resulted in the following table:

SmPC - section	PIL - section
1. NAME OF THE MEDICINAL PRODUCT	☐ 'name' in section '**1.** What X is and what it is used for'
2. QUALITATIVE AND QUANTITATIVE COMPOSITION	➢ 'active substance' if necessary in section '**1.** What X is and what it is used for - X contains the active substance Y' ➢ 'excipients' in section '**6.** What X contains'
<2.1 General description>	☐ 'for advanced therapy medicines description of cells or tissues and of their specific origin, including the species of animal in cases of non-human origin' in section '**1.** What X is and what it is used for'
<2.2 Qualitative and quantitative composition>	☐ 'for advanced therapy medicine a description of the contained medical devices or active implantable medicinal device and their specific origin' in section '**1.** What X is and what it is used for'
3. PHARMACEUTICAL FORM	☐ 'physical description' in section '**6.** What X looks like and content of the pack'
4.1 Therapeutic indications	☐ 'indication' in section '**1.** What X is and what it is used for'
4.2 Posology and method of administration	➢ 'dose' in section '**3.** How to <take> <use> X' ➢ 'route and/or method of administration' in section '**3.** How to <take> <use> X' ➢ 'duration of treatment' in section '**3.** How to <take> <use> X'

	➢ 'missed dose' and 'irregular use' in section '**3.** If you forget to <take> <use> X' □ 'withdrawal effects' in section '**3.** If you stop <taking> <using> X'
4.3 Contraindications	□ all 'contraindications' in the same order in section '**2.** Do not <take> <use> X'
4.4 Special warnings and precautions for use	➢ all 'warnings and precautions for use' in section '**2.** Warnings and precautions' ➢ 'warnings and precautions' related to side effects that occur while taking the medicine (eg. symptoms) in section '**4.** possible side effects' with appropriate cross-reference in section **2**. ➢ 'warnings on excipients' in section '**2.** X contains {name the excipient(s)}' □ 'withdrawal effects' in section '**3.** If you stop <taking> <using> X'
4.5 Interaction with other medicinal products and other forms of interaction	➢ 'interactions' in section '2. Other medicines and X' ➢ if interaction leads to 'dose adjustments' add cross-reference to section '**3.** How to <take> <use> X' ➢ 'interactions' not related to medicines in section '**2.** X with <food> <and> <,> <drink> <and> <alcohol>'
4.6 Fertility, pregnancy and lactation	□ 'conclusion summary' in section '**2.** Pregnancy <and> <,> breast-feeding <and fertility>'
4.7 Effects on ability to drive and use machines	□ if there is a 'cautionary advice' include it in lay language in section '**2.** Driving and using machines'
4.8 Undesirable effects	□ 'description of side effects' in section '**4.** possible side effects'
4.9 Overdose	□ 'symptoms of overdose and what to do' in section '**3.** If you <take> <use> more X than you should'
5.1 Pharmacodynamic properties	□ 'pharmacotherapeutic group' and/or type of activity, -> in section '**1.** What X is and what it is used for' e.g. "statins (SmPC) -> used to lower cholesterol (PIL)" other parts not in PIL
5.2 Pharmacokinetic properties	not in PIL
5.3 Preclinical safety data	not in PIL
6.1 List of excipients	□ 'excipients' in section '**6.** What X contains'
6.2 Incompatibilities	not in PIL
6.3 Shelf life	□ 'shelf life after reconstitution, dilution or after first opening' in section '**5.** how to store X'
6.4 Special precautions for storage	□ 'storage conditions' in section '**5**. how to store X'
6.5 Nature and contents of container <and special equipment for use, administration or implantation>	□ 'all pack sizes' incl. ancillary items included in the pack e.g. needles in section '**6.** What X looks like and content of the pack'
6.6 Special precautions for disposal <and other handling>	not in PIL

7. MARKETING AUTHORISATION HOLDER	☐ 'MAH name and address' in section '**6.** Marketing authorization holder and manufacturer'
8. MARKETING AUTHORISATION NUMBER(S)	not in PIL
9. DATE OF FIRST AUTHORISATION/RENEWAL OF THE AUTHORISATION	not in PIL
10. DATE OF REVISION OF THE TEXT	☐ 'date(s)' in section '**6.** This leaflet was last revised in …'
<11. DOSIMETRY>	not in PIL
<12. INSTRUCTIONS FOR PREPARATION OF RADIOPHARMACEUTICALS>	not in PIL

Table 6. Comparing SmPC and PIL structures – A challenge for Gravitate-Health

Comments from T1.3: We note that the congruency is not formally upheld between the two in that the header separation and numbering is non-congruent and that cross-referrals between sections are different in the two documents. Several sections included in the SmPC do not exist in the PIL. In most cases it relates to segments of little information value to the public. Since these documents will be the main basis for the ePI we note that the inherent problem that the same document will include overlapping information with internally different structure. This will limit the ability to focus the information for the end-user beyond the formal header level of the SmPC. Since the headers are far fewer in the PL the ability to focus that information will be limited by the ability to cross map between the SmPC and the PL. A first analysis suggests that this will not be possible to automate.

4.4 National availability of electronic PL and SmPC

We have mapped the national availability of WEB resources of PL and SmPC prior to the introduction of the ePI. We note that the demand for WEB based information has resulted in many variants of such information. We also noted that the propensity for distribution of SmPC for public dissemination varied across the countries.

Country	HA	Association	3rd Party	SmPC to public
Croatia	1	1	1	1
Greece	1	1	0	0
Estonia	1	1	0	1
Slovenia	1	1	0	1
Austria	1	1	0	1
Latvia	1	1	1	0
Belgium	1	1	1	1
Switzerland	1	1	1	0
Romania	1	1	0	0
Bosnia	1	1	0	1
Macedonia	1	1	0	1
Albania	1	1	1	0
Serbia	1	1	1	0
Montenegro	1	1	0	1
Kosovo	1	1	1	0
Finland	1	1	1	1

Lithuania	1	1	1	0
Sweden	1	1	1	1
Poland	1	1	1	1
Norway	1	1	1	1
Denmark	1	1	1	1
Spain	1	1	0	1
France	1	1	1	1
United Kingdom	1	1	1	0
Ireland	1	1	0	1
Bulgaria	1	1	0	0
Portugal	1	1	0	0
Germany	1	1	1	1
Italy	1	1	1	1
Netherlands	1	1	1	1

Table 7. National available HTML or pdf versions of PIL and SmPC

This preliminary list, made available by Health Authority (HA), by a Pharma Association or by a Third Party. The Third Party refers to secondary non-regulated web sources that include the PIL information. Also, the cultural difference with more paternalistic view not encouraging distribution of the SmPC to the public remain in some countries. Third party web republication of regulated information exist in about half of the EU countries. We conclude from table 6 that Gravitate-Health face a landscape where the end users have access to the content of the PILs on the web and most often in a “click by heading” presentation.

4.5 ePI

The following text on the ePI is derived closely from the EMA 'Key Principles for electronic product information (ePI)' published in 2020[9]. In those documents for the ePI it is stated (Quoted below)

> The ePI contains authorised, statutory product information for medicines (i.e. SmPC, PIL and labelling[10],[11]) in a semi-structured format created using the common EU electronic standard. ePI is adapted for electronic handling and allows dissemination via the web, e-platforms and print.
>
> ePI in the EU will cover all human medicines, including both centrally and nationally authorized medicines and will be created using a common electronic standard. The following definition of a common EU electronic standard for ePI is proposed:
>
> A common standard for ePI in the EU refers to the technical features of ePI (**including mark-up language, controlled vocabularies and interoperability specifications**) agreed by EMA, HMA, NCAs, EC, and representatives of the pharmaceutical industry, patients and HCPs. The standard will be used to generate ePI that fulfils the agreed key principles.

[9] https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/electronic-product-information-human-medicines-european-union-key-principles_en.pdf
[10] In certain procedures, Annex II of the marketing authorisation (manufacturer(s) responsible for batch release, conditions and requirements of the marketing authorisation, other conditions or restrictions as applicable) is provided electronically together with ePI.
[11] ePI does not include additional information specific to a Member State such as ‘blue box’ information (see: https://ec.europa.eu/ health/sites/health/files/files/eudralex/vol-2/2018_packaging_guidelines_en.pdf) or artwork of the marketed medicine package.

ePI is a public-health priority because it will expand the dissemination of unbiased, up-to-date, regulator-approved PI for all medicines in the EU. ePI will support, among other functions:

- provision of the latest information on a medicine's safety, benefits and conditions of use;
- better delivery of information so that the right information is available to the right HCP and patient / consumer at the point of need;
- informed decision-making by patients / consumers and HCPs.

ePI will facilitate creation of PI that is accessible to everyone, including users with print impairments, including physical impairments or learning difficulties, or for whom printed PI is difficult to access for other reasons. ePI allows the use of large fonts or high screen contrast for partially sighted users and audible formats for blind users and those with low literacy levels. ePI on the web will be accessible to screen readers, web and mobile applications, convertible to large font and amenable to other accessible formats. Accessible formats will provide the full and balanced product information to users in formats suitable for their needs.

ePI will enable increased efficiency in management of PI during regulatory procedures. By enabling PI changes to be made across all relevant PI annexes and products, ePI could eliminate many manually performed tasks and redundancies that are potential sources of error.

ePI will provide information on medicines that is amenable to analysis, and could be used to increase knowledge by facilitating study of characteristics of current EU medicines.

ePI will not supersede or negate the requirement of the pharmaceutical legislation (Article 58 of Directive 2001/83/EC1) to include a PL in the packaging of all medicines or directly convey all information required (by Articles 59 and 62 of the Directive) on the outer or immediate packaging.

Since the current legislation does not require the use of an electronic version of PI, the use of ePI will not constitute a new legal obligation.

ePI is intended for the delivery of the full and complete regulator-approved medicine PI only. ePI will not be used for delivery of promotional information.

ePI should always be published as freely accessible open data

ePI itself will not include any personal data.

In any event where processing (e.g. collecting or handling) of personal data may occur in relation to the implementation and use of ePI, for example in the context of a mobile application developed for the use of patients to access ePI, personal data processing must be in accordance with applicable European data protection legislation. This includes, in particular Regulation (EU) 2016/679 (GDPR) and Regulation (EU) 2018/1725 applicable to EU institutions.

It is envisaged that, eventually, ePI format will be used for the PI of all human medicines authorized in the EU through EMA and NCAs from the point of submission and throughout the evaluation process.

