

# D5.3 – Guidance for the implementation of common principles

# for practical governance of big data

# with a special focus on data to be used (and the implementation of data access and use) in public health, research and quality assurance in healthcare on a European scale

# WP 5 – Innovative Use of Health data

03-05-2021

Version 2.4

# 19<sup>th</sup> eHealth Network meeting, June 2021 For adoption

Grant Agreement nº 801558



Co-funded by the European Union's Health Programme (2014-2020)

To ease the uptake of innovative usage of data across the healthcare sector for the benefits of society, individuals and performance of Member State health systems, this guidance proposes suggestions for 20 actions to support the of creation of good governance principles, practices and methods in handling use of health data, including big data.





CONTROL PAGE OF DOCUMENT				
Document name	D 5.3 – Draft proposal for the eHealth Network on the guidance for the implementation of common principles for practical governance of big data with a special focus on data to be used (and the implementation of data access and use) in public health, research and quality assurance in healthcare on a European scale			
Work Package	WP 5 – Innovative Use of Health data			
Dissemination level	PU			
Status	Final			
Author(s)				
Beneficiary(ies)				

Dissemination level:

*PU* = *Public, for wide dissemination (public deliverables shall be of a professional standard in a form suitable for print or electronic publication) or CO* = *Confidential, limited to project participants and European Commission.* 

REVISION HISTORY						
Version Date Author Organisation Description						
0.1	01-04-2020	István Csizmadia, Robert Láng, Márton Kis, Anna Feller, Krisztina Sipos	NHSC, SU	First version of the document.		
0.2	13-07-2020	István Csizmadia	NHSC	Chapter on definitions, introduction added		
0.3	27-07-2020	Kathrin Trunner	GOeG	Results from FGM and first version of principles included		
0.4	04-09-2020	István Csizmadia, Robert Láng, Márton Kis, Anna Feller, Antal Bódi, Kornél Tóth, László Bencze, Claudia Habl, Kathrin Trunner, Alexander Degelsegger- Márquez, Cátia Pinto	NHSC, SU, GOeG, SPMS	Restructured version of v0.3 with details of principles, updated for QM		
0.4a	13-09-2020	Hugo Agius Muscat	MFH	QM review of version 0.4		
0.4b	16-09-2020	Anna Feller, István Csizmadia	NHSC, SU	Updated version after QM review		





D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

0.4c	20-09-2020	Hugo Agius Muscat	MFH	QM feedback on version 0.4b
0.4d	20-09-2020	István Csizmadia	NHSC	Updated version after QM feedback
0.4e	12-10-2020	László Bence, Kornél Tóth, Kathrin Trunner, Diogo Martins, Cátia Pinto, Arto Vuori, István Csizmadia	SU, GOeG, SPMS, THL, NHSC	Updated version addressing EC comments
0.4f	20-10-2020	Hugo Agius Muscat	MFH	QM review of v0.4e
1.0	22-10-2020	István Csizmadia	NHSC	v0.4f approved by SC
2.1	10-03-2021	lstván Csizmadia, Márton Kis, Antal Bódi, Kornél Tóth	OKFO, SU	v1.0 edited and draft recommendations for guidance added
2.1a	14-03-2021	Hugo Agius Muscat	MFH	QM review of v2.1
2.2	09-04-2021	Anja Laschkolnig, Claudia Habl, Kathrin Trunner, Cátia Pinto, László Bence, Kornél Tóth, Márton Kis, Antal Bódi, István Csizmadia	GOeG, SPMS, OKFO, SU	Updated and restructured version after QM feedback on version 2.1
2.2a	12-04-2021	Hugo Agius Muscat	MFH	QM review of changes in v2.2
2.3	13-04-2021	István Csizmadia	OKFO	Updated version after QM feedback
2.4	03-05-2021	Persephone Doupi, Kathrin Trunner, Anja Laschkolnig, Cátia Pinto, José Dias, László Bence, István Csizmadia	THL, GOeG, SPMS, SU, OKFO	Updated version after SC feedback

#### Disclaimer

The content of this deliverable represents the views of the author only and is his/her sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.





D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...]
 WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

#### **Table of Contents**

TABLE OF CONTENTS4	
ACRONYMS	
LIST OF FIGURES AND TABLES	
EXECUTIVE SUMMARY	
OVERVIEW ABOUT PRINCIPLES AND ACTIONS9	
HOW DID WE GET TO THE GUIDANCE?9	
TABLE 1 - OVERVIEW OF PRINCIPLE-SPECIFIC ACTIONS FOR GUIDANCE       11	
INTRODUCTION	
BACKGROUND14	
Working definitions14	
VISION, PURPOSE AND OBJECTIVES	
AUDIENCE AND SCOPE16	
Methodological Approach17	
Problems18	
Domains and use cases18	
Results from the Focus Group Meetings for identifying principles	
Conclusions from the FGMs on the design of the principles	
PRINCIPLES AND ACTIONS FOR GUIDANCE	
CLUSTERING OF PRINCIPLES AND PRINCIPLE'S MAPPING	
CLUSTER 1: INCREASE TRUST IN PRIVACY PROTECTION AND CYBERSECURITY	
Principle 1 – Work with anonymised or pseudonymised data23	
Principle 2 – Develop digital health literacy24	
Principle 3 – Disseminate data reuse purpose well25	
Principle 4 – Make a great effort to secure and protect data	
CLUSTER 2: INCREASE EFFICIENCY OF USING FUNDS AND FINANCIAL RESOURCES	
Principle 5 – Use existing technical standards	
Principle 6 – Involve stakeholders in co-creation	
Principle 7 – Consider current and future needs	
Principle 8 – Foster data reuse	
CLUSTER 3: FOSTER COMMON INTERPRETATION OF LEGAL AND ETHICAL RULES	
Principle 9 – Adapt model to the settings	
Principle 10 – Create code of conduct	
Principle 11 – Speak the same language	
Principie 12 – Aim for win-win of all stakeholders	
RELATIONS, INTERDEPENDENCES AND PRIORITISATION AMONG PRINCIPLES	
<sup>(</sup> PRINCIPLE - GUIDANCE - GOVERNANCE <sup>(</sup> ERAMEWORK 20	
Figure 8 – Approach to prepare guidance	
ACTIONS I AND 2, FOR PRINCIPLE I (CLUSTER I)	4
prepare stakeholders for adoption	r





Action 2 – Encourage Members states to dev repositories for the EHDS	elop anonymised and pseudonymised data 40
Actions 3 and 4, For Principle 2 (Cluster 1)	41
Action 3 - Adapt curricula to include focus on dig	ital health41
Action 4 - Monitor and measure Digital Health Li	teracy (DHL) level42
Action 5 For Principle 3 (Cluster 1)	42
Action 5 – Define and communicate the value of	data reuse for practical care benefits42
Actions 6 and 7, for Principle 4 (Cluster 1)	43
Action 6 – Use a data governance approach	
Action 7 – Use FAIRified data models and tailored	d solutions43
Actions 8 and 9, for Principle 5 (Cluster 2)	44
Action 8 – Consider the use of existing standards	and data repositories44
Action 9 – Prepare common requirements for 'se	cure processing environment'45
ACTIONS 10 AND 11, FOR PRINCIPLE 6 (CLUSTER 2)	45
Action 10 – Encourage national authorities and be	eneficiaries to identify and use good practices
for stakeholders involvement	45
Action 11 – Develop action plans to improve the	digital health ecosystem45
ACTION 12 FOR PRINCIPLE 7 (CLUSTER 2)	46
Action 12 - Engage key stakeholders in needs ass	essment46
ACTIONS 13 TO 15, FOR PRINCIPLE 8 (CLUSTER 2)	46
Action 13 – Provide clear view about the definition	on and features of 'data'46
Action 14 – Provide lists of free high-value datas	sets and assist internalisation of externalities 46
Action 15 – Identify data quality standards for a	'time-and-space-changing health data space' 47
Action 16 For Principle 9 (Cluster 3)	47
Action 16 – Identify different features of enviror	ment-specific ethical and behavioural norms
ACTIONS 17 AND 18, FOR PRINCIPLE 10 (CLUSTER 3)	47
Action 17 - Analyse legal basis of Codes of Condu	ct under GDPR and demonstrate added-value 47
Action 18 - Define scope and concept of CoC(s)	
ACTIONS 19 AND 20, FOR PRINCIPLE 11 AND PRINCIPLE 12 (CLI	USTER 3)48
Action 19 – Set up co-creation teams and identify	y people who can become 'translators'48
Action 20 – Develop common vocabulary for data	a sharing and access49
RECOMMENDATIONS AND NEXT STEPS	50
Recommendation 1:	Error! Bookmark not defined.
Recommendation 2:	Error! Bookmark not defined.
Recommendation 3:	Error! Bookmark not defined.
	51
APPENDIX 1 - MAIN FINDINGS D5.1 AND D5.2	51
APPENDIX 2 - WORKING DEFINITIONS	53
APPENDIX 3 - OVERVIEW WORK PACKAGE 5 'INNOVATIVE USE (	OF HEALTH DATA'58





D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...]
 WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

# Acronyms

Acronym	Description
AI	Artificial Intelligence
AP	Associated Partner
BDTI	Big Data Test Infrastructure
CEF	Connecting Europe Facility
CoC	Code of Conduct
CSA	Coordination and Support Action
D5.1	Deliverable of Task 5.1 in Work Package 5 of eHealth Action
D5.2	Deliverable of Task 5.2 in Work Package 5 of eHealth Action
D5.3	Deliverable of Task 5.3 in Work Package 5 of eHealth Action
DCF	Data Conversion Framework or 'the stakeholder value chain analysis framework for data conversion'
DHL	Digital Health Literacy
EDPB	European Data Protection Board
eHAction	eHealth Action – 3 <sup>rd</sup> Joint Action supporting the eHealth Network
eHDSI	eHealth Digital Service Infrastructure
EHDS	European Health Data Space
eHN	eHealth Network
elDAS	electronic Identification, Authentication and trust Services
EU	European Union
FAIR data	Findable, Accessible, Interoperable and Reusable data
FGM	Focus Group Meeting
GDPR	General Data Protection Regulation
НР	Health Professional
JA	Joint Action
ML	Machine Learning
MWP	Multiannual Work Programme
PA	Priority Area
PLA	Policy Level Action
SPE	Secure Processing Environment
Т	Task
TL	Task Leader
WP	Work Package
WPCo-L	Work Package Co-Leader
WPL	Work Package Leader





# List of figures and tables

Figure 1 – Innovative use of health data	15
Figure 2 - Root causes of major problems	18
Figure 3 - Use cases and domains	19
Figure 4 – Principle ratings before FGMs	20
Figure 5 – Principle ratings after FGMs	20
Figure 6 – Principle's mapping	22
Figure 7 - Accelerating and enabling relations between principles	38
Figure 8 – Approach to prepare guidance	<u>39</u>

Table 1 - Overview of principle-specific actions for guidance       12
--





# **Executive summary**

One of eHAction's main objectives is to improve the knowledge base for health and healthcare policy, envisioning development of new ways for data use and reuse across the healthcare sector. Tasks related to this objective are included in Work Package 5 (WP5): 'Innovative use of health data'.

This document (Deliverable 5.3 (D5.3)) is the proposal for the eHealth Network – Guidance for the implementation of common principles for practical governance of big data with a special focus on data to be used (and the implementation of data access and use) in public health, research and quality assurance in healthcare on a European scale.

The proposal provides information about the background, objectives, planned structure, working definitions and methods of D5.3, and contains suggestions on common principles for practical governance of big data.

According to the previous findings of WP5 deliverables D5.1 and D5.2,<sup>1</sup> several challenges and obstacles were identified: confidentiality issues, legal uncertainties, lack of trust, gaps in the legal environment, the level of digital literacy, lack of business models, interoperability issues, safety of health data, cybersecurity, intellectual property issues, unsolved issues of sharing benefits among stakeholders, and scarcity of resources including human resources (e.g. the number of data scientists). These hurdles will need to be solved or dealt with before the aim to create a European Health Data Space can be achieved.

WP5 has been looking for principles to mitigate or eliminate the root causes (and through them major problems) hampering Member States/countries from implementing measures to foster innovative use and reuse of health data in key use cases and domains.

Principles were drafted, based on literature, with three main pillars: The Study on Big Data in Public Health, Telemedicine and Healthcare (EU Big Data Study),<sup>2</sup> the OECD Ministerial Statement from 17 January 2017<sup>3</sup> and previous results from Work Package 5, namely deliverables D5.1 and D5.2. In the second phase of the work, the draft principles were discussed with actors and stakeholders in a two-step process: a survey, followed by a workshop held online on 23-25 June 2020 containing six focus group meetings with different stakeholder groups.

Twelve common principles in three clusters (as priority areas) for practical governance were introduced in the Draft Deliverable D5.3 submitted to (and endorsed by) the 18th eHealth Network meeting in November 2020. The three priority areas are:

- Cluster 1: Increase trust in privacy protection and cybersecurity
- Cluster 2: Increase efficiency of using funds and financial resources
- Cluster 3: Foster common interpretation of legal and ethical rules

<sup>&</sup>lt;sup>1</sup> For further details see Appendix 1

<sup>&</sup>lt;sup>2</sup> <u>https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata\_report\_en.pdf</u>

<sup>&</sup>lt;sup>3</sup> https://www.oecd.org/health/ministerial/ministerial-statement-2017.pdf





Co-funded by the European Union's Health Programme

(2014-2020)

The guidance for the implementation of common principles addressing the priorities contains twenty actions developed to assist and facilitate mitigation of challenges in handling big data in health.

