DOI: 10.1561/9781638282372.ch24

Chapter 24

The Interactive BigMedilytics Website: A Collection of Scientific Results, Best Practices, and Lessons Learned

By Holmer Hemsen, Roland Roller, Supriyo Chatterjea, Anne Marie Weggelaar-Jansen and Alexandra Muñoz Oliver

24.1 Introduction

The goal of BigMedilytics is, through the application of Artificial Intelligence (AI) technologies in combination with the analysis of Big Data, to transform Europe's healthcare sector and enhance patient outcomes and productivity. The European Commission (EC) funded BigMedilytics as a lighthouse project to foster this disruptive innovation effort, i.e., BigMedilytics should have a certain signalling effect for future projects.

One of the outcomes of BigMedilytics that stands out in comparison with similar Big Data Healthcare projects is the Interactive BigMedilytics Website. The Interactive BigMedilytics Website is a webpage that enables users to browse through a selection of aspects related to the BigMedilytics project. By navigating through the Interactive BigMedilytics Website, you cannot only find information on results and different aspects of the 12 study projects that participated in the EU lighthouse project but also read about lessons learned grouped by topics, such as ethical, legal, and privacy issues or with respect to validation, technology, and business impact.

The outcome of the BigMedilytics project consists not only of scientific articles describing concrete aspects of the application and its evaluation, e.g., system architecture, data (pre-)processing pipeline, application of (machine learning) methods, AI techniques, use of Big Data and evaluation, but it also contains more specific topics across these applications, such as ethical considerations and how compliance with relevant national and EU legislation (e.g., GDPR) can be achieved, as discussed in Chapter 3 (Using Causal Diagrams to Understand and Deal With Hindering Patterns in the Uptake and Embedding of Big Data Technology), Chapter 4 (Lessons Learned in the Application of the General Data Protection Regulation to the BigMedilytics Project), and Chapter 5 (Ethics). In addition, guiding principles, which facilitate the establishment of new healthcare projects in the domain, are incorporated by sharing practical experiences and lessons learned across the 12 studies and by disclosing certain pitfalls, for example, to establish a big data project in a hospital data scientists and medical staff have to work together. But working together requires a certain understanding of each other's field and methods and a common vocabulary to communicate with each other – an often time-consuming effort that cannot be avoided and therefore needs to be considered already in the project planning phase.

24.2 Related Work

Apart from the 12 BigMedilytics studies, various other projects have been carried out applying Big Data and AI in Healthcare. Several of them target a particular disease, for example, Big Data Analytics to improve cardiovascular diseases, [1, 2] kidney diseases, [3, 4] lung cancer, [5] or diabetes [6]. Some research also takes a specific Big Data in healthcare aspect under investigation, such as privacy and ethics or economic impact [7–10]. Hansen *et al.* present the results from a literature study on Big Data in Science and Healthcare [11]. Agrawal and Prabakaran studied Big Data initiatives in the United Kingdom and the United States and abroad to identify lessons learned and recommendations for general practice [12]. However, to the best of our knowledge, no similar website as the Interactive BigMedilytics Website exists to interactively explore the results from multiple Big Data and AI in healthcare projects including lessons learned and recommendations.

In the following, we discuss the selection of the content to present on the website and how we structured and organized it. We decided to use an existing web design template to present the content, so layout and interaction possibilities of course are predefined by this template. The choice of this particular template, among the many other templates available, is also briefly explained in the following paragraph.

24.3 Aim, Content, and Structure of the Interactive BigMedilytics Website

As already mentioned in Section 24.2, various projects have applied AI with Big Data analysis in healthcare projects. Some of them also reported lessons learned and recommendations for general practice. However, with its 12 study projects, the BigMedilytics project provides an exceptional opportunity to collect, compare, and report on lessons learned and promote recommendations across medical disciplines with an additional focus on business and privacy and ethical aspects.

The aim of the Interactive BigMedilytics Website is

- to provide users with an interactive way
  - to inspect and compare various aspects of the 12 studies carried out in BigMedilytics;
  - to read about lessons learned in BigMedilytics;
- to provide stakeholders with tailored information relevant to specific stakeholder groups;
- to link the new concepts and technologies developed within BigMedilytics to sections in a patient journey and to illustrate how these new technologies and processes can improve the healthcare system in the future;
- to support the development of future Big Data and AI in Healthcare projects.

In order to enable users of the website to compare certain aspects of the projects with each other, we asked each study participating in BigMedilytics to fill in a pre-structured template document. Via the template document, the following data on each study were collected:

- Key information: It provides information on involved partners and countries, keywords, and a task description.
- Building blocks: It contains information on system architecture including data flow, software components, and necessary hardware; processing of structured and non-structured data, containing information on data sources, data cleansing, processing of heterogeneous data streams, and real-time event detection; AI components used or developed including deep learning for NLP and image processing; and privacy and ethical issues;
- Learnings: It summarizes challenges and barriers, lessons learned, and main achievements.
- Output: A section providing information on published papers, open source, and demos.