However, in the short and medium term, some regulatory authorities may decide to continue to perform assessment as is done currently, and that ePI should be created once the regulatory procedure is complete.

The ePI implementation process will depend on the findings of feasibility analyses and will be described in a future roadmap to guide implementation. ePI will be made available to users, e.g. patients / consumers and HCPs, through websites at EMA level and if available, Member State level. ePI data will be made available for use in other e-health systems, such as electronic health records and e-prescribing systems. ePI will also be available for use by third-parties, who can reproduce ePI and make it available to patients and HCPs (as is already the case for PI today). The underlying documents are listed here[12]

[12] The consultation is published here: https://github.com/EuropeanMedicinesAgency/ePI-consultation.
The API for the ePI is here: https://github.com/EuropeanMedicinesAgency/ePI-consultation/blob/master/API%20specification/Draft-ePI-API-Specification-v1.pdf.
The template for XML is given here: https://github.com/EuropeanMedicinesAgency/ePI-consultation/blob/master/XML%20templates/ePI_template_instance.xml. An example is given here:

All stakeholders, including pharmaceutical companies and regulators, are expected to commit to implementation of the common electronic standard for creation of ePI for all EU medicines. However, timelines and processes for implementation will be flexible and amenable to the available resources and priorities at national level. A roadmap will be proposed by HMA and EMA to define the steps for development, which allows implementation in the EU on the basis of the key principles.

ePI shall support all official EU languages and Icelandic and Norwegian so that EU citizens will be able to read ePI in their preferred language when authorized ePI in that language is available.

ePI will interface and interact with many ongoing and foreseen eHealth initiatives. eHealth and related services should work together, within and across organizations or domains. ePI interoperability with cross-border prescription, electronic health records, the future European medicines web portal, pharmacovigilance systems, SPOR data management services, future ePI for veterinary medicines, a future European common data model, current electronic application procedures and national ePI systems must be considered in the design of EU ePI. Use of ePI in both an EU and global context should also be taken into account.

The EMA ePI set-up project was launched in 2021. The first phase of the set-up project is the creation of a FHIR base information standard for the EU (see figure below). The FHIR based ePI proof-of-concept standard is the output of a 1 year project with limited scope and specific objectives.

1. An ePI roadmap for Implementation is also one of the project objectives, due to complete end 2021. Details of this are not therefore available at the time of this report, but we do not know if the ePI roadmap will encompass any major revisions of the information structure of the documents that constitute the proof-of-concept ePI. However, the analysis of the structure of the SmPC and PL will probably not need any revision for the Gravitate-Health project.

2. In the scope of this work, ePI remains complementary to the paper and flexibility in its implementation at member state level is foreseen. Whilst the long term vision is for use of ePI through the regulatory process, this will not be so in the earlier phase of implementation in our porject.

3. The ePI contains, from an information point of view, partially overlapping information as different documents with disparate structure build the ePI. It will therefore be very difficult to use for downstream applications that require computability of the information.

4. There is limited integration with SPOR and SPOR remain in HL7 3 messaging standard, a standard with low uptake as HL7 FHIR has superseded it. Thereby, the structural elements that are necessary to reach the goals of the Gravitate-Health project can be reachable with a combined use of SPOR and ePI documents, although this may represent a possible complication for the Gravitate-Health project.

https://github.com/EuropeanMedicinesAgency/ePI-consultation/blob/master/XML%20templates/ePI_template_instance.xml.
The FHIR template is given here: https://github.com/EuropeanMedicinesAgency/ePI-consultation/blob/master/XML%20templates/ePI_template.xml

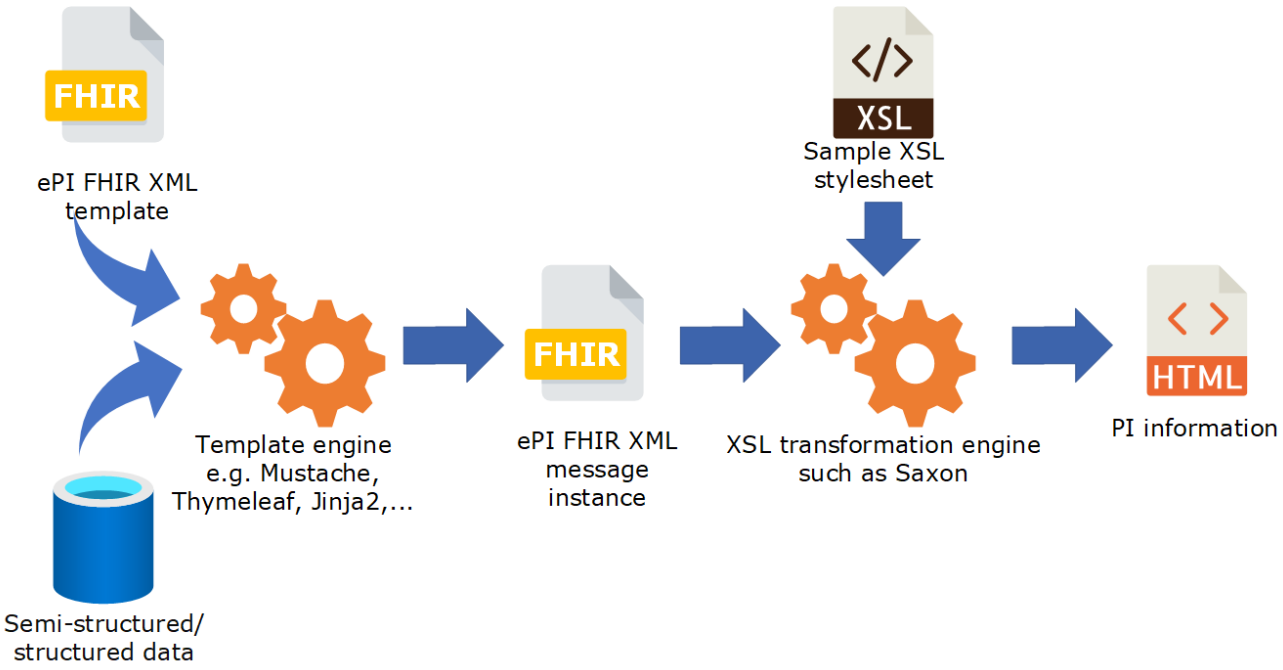

Figure 4. The conceptual drawing of the FHIR standardization for the ePI (Source EMA).

It is noted that no information handling is done beyond the formal header structure at this stage.

T1.3 comment on the ePI: The kick off of the EMA ePI set-up-project this year is a welcome development, although the current scope is quite limited. The ePI proof-of-concept standards contains the semi structured information from the SmPC, the PIL and in addition the information on labelling. Hence, it provides the same granulation and lack of congruency as the respective underlying documents. The current scope of activities does not include revisiting content requirements or templates. The long-term ambition of the ePI project will not be reached by the first FHIR standard project alone; while the immediate focus is to define a common standard in EU for ePI, it is anticipated that further work on implementation will progress, including further structuring. As the innate structure is lacking it will be difficult to refine the content in the G-Lens focusing procedure. This is an unresolved issue at the time of the writing of this deliverable, and should be considered in further development of the ePI to maximize value that can be derived over time to deliver on the vision for Gravitate-Health.

Summary of the ePI:

Regulated status	Trustwothyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	x	1	+++	+(++)	pan EU

4.6 SPOR - The EMA implementation of IDMP standard

ISO IDMP came from a need to standardize the definition of medicinal product information to facilitate the identification and exchange of such information in the context of pharmacovigilance activities (e.g. identifying medicines causing Adverse Events (AEs)). In the EMA version it runs under the name of SPOR[13].

The International Organisation for Standardisation (ISO) Identification of Medicinal Products (IDMP) standards specify the use of standardized definitions for identification and description of medicinal products for human use. The purpose of these standards is to facilitate reliable exchange of medicinal product information in a robust and consistent manner, by providing a **common product 'language' for stakeholders** to use in their interactions. The use of these standards is a regulatory requirement as mandated by the EU legislation (Commission Implementing Regulation (EU) No 520/2012 [articles 25 and 26]).

Five separate standards establish definitions and concepts, and describe data elements and their structural relationships. They cover the following aspects to describe a medicinal product:

- Medicinal product name;
- Ingredient substances;
- Pharmaceutical product (route of administration, strength);
- Marketing Authorization;
- Clinical particulars;
- Packaging;
- Manufacturing.

ISO IDMP covers the entire product lifecycle: products in development, investigational products, products under evaluation and authorized products.

ISO IDMP has multiple use cases within the regulatory context. For example:

- **Pharmacovigilance:** Adverse event reports are based on a harmonized set of product definitions, improving the quality of data used for signal management, and speeding up communication, decision-making and actions;

- **Regulatory submissions**: Submissions use a consistent standard to capture and manage data, allowing information on medicinal products to be shared and re-used across different procedures and among various regulators (subject to confidentiality restrictions);

[13] https://www.ema.europa.eu/en/documents/other/introduction-iso-identification-medicinal-products-spor-programme_en.pdf

- **Clinical trials**: Stakeholders can access Clinical Trial data using agreed and well-supported standards, improving the assessment and scientific evaluation of medicines as well as communication and transparency;

- **Good Manufacturing Practices** (GMP) inspections: Inspections on manufacturing sites are based on accessible information, which streamlines inspections particularly for urgent situations involving defects. Faster detection of falsified medicines can also be supported as a result of consistent data standards.