The adoption of this policy proposal at the level of the eHealth Network would encourage Member States, stakeholders and the Commission that the recommendations within this document should be considered for implementation where relevant, in line with European and national strategies and priorities. Member States, stakeholders and the Commission should critically assess which policy actions are most needed in their context.

# **Overview about Principles and Actions**

To suggest principles that can eliminate or mitigate root causes of challenges for data reuse, WP5 formed the following three clusters (representing priority areas) to group the 12 principles, each related to one root cause:





Cluster 1: Increase trust in privacy protection and cybersecurity

- Principle 1 Work with anonymised or pseudonymised data
- Principle 2 Develop digital health literacy
- Principle 3 Disseminate data reuse purpose well
- Principle 4 Make a great effort to secure and protect data

- Cluster 2: Increase efficiency of using funds and financial resources

- Principle 5 Use existing technical standards
- Principle 6 Involve stakeholders in co-creation
- Principle 7 Consider current and future needs
- Principle 8 Foster data reuse
- Cluster 3: Foster common interpretation of legal and ethical rules
  - Principle 9 Adapt model to the settings
  - Principle 10 Create code of conduct
  - Principle 11 Speak the same language
  - Principle 12 Aim for win-win of all stakeholders

These principles were also discussed with members of the eHAction and selected stakeholders and the European Commission Services.

Finally, altogether 20 actions (see table 1 below) were developed to give guidance to the EU Member States and stakeholders likewise on how to potentially implement the 12 principles.

## How did we get to the Guidance?

In a first step 30 actions were derived by the task leader from the ideas, suggestions and conclusions of the Focus Group meetings organised by T5.3 team in June 2020. The content of these suggestions were overviewed and restructured into a final set of 20 actions by T5.3 contributors. This set contains at least one guidance for each principle. Generally, actions belong to a specific principle, excluding Actions 19 and 20, which provide suggestions for assisting implementation of Principles 11 and 12. In addition two general suggestions for action have been dedicated to all principles:

- Go on preparing focused actions, e.g. Coordination and Support Action (CSA), Joint Action (JA) or procure services, as well as national and/or regional actions (e.g. open calls or predefined projects), to assist the implementation of specific actions (e.g. for the identification of use cases, elaboration of protocols and standards, etc.).
- Develop the regulatory framework for secondary use of data through guidelines and plans to guide action on governance models, data quality and legal compliance requisites.



D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

# Table 1 - Overview of principle-specific actions for guidance

Priority area	Principle		Guidance		
(Cluster)	Name	Short description	Principle-specific actions		
Cluster 1: Increase trust in privacy protection	Principle 1 – Work with anonymised or pseudonymised data	Work preferably with anonymised or pseudonymised data for the purpose of secondary use of data.	Action 1 - Elaborate protocols and standards for anonymisation and pseudonymisation and prepare stakeholders for adoption Action 2 – Encourage Members states to develop anonymised and pseudonymised data repositories for the EHDS		
cybersecurity	Principle 2 – Develop digital health literacy	Make efforts to improve digital health literacy in the population and healthcare providers.	Action 3 - Adapt curricula to include focus on digital health Action 4 - Monitor and measure Digital Health Literacy (DHL) level		
	Principle 3 – Disseminate data reuse purpose well	Communicate in a socially acceptable and timely manner the purpose of reusing data.	Action 5 – Define and communicate the value of data reuse for practical care benefits		
	Principle 4 – Make a great effort to secure and protect data	Make great efforts for and invest in data protection and cybersecurity.	Action 6 – Use a data governance approach Action 7 – Use FAIRified data models and tailored solutions		
Cluster 2: Increase efficiency of using	Principle 5 – Use existing technical standards	Choose appropriate existing technical standards instead of development of new ones from scratch.	Action 8 – Consider the use of existing standards and data repositories Action 9 – Prepare common requirements for 'secure processing environment'		
funds and financial resources	Principle 6 – Involve stakeholders in co- creation	Identify and involve stakeholders in the secondary use of health data, big data and value-based solutions.	Action 10 – Encourage national authorities and beneficiaries to identify and use good practices for stakeholders involvement Action 11 – Develop action plans to improve the digital health ecosystem		
	Principle 7 – Consider current and future needs	Learn and systematically integrate current, emerging and future needs of both, decision makers and stakeholders.	Action 12 - Engage key stakeholders in needs assessment		



D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

	Principle 8 – Foster data reuse	Find means to make existing data sources visible and preferably use open source data models to allow for reuse.	Action 13 – Provide clear view about the definition and features of 'data' Action 14 – Provide lists of free high-value datasets and assist internalisation of externalities Action 15 – Identify data quality standards for a 'time-and-space-changing health data space'		
<b>Cluster 3:</b> Foster common interpretation of legal and ethical rules	Principle 9 – Adapt model to the settings	Treat health, health-related and non- health data with a special use case / business model for data management & data governance.	Action 16 – Identify different features of environment-specific ethical and behavioural norms		
	Principle 10 – Create code of conduct	Create international terms of data ownership, sharing, exchange, access and control to mitigate risks regarding sensitivity and privacy	Action 17 - Analyse legal basis of Codes of Conduct under GDPR and demonstrate added-value Action 18 - Define scope and concept of CoC(s)		
	Principle 11 – Speak the same language	Share common vocabulary and meta-data to avoid misunderstandings and lead to exploit the value of data.	Action 19 – Set up co-creation teams and identify people who can become		
	Principle 12 – Aim for win-win of all stakeholders	Measure, understand and communicate why stakeholders of data sharing and access are important to each other, as well as their value for them.	'translators' Action 20 – Develop common vocabulary for data sharing and access		





D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

# Introduction

eHAction is a Joint Action that is supporting the eHealth Network (eHN) on specific objectives related to the priorities of the current Multiannual Work Programme of eHN (MWP 2018-2021). The main eHAction work is based on establishing interoperability for enhancing safe cross-border exchange of health data between health professionals and healthcare providers. This kind of initiative can generate access for patients to their health data, wherever they are located. It is of utmost importance to be able to sustain a mechanism after the present Joint Action, as well as after the end of the Connecting Europe Facility (CEF) instrument. Further co-funding mechanisms could be designed and proposed, including those regarding modernisation of digital infrastructures (at European, national or regional level) to enable and maintain the cross-border exchange of health data, as well as the access of patients and professionals to this data for primary and secondary use.

One of the eHAction's main objectives is to improve the knowledge base for health and healthcare policy, visualising development of new ways in the usage of data across the healthcare sector. Tasks related to this objective are included in Work Package 5 (WP5): 'Innovative use of health data'. WP5 focuses on priority areas B.1, B.2 and B.3 of MWP 2018-2021. Secondary use of data and big data can provide value for research, teaching, managing and planning healthcare systems. It can also be a great opportunity for the development of personalised medicine, the improvement of the effectiveness of medicines, efficiency of health systems, and continuity of care. It is also a much-needed opportunity for improvement of traditional public health surveillance systems, prevention and control strategies. The COVID-19 pandemic stressed the need to use data to improve early warning and response systems, adding precision and timeliness to disease surveillance. WP5 faces the problem of lack of awareness of these potential benefits and the need to share expertise.

This proposal provides information about the background, objectives, planned structure, working definitions and methods of Deliverable 5.3, and contains common principles for practical governance of big data. The final milestone of Work Package 5, 'Discussion Paper for the eHN on the implementation of common principles for practical governance of big data with a special focus on data to be used in public health, research and quality assurance in healthcare on a European scale with guidance on implementation of data access and use (Deliverable D5.3)', is hereby submitted as the final report, containing a proposal for the guidance for implementation of the principles as well.

The document builds upon:

- Deliverable D5.2 'Report on identified cross-border use cases, including assessment of pros & cons of stakeholders, and practical solutions with potential for European scale benefits', which was delivered and adopted during the 16<sup>th</sup> eHealth Network meeting in November 2019;
- Deliverable D5.1 'Report on policy-level actions on innovative use of big data in health', which was adopted at the 17<sup>th</sup> eHealth Network meeting in June 2020.





# Background

WP5 had its kick-off meeting in Brussels in October 2018, met stakeholders at DIBSS Conference in Dubrovnik in May 2019 and held a workshop in Prague in September 2019 with the participation of Member State representatives, where experts tested and validated a canvas tool intending to capture the rationale behind the lack of use of already collected health data for better health outcomes. Using this canvas tool, named 'the stakeholder value chain analysis framework for data conversion' (in short: Data Conversion Framework or DCF tool) WP5 mapped privacy aspects, as well as identified obstacles that prevent Member State/country policies from being replicable either in other Member States/countries or on an EU level. The workshop contributed to finalising the DCF tool to identify key stakeholders, their needs, goals, offers/evaluation, behaviour and attitude, as well as interdependencies, obstacles and possible interventions and expected results.

To make a smooth transition from tasks T5.1 and T5.2 to task T5.3, WP5 joined the workshop on health data governance for secondary use, named 'Towards the European Health Data Space - National Strategies for Secondary Use of Data in the Context of National and EU Digital Health Networks', organised by the co-leader of WP8 in Lisbon in January 2020. From 23<sup>rd</sup> until 25<sup>th</sup> of June six focus group meetings (FGMs) on common governance principles for the reuse of health data were held by the T5.3 team as virtual workshops.

Furthermore, the opportunity to remove limits to scale, scope and learning opens the door to a fairly unlimited set of use cases for innovative use of health data and big data too. This trend has been highlighting the importance of innovative use of health data, since the 2018-2021 Multiannual Work Programme (MWP) of the eHealth Network was adopted. In parallel, the European Commission has identified the need and importance of EU level data spaces and repositories, as potential enablers and catalysts of innovative use of health data.

# Working definitions

Working definitions were elaborated by WP5 members at the kick-off meeting in order to lay down statements of the meaning of certain phrases which have been explained in various ways in the literature. These working definitions were introduced to the eHealth Network in November 2018, and are considered as continuously evolving definitions based on the evolutionary and constantly changing nature of the field. Definitions reflect the value-based approach followed in WP5, where value refers to satisfaction of a specific need and replicability at an economical cost. When elaborating deliverable D5.3, WP5 contributors added definitions for some new terms to the list.

The full set of WP5 working definitions are available in Appendix 2.

## Vision, purpose and objectives

**Work Package 5 vision** is to develop approaches for innovative usage of data across the healthcare sector and the knowledge base for healthcare policy and related purposes.



D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

WP5 looks at data as key driver of disruptive innovation in the economic, societal and social systems. Therefore, WP5 is committed to make an impact on the use of health data, as well as assisting data-driven innovation leading to patient-centred health systems, evidence-based health policies and decision-making.

WP5 is also committed to assist decision makers understanding and exploiting main advantages of the innovative use of health data.



Innovative use of health data, regardless of whether the use is primary or secondary, fosters innovation in the field of public health interventions, prevention strategies and health system management, as well as in the organisation and provision of health services and medical care, including health promotion and disease prevention interventions.

It has the potential to improve public health outcomes, enhance the quality of care to patients and respond to unmet needs, and also to foster the competitiveness of stakeholders and to improve the cost-efficiency and sustainability of health services and medical care. Innovative use of health data will also uncover new ways for surveillance and control of global health threats and emerging infectious diseases, upscaling health systems surveillance and response capacity.

**The purpose** of this document is to assist in bringing closer altering approaches, adherence and empowerment of decision makers. There is much uncertainty on how to go forward on benefiting from reuse varying datasets and apply artificial intelligence on the practical level.

On the policy level it is important to settle common approach and strengthen the awareness, adherence or empowerment to identify the strengths, opportunities, weaknesses and threats





of the reuse of data, in order to be able to realise possibilities to deliver societal and economic benefits of innovative use of data and big data.

It is also important to raise awareness not only of the value big data offers to improve health, but also of the several operational barriers to reuse data to realise its potential. Data collections, extractions, quality, semantics, interoperability – to name a few – are not a given in the big data era. There is a need to raise health policy makers' awareness of the data related services needed, as digital transformation changes current data assets.

**The overall objective** of WP5 is to support the application of good practices in Member States and provide guidance at European Union level on handling big data in health within the existing EU regulatory framework, on secondary use of personal health data, and consequently to ease the uptake of innovative usage of data across the healthcare sector for the benefits of society, individuals and performance of Member State health systems.

Task 5.3 'Towards an attempt to define common principles for practical governance' foresees to deliver document on:

 Guidance on practical governance for eHN and Member States, providing a framework for implementation of common principles for practical governance of big data, including privacy protection and security, aiming at improving health data transferability across borders, with a special focus on data to be used in public health, research and quality assurance in healthcare on a European scale.

The guidance includes a set of actions to the eHN, Commission and Member States/countries on how to implement principles elaborated to address issues of increasing trust and utilising resources more efficiently, as well as fostering common interpretation of legal and ethical rules.

 Recommendations to the eHN, Commission and Member States/countries for next steps related to the priority areas addressing the three clusters of the principles. Recommendations are taking into account ongoing activities within the EU in order to target most pressing issues. Recommendations are plotted accordingly the results of consultation with a broad list of stakeholders.