In this way, we ensured that the projects provided comparable blocks of information.

The input for a chapter on general learnings (see also Chapter 27 – General Learnings From the Horizon 2020 Project BigMedilytics) is a direct result of the experiences of the 12 studies carried out in BigMedilytics as described in filled-in template documents, additional items that have been added manually are a result of manifold discussions within the BigMedilytics consortium. To present each stakeholder group with only relevant items, the information has been manually selected and clustered. To be able to link the new concepts and technologies developed within BigMedilytics to different stages in a patient journey, we first sketched and described a fictitious but realistic patient journey throughout four stages: 'at the GP', 'at the Hospital', 'at Home', and 'Follow-Up GP'.

In order to view the content of the website, five different entry points to inspect results, findings, and lessons learned gathered from BigMedilytics exist: 'Pilot View', 'Transversal Aspects', 'Stakeholder Perspective', 'Patient Journey', and 'General Learnings'. All views consist of a structured hierarchy of layers to be able to quickly navigate to the content of interest. In Table 24.1, the different views are presented in more detail.

24.3.1 Web Design of the Interactive BigMedilytics Website

The goal of the Interactive BigMedilytics Website is to present the selected and curated information in an interactive but also appealing and modern way. In the beginning, we discussed several web design options and prototyped some of them. A traditional, old-style, web design by simply transforming the study documents into HTML, adding menus and links between sections and studies' descriptions was quickly ruled out, because 'flexibility and efficiency of use' (Nielsen Usability #7 of usability heuristics) could not be achieved for both navigation within studies (Pilot View) and across studies (Transversal Aspects) [13]. We finally decided to use an existing web design called zoomable circular packing[i] to present study-specific information, transversal aspects between studies, such as system architecture, information relevant for a specific stakeholder group, patient trajectory-related information, and, last but not least, lessons learned throughout the BigMedilytics project. The chosen web design uses circles inside circles to represent hierarchies of information. While the outer circle and its label represent items belonging to the current level, circles inside circles represent the next level in the navigation hierarchy. Figure 24.1 shows the top layer of the Interactive BigMedilytics Website with its five different entry points for navigation.

i. See, for example, https://observablehq.com/@d3/zoomable-circle-packing for an implementation.

Table 24.1. The five views of the Interactive BigMedilytics Website.

Pilot View	The pilot view provides a theme-centric perspective on study-specific data. The elements of the first navigation level are arranged by study category, which is either population health and chronic disease management, oncology, or industrialization of healthcare services. At the second level, the user can choose a specific study to inspect, such as kidney disease or asset management. After selecting this study, the user can in the third level choose a particular aspect of the study to read about, which in the fourth level is shown. All studies provide an identical set of aspects that can be explored, for example, Business Aspects, Challenges and Learnings, or Achievements.
Transversal Aspects	In contrast to the Pilot View, as discussed above, the Transversal Aspects view shows the user a horizontal view (across studies) of the content. The elements of the first navigation level in this view are arranged by aspect and are similar to the third level in the pilot view but with the difference that the aspects building blocks, ethical and privacy issues, and security issues are further subdivided, so the chunks of information presented to keep a reasonable size to be inspected interactively. Building blocks, for example, is further split into the aspects data processing, system interaction, auditory and logs, Natural Language Processing, image processing, system architecture, and prediction.
Stakeholder Perspective	The Stakeholder Perspective organizes the information in such a way that relevant information for the specific target group is shown. In total, eight target groups are addressed: policymaker, hospital decision-maker, privacy officer, hospital IT/equipment, clinical staff, data scientist, health insurance, and patient. At the first navigation level, the user selects the target group of interest, and at the second level, the topic of interest. For example, policymakers can read about trustworthy AI, privacy measures, or learnings. In contrast, the information presented to data scientists is of wider scope. In addition to information about policies, also information about more technical issues such as natural language and image processing or information about access control is included for this group.
Patient Journey	The Patient Journey view organizes the information according to a realistic but fictitious patient journey and shows in the first navigation level four different stages in this journey: at the GP, at the Hospital, at Home, and Follow-up GP. After selecting a particular stage, the user can either read about studies including this stage in the patient journey or about applying BigMedilytics Technologies to this specific patient Journey stage.

(*Continued*)

Table 24.1. Continued

General Learnings	The General Learnings view presents lessons learned from the BigMedilytics project that are of a broader and less study-specific nature. The information presented in the first navigation level is grouped into eight thematic categories, such as Data, Technology, or Business Impact. For example, under the category Ethical, Legal, and Privacy, it is pointed out that consent gathering to use patient data is of major importance before the data can be used and analysed. Category Big Data Blueprint differs because it (a) links to the Blueprint document (see Chapter 25 – Data Processing in Healthcare Using CRISP) and (b) to a Blueprint Matrix table listing for a selection of aspects, e.g., Natural Language Processing, the studies contributed to.