The IDMP holds a lot of specific drug information that is potentially usable for the G-Lens project summarized in the following figure:[14]

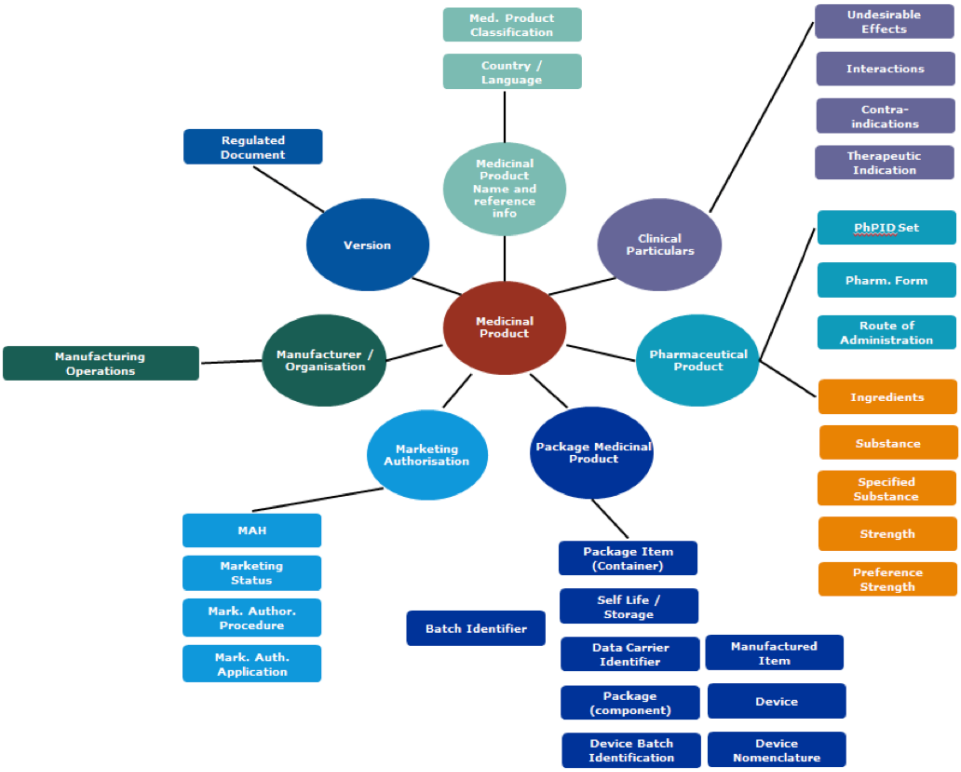

Figure 5. The content of the SPOR. Most information is not directed towards the end user.

Implementation of the ISO IDMP standards is governed by the following specifications:

- ISO IDMP Implementation Guides (Technical Specifications): Define the technical details on how to implement the standards, such as specific fields, their formats, and business rules describing their use;

[14] https://www.ema.europa.eu/en/documents/other/introduction-iso-identification-medicinal-products-spor-programme_en.pdf

- EU Implementation Guide: Provides guidance on the interpretation of data fields specifically for the EU regulatory environment as well as guidance on the processes for submitting and updating data.

- HL7 Messaging Specifications: Define the messages that will be used to exchange IDMP information, which are based on HL7 Standards;

The ISO IDMP standards will be implemented in phases, through a set of projects known as SPOR data management services (Substances, Products, Organisations, Referentials). They will establish ISO IDMP compliant business services for the central management of data in each of the four SPOR areas: The SPOR data management services are:

- Substance Management Services (SMS);
- Product Management Services (PMS);
- Organisations Management Services (OMS);
- Referentials Management Services (RMS).

The phased implementation of the ISO IDMP standards has been endorsed by the European Union Network Data Board (EUNDB) and the EU ISO IDMP Task Force (aka SPOR Task Force). The first two projects that EMA will deliver are RMS and OMS. They will lay the foundations for the subsequent delivery of SMS and PMS.

Comment from T1.3: The SPOR/IDMP is under implementation. It is issued in a different standard for communication than the ePI, however interaction with ePI is foreseen in the ePI key principles. The aim of SPOR is to establish safe communications on drugs across the full product cycle. It has no ambition to be part of consumer information and is therefore not planned as to meet the surface demands of the consumer (easy access, readability etc.).

Summary of SPOR characteristics

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	+++	2	0	+	pan EU

4.7 EPAR

A European public assessment report (EPAR) is published for every human or veterinary medicine application that has been granted or refused a marketing authorization. This follows an assessment by EMA of an application submitted by a pharmaceutical company in the framework of the Central authorization of medicines.[15]

[15] European public assessment reports: background and context | European Medicines Agency (europa.eu)

An EPAR provides public information on a medicine, including how it was assessed by EMA. The EPAR is referred to in Article 13(3) of Regulation (EC) No 726/2004, which requires EMA to publish a public assessment report for each centrally authorized medicine together with a public-friendly overview.
EMA has developed the EPAR concept over time to ensure that it delivers a usable, transparent and appropriately detailed body of information. The EPAR content and structure have therefore evolved over time and may be further developed in future.
An important role of the EPAR is to reflect the scientific conclusions of the relevant EMA committee at the end of the assessment process, providing the grounds for the committee opinion on whether or not to approve an application.
All EPARs are published on the EMA website and can be viewed under human medicines and veterinary medicines.
An EPAR is not a single document but an information resource containing several components, including a core set of regulatory documents. EPARs are displayed on the EMA website and the individual components can be viewed online, downloaded and/or printed out.
Information handled during the scientific assessment which is considered confidential is removed before an EPAR is published.
EPARs are updated periodically to reflect the latest regulatory information on medicines. If the original terms and conditions of a marketing authorization are varied, the EPAR is updated to reflect such changes with an appropriate level of detail.
EPARs are displayed on the EMA website using four different sections containing different components of the EPAR. The below table provides an overview.

Section	Type of information
Overview	Public-friendly overview in question-and-answer format.
Authorization details	Key details about the product and the marketing authorization holder.
Product information	Package leaflet and summary of product characteristics; labelling; list of all authorized presentations; pharmacotherapeutic group; therapeutic indications.
Assessment history	Public assessment report for the initial authorization; public assessment report(s) for any variation concerning major changes to the marketing authorization; orphan maintenance assessment report or withdrawal assessment report (as of 17 January 2018); tabulated overview of procedural steps taken before and after authorization.

Table 8. An overview of the EPAR content

Some components of the EPAR are always published in all official EU languages:

- Public-friendly overview;
- Labelling;
- Package leaflet and summary of product characteristics;
- List of all authorized presentations.

The other elements of the EPAR are published in English only:

- Public assessment report(s);
- Tabulated overview of procedural steps taken before and after authorization;
- Other content available only as a web page (e.g. information under the 'Authorization details' tab).

Comment from T1.3: It is an elaborate procedure to publish and maintain an EPAR on a drug[16]. It should be noted that both the Product leaflet (PIL) and the summary of product characteristics (SmPC) are components in the EPAR. The EPAR is published on the web but has no formal structure above and beyond the needs of HTML. At this point it is deemed not to add information of value to the end users beyond the SmPC/PIL.

Summary of EPAR characteristics:

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	++	0	+	+++	pan EU

4.8 The considerable overlap of content (but not structure) between the sources of information

A number of sources of regulated information have been mentioned above. It should be noted that the information landscape in general holds a low level of structure in combination with overlap of information, however in most cases without a corresponding structure. In the figure below, some of the information mass overlaps are depicted. As can be seen, it is a de-facto landscape that has emerged, where different sources of information have different purposes and target groups. They have all (except the ePI being developed) been created as final human readable information with no ambition for further refinement or sophisticated user adjusted presentations. Hence, the regulatory documents have real language characteristics presuming a considerable ability to parse the text on the receiving end (= human reading). Thereby, such documents share characteristics of common language in that they use a flexible and context variable structure for categories and hierarchies. Such information landscapes create substantial difficulties before automatic computations may be made on the content (such as the highlighting function in the G-Lens).

[16] https://www.ema.europa.eu/en/documents/sop/standard-operating-procedure-preparation-european-public-assessment-report-human-medicinal-product_en.pdf

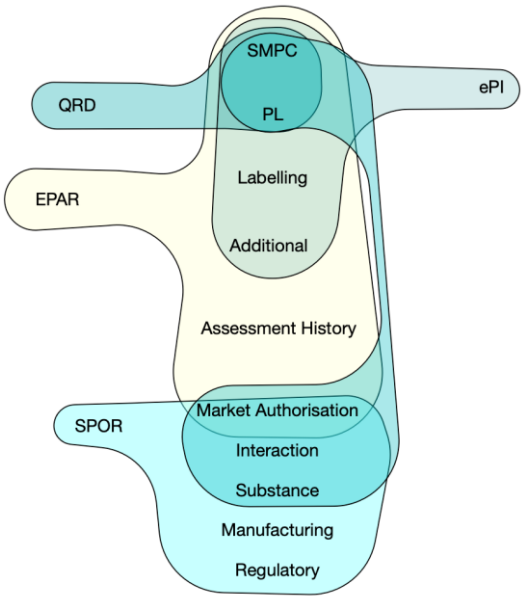

Figure 6. A schematic depiction of the overlaps of content between the different regulated information sources that have been identified.

The QRD, mentioned here in figure 6, is only to point out that it is a common template document for both PIL and SmPC. The fact that there is an overlap in the information content does by no means indicate a common information structure. Hence, the QRD regulates the PIL and the SmPC, but in spite of overlapping information there is not a common headings structure.

4.9 Sources of medicinal product information with more limited geographical reach:

Our ambition in this section is to give some examples rather than an exhaustive list. There are a few considerations needed when including local sources:

Consideration	Comment
A local source can be reached outside the local jurisdiction	Using local sources requires the G-Lens to convey full transparency on where data has been obtained
A local source may demand licensing for use beyond simple access to single articles	This is a consideration that should be done prior to inclusion of such sources
A local source may have a limited coverage of the available options or be updated on an irregular basis	Deriving drug information from an incomplete and dated database is considered to be a risk
A local source may have a technical I/O that requires extensive coding	This should be clarified before investment in bringing on a new source is started
A local source may have a limited number of languages deployed	If the aim is that G-Lens must reach all citizens in an area where it is deployed this would pose a severe limitation of capability

Table 9. Considerations on local sources

4.9.1 Datapharm

Datapharm[17] has been in business almost half a century to assist in the provision of medical information by technology-enabled solutions to the global life sciences and wider healthcare sectors with main geographic target being the UK. Its leading platforms improve the accessibility and usability of medical information. Datapharm provides the EMC (electronic medicines compendium) suite of product directed to serve industry, public and health care professionals. The EMC contains up to date, accessible information about medicines licensed for use in the UK. The EMC website has 10,000 medicines published by over 300 pharmaceutical companies, all of which have been checked and approved by either the UK or European government agencies that license medicines. The post preparation of information checking by authorities creates **a special status for the EMC as regulated information** in spite of a very elaborate refinement of the available industry and authorities derived information. It should be noted that the effort to maintain the regulated status is highly laborious but also guarantees that the information does not hold any promotional content.

Comment from T1.3: The admirable groundwork that Datapharm has undertaken during its long existence has demonstrated the need for consistency and structured work when dealing with primarily semi-structured information of limited granularity. Understanding data structure has made Datapharm also a supplier to Pharma where support is offered for the regulatory process and at the same time the EMC web page provides information to the public and health care professionals.