# Audience and Scope

Data governance entails defining, implementing and monitoring strategies, policies, measures and shared decision-making over the ability, skills and actions of the management and use of data assets.<sup>4</sup>

Therefore, governance of big data in health is the process of managing the availability, usability, integrity and security of the consolidated data from existing fragmented data sources for understanding, forecasting and improving health and health system status, needs and

eHAction - Joint Action supporting the eHealth Network - www.ehaction.eu

<sup>&</sup>lt;sup>4</sup> <u>https://ec.europa.eu/info/sites/info/files/summary-data-governance-data-policies\_en.pdf</u>



performance. It is based on standards and policies that regulate secondary use of data, including respect for privacy, quality and integrity of data, and access to data.

These principles and guidance for their implementation are addressed to national and Community legislators and policy makers. They have a wide focus on data to be used (and the implementation of data access and use) in public health, research and quality assurance in healthcare on a European scale.

The scope of this document entails Governance of data reuse for:

- assisting academic and sponsored research
- providing tools and evidence in public health
- strengthening quality assurance in healthcare:
  - supporting institution management and system governance
  - improving patient care by assisted decision making, diagnostic analytics, robots (software and equipment), smart devices, telehealth, mHealth and remote care
  - reducing (hospital) re-admission rate
  - o forecasting patient demand, attitude, behaviour and need
  - o predicting outbreaks and spread of the epidemics
  - o optimising administration, finance, procurement, inventory, infrastructural capacities, workforce management and investment
  - o etc.

## Methodological Approach

Principles were drafted, based on three main pillars: The Study on Big Data in Public Health, Telemedicine and Healthcare (EU Big Data Study),<sup>5</sup> the OECD Ministerial Statement from 17 January 2017<sup>6</sup> and the results from Work Package 5, namely the deliverable D5.1 and D5.2.<sup>7</sup> This was followed by a large-scale discussion with actors and stakeholders in the field (e.g. public health, legal experts, digital health, technical persons, EU-wide networks and EU services). This was conducted in a two-step process: 1) A survey with a first prioritisation was conducted, and 2) a series of six workshops took place, with altogether attendees clustered as follows: 50% from public bodies, 22% from European Commission and Agencies, 20% from international organisations, and 8% from other sectors (academia, R&D). WP5 has been looking for principles to mitigate or eliminate the root causes (and through them major problems) hampering Member States/countries to implement measures to foster innovative use and reuse of health data in key use cases and domains.

<sup>&</sup>lt;sup>5</sup> <u>https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata\_report\_en.pdf</u>

<sup>&</sup>lt;sup>6</sup> https://www.oecd.org/health/ministerial/ministerial-statement-2017.pdf

<sup>&</sup>lt;sup>7</sup> LINK to D5.1 and D5.2 (ehaction.eu)





#### Problems

Major issues holding the EU back from realising its potential in the data economy were introduced in the latest EU Data Strategy.<sup>8</sup> These problems have their impact on healthcare and health economy, too. According to the assessment made by WP5, the root causes beyond the obstacles faced in the governance of secondary use of health data, identified by the DCF tool, have been considerably affecting the problems listed in the EU Data Strategy (see Figure 2).

	Root causes				
<u>Problems</u>	Lack of trust	Lack of resources	Legal uncertainties		
Not enough data available for reuse	XX	Х	Х		
No real user empowerment	ХХ				
Lack of data processing and storage solutions	Х	XX			
Absence of comprehensive data governance approaches (legal & technical barriers)		х	ХХ		
Skills shortage and low data literacy	х	XX	х		
Fragmentation of the single market		Х	ХХ		
Healthcare systems are separated, what is not the case in other systems	X		ХХ		

Figure 2 - Root causes of major problems

#### Domains and use cases

Successful implementation of the principles to develop and sustain a friendly environment for secondary use of health data and big data depends on the recognition of the specificities of the domains and use cases where they will be used. Therefore, development of guidance for implementation is based on having these specificities into consideration.

The term 'domain' refers to one or several data sets that covers particular topics.<sup>9</sup> Principles recommended by WP5 affect practical governance of big data in the following topics (domains):

- 1. Policy, Governance and Management (PGM)
- 2. Education and Training (ET)
- 3. Legal, Ethical and Regulatory Issues (LER)
- 4. Interoperability and Technical Standards (ITS)

<sup>&</sup>lt;sup>8</sup> A European strategy for data, Brussels, 19.2.2020 COM(2020) 66 final (<u>https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN</u>))

<sup>&</sup>lt;sup>9</sup> https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1596807664480&uri=CELEX:32019R1700





#### 5. Dissemination and Stakeholder Involvement (DS)

WP5 identified 14 use cases of secondary use of health data and big data. Each of these use cases can be related to two or more domains. WP5 identified following relations between use cases and identified domains (see Figure 3).

USE CASES	PGM	ET	LER	ITS	DS
<ol> <li>Forecasting patient demand, attitude, behaviour and need</li> </ol>	х				х
2. Hospital re-admission reduction	х	х	x		
3. Public health, surveillance and prediction of health threats and diseases		х	х		х
4. Clinical decision support, assisted diagnostic analytics and robots		х	х	х	
5. Personalised medicine	х		x		х
6. Robots, smart devices, telehealth, mHealth and remote care	х	x	x	х	х
7. Administration, finance, inventory and investment optimisation	Х		х		
8. Prediction of procurement demand and workforce needs	х	х	x		
9. Big data assisting legislation, governance and regulatory activities (e.g. disease epidemiology or product related)	х		x	x	x
10. Hidden citizen behaviour patterns analysis		х	х		х
11. Insight and evidence for policy-making and investments	Х		x		
12. Prediction of health and health related needs, expenditure and costs		х	х	х	
13. Academic, clinical and translational research and sponsored research			х		х
14. Provision of tools and evidence			Х	Х	

Figure 3 - Use cases and domains





#### **Results from the Focus Group Meetings for identifying principles**

From 23<sup>rd</sup> to 25<sup>th</sup> June 2020, six focus group meetings (FGMs) on common governance principles for the reuse of health data were held by the T5.3 team as virtual workshops. The three days were organised for three different target groups/audiences:

DAY 1. Technology and Interoperability, Tuesday, 23 June 2020

- DAY 2. Public Health and Research, Wednesday, 24 June 2020
- DAY 3. Policy and Regulation, Thursday, 25 June 2020

A total of 101 participants from 26 European countries attended the 6 events.

FGMs helped WP5 to find replies to questions regarding policy level unmet needs for common principals and practical guidance why and how to foster innovative use of health data, big data and artificial intelligence (AI), as well as how to implement available relevant policy recommendations. In addition, 12 draft principles, developed by WP5 prior to the FGMs, were discussed allowing the WP team to further develop and reshape them in order to prepare their submission to the eHealth Network. Principles were also rated and prioritised by participants putting them in order of importance.

Participants could rate the preliminary principles during the registration period, and the result of the rating was introduced at the PGMs (see Figure 4).

Participants were asked to prioritise the further developed and reshaped principles at the wrapup of the FGMs. The aggregated result of the voting is introduced in Figure 5.



The final result of the voting highlights that implementation of five principles (reflecting most pressing issues) need actions first: Principle 1 in priority area (cluster) 1, Principle 6 and 8 in priority area (cluster) 2 and Principle 10 and 11 in priority area (cluster) 3.





#### Conclusions from the FGMs on the design of the principles

- Principles shall be designed for policy makers Policy/macro level + Best practices
- Principles shall support policy makers to offer win-win data reuse governance model for their key stakeholders Interoperable and aligned with stakeholders' principles and programmes (e.g. Data Saves Life)
- Principles shall be easy to start and implement Low hanging fruit
- Principles shall be effective Eliminate or mitigate root causes of challenges
- Principles shall be efficient Effective + no side effects + minimised costs

# **Principles and Actions for guidance**

## Clustering of principles and principle's mapping

Following the idea to recommend principles that eliminate or mitigate root causes of challenges, WP5 formed 3 clusters (as priority areas) to group the principles, each related to one root cause.

Figure 6 shows the allocation of principles in the map of the following 3 clusters:

- Cluster 1: Increase trust in privacy protection and cybersecurity
- Cluster 2: Increase efficiency of using funds and financial resources
- Cluster 3: Foster common interpretation of legal and ethical rules

Some principles can be related not only to one root cause, but several.

When grouping principles into clusters, each principle was linked to the root cause that could be most influenced by that principle.

E.g. Principle 2 (Develop digital health literacy) is most linked to the goal to increase trust in privacy protection and cybersecurity, but it also contributes to increase efficiency of using funds and financial resources on the one hand, and it helps to foster common interpretation of legal and ethical rules on the other hand.

Principle 5 (Use existing technical standards) recommends develop new standards if the use case requires, otherwise use available ones to save time or costs and make use of global or European interoperability.

In this way, if one deploys existing standards, your system would be more efficient. In addition, you can communicate that you chose standards that have proven to strengthen privacy, data security and safety. Thus implementation of Principle 5 can contribute to increase trust in privacy protection and cybersecurity as well.

Another aspect of the grouping was whether the elimination of a root cause by the principle also contributes to mitigating the effect of another one. E.g. Principle 3 (Disseminate data reuse





purpose well) could contribute to increase efficiency of using funds and financial resources if the level of trust in privacy protection and cybersecurity had been raised.

Two or three or even more principles, from different clusters, can target the same goal (from the point of view of their clusters)! For example, P3 (Disseminate data reuse purpose well) and P12 (Aim for win-win of all stakeholders) promote the personalised and socially important utilisation of shared, accessed, processed data. P3 is for trust, while P12 is for mitigating legal uncertainties. P7 (Consider current and future needs) is also related to these ones in terms of optimal utilisation of resources. Even P6 (Involve stakeholders in co-creation) can be connected to them.



Figure 6 – Principle's mapping



# **Cluster 1: Increase trust in privacy protection and cybersecurity**

The principles presented in Cluster 1 target the 'lack of trust' that was introduced in D5.1. These principles intend to increase trust in privacy protection and cybersecurity.

#### Principle 1 – Work with anonymised or pseudonymised data



Work preferably with anonymised or pseudonymised data for the purpose of secondary use of data.

Governance of big data shall support measures to use either anonymised or pseudonymised data. It needs, however, to be analysed for each specific use case separately and adapted to the particular national context and legal framework.

Cross-sectoral and cross-country data is needed for research purposes (a typical example is cooperation within European Reference Networks). Better protection must be ensured for individuals in terms of data processed about them and so, less risks to the right to data protection and privacy when re-use of personal data occurs, in compliance with Art. 9 and Art. 89 of GDPR. It is recommended to preferably work with anonymised or pseudonymised data for the purpose of secondary use of data.<sup>10</sup>

The decision to use either anonymised or pseudonymised data needs to be analysed for each specific use case separately, and depends on the field of data reuse or the particular national situation and legal framework.

In the European Parliament resolution<sup>11</sup> of 14<sup>th</sup> March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement, the Parliament under point 10: Underlines that the intrinsic purpose of big data should be to achieve comparable correlations with as few personal data as possible; [and] stresses, in this regard, that science, business and public communities should focus on research and innovation in the area of anonymisation. Under point 11 in the same resolution, the European Parliament 'Recognises that the application of pseudonymisation, anonymisation or encryption to personal data can reduce the risks to the data subjects concerned when personal data are used in big data applications; further highlights the advantages of pseudonymisation provided for by the GDPR as an appropriate safeguard; recalls that anonymisation is an irreversible process by which personal data can no longer be used alone to identify or single out a natural person; takes the view that contractual obligations should ensure that anonymised data will not be reidentified using additional correlations by combining different data sources; calls on the private and the public sector and other actors involved in the analysis of big data to regularly review

<sup>&</sup>lt;sup>10</sup> See our working definitions section below for a definition of pseudonymisation and the difference to anonymisation <sup>11</sup> European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI)) <u>https://eur-lex.europa.eu/legal-</u> <u>content/EN/TXT/?qid=1596809628647&uri=CELEX:52017IP0076</u>





such risks in the light of new technologies and to document the appropriateness of the measures adopted; calls on the Commission, the European Data Protection Board and other independent supervisory authorities to prepare guidelines on how to properly anonymise data in order to avoid future abuses of these measures and to monitor practices'.

While data anonymisation provides the highest degree of data protection for the data subjects, anonymous data will not be useful in many big data applications. Hence the relevance of pseudonymisation as a strategy towards data minimisation and data protection in the regulatory context of GDPR. In its European Strategy for Data, the European Commission commits to developing standards, tools and best practices on pseudonymisation.<sup>12</sup>

#### Principle 2 – Develop digital health literacy



In the digital era, it is expected that being helped by technology, patients will eventually become their physicians' partners in medical decision-making.<sup>13</sup> Decision-making in public health, research and improving quality of healthcare provision and systems is expected to be more assisted by artificial intelligence, and the role of digital technology is continuously and rapidly growing. Thus, it is obvious that stakeholders which generate, capture, share, access, process and use data have to be partners in developing and using digital health technology and services. Therefore, it is inevitable that digital health literacy (DHL) of all stakeholders, especially health professionals, shall be improved by planned actions.

DHL helps people to assimilate the information they obtain regarding new technology and protocols, and it enables them to put up questions, reply them and understand the answers. In this way, DHL is a key tool to increase trust and action must be taken to develop DHL skills.

<sup>&</sup>lt;sup>12</sup> COM(2020) 66 final

<sup>&</sup>lt;sup>13</sup> Mesko, B., Drobni, Z., Benyei, E., Gergely, B., Gyorffy, Z.: Digital health is a cultural transformation of traditional healthcare. mHealth 3, 38 (2017). <u>https://doi.org/10.21037/mhealth.2017.08.07</u>





#### Principle 3 – Disseminate data reuse purpose well



Communicate in a socially acceptable and timely manner the purpose of reusing data.

Trust in use and reuse of data can be easily destroyed or strengthened by communication to different stakeholders. The way of communicating purposes and results to data controllers is a key aspect. However, it may be difficult to articulate, identify or predict the role and importance of a specific piece of information or even a data subject (a concrete person), thus the purpose of asking consent for controlling one's data can face difficulties. Therefore, to make it obvious that the data subject has consented to the particular processing, data controllers have to develop or use proper methods to disseminate and communicate their general and specific scientific, social or business goals. The more difficult it is to characterise and specify a purpose in advance, the more important it is to find and disseminate socially acceptable purposes and make it clear that these purposes are serving 'public good'. WP5 accepts the economic definition of this term: 'Public good, in economics, a product or service that is non-excludable and nondepletable (or 'non-rivalrous').'<sup>14</sup> In this context, data as 'public good' can be: 'Using data for public benefits by applying tools and techniques that are generally used in business applications'.<sup>15</sup> In addition we also have to take into consideration that a 'public good' has two characteristics: 'non-rivalry' and 'non-excludability'. This means that it must be checked if the amount available for others could be reduced when data is consumed, and/or it can be possible to provide data without it being possible for others to enjoy.<sup>16</sup>

In order to be able to develop successful dissemination activities to communicate socially acceptable purposes providing public benefits, data controllers, as well as policy makers who intend to improve innovative use of health data and big data, are recommended to identify:

- who the key stakeholders, having influence over the decision on using results of data processing, are. Either because they're paying for it, or because they can persuade whoever is paying;
- who can end up experiencing the product or service developed or improved by using or re-using data;
- whose life would improve because of the product or service developed or improved by using or re-using data.

These three groups of stakeholders are the customers, end-users and beneficiaries for non-forprofit organisations.<sup>17</sup> Creation of public good, e.g. using data for public benefits, may not be

<sup>&</sup>lt;sup>14</sup> https://www.britannica.com/topic/public-good-economics

<sup>&</sup>lt;sup>15</sup> https://towardsdatascience.com/data-for-public-good-1414cbc99335

<sup>&</sup>lt;sup>16</sup> <u>https://www.economicshelp.org/micro-economic-essays/marketfailure/public-goods/</u>

<sup>&</sup>lt;sup>17</sup> https://isaacjeffries.com/blog/2016/3/1/customers-end-users-and-beneficiaries





paid directly but financed (e.g. from taxes or donations), therefore, the customer is not a payer but a financier who also decides or influences decisions on the use of money.

Communication of the public good to the three groups requires development and use of different strategies, ways and content.

Strategies can partly build on data altruism in order to facilitate decisions on which data generated and held by data subjects can be used, how and by whom can these data be used, and what the public good purposes are.

Strategies can also build on the identification and introduction of the value that the public good offers to the data subject personally, therefore personalisation of the public good is also important.

It also has to be ensured that the way how data will be processed is transparently communicated. This applies for safety, security and privacy solutions, as well as the use of interoperability standards and the efficiency issues.

Communication strategies and implementation plans shall be elaborated in line with developing, measuring and updating digital health literacy of these groups, as well as the number and the level of acceptance of codes of conduct in different domains and use cases.

#### Principle 4 – Make a great effort to secure and protect data



This principle builds on the so called 'comprehensive data governance model' where system/organisation-wide data governance aggregates historic and the newest data together. It requires full engagement and adherence at leadership level, and clear management of data control and access rights.

Data security and data protection is a multidimensional model. It is a functional space where it is necessary to ensure the physical protection, administrative protection, technical protection of the data, as well as compliance with legal regulations.

There are general rules to:

• ensure the confidentiality, integrity, and availability of all data they create, receive, maintain or transmit;





- identify and protect against reasonably-anticipated threats to the security or integrity of the information;
- protect against reasonably-anticipated impermissible uses or disclosures;
- ensure compliance by their workforce.

Responsible data governance approaches must ensure:

- Integrity All entities involved in data governance must have integrity in their communication with each other.
- Transparency Data governance processes need to be transparent in order to make it clear when and how data decisions and controls are implemented in the processes.
- Auditability Data decisions and processes must be auditable, with documentation of compliance-based and operational control requirements
- Accountability One must define who is accountable for data-related decisions and controls, and for data management activities.
- Breaks and counterweights One must have a data-governance-based data lifecycle, and one needs to define data source and data cryptography protection, and who makes data, collects data, uses data and manages data.
- Standardisation Data governance supports quantum resistance standardisation of data.
- Change management Data governance supports proactive and reactive change management activities.
- Risk management Data governance requires security rules, what covers administrative, technical and physical dimensions, as well as the data owner and data controller's activities. Artificial Intelligence should be mentioned as a risk factor, because we do not know its limits and effects. It can be managed through understanding the 'transparency paradox'<sup>18</sup> (transparency makes methods vulnerable while it can help mitigate issues of fairness, trust and discrimination).
- Resilience solution One needs data governance to be able to manage disaster and force majeure situations.
- Confidentiality Data governance must ensure data security, to serve privacy, protect individuals' medical records and other personal health information.

<sup>&</sup>lt;sup>18</sup> Andrew Burt: The AI Transparency Paradox, published on hbr.org, December 13, 2019



# **Cluster 2: Increase efficiency of using funds and financial resources**

#### Principle 5 – Use existing technical standards



The intention behind this principle is to foster the use and harmonisation of existing technical standards and to avoid duplication of efforts by inventing new standards when it is not justified. Interoperability is key when it comes to big data and secondary use of health data. In particular, relying on the definitions of the 'Refined eHealth European Interoperability Framework', we refer to the interoperability level of *information*, i.e. the data models, their functional description and links to terminologies.

The reuse of existing technical standards and exploitation of means already available like terminologies and data models allows an efficient use of resources and contributes to further increase interoperability and the efficient use of resources.

#### Principle 6 – Involve stakeholders in co-creation



Identify and involve stakeholders in the secondary use of health data, big data and value based solutions

Efficiency of using funds and financial resources for the innovative use of health data and big data in health is closely related to who is taking part in this process and how.

On the one hand, acceptance of an innovation depends on how key stakeholders evaluate it. Therefore, focusing on the needs and expectations of only one stakeholder is likely to result in failure and waste of time and money.

On the other hand, innovative use of health data and big data in health means various data sources that must be reliable, safe, findable, accessible, etc. This requires co-operation with varying data subjects and controllers.

Thus, identification of key stakeholders, as well as their needs and expectations, is essential. WP5, using the Data Conversion Framework tool, identified the following key groups of





stakeholders for innovators and researchers: 1) patients and citizens; 2) care providers, including professionals, organisations and authorities; 3) payers. In this context, innovators and researchers are those data controllers who use or reuse health data in order to create something new or to renew or develop existing things, services, solutions, protocols or systems.

Asking them about their needs is the first step, however, success needs continuous cooperation along the whole process of translating an idea or invention into goods or services that create value, or for which customers will pay. Innovators and researchers shall assess how stakeholders feel about key features of existing and innovated or invented healthcare services and protocols or solutions, devices or infrastructure, etc., and shall ask which of the following three categories each feature falls into:<sup>19</sup>

- **Non-negotiables:** Performance features that make an offer minimally acceptable.
- **Differentiators:** Features that positively distinguish an offering from the competition.
- **Dissatisfiers:** Stakeholders are not happy about these attributes but may be willing to endure them for a time if compensating differentiators exist.

Innovators and researchers are recommended to involve their stakeholders in looking for a differentiator to add at each link in the consumption chain (from suppliers or resource providers to the end-users) and try to find at least one differentiator for each of the stakeholders. Removing a dissatisfier, which effectively creates differentiation from competitors, is also an option. Once they've come up with new differentiators to improve how to satisfy a given need, they have to assess how they might affect other stakeholders — particularly whether they might conflict with a dominant non-negotiable. Naturally this is work in progress which can be updated based on the changes or findings in all stages of the innovation process.

Involvement of all the key stakeholders, in most cases, faces the challenge that people who need to be involved in the co-creation process or data governance are really busy. To be successful, a data governance initiative requires time and input from subject matter experts, with a deep understanding of the data or the purpose and opportunities of data reuse, and key managers who can make operational changes happen. They will be the innovation scouts<sup>20</sup> or data stewards.<sup>21</sup>

Involvement of stakeholders in co-creation processes can be carried out anywhere, however, innovation hubs (places and tools for professional-user led co-creation) and living labs (centres for individual-user led innovation) offers more friendly environment for value creation.

Implementation of Principles 3, 6 and 12 can be carried out together. While Principles 3 and 12 assists to increase level of trust and eliminate misunderstandings around data sharing and access in a more effective way, Principle 6 can be a tool to use data to innovate something in a more efficient way.

<sup>&</sup>lt;sup>19</sup> How To Get Ecosystem Buy-In - A tool kit for assessing the way an innovation will affect each stakeholder, by Martin Ihrig and Ian C. MacMillan, Harvard Business Review, March-April 2017

<sup>&</sup>lt;sup>20</sup> <u>https://www.innovationagencynwc.nhs.uk/media/Images/Innovation%20Scout%20Programme%20Overview.pdf</u>

<sup>&</sup>lt;sup>21</sup> <u>https://www.ehidc.org/sites/default/files/resources/files/Data-Gov-in-Healthcare\_092016.pdf</u>





WP5 recommends the policy makers and implementing bodies of funding programmes in the Member States/countries and the European Union to assess which actions or measures could be used.

#### Principle 7 – Consider current and future needs



Learn and systematically integrate current, emerging and future needs of both, decision makers and stakeholders.

A long-term vision of big data use in health sector requires that a continuous process is established to develop big data services as a new asset for health and health services (policy making, management and care provision) based on scalable and sustainable governance. Particularly, in the health sector, data governance mechanisms should be identified to use data for the benefit of the society and to support different organisations in improving their decisions, while minimising societal risks, and be continuously reviewed to ensure adequacy.

There is a growing acknowledgement that governance is a key aspect for managing and getting value out of big data, to more effectively implement big data tools and to maximise the rewards for the organisational and institutional point of view. The framework for implementing data governance can have different components: policy and organisational, technological and standards related. In fact, due to the number of different health systems models and organisations, with their individual culture, numerous possibilities to implement data governance exists. However, any data governance framework to deal with big data in healthcare organisations, and in the health system as a whole, must ensure stakeholders have access to timely, consistent and accurate data resulting in quality data-driven decisions.

One of the main drivers for implementing data governance must be a proactive, rather than a reactive, approach to deal with ever-growing needs of data and the dynamic nature of healthcare, public health and research needs. As such, having an organisational structure within the healthcare system that aligns all stakeholders on value (including matters experts, managers, health professionals, IT and data specialists and citizens), through an efficient leadership and communication strategy, is a requirement for an organisation-wide approach and dedicated organisational initiatives to optimise data assets to deliver both sophisticated and practical information (NOESIS,<sup>22</sup> Novel Decision Support tool for Evaluating Strategic Big Data investments in Transport and Intelligent Mobility Services).

<sup>&</sup>lt;sup>22</sup> <u>https://cordis.europa.eu/project/id/769980</u>





Engaging stakeholders in a data-driven approach to health requires a trusted environment in the healthcare system to perform its functions effectively, based on data insights. Maintaining a trusted reputation amongst its many data providers, data recipients and stakeholders engagement is also a critical requirement to achieve a continuous governance process that is sensitive to the ever-growing needs of data for decision making in different settings and for different purposes. Data needs of citizens, health professionals, healthcare institutions and central agencies must be accounted for since the beginning through a customer-centred approach to secondary use of data.

As such, a common principle for practical implementation of (big) data governance is to consider the dynamic and evolving nature of the healthcare system. This can only be achieved by ensuring stakeholders play key leadership roles in the development of governance frameworks and are empowered to participate in the development of innovative data-driven technologies when they arise.

#### Principle 8 – Foster data reuse



Find means to make existing data sources visible and preferably use open source data models to allow for reuse.

It is also important to include data to support the decision-making process.

One must address the right decision makers, users and beneficiaries of the business model for open data and data sharing to foster data reuse.

The main purposes of reuse of health and healthcare data are to speed up medical innovation and improve citizens' health. Reusing data can provide important insights into results of innovation in technology, development of new tools and also new skills.

There are several aspects of reuse of health and healthcare data:

- focusing on strictly health and health related outcomes;
- economic aspects;
- social impacts; and
- the mix of all above aspects.

These considerations top the notion of sole access to information by citizens.