Summary Datapharm characteristics

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Partially	+++	+++	2	+	+++	pan EU

4.9.2 Nationally authorized sources (exemplified by FASS (Sweden))

This source of information publishes the SmPC and the PL of all marketing permit holders in Sweden. It used to be published in print once a year but as of a decade ago it is available on the WEB (www.fass.se) where information to the public and the profession is published back to back. It uses WHO indication-oriented ATC codes[18] to formalize the classification of the drugs. We have identified similar arrangements in many EU countries (See Table below). We note that some are hosted by associations of drug companies and some by authorities. Also, there is a variability to what degree they contain information with regulated status and also regarding the reach on the market. While some contain all drugs others may exclude generica or have a voluntary participation.

[17] Datapharm is also a partner in the Gravitate-Health project
[18] ATC code, is a classification system for medicines. The drugs are divided into different groups according to the indication area. The ATC code is used by the WHO for reporting side effects and is recommended by the WHO.

In general, these web-sites have a stylized header structure close to the one in SmPC. A common feature is that the information is presented as human readable and clickable headers that present the underlying information upon click. This very simple feature yields a sense of interaction and reduce the cognitive load for the end user.
As an example: The Swedish contains separately the PIL, SmPC, the formatted FASS text, Important patient information, labelling and packaging variants, availability in pharmacies, pictures on how to split tablets and other graphics, IPS for use, environmental info. All information may upon a click be read aloud and presented in large font/high contrast versions to accommodate accessibility requests.

Country	Name of catalogue	Link
Belgium	Compendium	www.pharma.be
Denmark	Indlægssedler	www.indlaegssedler.dk/
Finland	Pharmaca Fennica	www.pif.fi
France	Dictionaire Vidal	www.vidal.fr
Ireland	Medicines	www.medicines.ie
Italy	L'Informatore Farmaceutico	www.informatorefarmaceutico.it
Netherlands	Pharmacoterpautisch Kompas	Home | Geneesmiddeleninformatiebank | College ter Beoordeling van Geneesmiddelen
Norway	Felleskatalogen	www.felleskatalogen.no
Switzerland	Arzneimittelkompendium der Schweiz	www.documed.ch
Spain	Agencia Espanola de Medicamentos y Productos Sanitarios	www.aemps.gob.es
Great Britain	Medicines.org.uk including the electronic Medicines Compendium	www.medicines.org.uk
Czech Republic	Stétni ustav pro kontrolu lééiv. SUKL	www.sukl.cz
Germany	Rote liste Gelbe liste Pharmindex	PharmNet.Bund - Arzneimittel-Informationssystem (pharmnet-bund.de)
Hungary	Gyogyszer Kompendium	Országos Gyógyszerészeti és Élelmezés-egészségügyi Intézet (gov.hu)
USA	The United States Pharmacopeia Drug	www.usp.org
Sweden	Fass	www.fass.se

Table 10. National sources for the PL and SmPC

4.9.3 Janusinfo (Sweden) as an example of information on interactions

Janusinfo[19] is a non-commercial website providing drug information to support healthcare professionals in their everyday work. The website is the electronic means of communication of the Drug Therapeutic Committee and the Health and Medical Care Administration of the Stockholm County Council, Sweden. Contents and functions of the website should contribute to evidence-based and

[19] https://janusinfo.se

cost-effective drug treatment. Their aim is to become the main website of drug information for healthcare professionals in the Stockholm area.

The information on the website has been developed in collaboration with clinical pharmacologists and other experts in specific therapeutic areas.

Contents in brief
- News
- Therapeutic guidelines
- Kloka listan (Wise List; recommended drugs)
- Evaluation of new drugs
- Drugs and birth defects
- Drugs and breastfeeding
- Sex, gender and drugs
- Drug interactions
- Education
- Drug statistics.

The editorial board systematically follows developments in the field of drug treatment and publishes brief medical as well as general news articles related to drugs. An important criterion for news selection is that it should be considered to be or to become relevant in patient care. An important part of Janusinfo consists of evidence-based guidelines produced by Drug Expert Panels. The panels also produce the Kloka Listan (Wise List), which contains those drugs that have been recommended and Informatics. Drug statistics offer sales returns and sales statistics of drugs in the Stockholm County with analysis and comments.

Comment from T1.3 on Janusinfo.se: The site was mentioned as it has the nice feature of providing a possibility to do on line interaction query which is rather effective since the underlying structure is of good quality and include detailed ATC codes.

5 Sources of information – Enrichment and education resources

It is foreseen that the Gravitate-Health community will over time assemble enrichment resources as well as educational material. Such work would be substantially be supported by the establishment of a Gravitate-Resources-with-Enhancement-Database. By inclusion of indexing features corresponding to those described in chapter 7 (the information model). This entails that all added material should be labelled with a characterizing vector that quantitatively describe a number of dimensions such as knowledge domain, level of complexity, required health literacy level, primary target group, issuer, languages supported, relation to certain drug classes etc.

5.1 CareAnimations (www.careanimations.com)

The example CareAnimations is a commercial entity that aims to help care providers to provide patients with easy-to-understand and tailored information to improve treatment adherence and self-management, i.e., goals that overlap the Gravitate-Health project. With a team of medical and pharmaceutical experts, medical writers, animators, and IT specialists, CareAnimations offer different communication tools for pharmacies, hospitals and doctors that are easy-to-use and easy-to-implement in existing workflows. For patients, CareAnimations create information that is easy-to-access and easy-to-understand for almost everyone, including people with limited health literacy, by using animated videos and pictograms. CareAnimations has built substantial information resources and have now more than 11000 animations in a library.

Comment from T1.3: Importantly, CareAnimations have both material on drugs but also in the same library health literacy supporting material on diseases. They use a strategy of re-confirming the content with the regulatory bodies. CareAnimations have built a solid reputation and is endorsed by e.g., CBG MEB, NHS and others. They are operating in Belgium, Denmark, France, Germany, India, Spain, Sweden, and The Netherlands. The material is created to seamlessly serve different languages, which makes it attractive. Products such as those exemplified by CareAnimations offer a clear potential for the enrichment process of G-lens focusing of information.

Summary CareAnimations:

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
Yes	+++	+++	2	+++	++	Select EU countries

6Sources of information – The end user contribution

Patient reported information is considered to be of high trust and relevance provided that the origin of the information is verified. Therefore, citizen/patient reported information has different trustworthiness pending on the issue of end point identification. General information may distributed irrespective of ID status whereas the id-status must be secure should the G-Lens distribute any information that provides personalized advice. Currently, the aim is to provide only highlights of regulated text and authorized enhancements in response to end user queries. That means that the general user with a general question does not need to be identified. However, if somebody would like to authorize the use of information from a protected source with sensitive information a full verification of identity is needed. The same conclusion regards any use of the information collected with the G-Lens for research purposes. Here, a proper informed consent must be obtained and this requires that all contacts with the subject need fully proven identity of the end user.

6.1 The electronic health record/EHR and EMR

The European Commission has stated that the interoperability of the existing EMRs is at a very low level and taken an initiative[20] in order to make it possible to at least at a very basic level exchange information based on standard components. The European Commission mentioned specifically five components. No country has as of yet implemented all five of these and a standard for the proposed communication capabilities has not been proposed. The combination of strong proprietary formats for data storage, the lack of communicated information and data models from major vendors and the legal/responsibility issues suggest that using direct EMR data is not an option for G-Lens as formal release and handover of the data is necessary according to the GDPR. The poor structure of the data means that a considerable 1-to-1 mapping needs to be done and also that a lot of information needs to be harvested and digested in order to recover anything worthwhile. In order to circumvent the poor EMR data structures large scale attempts have been launched to probe for structure in the masses of information. In summary, most projects have not reached the goals, have required massive amounts of information and also concluded that the inherent structure in general medical records is so noisy that individually predictive data can never be reached. This is based on the fact that the completeness of the EMR cannot be assessed and that the data coverage is highly varying across individuals[21]. Large scale NLP extraction of information often use matrices (such as ICD10 or highly specialized local constructs) to reduce the information space and thereby reduce the variability. In spite of such efforts the precision almost never yields clinically valid and safe data on the individual level in spite of old claims in the initial communications of methods[22]. Almost without exception the data annotation in the EMR is of poor reach and almost all vector power of the single data point is lost[23]. This rests on the fact that the EMRs were created for record keeping to protect the professional structures and provide a basis for the single doctor to do the work. Most of the information is based on freely dictated free text. Lately the EMRs have been amended with a limited capability of handling structured information. However, uneven coverage and the lack of between vendor interoperability creates major obstacles.
Large scale projects have attempted to circumvent the lack of interoperability by republishing data extracted from the EMR internally and create data bases for clinical science purposes. One example is EH4CR (https://www.imi.europa.eu/projects-results/project-factsheets/ehr4cr) that in spite of very ambitious efforts showed marked resistance to scale for general use, mainly because of the lack of inherent structure in the EMRs.
Therefore, the EMRs are seen as distinctly difficult to use as source in any procedure that rests on automatic extraction of information.

[20] https://digital-strategy.ec.europa.eu/en/library/recommendation-european-electronic-health-record-exchange-format
[21] https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(21)00041-9/fulltext
[22] Sheng-Feng Sung, Kuanchin Chen, Darren Philbert Wu, Ling-Chien Hung, Yu-Hsiang Su, Ya-Han Hu, Applying natural language processing techniques to develop a task-specific EMR interface for timely stroke thrombolysis: A feasibility study, International Journal of Medical Informatics, Vol 112, 2018, 149-157,
[23] Martin Ingvar, Mathias Blom, Casper Winsnes, Greg Robinson, Lowie Vanfleteren and Stan Huff, On the annotation of health care pathways to allow the application of care-plans that generate data for multiple purposes Methods, Front. Digit. Health - Health Informatics, in press 2021

The emerging EU Commission proposed route to standardization revolves around 5 areas regarding interoperability on patient information:

1. (a) Patient Summary;
2. (b) ePrescription/eDispensation;
3. (c) Laboratory results;
4. (d) Medical imaging and reports;
5. (e) Hospital discharge reports.

The one that is most developed is the International Patient Summary (IPS) and it is also the component that is of most interest for the G-Lens development in the Gravitate-Health project.