To improve the reuse of health and healthcare data, WP5 proposes the following actions (over and above the actions of primary use):

Reuse of health data should be incentivised by public entities on the regional and EU levels (most relevant fields: continuity of care delivery, personalised medicine, clinical research,





pharmacovigilance, quality improvement, infectious disease, bio-surveillance, financial analysis). Outcome-oriented incentives are preferable which should contribute to results (above direct healthcare-related benefits) in the field of science and informatics. Those can improve patient-clinician partnership and cross-sectoral collaboration (healthcare, information technology, manufacturing etc.). It should be a dynamic process with a continuous learning curve.

When setting up a framework for reusing health data, key – and relevant – stakeholders should be included and brought together in the process (i.e. public authorities, healthcare policymakers, professionals, pharmaceutical companies, research organisations, patient organisations, enterprises, etc.).

Reusing health data should facilitate actions in setting up research, quality and public health strategies. Benefits of reuse of data should be estimated if possible before taking action.

Business models can be constructed taking into consideration health data governance frameworks. Charges for the reuse of such data can be considered depending on the purpose of reuse (i.e. marginal costs for usage and dissemination). Externalities should be internalised in models if reasonable and possible.

Application of the Open Data Directive<sup>23</sup> should be considered in the field of health data. The Open Data Directive requires the adoption by the Commission a list of high-value datasets to be provided free of charge. These datasets, have a high commercial potential and can speed up the emergence of value-added EU-wide information products. They will also serve as key data sources for the development of Artificial Intelligence. Reuse of these datasets is associated with important benefits for the society and economy. They are subject to a separate set of rules ensuring their availability free of charge, in machine readable formats, provided via Application Programming Interfaces (APIs) and, where relevant, as bulk download.

Data quality should be continuously improved to get more usable health data for reuse. Therefore, WP5 recommends to assess the opportunity of adopting (the approach and solutions of) the data quality model provided by the ISO/IEC 25012 standard<sup>24,25</sup> which links data quality to the degree to which a set of data characteristics meets the requirements. It helps not only to disseminate data reuse purposes (recommended in Principle 3), but can be a proper tool to increase efficiency of using funds and financial resources for innovative use of health data and big data as well.

<sup>&</sup>lt;sup>23</sup> Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. <u>https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information</u>

<sup>&</sup>lt;sup>24</sup> <u>https://www.iso.org/standard/35736.html</u>

<sup>&</sup>lt;sup>25</sup> <u>https://iso25000.com/index.php/en/iso-25000-standards/iso-25012</u>





Data and big data quality is an essential topic for public health and research, healthcare management and governance and 'businesses, providing accurate information in order to make correct decisions accordingly.'<sup>26</sup>

It is also elementary and an important prerequisite for reuse to build a trusted digital health data sharing environment. Trustworthy and transparent use of health data is essential.

Security, legal and ethical aspects could be different from those applied at primary use due to:

- different stakeholders,
- different purposes and fields of application,
- different processes, and
- different space-time structures.<sup>27,28,29</sup>

It should be considered to launch a 'time-and-space-changing data space' for European healthcare, which would be able to satisfy artificial intelligence (AI) requirements to show changes and deviations in the space-time structures and would be able to predict future trends and assist decision-making related to varying purposes of different stakeholders. It is, however, especially important to protect and manage sensitive personal health data in accordance with the GDPR.

## **Cluster 3: Foster common interpretation of legal and ethical rules**

#### Principle 9 – Adapt model to the settings



*Treat health, health-related and non-health data with a special use case / business model for data management & data governance.* 

Shall health challenges in different environments be treated with special strategies and models for data reuse management or data governance? If you understand where people live, you can develop and implement data reuse governance and financing models and communication strategies which are adopted to the settings of people, and in this way you can develop schemes to foster common interpretation of legal and ethical rules that help you raise the level of trust and improve resource management as well.

<sup>&</sup>lt;sup>26</sup> <u>https://www.sciencedirect.com/science/article/pii/S1877050919305915</u>

<sup>&</sup>lt;sup>27</sup> https://link.springer.com/article/10.1186/s40537-018-0116-9

<sup>&</sup>lt;sup>28</sup> Space-time Data Science for a Speedy World, Harvey J. Miller, Published 2015

<sup>&</sup>lt;sup>29</sup> <u>https://www.publichealth.columbia.edu/research/population-health-methods/spatiotemporal-analysis</u>





Urban environments, for example, play an essential role in shaping human health and wellbeing, as 55% of the population worldwide live in urban areas. 97% of cities in low-and middleincome countries with more than 100,000 inhabitants do not meet air quality guidelines (49% in high-income countries).<sup>30</sup>

Rural areas often have the highest rates of premature death due to tobacco, nutrition and alcohol issues. Almost half the world's people live in rural and remote areas. The problem is that most health workers live and work in cities.<sup>31</sup>

Despite the significant advances that have been made at the EU level in terms of health inequality measurement, there is still a lack of comparable health-related data across countries and regions. Existing measures and indicators of health status and quality of life are considered inadequate to capture the holistic understanding of population health, with multiple determinants involved.<sup>32</sup>

Therefore, following the logic and findings of the PULSE project,<sup>33</sup> shifting public health from a surveillance-based system to an inclusive and collaborative system via citizen engagement through a citizen science approach using digital technology is strongly recommended. This requires creation of environment specific health and wellbeing datasets.

Data generated and held (1) in health systems, (2) in research activities and (3) by citizens shall be federated for secondary use purposes and, together with (4) cross-sector data about determinants of health and geographic features, shall be shared and could be accessed.

New investments shall be made in combined information sources deriving greater context. 'This is needed to more effectively and securely leverage patient data, recognise patterns faster, and manage the health of populations more effectively. (...) Population health management is fundamental to the transformation of healthcare delivery. For providers, the term translates to knowing what's going on with your patients and taking evidence-based, standardized and, where appropriate, automated actions to proactively achieve the best outcomes.'<sup>34</sup>

This requires use-case-specific decisions on following issues in policy and implementation levels:

- which data generated and held by the four sources above can be used;
- who can use the data and for what purposes;
- what the common interpretations of different legal and ethical rules can be;

 <sup>&</sup>lt;sup>30</sup> María Fernanda Cabrera Umpiérrez, PULSE Coordinator, UPM, <u>http://www.project-pulse.eu/events-2/pulse-final-workshop/</u>
 <sup>31</sup> WHO Global policy recommendations on Increasing access to health workers in remote and rural areas through improved retention (2010)

https://apps.who.int/iris/bitstream/handle/10665/44369/9789241564014\_eng.pdf;jsessionid=928CD4C30FC14287A105522A 5B93DED7?sequence=1

 <sup>&</sup>lt;sup>32</sup> <u>https://www.researchgate.net/publication/321361767</u> <u>Atlas\_of\_Population\_Health\_in\_European\_Union\_Regions</u>
 <sup>33</sup> See footnote 20

<sup>&</sup>lt;sup>34</sup> Femi Ladega and Gurdip Singh, Digital Revolution Enables Population Health Management <u>https://www.himssasiapac.org/sites/default/files/HIMSSAP\_Whitepapers\_CSCDigitalRevolutionEnablesPopulationHealthManagement.pdf</u>





- how to communicate the value of the use of artificial intelligence and machine learning solutions that can easily remove limits to scale, scope, and learning;<sup>35</sup>
- how to transform public health from a reactive to a predictive system focused on both risk and resilience by the digital population health model.

The most concerned use cases are<sup>36</sup> '1. Forecasting patient demand, attitude, behaviour and need' and '3. Public health, surveillance and prediction of health threats and diseases', as well as '10. Hidden citizen behaviour patterns analysis', '11. Insight and evidence for policy-making and investments' or '12. Prediction of health and health related needs, expenditure and costs'.

Principle 9 does not concentrate on technical interoperability but on how to develop proper models and use terms of co-operation and codes of conduct to predict, mitigate and manage public health problems, and promote community health in different environments where people live.

#### Principle 10 – Create code of conduct



Create international terms of data ownership, sharing, exchange, access and control to mitigate risks regarding sensitivity and privacy.

The European Parliament, in its resolution adopted in 2017, urged 'the private and public sectors and other data controllers to make use of instruments provided for by the GDPR, such as codes of conduct and certification schemes, in order to seek greater certainty over their specific obligations under Union law and to bring their practices and activities into compliance with the appropriate EU legal standards and safeguards.'<sup>37</sup>

In the health sector, a code of conduct or a network of codes of conduct can foster common interpretation of the legislation applicable to personal data, including secondary use of data for the purposes of public health, health research, and quality assurance in healthcare.

<sup>&</sup>lt;sup>35</sup> Competing in the Age of AI - How machine intelligence changes the rules of business by Marco Iansiti and Karim R. Lakhani: <u>https://hbr.org/2020/01/competing-in-the-age-of-ai</u>

<sup>&</sup>lt;sup>36</sup> See Figure 3 - Use cases and domains

<sup>&</sup>lt;sup>37</sup> European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))



D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

#### Principle 11 – Speak the same language



Share common vocabulary and meta-data to avoid misunderstandings and exploit the full value of data.

People who need to collaborate on data governance speak different languages. Years of traditional information silos in healthcare have resulted in language barriers. Finance speaks one language, clinicians another, supply chain another, and IT yet another. Terms and definitions are provided for the by the legislation but there are gaps in their interpretation and application in practice. Breaking down these language barriers takes time and culture change. Because of these silos, people don't always understand the other sides. Each department faces unique challenges and constraints to make the cultural and operational changes needed to develop a more data-driven culture and to support data governance.

A shared data-related vocabulary (information, technological and data standards) is a key achievement to render operational any data governance model. In fact, one of the main challenges healthcare systems are facing is data quality due to the complexity of the clinical-systems' data structure, massive growth in clinical data volume and the lack of standardisation between the clinical systems in terms of naming and modelling.<sup>38</sup>

The following actions should be taken:

 Implement a multidisciplinary team approach: To create a strong foundation for practical implementation of data governance across the healthcare sector, there is a need for people from different organisations and departments to collaborate. These people could then act as 'translators' to facilitate conversations and break silos, helping to tackle data problems arising from organizational silos. A goal of the data governance programme should be to develop more of these 'multilingual people'.<sup>39</sup>

<sup>&</sup>lt;sup>38</sup> Alofaysan S., Alhaqbani B., Alseghayyir R. and Omar M.. The Significance of Data Governance in Healthcare - A Case Study in a Tertiary Care Hospital. DOI: 10.5220/0004738101780187 In Proceedings of the International Conference on Health Informatics (HEALTHINF-2014), pages 178-187)

<sup>&</sup>lt;sup>39</sup> Paula J Edwards; Why is Data Governance in Healthcare so Difficult? Common challenges and ideas for overcoming them; himinformatics. <u>https://www.ehidc.org/sites/default/files/resources/files/Data-Gov-in-Healthcare\_092016.pdf</u>





- Create data quality management teams (in which data-related vocabulary and standards are included) that emerge from this broad integration and collaboration strategy. This includes having data stewards in every organisation and/or department of the healthcare system to spend time on data governance work: proactive data quality management, definition/adaptation, review and enforcement of data standards, and updating of role-based access policies, as requirements (legal, social and technological) inevitably change.
- Define and disseminate a common data-related vocabulary to ensure that data management activities are in line with the overall objectives of the healthcare system, are understood by stakeholders, and reflect and preserve the value to society from the sharing and analysing datasets as a collective resource. It is an operational requirement for practical implementation of data governance, as data quality refines the basis of how data are interpreted (metadata) as well as accessed (data access) by users.

A common vocabulary can be also included in the code of conduct described under Principle 10, mainly its section clarifying key terms and definitions.

# Measure, understand and communicate why stakeholders are important to each other

## Principle 12 – Aim for win-win of all stakeholders

With progress comes new challenges, such as maintaining patient privacy in a data-driven healthcare world and using data to predict rather than just diagnose and treat patients' diseases.<sup>40</sup> Legal uncertainties related to privacy issues can be mitigated by codes of conduct and other tools. However, developing and using them can be made more effective if stakeholders could find win-win solutions to their problems.

Implementation of Principle 12 can be carried out together with Principles 1, 4 and 6. While Principles 1 and 4 help to increase trust (mainly in data controlling and processing) and Principle 6 is a tool for using data to innovate something in a more efficient way, Principle 12 assists to eliminate regulatory, contractual or ethical misunderstandings around data sharing and access in a more effective way.

eHAction – Joint Action supporting the eHealth Network - www.ehaction.eu

<sup>&</sup>lt;sup>40</sup> <u>https://hbr.org/sponsored/2017/11/three-keys-to-unlocking-data-driven-health-care</u>





# Relations, interdependences and prioritisation among principles

Co-funded by the European

Union's Health Programme

(2014-2020)

Any single principle cannot solve all issues related to the major problems but should be part of a complex toolbox. They should also have links to some other principles in the same cluster or another.

Visualising how our efforts with realising one principle affect the value/success factor of another one, helps to decide which principle should be implemented first. Figure 7, 'accelerating relations' displays the relation between two principles. Green indicates that a principle listed in the rows have a quick positive impact on the other one listed in columns. E.g. creating a Code of Conduct (CoC) quickly could have a positive effect on carrying out a lot of other principles. Red colour reflects 'enabling relations', where principles listed in the rows eliminate obstacles to implement other ones listed in the columns. For example, trying to aim for win-win for all stakeholders might result in slowing down the process of a lot of other principles to take effect. Therefore, not all the 'accelerating relations' are low hanging fruits. For instance, it might need a long time to develop a CoC because stakeholders do not aim for win-win.