Comment from T1.3 on the use of the EHR as a source: The issue of the poor structure, the abundance of free text notes and low coverage suggest that this source is highly complex to use, has poor access and is highly laborious as e.g., FHIR APIs to EMR records are almost non-existent. Also, it is noted that health care providers that allow access to the EMR do it for two reasons. In some countries it is a legal obligation to provide the information to a central repository.

EMR/EHR summary:

Regulated status	Trustworthyness	Consistent structure	Granularity	General accessability	General coverage	Geographic reach
No	++	0	0	(+)	+	Provider border

6.2 The IPS

An **International Patient Summary (IPS) document** is an electronic health record extract containing essential healthcare information about a subject of care. As specified in EN 17269 and ISO/DIS 27269, it is designed for supporting the use case scenario for 'unplanned, cross border care', but it is not limited to it. It is intended to be international, i.e., to provide generic solutions for global application beyond a particular region or country. The specific purpose is for other caregivers to quickly grasp the health care needs of a patient they have never seen before. Hence, the information profile is well adjusted for use in the Gravitate-Health setting specifically when compared to the EHR.
The IPS is seen as a toolbox that may develop, and an Implementation guide is under way for HL7 FHIR and there exist also a guide for HL7 CDA. The use of the word minimal reflects the ideas of 'summary' and the need to be concise, but also alludes to the existence of a core set of data elements that all healthcare professionals can use; it is intended to be a speciality agnostic and condition independent set. It does not imply that all the items in the data set will be used in every summary. It is also possible to refine the extract from a record such that the content of the summary is more relevant to a particular condition (e.g., asthma) but no asthma-specific elements will be specified in this standard. The

IPS Document or IPS can be extended by non-IPS standard condition-specific data. 'Non-exhaustive' recognizes that the ideal data set is not closed, and is likely to be extended, not just in terms of requirement evolution, but also pragmatically in instances of use. [EN 17269; ISO/DIS 27269].

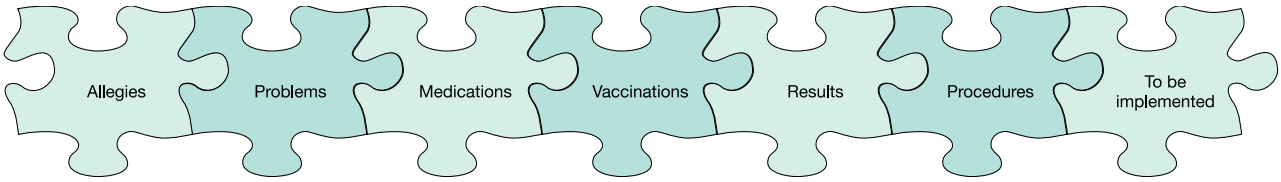

Figure 7. An illustration of the components that are described so far in the HL7 FHIR implementation guide

The compositions of the IPS represents a set of modules.

The IPS document is composed by a set of robust, well-defined and potentially reusable sets of core data items (indicated as IPS library in the figure below). The tight focus of the IPS on unplanned care is in this case not a limitation, but, on the contrary, facilitates their potential re-use beyond the IPS scope.

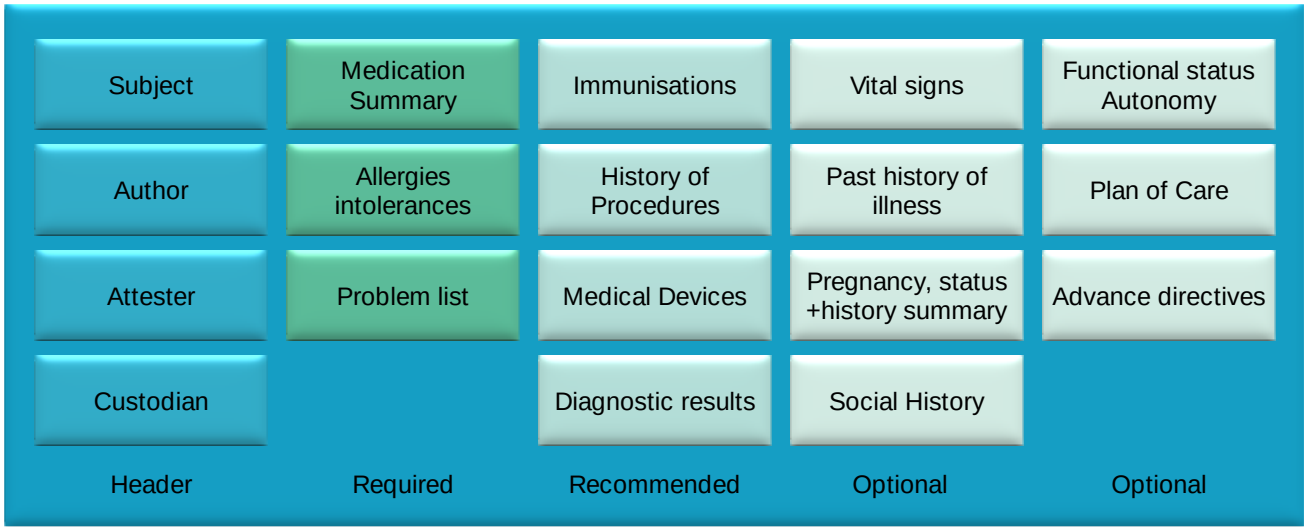

Figure 8. The content of the IPS modules

6.2.1 Inherent structure of the IPS:

The International Patient Summary is specified as a templated document using HL7 CDA R2. The expressiveness of SNOMED CT and other primary terminologies enables this specification to represent the two general categories "condition/activity unknown" and "condition/activity known absent" in a style which is more independent of the underlying syntax (CDA R2 or FHIR), as explained in detail in section 4.2.

To be universally exchangeable and understood, a patient summary must rely as much as possible on structured data and multilingual international reference terminologies that are licensed at no cost for global use in the International Patient Summary. In the case of SNOMED CT, it is envisioned that SNOMED International could embrace the idea of a globally accessible open and free specification for the International Patient Summary that references a core set of

globally accessible and usable value sets licensed at no-cost with the aim to serve the public good. In this spirit, this version of the International Patient Summary defines SNOMED CT as a primary terminology (the meaning of "primary terminology" is explained in section 4.1) and it is used in many of the value sets. To allow, however, a global and free implementation of the IPS this guide does not impose the usage of these SNOMED CT-based value sets. This choice may be revised in future versions. Other primary terminologies used in this specification are LOINC for observations (e.g., laboratory tests) and document sections, UCUM for units of measure, and EDQM Standard Terms for dose forms and routes. Looking at the availability of other globally usable reference terminologies and toward alignment with a future FHIR version of the IPS, in some selected cases FHIR-defined terminologies are recommended.

This specification adopts ART-DECOR® as the specification platform for this Implementation Guide and uses the HL7 template exchange format. This tool and format are increasingly used by several regions, including European countries, and have been adopted by the EU eHealth Digital Service Infrastructure (eHDSI) project for the operational deployment of the EU cross-borders patient summary and ePrescription services. Users of the specification can visit the IPS project page in ART-DECOR® to browse the specifications and review examples. Users may also use the tool to validate their IPS instances.

Comment from T1.3 on the IPS: The initial ambition of international interoperability of the IPS can certainly be reached with the current roadmap for the IPS. There are a set of limitations when using the IPS in the Gravitate-Health environment that need to be considered.

1. The choice of recommending a vast canonical reference set of terms (SNOMED CT) but not enforcing it, limits the ability to compute on data from the IPS. An example could be allergies that internationally has a very well established taxonomy with exemplars at each categorical level and still anchored in SNOMED CT. Without taxonomic anchoring and a defined subset of SNOMED CT terms that data becomes non-computable.

2. The rate of implementation is low and to date no examples of national or regional repositories have been identified. Hence, it is unknown if the lack of availability will be limiting the development of the G-Lens functionality

3. The choice of the ART-DÉCOR is probably the best available option; however, it requires high end knowledge of the user to appreciate the functionality of this platform.

We conclude that the computability of the IPS for use in the highlighting of the G-Lens response is limited with the present specifications and speed of implementation. Also, the general availability for patients using G-Lens to submit their IPS as a basis for the focusing procedure remains in question.

6.3 Questions and personal responses/User profiles

The initial prompt to start a G-Lens query is an initiative from the citizen/patient. This will probably be a two-step procedure with follow-up for clarifications on the posed query for information. The procedure may also include measure to detect preferences and user profiling by target questions. However, clumsy and laborious on-boarding procedures may limit the user experience. Therefore, an interactive profiling against a persona dimension gallery may constrain the length of the procedure. User profiling necessitates the use of stored sensitive data either locally or centrally and must be included in the early information to the user either as information or as part of the consent procedure.
Upon the final query response from the perceived user experience and other KPIs are evaluated. While the G-Lens have the user on-line it is quite uncomplicated to pose some evaluating questions. Once logged off, it is a different ballgame including posing questions that are formulated on the basis of sensitive information. Here, procedures for identification and matching of earlier interactions with later is central. From a sensitivity point of view a full GDPR safe machinery needs to be available also for the evaluation of "neutral information".
From an architectural point of view, it is of essence that an agreement exists on where the personal data are stored. The consensus in T1.3 work group is that individual data are stored as peripherally as possible as there is little variation in the core central G-Lens procedures and variability belongs in the microservices. This does not preclude a prescribed functionality for securing identity of individuals and measures to secure data integrity.

Aim	Means that	Comment
Provide health information to meet the subject's preferences and needs	Onboarding, profiling to persona matrix, Secondary request for more information or confirmation of findings.	Well structured information
To know user satisfaction	Divided between within G-Lens operations and general evaluations for chosen KPIs.	A general output of assisting information to external evaluations
To know if behavior related to health is influenced and if compliance to treatment is affected	This is generally derived outside the core G-Lens operations. However. The G-Lens should provide the service to handle consent and of course the persona-vector for the individual based on consent.	A general output of assisting information to external evaluations
To securely know the identity of the end-user	This has yet to be explored further given the national solutions that vary across EU countries	National solutions necessary
To secure that the end-user always has the latest version of the information	A centrally devised mechanism that allows for subscription of the basic regulated information in latest version.	A subscription possibility of equal design across all the instances of the G-lens.

Table 11. Considerations on manually obtained information from the end-user

6.4 What can we compute with the available information?

As outline above the lack of structure in the targeted information sources seems to be the major difficulty for reaching the goals of the Gravitate-Health project. We plan on using three different streams of information as a basis to compute an individually focused result. Such computations are always limited by the stream that has the poorest structure. A general analysis demonstrates that the header/text based repositories of regulated information provide the lowest level of structure. Hence, the conclusion that a dialogue with regulators regarding the future strategy for the provision of e-based information such as the ePI is highly warranted by the Gravitate-Health project in order to work towards the fullest realization of the potential benefits of ePI.

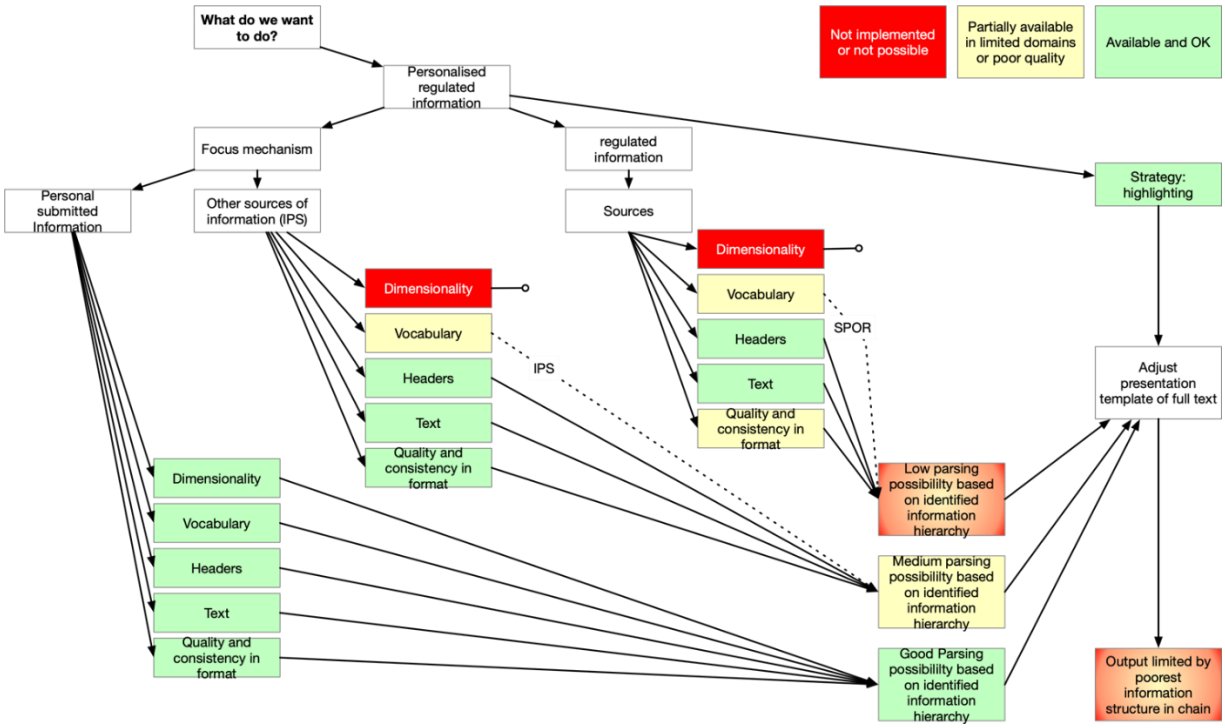

Figure 9. The general structure of the information from the three streams of information demonstrates that the regulated sources provide a hurdle for reaching individually focused information.

7 G-Lens Information Process Model

The G-Lens operation involves end user queries, acquisition of personal information, fitting of data to the persona archetypes, extracting and refining the information from trusted sources. As a complication EU national laws, languages and local information sources that may be employed in the G-Lens function. The discussion has therefore led to split of the G-Lens into one pan-EU part paired with multiple national instantiations. Such an architecture is dictated by the information landscape and at the same time underscores the need for a consortium wide agreement on the principles of information handling, i.e., an information model and information process model. Such models can have a very

high level of detail or alternatively be used as a framework during a development phase that serves the common development of a technical solution.

A simplified G-Lens process model can be illustrated as follows in a standard BPMN 2.0 format:

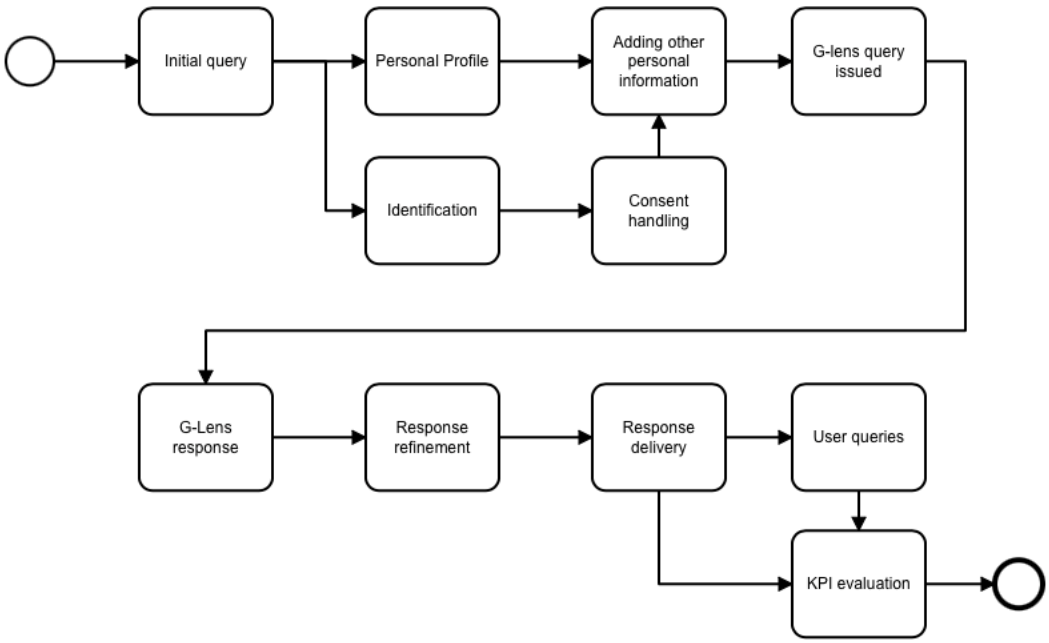

Figure 10. A low resolution process model for the G-lens depicted in standard BPMN 2.0 format

The process model illustrates the schematic process components from user initiation to response including the evaluation of the response. As a user-centered service tool the G-Lens process can only be initiated by the user. Several of the following steps necessitate a dialog with the user.

Progressing from the information model to an information process model means to add the necessary information sources and map where computations are made. Given the information landscape the information process model must separate processes that contain information with regulated status and information that is trusted.

The information process model has been designed in dialogue with the work group. Importantly it was deemed not necessary for the model to be exhaustive but rather provide a common understanding of the information steps towards the G-Lens function. The model is outlined below:

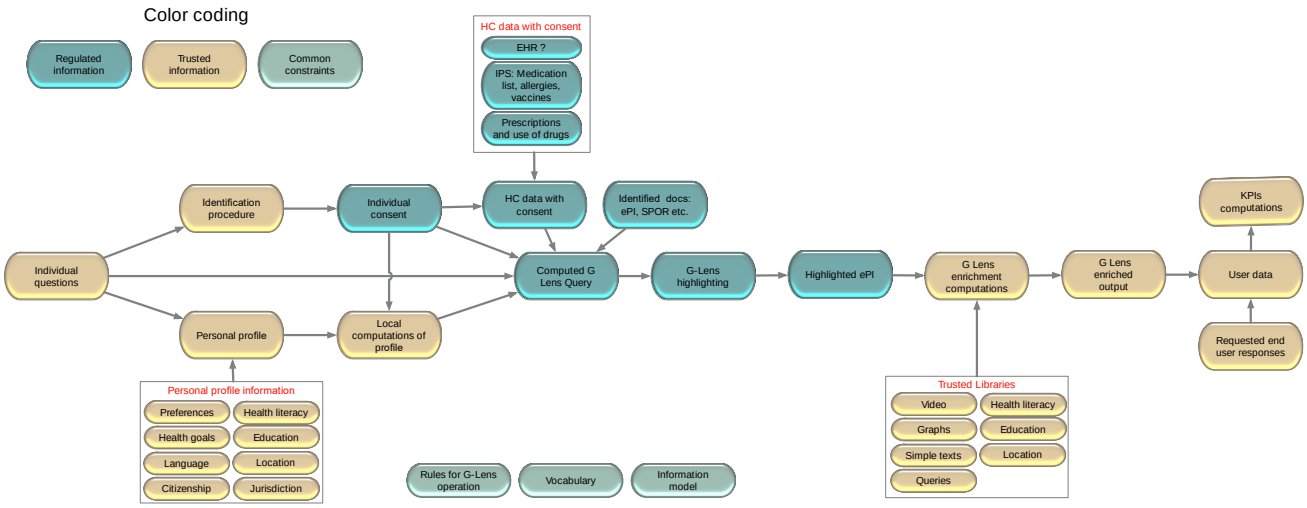

Figure 11. A more detailed core information process model for the G-Lens project

Beyond the process model the point in the process where the sources of information are added are outlined.

7.1 Common constraints for the G-Lens information process:

Three common constraints are noted. These are the rules for the use of the G-Lens, the common governed vocabulary and the information model. These constraints are given as to secure a smooth distributed development of the G-Lens.

7.1.1 Rules for the G-Lens

Given the distributed development and implementation of the G-Lens model common rules have to be met regarding choice of standards, language for queries, handling of consent and adherence to the general regulatory framework. One such rule that has emerged as a likely candidate is that all modules adhere to HL7 FHIR for any communication between modules.

7.1.2 Common vocabulary

A common vocabulary that has a central governance is of essence in order to create possibilities for intelligent use of the available information ad for the computation of the personal profile. Hence, an early workgroup should determine what dimensions will be employed in the persona based profiling procedure.

7.1.3 Information model

The common information model is particularly important during the development phase and given the moving landscape regarding external information and technological advances it should be version handled and governed centrally in the project.

7.2 Personal profiling

A core function that is projected for the G-Lens is the ability to focus (fig 12, a) the regulated information (fig 12, a) and this entails the use of distilled dimensions from the ground work on the personas (fig 12,c). If a common dimensional space is used both for the personas and for the classification of the augmentation material (fig 12, d) this process can be highly effective to the point of full automaticity. This entails an extensive cataloguing of the content enhancement resources. Also, the initial presentation of the raw ePI could intelligently use focus on the header level (Fig 12, e) to constrain the amount of information presented in the first instance.