It is quite interesting to observe that, while all principles are quite important, still the relation between them is quite dynamic. This relationship matrix in Figure 7 below offers the opportunity to try to establish an order of the principles realisation.



Figure 7 - Accelerating and enabling relations between principles





# Guidance for implementation of the principles in form of actions

# **'Principle – Guidance – Governance' Framework**

The guidance consists of suggestions for practical actions to illustrate how policy makers could implement and make use of the implementation of the 12 principles to eliminate root causes of the obstacles to utilise the potential of harnessing new opportunities arising from big data and improved data analytics capabilities (see Figure 9).



Figure 8 – Approach to prepare guidance

In Figure 8 it is illustrated how implementation of clustered principles can assist eHN and Member States/countries to foster innovative use of big data in health by making use of suggestions for practical fields of governance. The set of actions, aiming to support both high level and operational level measures, consists of cluster and principle-specific actionable items. Table 1, in the Executive summary, gives an overview of all the actions for guidance.

# Actions dedicated to all principles:

- Prepare focused actions, e.g. Coordination and Support Action (CSA), Joint Action (JA) or procure services, as well as national and/or regional actions (e.g. open calls or predefined projects) to foster the implementation of specific actions (e.g. for the identification of use cases, elaboration of protocols and standards, etc.).
- Develop the regulatory framework for secondary use of data through guidelines and plans to guide action on governance models, data quality and legal compliance requisites.

# Actions 1 and 2, for Principle 1 (Cluster 1)

Action 1 - Elaborate protocols and standards for anonymisation and pseudonymisation and prepare stakeholders for adoption

Stakeholders need to be confident that they are prepared to understand and apply concepts of anonymisation and pseudonymisation. This can be achieved by complying with commonly





agreed protocols. Rules of measuring, verifying, registering and certifying data anonymisation procedures, shall be adapted to the provisions of the (proposed) Data Governance Act, and should be addressed in the legislation proposed for the EHDS.

Action 11 suggests to the eHN, Commission and Member States/countries to:

- Identify:
  - use cases on when, why and how data should be anonymised or could be pseudonymised;
  - priorities for use cases, harmonised to relevant common EU strategies and work programmes;
  - o focused protocols for different use cases of anonymisation or pseudonymisation;
- Based on GDPR, provide common legal background for the identification of use cases and protocols in regulation for the European Health Data Space (EHDS);
- Call on national, regional and local authorities to translate common European guidelines into national, regional or local protocols on how to properly anonymise or pseudonymise data, and provide a common roadmap for implementation ;
- Introduce protocols to stakeholders, consult them about the protocols, and encourage them to prepare for and adopt the protocols appropriately;
- Ensure transparency regarding the choice and differences between anonymisation or pseudonymisation;
- Define the potential roles and mandates of different data authorities (e.g. 'data permit authority').

Action 2 – Encourage Members states to develop anonymised and pseudonymised data repositories for the EHDS

Action 2 suggests to the eHN, Commission and Member States/countries to:

- Support national public bodies and recognised data altruism organisations to prepare anonymised and pseudonymised repositories for secondary use purposes by mirroring primary data sets (as soon as possible after capturing primary data). Combining actions with Principle 3 'Disseminate data re-use purpose well'! can increase effectiveness of the support.
- Ensure timeliness and continuous update of data for reuse:
  - Physically transform anonymised data, logically separate it from the original data(base) and store it in national/regional centralised data warehouses which could be prepared and authorised for connection to federated networks in the frames of common data governance regulations;
  - Encrypt the data when captured / at the same time it is created and recorded;
  - Pseudonymise data, when new information is generated, the system must be able to feed back the new information to the original source;





# Actions 3 and 4, for Principle 2 (Cluster 1)

#### Action 3 - Adapt curricula to include focus on digital health

Curricula for various stakeholders should be adapted or elaborated to include eSkills/eAbility in all EU countries to develop digital health literacy and to strengthen the understanding of digital health solutions as well as adherence to data sharing and reuse principles.

Start teaching digital health literacy (DHL) at the age of 14! Different digital health services, as well as data generated by citizens, are becoming more and more important. Therefore, improving basic DHL of young people is essential to pave the way for the use of varying digital solutions for health promotion, disease prevention and treatment, higher level and/or specialised education and training, as well as research and innovation.

eHAction Task 4.2 and Task 6.3 have already delivered recommendations regarding people empowerment, e.g. by improving digital health literacy (DHL) of patients, and e-skills for professionals (e.g. by implementing the eHealth competence model). These deliverables can be utilised during preparation of DHL-curricula for healthcare providers; however, additional parts shall be developed to provide complete curricula.

Action 3 suggests to the eHN, Commission and Member States/countries to:

- Include DHL in education from 14 years of age;
- Build on the recommendations of eHAction deliverables D4.2 and D6.3, and identify additional parts for the complete DHL-curricula for healthcare providers;
- Develop proper parts of DHL-curricula to provide a clear view about the definition and features of 'data', how data saves life and creates value, and why data differs from general commodities in legal aspects (e.g. ownership, property rights or excludability of non-rival goods);
- Consider special features of DHL-curricula:
  - Special programmes for management and clinical/professional staff, including teach-the-teacher education and training to improve citizens' adherence and empowerment;
  - Focused trainings to understand how adoption of new digital technology changes disease prevention and treatment or rehabilitation protocols, staff functions and professional co-operation, and how new/innovated care services create new unmet needs for research and technological innovation;
  - Digital health and innovation scout programmes and trainings to prepare stakeholders for participation in co-creation and open innovation co-operations (triple/quadruple helix co-operations).





#### Action 4 - Monitor and measure Digital Health Literacy (DHL) level

DHL of healthcare providers and citizens should be monitored and measured by evidence-based tools adapted to EU Member States and European countries to update or further develop curricula and design interventions appropriately, measure their impact, and ensure no one is left behind.<sup>41</sup>

#### Action 44 suggests to the eHN, Commission and Member States/countries to:

- Measure the complete spectrum of information gathering as well as stakeholders interactivity with digital tools, smart devices and protocols of digitally-enabled disease prevention and treatment (intra- and inter-organisational or cross-border eHealth Digital Service Infrastructure, equipment, mobile applications or the web);
- Develop a common instrument for monitoring this complete spectrum of DHL to measure operational skills, navigation skills, information searching, reliability, relevance, self-generated content, and relevant basic knowledge about ethical and legal principles;

# Action 5 for Principle 3 (Cluster 1)

#### Action 5 – Define and communicate the value of data reuse for practical care benefits

The role of data reuse in health and healthcare quality must be communicated to all stakeholders, from policy makers to health professionals as an *easy-to-use* tool, expressed in practical care benefits, recognised by all to ensure clear understanding of the purpose. Especially, identifying privacy protection and cybersecurity as a fundamental part of the data lifecycle.

Action 5 suggests to the eHN, Commission and Member States/countries to:

- Develop technical guidelines for system-level purposes that create value for data subjects and users, that all stakeholders can understand easily and relate to; especially:
  - Introduce how patients, as key stakeholders, can gain influence over decisions by using results of data processing;
  - Identify and articulate practical care benefits that can be personalised or individualised to engage data subjects. These care benefits should be communicated, disseminated and used in a clear, easy and personalised way to engage patients to co-operate in identifying and filling gaps in data collection for secondary use;<sup>42</sup>
  - Develop technical guidelines that define, at system level, how data can be used to improve the experiences of citizens as customers, end-users and beneficiaries of healthcare services. Add information on how these experiences can be

<sup>&</sup>lt;sup>41</sup> Zrubka, Z., Hajdu, O., Rencz, F. *et al.* Psychometric properties of the Hungarian version of the eHealth Literacy Scale. *Eur J Health Econ* 20, 57–69 (2019). <u>https://doi.org/10.1007/s10198-019-01062-1</u>

<sup>&</sup>lt;sup>42</sup> https://www.cerner.com/perspectives/3-things-to-know-consumer-directed-health-record





improved by data to establish more personalised services and care, as well as to increase social wellbeing.

# Actions 6 and 7, for Principle 4 (Cluster 1)

#### Action 6 – Use a data governance approach

Data governance is a fundamental tool to realise the potential of big data. Ensuring that data is reused in a responsible manner is a health policy imperative to maintain citizen's trust and meaningful investments in secondary use of data.

Action 6 suggests to the eHN, Commission and Member States/countries to:

- Identify the possible vulnerabilities of systems and identify weaknesses and risks in legal, technical and operational domains;
- Aim to audit/validate/certify not only the systems as a sovereign entity, but also all related (sub)systems and interlinked systems as well as medical (smart) devices in an integrated approach;
- Use 'zero trust' approach in data capturing in a controlled and audited/secured manner, and use measures to follow auditable/certified protocols;
- Define a risk management plan focused on prevention (instead of troubleshooting);
- Build, maintain, ensure and develop high availability services to provide a georedundant architecture, a second (preferably third) node for continuously sustained data flow;
- Follow and assess implementation of relevant trust frameworks for the conduct of interoperable certificates on how regulations, rules, policies, protocols, formats and standards are defined as well as certificates are issued, and make use of the learnings to strengthen trust in the common heath data space. All actors should be required to reach a specified level of (cyber)security, which includes the definition and audit of the systems' parameters;

Quantum resilience must be taken into account for further development to prepare and assist actions **now** to protect health data, since it can be encrypted and kept confidential for more than 10 years and an attacker could gain access to the cipher-text. In this context, 'quantum resilience' refers to the development and use of quantum resistant cryptographic algorithms that are proposed as potential candidates to provide post-quantum security resilience.<sup>43</sup>

43

#### Action 7 – Use FAIRified data models and tailored solutions

Action 7 suggests to the eHN, Commission and Member States/countries to:

<sup>&</sup>lt;sup>43</sup> https://www.enisa.europa.eu/publications/post-quantum-cryptography-current-state-and-quantum-mitigation





- Establish adequate control and certification systems to manage risks caused by the 'transparency paradox'.<sup>44</sup> Foster the use of AI in these systems;
- Align all related sub- and vendor systems to the specific system architectures of the complex integrated system, and take into consideration changes during the lifecycle;
- Examine, analyse and ensure the network infrastructures and circumstances for the secured integration of the whole big dataflow into the complex systems;
- Develop 5G/6G network and foster development of tailored solutions.

# Actions 8 and 9, for Principle 5 (Cluster 2)

#### Action 8 – Consider the use of existing standards and data repositories

The use of existing data-related standards and data repositories needs to be considered to ensure data quality and semantic interoperability. A common approach is necessary to achieve European wide consensus in the selection of data quality and semantic standards and requires strong leadership roles.

Action 8 suggests to the eHN, Commission and Member States/countries to:

- Adopt a common approach to identify relevant data-related standards and repositories that:
  - Encourages stakeholders to consider whether existing standard data models are useful for big data application, e.g. if you are aiming to combine administrative data with data from health records (e.g. OMOP Common Data Model,<sup>45</sup> CE marking, ISO standards, releases of ENISA or other organisations, Cyber Act, NIS 2.0);
  - Encourages stakeholders to consider whether their data can be coded according to widely used terminologies that allow linking to other healthcare data sources (e.g. SNOMED CT);
  - Supports stakeholders in considering standards for describing their metadata to comply with FAIR principles;
  - Considers making use of/contributing to existing data repositories (e.g. in the context of the European Health Data Space and the European Open Science Cloud);
  - Supports efforts to provide health-related data in a machine-readable format;

<sup>&</sup>lt;sup>44</sup> Andrew Burt: The AI Transparency Paradox, published on hbr.org, December 13, 2019: <u>https://hbr.org/2019/12/the-ai-transparency-paradox</u>

<sup>&</sup>lt;sup>45</sup> <u>https://www.ohdsi.org/data-standardization/the-common-data-model/</u>



#### Action 9 – Prepare common requirements for 'secure processing environment'

Action 9 suggests to the eHN, Commission and Member States/countries to:

- Foster use of and further development of existing technical standards and specifications by preparing common requirements for setting 'secure processing environment'<sup>46</sup> (SPE) in place, as well as common set of obligations regarding certification of SPEs;
- Foster the use of industry-independent building blocks that allow digital health transition in general and speed up the process of creating the EHDS by utilising specific ones for secondary use of health data, and coordinate their inclusion in health-specific services systematically.<sup>47</sup>

## Actions 10 and 11, for Principle 6 (Cluster 2)

Action 10 – Encourage national authorities and beneficiaries to identify and use good practices for stakeholders involvement

Action 10 suggests to the eHN, Commission and Member States/countries to:

- Encourage beneficiaries of grants to use models for identification and involvement of stakeholders in the secondary use of health data and big data. Organise national and regional events to introduce good and best practices and share learnings, e.g. innovation hubs, innovation scout programmes for healthcare organisations, living labs or cooperation based on innovation procurement;
- Combine measures for the implementation of Principle 6 with actions conducted according to Principles 3 and 12;
- Foster national and regional use of good and best practices gathered by EU level research, innovation, health and digital programmes.<sup>48,49,50</sup>

#### Action 11 – Develop action plans to improve the digital health ecosystem

Action 11 suggests to the eHN, Commission and Member States/countries to:

- Foster the creation of open innovation platforms for digital health by empowering healthcare providers.
- Develop national and regional action plans to create a network of innovation hubs and living labs to improve the digital health ecosystem;

<sup>&</sup>lt;sup>46</sup> Defined as Article 2 of Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act), COM(2020) 767 final, <u>https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767</u>

<sup>&</sup>lt;sup>47</sup> https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Big+Data+Test+Infrastructure

<sup>&</sup>lt;sup>48</sup> https://webgate.ec.europa.eu/dyna/bp-portal/

<sup>&</sup>lt;sup>49</sup> https://www.interregeurope.eu/policylearning/good-practices/

<sup>&</sup>lt;sup>50</sup> https://ec.europa.eu/programmes/horizon2020/en/newsroom/551/



- Identify roles for digital health innovation hubs in your country or region to launch focused actions;
- Develop and run innovation scout and data stewardship programmes at healthcare organisations.
- Assess possible roles of data altruistic organisations in involving stakeholders in your digital health innovation cooperation ecosystem.