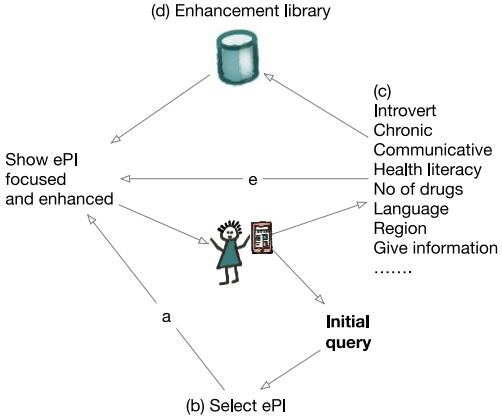

Figure 12. The profiling of the end user is a separate process from the selection of the ePI. It is used in the steps of enrichment and communication with the end user

8G-Lens Information Model

This Deliverable (D1.4) from Task 1.3 specifies a prototype for an information model that could provide a basis for across consortium agreement on interoperability and data architecture. As in the case of the information process model the information model represents a first version that should be governed and continuously be updated along with the project development. A number of underlying assumptions and findings have guided the development of the information model.

Assumptions

- An information model may be developed gradually with increased granularity
- Information with status as regulated should retain that status through the G-Lens if possible
- The computability of results depends on the ability to make structure and semantic content explicitly available.
- All information has layers with content, structure and concepts.

- There are four main streams of information:
    - Sources with regulated information
    - Personally submitted information
    - Other sources of patient information
    - Library of enhancement resources
- All information manipulations should be based on automatic functions
- A guiding information model is a basis for the creation of semantic interoperability and definition of computability of the information mass.

Findings

- Direct access to EHR/EMR is only possible in a limited way for the G-Lens and is therefore not included in the information model
- The limited structure and overlap of content but not structure between different regulated document sources is a fundamental problem for the Gravitate-Health project.

8.1 The task of creating an information model

The development of the hypothesis of an information model has been core of the discussion in Task 1.3 meetings. We have reached group acceptance of the proposed information process model and information model described and the current model is fully suited to serve as a basis for reaching consensus across the consortium. We foresee that next versions of the information model will evolve to increased granularity. For the ease of interpretability the information model has a certain resemblance with the information process model.

8.2 A proposed information model

Based on the discussions above the final version of a low resolution information model is given below in standard UML format. The focus has been to delineate contingencies between different sub-domains.

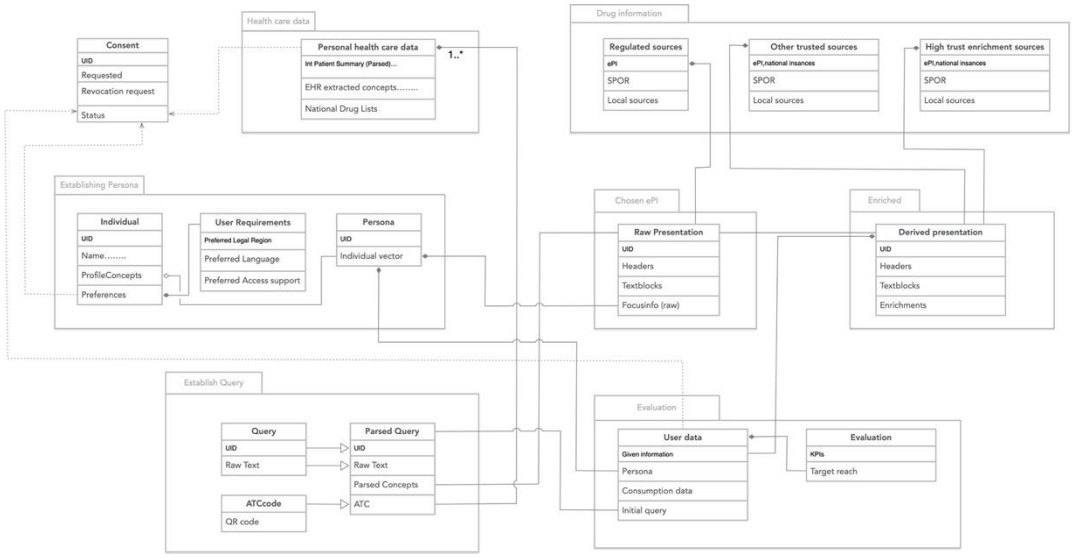

Figure 13. A hypothesis for a low resolution information model showing the subdomains that have emerged during the work in T 1.3.

The model is given in UML format. The model meets the demand listed above including the separation of all information that is classified as regulated from other information. The proposed drawing represents a hypothesis and may serve as a basis for further refinement during the Gravitate-Health project.

The tentative list of identified information sub-domains or points of manipulation of information are given below:

Entity in information model	Explanation	Vocabulary dependent	Comment, information elements
Free text query	The initial query with low or no structure	no	Free text
Query content	Free text mass of question	no	Free text
Query structure	By formatting follow up question initially field specific information may be entered	yes	Structure guided by input mechanism
Query concepts	Identified objects from the vocabulary	yes	Identified concepts in free text
Reformatted query	Reformatted with available structure and concepts		Extracted information, identified drug
Status sensitive data	Concept of sensitivity determined	yes	Toggle Y/N
Collect consent	Consent query and formatted	yes	Structured query, all aspects of interest listed
Compute consent	Determine what consents are given including research studies	yes	Individual consent vector
Revoke consent function	Function to submit structured and vocabulary correct revoking of consent	yes	Ability to selectively modify the individual consent vector
Repository consents	Stored consents in full matrix	yes	Repository for individual consent vectors

Structured persona profiling query	Standard query to obtain data for profiling according to the persona dimensions	yes	Query results in personal profile vector encoded according to persona vector
Language region	Preferred language	yes	Active choice of predetermined array
Geographical region	Region, legal jurisdiction	yes	Active choice of predetermined array
Accessibility needs	Structured representation of requested enhancements in the presentation	yes	Accessibility vector determined
Requested persona dimensions	Semi quantitative profiling along the identified persona dimensions including health literacy, medications, communication preferences	yes	Individual data collected based on the persona dimensions
Persona dimensions	Comprehensive list of persona dimensions of interest	yes	Identified persona dimensions and value space for each
Computed persona	Matching of query response with the identified dimensions. Matrix persona fingerprint established	yes	Individual result vector based on the persona dimensions
Prescription register	National prescription/dispense registers that give IDMP based identification of drugs	no	Valuable structured source of information
Drugs of interest	List of drugs of interest as to make the choice later on from the ePI repository	no	Identified IDMP IDs established as a list
Initial ePI selection	Based on extraction of the initial information a suggested ePI/PL is selected for presentation	yes	Probability based list of ePIs of interest
ePI	EMA repository of ePI	no	Structure for search unknown at present
SPOR	EMA repository of SPOR	no	IDMP structured data available
Raw focused presentation	Clickable focused presentation of partial PL from ePI	yes	Selected PIL displayed by headers and opened single segment based on primary query
Request of clarifications	Secure the initial choice is the requested one. Secure more persona information if needed	yes	Verification of correct choice, follow up questions
Response to structured query for more information	Structured response given	yes	Structured query
Collected structured clarifications	Improved matrix for persona and secured choice of drug of interest	yes	Refined vector of additional guiding information and verification of ePI choice

IPS	When available use IPS for patients.	no	Mechanism of use not determined at this stage
Local sources	National resources with drug and health information that are either regulated or highly trusted.	unknown	Here it is foreseeable that local solutions will have very different architectures. Mechanism of use not determined at this stage
Enrichment resources	Repositories of highly trusted enrichment material such as text, pictures, movies	no	Regulated but largely unclassified material
Codified enrichment resources	Codified according to a vocabulary matrix in order to make choices computable	yes	All source material classified with Gravitate-Health classification vector
Secondary computation	Based on final collected information full focused is presented	yes	All computable elements are used
Request user experience	Structure query on PREM, PROM, adherence, behavior	yes	Structured query according, consent used
KPI definitions	KPI defined in a vocabulary matrix and fully dimensionalised	yes	KPI defined in a vocabulary matrix and fully dimensionalised
KPI computations	Within G-Lens computations according to the matrix	yes	computations according to the matrix matching query responses

Table 12. List of identified masses of information that the information model entail. Some of them are depicted in fig 13.

9 Mapping to KPIs (T1.4)

An information model and information process model should in principle not constrain any of the decided KPIs. This necessitates a design where the generality of the dimension KPI is built into the model and that the information process model includes information flows that allow for the collection of data from the information process as well as directed queries to the users.
A mapping of the KPI resulted in some main groups of KPIs where only those belonging to the class “user evaluation” seem to have a clear dependency on the information process model. The other classes are either general group level measures or surface measures.

Group of KPIs	KPIs	Information model/Process model
User evaluation	Digital solution provides notifications and updates on prescription/OTC ePI Patients understand medication benefits, how and why to take medication	In the post G-Lens communication step with the users

	Provider experience with patient use of G-Lens Patient empowerment and activation User empowerment through digital solution features Patient empowerment and activation Patient assessment of G-Lens: no information overload or missing data Trust index for patients Outcomes Achieved lower risks across population	
Directed queries	Health provider satisfaction Trust index for providers Better medication compliance Safer use of medication/ therapy administration Legal and Privacy requirements balance “need to know”, usability, accessibility, and no harms policy KOLs in Pharma and policy maker for experience User empowerment through digital solution features	Directed queries external to the G-Lens
Surface and technical measures	EU Citizen awareness about the project Citizen awareness of importance of self-management Platform availability, usability, and accessibility Platform availability and accessibility Platform addresses physically, auditorily, visually, challenged, dyslexia requirements Maturity of the technology platform at end of project	General evaluation outside of the G-Lens information model
Sources and refinement	Multilingual capability Digital literacy level of educational material	Directed queries external to the G-Lens
Mixed dependencies	User Preferences / Co-creation Provider awareness of own bias and understanding of patients’ attitude towards medication adherence	Information model dependent and technically dependent

Table 13. Grouping of the preliminary KPIs and their respective dependency on the information model. The information model provide support for all KPIs that are measured within process. The others are considered orthogonal.