# Action 12 for Principle 7 (Cluster 2)

#### Action 12 - Engage key stakeholders in needs assessment

Action 12 suggests to the eHN, Commission and Member States/countries to:

- Establish a framework for data governance that ensures continuous engagement of key stakeholders in needs assessment, monitoring and review process of data assets;
- Combine your actions with implementing actions for Principles 2, 6 and 12. On one hand, identification of needs helps to develop and implement successful trainings and/or involvement in co-creation, and on the other hand, the use of gamified learning tools and/or an experience of participating in co-creation bring tangible advantages in engaging someone in learning and understanding the needs of other players.

# Actions 13 to 15, for Principle 8 (Cluster 2)

#### Action 13 – Provide clear view about the definition and features of 'data'

Action 13 suggests to the eHN, Commission and Member States/countries to:

 All data protection statements need to be written in an easy-to-understand way. Differences between national rules for interpretations of GDPR can be treated more easily if stakeholders understand that 'the idea of owning data is challenging because data is not like other goods that we can own. It is non-rivalrous – I can both give it to you and still have it myself without it costing me any of the original good'. Exploring these differences helps to better understand and learn the common EU rules laid down in GDPR itself. Therefore, key technical, legal and ethical features of data, preparing data protection statements, asking for consent or data access (as well as providing it) shall be taught.<sup>51</sup>

Action 14 – Provide lists of free high-value datasets and assist internalisation of externalities

Action 14 suggests to the eHN, Commission and Member States/countries to:

<sup>&</sup>lt;sup>51</sup> <u>https://royalsociety.org/~/media/policy/projects/data-governance/data-ownership-rights-and-controls-October-2018.pdf</u>



- Speed up the emergence of value-added EU-wide information products by preparing and managing a list of high-value datasets to be provided free of charge by data users who have 'lawful access to certain personal or non-personal data and is authorised to use that data for commercial or non-commercial purposes',<sup>52</sup>
- Foster development of sustainable services by engaging stakeholders to acknowledge internalisation of externalities in their models if reasonable and possible.

Action 15 – Identify data quality standards for a 'time-and-space-changing health data space'

Action 15 suggests to the eHN, Commission and Member States/countries to:

- Assess the opportunity of adopting (the approach and solutions of) the data quality model provided by the ISO/IEC 25012 standard to foster continuous improvement and maintenance of data quality;
- Engage stakeholders (data subjects, data holders, data users, data altruism organisations and public sector bodies) to contribute to the creation of a 'time-andspace-changing health data space' being able to assist prediction of future trends and decision-making related to varying purposes of different stakeholders.

# Action 16 for Principle 9 (Cluster 3)

# Action 16 – Identify different features of environment-specific ethical and behavioural norms

Action 16 suggests to the eHN, Commission and Member States/countries to

 encourage stakeholders to participate in the creation of environment-specific health and wellbeing datasets of cross-sector data about determinants of health and geographic features, and federate them (in data spaces) with data generated and held in health systems, in research activities and captured by citizens, by identifying different features of environment specific ethical and behavioural norms.

# Actions 17 and 18, for Principle 10 (Cluster 3)

Action 17 - Analyse legal basis of Codes of Conduct under GDPR and demonstrate addedvalue

The legal basis for the adoption of codes of conduct is provided by GDPR. Drawing up codes of conduct shall be encouraged by the EU and national authorities, under Article 40.

Action 17 suggests to the eHN, Commission and Member States/countries to:

<sup>52</sup> COM(2020) 767 final

eHAction – Joint Action supporting the eHealth Network - www.ehaction.eu



- Analyse and regularly review the legal basis of Codes of Conduct (CoCs) under GDPR and possible future revisions (plus Member State legislation if applicable): aim, general content, personal and material scope, legal impact;
- Analyse European Data Protection Board (EDPB) guidelines on Codes of Conduct from the perspective of possible aims of CoCs and potential to bridge harmonisation gaps between Member States, and align data protection rules among specific sectors; identify relevant elements for health data;
- Carry out cost-benefit analysis.

#### Action 18 - Define scope and concept of CoC(s)

Action 18 suggests to the eHN, Commission and Member States/countries to:

- Define the scope and key elements of one overarching CoC or a set of CoCs. It includes legal safeguards to ensure compliance with data protection requirements, and liability issues.
- Organise stakeholder consultations involving all potentially interested organisations to ensure a wide-range acceptance. Data protection authorities need to be consulted where relevant under GDPR or other applicable legislation. Communicate adopted CoC(s) to a wide range of stakeholders and the general public and support dissemination, education and training activities.
- Monitor and evaluate implementation by regular evaluation and review, as key elements of the follow-up activities, and keeping CoC(s) up-to-date in a changing societal and economic environment.

A code of conduct cannot solve all issues related to legislation but should be part of a complex toolbox, and address different challenges. It should also have a link to the vocabulary described under Principle 11.

# Actions 19 and 20, for Principle 11 and Principle 12 (Cluster 3)

#### Action 19 – Set up co-creation teams and identify people who can become 'translators'

To implement the concepts laid out in Principle 11 'Speak the same language' and Principle 12 – 'Aim for win-win of all stakeholders', Action 19 suggests to the eHN, Commission and Member States/countries to:

 Encourage organisations and stakeholders to set up co-creation teams to learn about the needs, motivations and expectations of the others by identifying key features and attributes (e.g. non-negotiables, differentiators and dissatisfiers) of data sharing, access and control in order to scale up a proof of concept;





- Encourage organisations and stakeholders to avoid pigeonholing data scientists.<sup>53</sup> Data scientists can be in the centre of the co-creation processes with their colleagues at organisational level and stakeholders (such as patients, care providers and payers) at the level of the ecosystem, where they can scale up a proof of concept or participate in the development of a common vocabulary;
- Find people who can become 'translators' to facilitate conversations and break silos, helping to tackle data problems together;
- Combine the action with implementation of Action 11 Develop action plans to improve the digital health ecosystem. Innovation scouts, can take part not only in translation, but in dissemination of the common language when involving people in co-creation. After proper training, they can participate in organising and conducting further trainings for stakeholders to learn and use the same language;
- In addition, combine your efforts with your actions to increase DHL.

#### Action 20 – Develop common vocabulary for data sharing and access

An established and recognised vocabulary for data related activities is needed to support EU and national levels collaboration in a data-driven health ecosystem.

Action 20 suggests to the eHN, Commission and Member States/countries to:

- Develop data vocabulary for common understanding of quality initiatives (in which datarelated vocabulary and standards are included) overarching technological, ethical, business and legal features of digital health literacy;
- Involve stakeholders in the development of the vocabulary and engage them to share how they express their needs, requirements and offerings in negotiations and agreements; Find not only the differences, but the objectives and/or other reasons as well;
- In addition, combine these efforts with your actions to promote development of codes of conduct.

<sup>&</sup>lt;sup>53</sup> <u>https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture</u>





Co-funded by the European Union's Health Programme

(2014-2020)

# **Recommendations and next steps**

WP5 intends to recommend 'next steps of national and common implementation' as well. The Principles and Actions contain a suggestion for <u>endorsement by the eHN</u> on the way that the eHN members will commit themselves to:

- 1. Go into a national dialogue on how to use the final report (common principles + implementation guidance) with appropriate actors in the field,
- 2. Use as input to future work in the field (e.g. joint actions or coordination and support actions),
- 3. Convey principles to decision makers of local, regional, national and international funding programmes, and
- 4. Consider the principles in their decision making in planning and implementing development programmes or integrating results in healthcare system or health economy.

#### Recommendations for adoption by the eHN:

Having in mind the final result of the voting of stakeholders participating in the focus group meetings in June 2020, WP5 suggests the eHealth Network to support and facilitate Member States, stakeholders and the Commission

- to prepare focused actions, e.g. Coordination and Support Action (CSA), Joint Action (JA) or procure services, as well as national and/or regional actions (e.g. open calls or predefined projects) for the implementation of measures in the actions related to Principles 1, 6, 8, 10 and 11,
- to develop regulatory framework for secondary use of data through guidelines and plans to guide action on governance models, data quality and legal compliance requisites,
- to critically assess which policy actions of the Guidance are most needed in their context.





# **Appendices:**

- 1. Main findings
- 2. Working definitions

# Appendix 1 - Main findings D5.1 and D5.2

# 1. What are the main challenges regarding big data and artificial intelligence in health?

There are many policy and operational level recommendations in this topic. Most of them are still relevant in most Member States and organisations. The result of our mapping showed that three general obstacles appeared as reasons slowing down or hampering translation of policy-level recommendations into actions: lack of trust, legal uncertainties, and lack of funding and financial resources.

All these obstacles can be traced back to a general lack of data governance related priorities in health policy at Member State/country level. Many recognise that implementing effective data governance is critical to meet increasing demand for information to support value-based care and population health, however most countries are only at the beginning of a complex journey to encourage the development and safe use of health data. Health Ministry leadership is necessary to ensure that delivering the data to manage this important sector is at the forefront of government policy and action. Optimal decision making about potential statistical and research uses of personal health data can only be achieved if there is an overarching data governance framework in the country that has been aligned to minimise societal risks and to maximise societal benefits from data uses. Mechanisms of collaboration must be designed to support countries in developing data governance frameworks and engaging in regulatory and legislative reforms, including those necessary as the result of the EU Data Protection Regulation).<sup>54</sup>

# 2. Lack of trust is one of the identified barriers in the D5.1 Draft report on policy-level actions on innovative use of big data in health. What can be done to overcome this?

Common efforts are needed to eliminate obstacles caused by lack of willingness, unavailability of comprehensive data or fear of abuse. The same applies to resistance to learn, lack of motivation for further education and lack of expertise or resources (not only financial, but trained personnel as well).

In addition, the unmet need about transparency in capturing, cleaning, storing, sharing or using data shall be satisfied, and a significant amount of data still kept somewhere in paper format, shall be traced and digitised. At the same time, key stakeholders have to understand the difference between analysis and reporting.

Last, but not least, stakeholders have to confront the 'transparency paradox'. To do so, they need assistance in finding or updating solutions on how transparency could make AI methods less vulnerable, while they can help mitigate issues of fairness, trust and discrimination.

<sup>&</sup>lt;sup>54</sup> OECD (2015), Health Data Governance: Privacy, Monitoring and Research, OECD Publishing, Paris



3. Within the scope of D5.2 - Report on identified cross-border use cases, WP5 conducted interviews to collect information from key stakeholders. How do you assess the achieved feedback?

The in-depth interviews enlarged the knowledge pool on innovative cross-border use of health data. Even though the projects selected for in-depth analysis had their foci on the intervention and implementation of innovative health technologies from very different disease areas (schizophrenia, rare diseases, MCC, etc.) most of them faced rather similar challenges when implementing their project.

In particular, issues of privacy protection, ethics, data security, health assessments, data quality, interoperability of health data systems, and demonstrating added value to the key actors (such as citizens, patients and professionals) were brought up.

The analysis of the conducted interview results shows that there already exist a plethora of experiences and project results that are accessible and available from different projects. The projects could be consulted to develop a knowledge base and a framework for continuous exchange of best practices at the EU level.

# 4. Considering the related challenges and opportunities, what should be the next steps to advance the innovative use of health data in Europe?

It is quite important to define priority areas and compose the optimal set and order of the most challenging and most promising use cases. Finding, implementing or developing use cases that provide optimal set of gains at different levels shall be fostered by launching dedicated regulatory and specific funding programmes to:

- reuse and further develop existing and technical capacities for exchange data
- foster empowerment, adherence and grit to increase level of knowledge, skills and competences related to sharing and accessing data for analysis and innovation purposes
- process good/best practices for optimised use cases
- develop and implement new curriculum and training programmes
- prepare and run communication, and dissemination and CSR programmes.
- accelerate product, procedure and service development and bring new innovations and advances to patients
- mitigate uncertainties about implications and unclear impacts on intellectual property issues, fairness, safety, security, trust, liability, algorithmic transparency, social inclusion
- offer clarity about the roles of varying agencies in the Member States/countries and the EU.

For example, the creation of a European Health Data Space (EHDS) may help to launch these dedicated regulatory and specific funding programmes, and in this way, it can foster innovative use of health data. It can be built not only on the network of national, regional, local or organisational data spaces, but on the cooperation of key stakeholders throughout Europe as well. It is likely that EHDS could be an important, fundamental part of the digitalised healthcare ecosystem, therefore it requires EU level co-ordination.