We conclude that the proposed information model with its limited granularity does not pose any constraints on the KPIs so far listed in the project. However, given that the G-Lens will be installed in multiple national instances we suggest adding a dimension in the KPIs where comparisons between the different instances are included.

10 A mapping of each proposed scenario against the information process model

The task of mapping the scenarios regarding their fit to the developed information model has to be labelled with the precaution that it is a preliminary evaluation for two reasons. Firstly, the scenarios are in very different stages of maturation, and secondly, the available information on the data architecture and data design within each of the scenarios is virtually non-existent. Thus, we have

restrained the analysis to an individual guesstimate of the fit to the developed information process model (addendum 1). It is not meaningful to map towards the information model because of scarcity of basic information.
The overall map resulted in the following figure:

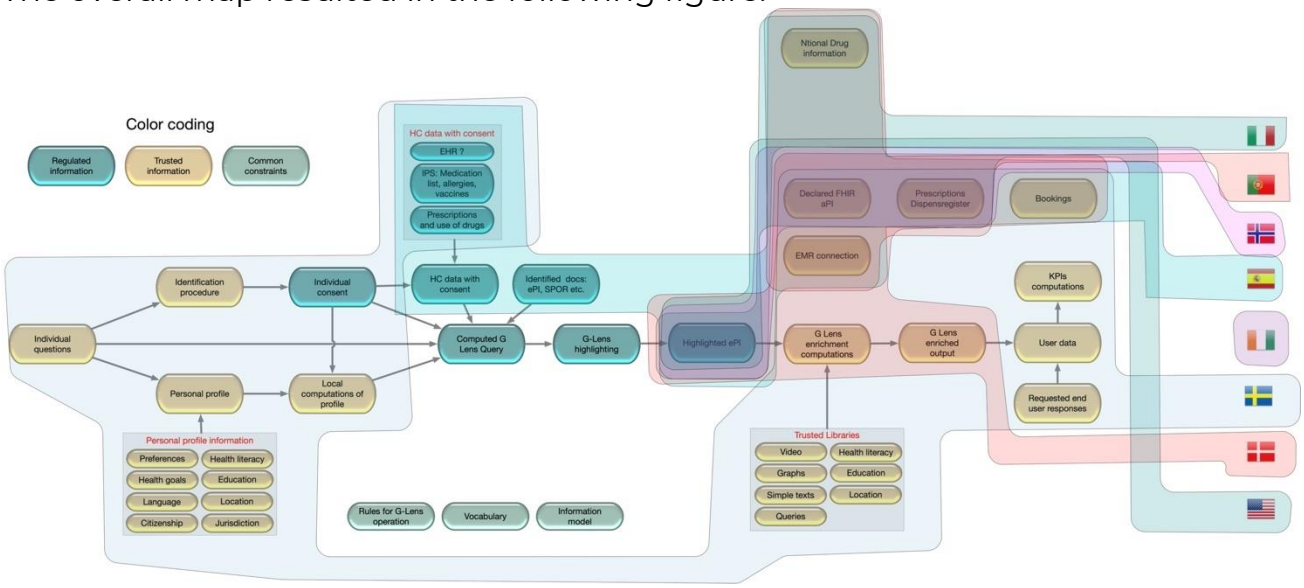

Figure 14. The information process mapped with the scenario provisions of services.

The overall analysis demonstrates that the majority of scenarios are focused on delivering PIL type of information and almost exclusively starting from a national source. The Swedish have explicit enrichment and effect evaluation built into the scenario. The Portuguese, Spanish and Swedish scenarios discuss interaction with health care information. None of the scenarios have revealed plans for receiving information from EMA ePI, please note that the roadmap for ePI implementation is not yet available. Several of the scenarios have a detailed FHIR commitment, which means that it will be a manageable task to include such procedures in most of the scenarios. We have at this point not found any scenarios that are explicitly incompatible with the information process model. However, information is limited at this time for several scenarios. Each of the scenario mappings are given in addendum 1.

11 A strategy to consecutively include new developments in the information structure provided from trusted sources

It has become evident during the preparation of this report that the information landscape in the regulated sphere provides a magnificent challenge in that it has a pre-digital age logic defined as:

- No limited vocabularies
- No fully enforced use of canonical vocabularies such as ICD10, IDMP or SNOMED CT

- No defined hierarchical taxonomies (limited formal structure)
- Manual procedures (word templates often) for obtaining regulatory status
- Manual re-confirmation following any modification of the content
- Overlap of information content without corresponding structure
- No document instructions for interoperability or “structure at source” instructions
- Of the regulated sources only the PIL has the ”informed layman” as a target group.

This means that the intended use of the regulatory texts is that they should be read and understood by the intended target population and they are not intended as information sources that can be further refined (without following a regulated change management process) without losing the regulated status.

The most compelling obstacle that has been identified is the ePI which, while being under development, at the same time suggests that it will become a source of information for further refinement and public use by third parties but yet retains the characteristics of the underlying documents as defined by the QRD templates. This means that is spite of the surface level standardization all of the points above remain in the ePI.

There is a roadmap for the further development/Implementation of the ePI under way. While little information has been released on this as yet, it is not anticipated that plans would include reformation of the underlying documents. Within the EMA action plan[24] related to the European Commission’s recommendations on product information the need for improvements in the QRD templates was identified but this has not yet been progressed. It is therefore unlikely that the ePI will develop the internal information structure more than marginally during the project time for Gravitate-Health.

The analysis of the sphere of regulated documents including the overlap of information content (Chapter 2) suggest that the shortest route to deliver better structured information in the regulatory documents would be two long term commitments.

Suggested commitment number 1: Reformation of the QRD template to ensure that the headers and content of corresponding SmPC/PIL have a 1:1 relation. This would also entail a removal of the ability to cross reference between different headers differently in the SmPC and PIL.

Suggested commitment number 2: Prioritize the work on the SPOR that use the IDMP standard and insert a dependency with SPOR as a parent and PIL/SmPC/LABEL as children. This would rationalize the development of

[24] European Medicines Agency action plan related to the Commission recommendations on European Union product information (europa.eu)

regulatory documents as there is a partial doubling of the information in those sources. Obviously, the information published according to a structural standard like the IDMP is of better use for later refinement that information published in a simple header/text format.

Given the dynamic nature of the current EU telematics landscape, close collaboration between stakeholders over the coming years will be key to enable alignment across initiatives and work towards the establishment of an enhanced health information landscape, for example an ePI with high levels of structuring, which will maximize the value to users of digital health information tools such as those that Gravitate-Health will be developing.

The strategy for being prepared for new developments on the structure in regulated information is quite simple – the aim is to have an adaptable model built on general principles rather than continuous special cases. This principle has guided the presented information model (chapter 6). The G-Lens environment is defined in categories with external sources of regulated as a single category with a aim to extract **computable elements, structure and enclosed information**.

The developed construct makes the G-Lens adaptable to any change in the structure of the information in the regulated sphere. Also, any addition of a new source may be analyzed in the same way. If a source changes implementation of a standard of adds new structural elements to its content it will not influence the validity of the previous information model. The change may be implemented by a simple amendment.

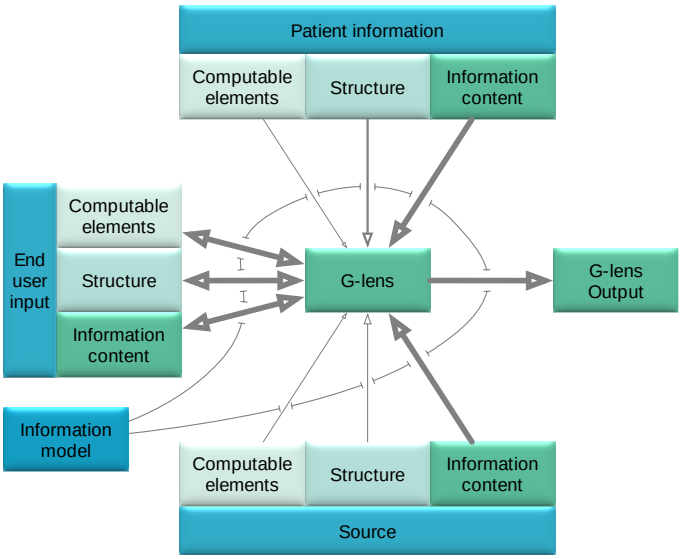

Figure 15. Computing the G-Lens output pertains to matching three different classes of information. It is the class with the lowest structure that limits the computability

12 Concluding discussion

It should be noted that a number of web applications exist in several countries within the EU where drug information is made available to the public that use clickable headers. This means that without any computations for highlighting a far better presentation can be reached as compared to the presentation of a full document. On the other end extensive personal profiling entails using sensitive information (GDPR) but also run the risk of becoming personal advice and not general information. That would take the G-Lens from being a source of information to become a medical device under MDR. This means that the projected ability for the G-Lens to provide individually tailored information in a focusing procedure has identified challenges and these areas will require appropriate consideration.
Above and beyond these challenges stands the structural quality (not the content quality) of the regulated sources of information at the present time. The lack of an explicit computable structure and the fact that a vocabulary is not enforced caps the reachable capability of the G-Lens concept. We conclude that a project wide development of a common strategy to reach the goals to provide personally tailored useful and attractive information is warranted, including collaboration with key stakeholders in order to align on the key priorities for future development that will allow for maximal value of ePI to be delivered to end users.
This leads us to summarize:

1. Most currently available sources of information have a very limited structure and granularity of the information.

2. An ePI Set-up-project was launched by EMA in 2021. Deliverables due by the end of the year include a FHIR-based EU proof- of-concept standard for ePI and roadmap for future implementation. As of yet, no information suggest that the ePI implementation will entail a reformation of the information structure of the constituting documents.

3. Implementation of ePI within Europe is likely to take some years, however, with different countries proceeding at different rates. ePI will not have full reach before the end of the Gravitate-Health project. Gravitate-Health will focus on use of the available resources in the testing scenarios.

4. The ability to automatically focus information in an individualized way is put in question given the lack of structure beyond a limited number of headings in the key available information sources. This will need to be further considered as the project continues.

5. An implementation of a pan European central G-Lens mechanisms in combination with national instances is recommended as incorporation of nationally available information is best served that way. The information model serves as a guideline for this more complex implementation

6. Strong collaboration across stakeholders/different telematics initiatives (e.g., relating to ePI, to SPOR) will be key to optimize development of ePI and digital tools to be of the greatest benefits to end user.