# **Appendix 2 - Working definitions**

Working definitions were elaborated by WP5 members at the kick-off meeting in order to lay down statements of the meaning of certain phrases which have been explained in various ways in the literature. The kick-off meeting of Work Package 5 was held on 16 October 2018 in Brussels. The meeting included a workshop to review available definitions and to propose for adoption those that are important to empower patients, policy makers and professionals about the innovative use of health data. Experts agreed on defining health data, big data in health, big data analytics in health and innovative use of health data. These working definitions were introduced to the eHealth Network in November 2018, and are considered as continuously evolving definitions based on the evolutionary and constantly changing nature of the field. Definitions reflect the value-based approach followed in WP5, where value refers to satisfaction of a specific need and replicability at an economical cost. The following definitions were used in deliverable D5.1, endorsed by the eHN in June 2020.

- 1. 'Data' means information, especially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer.
- 'Pseudonymisation' means the processing of personal data in such a manner that the 2. personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.<sup>55</sup> In a GDPR context, pseudonymised data is considered personal data. GDPR does not apply to anonymised data, which is not considered personal data. (GDPR whereas: (26) 'The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore

<sup>&</sup>lt;sup>55</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Article 4 (5) Definition of Pseudonymisation. https://eurlex.europa.eu/eli/reg/2016/679/oj



concern the processing of such anonymous information, including for statistical or research purposes.')

- 3. 'Anonymisation' means the process of changing data into anonymous data which do not relate to an identified or identifiable natural person, or the process of rendering personal data anonymous in such a manner that the data subject is not or no longer identifiable.<sup>56</sup>
- **4.** *'Machine-readable format'* means a file format structured so that software applications can easily identify, recognise and extract specific data, including individual statements of fact, and their internal structure.
- **5.** *'Health data'*: Patient data in health records (records kept by health professionals and care providers, as well as self-reported health data), data from apps and wearables, any background data that will give insights on the social determinants of health.
- 6. Primary and secondary use of health data: In many cases the term 'innovative use of health data' is defined as secondary use of health data and big health data. However, it is worth considering whether there was a tangible border between primary and secondary use, or there are other data usable for primary and/or secondary use.

In our interpretation, primary use of health data is related to the care or treatment of a person (a data subject<sup>57</sup>), while secondary use covers every other case related to any goals regarding policy making, system governance or planning, regulation, authorisation, control, monitoring, governance, management, research, innovation, development, etc. In this context, secondary use of health data is not only use (reuse) of health data again or more than once, but the use of the data of one data subject for the benefit for more or other data subjects other than the initial purpose.

The pellucid border between primary and secondary use is well reflected in the introductory description of the UK NHS Innovative Uses of Data Team: 'Our Innovative Uses of Data (IUoD) team aims to improve our information analysis and reporting, by using novel data science techniques. This will enable new insights from data that work to improve health and social care. Products will be focused on the needs of patients, clinicians and organisations within the health and social care sector, to increase the likelihood of delivering real benefits that will improve patient outcomes.' <sup>58</sup>

7. **'Open Data**' as a concept is generally understood to denote data in an open format that can be freely used, reused and shared by anyone for any purpose. Open data policies which encourage the wide availability and reuse of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, and

<sup>&</sup>lt;sup>56</sup> Based on the definition of Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, Article 2 (7).

<sup>&</sup>lt;sup>57</sup> One shall be very careful to use the notion of 'owner' or 'ownership' in relation to any personal data without deeper legal elaboration. There is an ongoing professional debate about the concept of ownership in this regard. In this context ownership refers to the origin and the beneficiary of data. Therefore, the working definition uses the term 'data subject' who can be an identified or identifiable natural person accordingly the GDPR.





which promote the circulation of information not only for economic operators but primarily for the public, can play an important role in promoting social engagement, and kick-start and promote the development of new services based on novel ways to combine and make use of such information.<sup>59</sup>

- 8. **'Health literacy'** is defined as 'the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions'.<sup>60</sup>
- 9. 'eHealth literacy', according to the Lily model of Norman and Skinner, covers a broader concept, encompassing traditional literacy (basic ability to read and comprehend written text), information literacy (the ability to find and use information), media literacy (the ability to think critically about media content and context), computer literacy (the ability to use computers for problem solving) and scientific literacy (understanding how knowledge is created with its aims, methods, limitations, and politics), in addition to traditional health literacy. eHealth literacy has been defined as 'the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem'.<sup>61,62</sup>
- 10. **'Digital Health Ecosystem'** integrates mHealth, telehealth and telemedicine, eHealth governance, Digital health literacy, Electronic health records, Social networks, Standards and interoperability and big data.<sup>63</sup>
- 11. **Digital health literacy:** 'is the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to preventing, addressing or solving a health problem'.<sup>64</sup> Better digital health literacy can lead to enhanced prevention models, better observance of healthier behaviours and improved wellbeing. Digital health literacy is one of the characteristic elements of 'Digital Health Ecosystem'<sup>65</sup> *eHAction Work Package 4* dealing with digital health literacy of people has recently emphasised that 'People who are digitally health literate can find, appraise and understand health information from electronic sources, before applying it in their own lives to achieve better health outcomes. Improving the level of digital health literacy empowers people to take a more active role in the management of their health and wellbeing.' WP5 adds that 'as the proliferation of wearables, connected medical devices, personal health records and mobile apps continues, consumers are becoming increasingly interested in capturing, analysing and sharing their own health data. While providing both challenges and opportunities for physicians, this information, also known as patient-generated health data (PGHD), can help people become more engaged in their own

<sup>&</sup>lt;sup>59</sup> <u>https://eur-lex.europa.eu/legal-content/HU/TXT/?uri=CELEX:32019L1024</u>

<sup>60</sup> https://www.ncbi.nlm.nih.gov/books/NBK216035/

<sup>61</sup> https://www.jmir.org/2006/2/e9/

<sup>62</sup> https://link.springer.com/article/10.1007/s10198-019-01062-1

<sup>63</sup> https://www.who.int/global-coordination-mechanism/working-groups/digital\_hl.pdf

<sup>&</sup>lt;sup>64</sup> <u>https://www.who.int/global-coordination-mechanism/activities/working-groups/17-s5-rowlands.pdf</u>

<sup>&</sup>lt;sup>65</sup> https://eurohealthnet.eu/sites/eurohealthnet.eu/files/publications/PP\_Digital%20Health%20Literacy\_LR.pdf



care.'<sup>66</sup> Therefore, WP5 stresses that digital health literacy also covers ability to share health data and understanding the importance and possibilities of innovative use of health and health related data.

- 12. **Big data in health**: Consolidated data from existing fragmented data sources for the purpose of understanding, forecasting and improving health and health system status, needs and performance. (This working definition was developed at the kick-of meeting of WP5, having into consideration the following definition of the Study on Big Data in Public Health, Telemedicine and Healthcare (hereinafter referred as 'EU Study'):<sup>67</sup> 'Big Data in Health refers to large routinely or automatically collected datasets, which are electronically captured and stored. It is reusable in the sense of multipurpose data and comprises the fusion and connection of existing databases for the purpose of improving health and health system performance. It does not refer to data collected for a specific study.')
- **13.** Big data analytics in health: Statistical learning methods and algorithms applied to big data in health, which include descriptive analytics, mining/predictive analytics to support evidence-based decision making, analytical techniques that are ideal for analysing a large proportion of text-based health documents and other unstructured clinical data (e.g. physicians' written notes and prescriptions and medical imaging).
- **14.** Artificial intelligence: Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions with some degree of autonomy to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems), or can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications). Many AI technologies require data to improve their performance. Once they perform well, they can help improve and automate decision making in the same domain.<sup>68</sup>
- **15.** 'Innovation' is the process of translating an idea or invention into goods or services that create value, or for which customers will pay. To be called an innovation, an idea must be replicable at an economical cost and must satisfy a specific need. Innovation involves deliberate application of information, imagination and initiative in deriving greater or different values from resources, and includes all processes by which new ideas are generated and converted into useful products.<sup>69</sup>
- **16.** *'Innovative use of health data'*: The use of health data is considered 'innovative' if this use results in better patient outcomes and/or higher quality of healthcare delivery and/or higher productivity and performance. Our approach in defining innovative use of health
- <sup>66</sup> https://letstalkrespiratory.eu/the-opportunities-and-challenges-of-patient-generated-health-

<sup>&</sup>lt;u>data/?gclid=Cj0KCQjwhb36BRCfARIsAKcXh6HHdL7nQfWTcVLPrCXsjs5Mm3NdTdowzrUMoZutsJJ0ZiihVlzlQMcaAqX5EALw\_wc</u> <u>B</u>

<sup>67</sup> https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata\_report\_en.pdf





data is also based on the definition of innovation of the World Health Organization (WHO): 'Health innovation identifies new or improved health policies, systems, products and technologies, and services and delivery methods that improve people's health and wellbeing. Health innovation responds to unmet public health needs by creating new ways of thinking and working with a focus on the needs of vulnerable populations. It aims to add value in the form of improved efficiency, effectiveness, quality, sustainability, safety and/or affordability. Health innovation can be preventive, promotive, curative and rehabilitative and/or assistive care. The WHO engages in health innovation to achieve universal health coverage within the context of the Sustainable Development Goals.'<sup>70</sup> Innovative use of health data is determined by the ways of converting unstructured, separated datasets into new or renewed things, services, solutions, organisations or systems.

- 17. **Data governance** entails defining, implementing and monitoring strategies, policies and shared decision-making over the management and use of data assets. It is performed by the staff of public sector bodies (of the Member States/countries and the Commission) with established data-related roles.<sup>71</sup>
- 18. **Data policies** are a set of broad, high-level principles which form the guiding framework in which data assets (in public sector bodies of the Member States/countries and the Commission) can be managed. More specifically, data policies govern data management, data interoperability and standards, data quality, data protection and information security. Where needed, data policies will be complemented by detailed guidelines and processes. Such guidelines and processes may be developed either at corporate level, interservice level or local level. This remains outside the scope of this document.<sup>72</sup> See also definition of 'data policy' in the European Interoperability Reference Architecture: 'Data Policy: A set of broad, high level principles which form the guiding framework in which data management can operate.' <sup>73, 74</sup>
- 19. A **data asset** is any collection of data, any data set or any information that is somehow linked, e.g. by common codes or metadata, which has been created by public sector bodies of the Member States/countries and the Commission, collected from Member States or other stakeholders, or acquired from third parties in the context of projects, policy or administrative processes. Data assets may be structured or unstructured, static or dynamic, raw or curated. Data assets are in digital formats.

Structured data assets are organised according to a predefined data model or schema and the content of each field/variable can assume only predefined values. Unstructured or

<sup>&</sup>lt;sup>71</sup> Data governance and data policies at the European Commission, EUROPEAN COMMISSION, Secretariat-General, July 2020

<sup>&</sup>lt;sup>72</sup> Data governance and data policies at the European Commission, EUROPEAN COMMISSION, Secretariat-General, July 2020

<sup>73</sup> OECD: https://stats.oecd.org/glossary/detail.asp?ID=4454

<sup>&</sup>lt;sup>74</sup> <u>https://joinup.ec.europa.eu/taxonomy/term/http\_e\_f\_fdata\_ceuropa\_ceu\_fdr8\_fDataPolicy</u>



semi-structured data assets are not structured via predefined data models, schemata or code lists.<sup>75</sup>

20. For **further working definitions** see Article 2 of the Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act) COM/2020/767 final.<sup>76</sup>

# Appendix 3 - Overview Work Package 5 'Innovative use of health data'

Task	Deliverables	MWP 2018-2021
T5.1 Mapping, awareness raising and policy relevant actions on innovative use of big data in health (Lead: NHSC)	D5.1 Report for the information of the eHN on policy level actions (M24)	B.1 Awareness raising of using Big Data in healthcare Expected outcome: Increase awareness on the possible impacts, challenges, risks and directions of Big Data in healthcare.
T5.2 Sharing and learning best practices on European level (Lead: THL)	D5.2 Report on identified cross-border use cases, including assessment of pros & cons of stakeholders, and practical solutions with potential for European scale benefits (M18)	B.2 Develop common vision of innovative use of data in healthcare Expected outcome: Common vision and priorities for innovative use of data in healthcare. Clear criteria are needed to define which use cases / good practices of Big Data use in healthcare have the best potential to improve the sustainability of healthcare system and provide examples for benchmarking.
T5.3 Towards an attempt to define common principles for practical governance (Lead: NHSC)	D5.3 Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data with a special focus on data to be used (and the implementation of data access and use) in public health, research and quality assurance in	<ul> <li>B.3 Governance and methodologies for innovative use of health data, including big data</li> <li>Expected outcome: Common principles to facilitate the development of innovative use of data projects at European Level.</li> <li>(+ Guidance for the implementation of common principles)</li> </ul>

<sup>&</sup>lt;sup>75</sup> Data governance and data policies at the European Commission, EUROPEAN COMMISSION, Secretariat-General, July 2020

<sup>&</sup>lt;sup>76</sup> <u>https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767</u>





D5.3 – Proposal for the eHN on the guidance for the implementation of common principles for practical governance of big data [...] WP 5 – Innovative Use of Health data Version 2.4, 03-05-2021

healthcare on a European scale (M36)