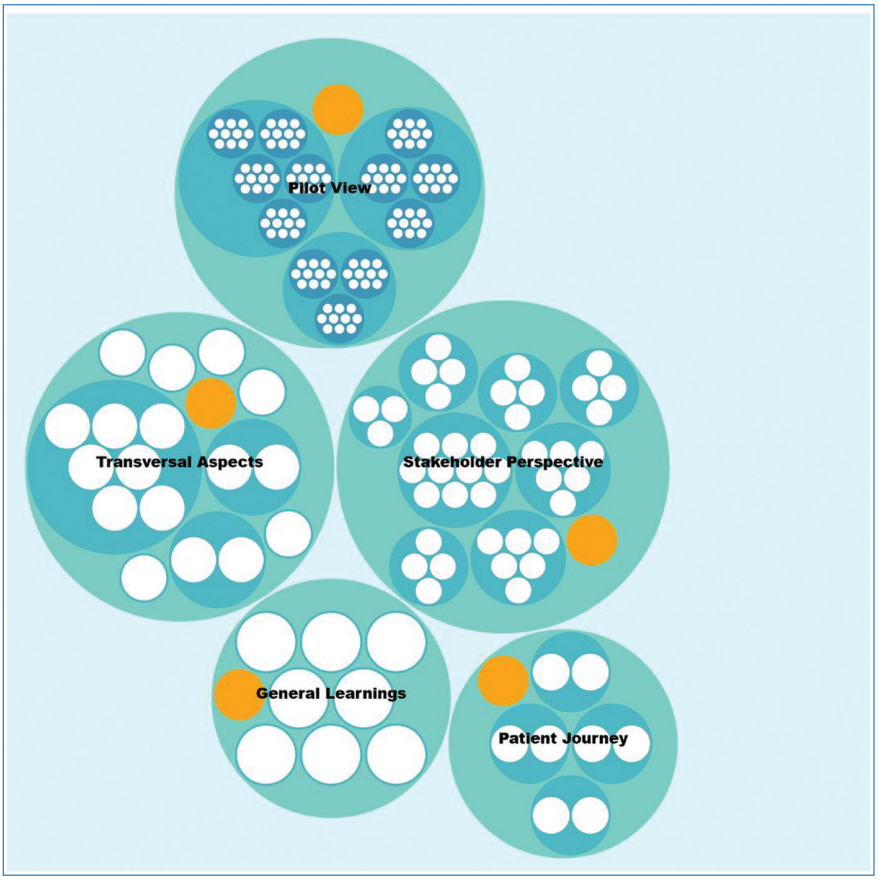

Figure 24.1. Interactive BigMedilytics website – the five different entry points to view the data.

During early tests of the Interactive BigMedilytics Website, we observed that the users anticipated the chosen web design quickly and rapidly learned how to navigate through the hierarchy to find the content of interest. But we also noticed that the web design is, especially for new users, not fully self-explanatory. Therefore, we added upfront, i.e., before showing the circular packaging design, a section with help texts, explaining the following aspects: "What is the Interactive BigMedilytics Website?", "How to navigate it?", "Who is it for?", and "What will you find?". To keep the help texts close to the interaction area (see Figure 24.1), initially, these help texts are collapsed and only the section title is visible to the user. In addition, the "How to navigate the Interactive BigMedilytics Website?" subsection provides two videos to show both the general structure of the webpage and second a tutorial on how to navigate the Interactive BigMedilytics Website. Furthermore, we slightly modified the standard zoomable circular packing web design as defined by the JavaScript library and introduced a circle with a special meaning and a distinct colour – the "About" circles. These circles only appear in the top-level view and, if selected, explain the view and its structure.

24.3.2 Structure and Content of the Interactive BigMedilytics Website

When a user clicks within a circle (see Figure 24.2), it reaches the next hierarchy level in the chosen view. Users can go back to one hierarchy level at a time by clicking into the surrounding circle (see Figure 24.3) or reach the top level of the hierarchy by clicking outside the circles area. The leaves of the navigation tree point to the content, presented in a classical, text-only, design (cf. Figure 24.2: No. 5 and 6). The web design therefore enables "user control and freedom" as Nielsen requires for good web design [13, #3].

Figure 24.2 shows a sample interaction path starting with the overview page (1), then selecting the Pilot View (2), choosing heart Population Health and Chronic Disease (3), and selecting Hearth Failure (4) and Business Aspects of this study to see the final content in (6). Please note that navigation is possible in both directions as shown in Figure 24.3. Starting with the pilot view (cf. Figure 24.1), a user can, for example, choose Oncology (2) and then inspect the next layer (3b) or go back to (2) and instead choose Population Health or Chronic Disease (3a).

The content presented in Pilot View and Transversal Aspects is based on the documents describing aspects of each study. The predefined structure has been mapped to the hierarchical web design layers, so equal sections, such as Key Information, can be accessed in a similar way within Pilot View (aspect shown for a single study) and Transversal Aspects (aspect shown for multiple studies) view. The structure presented in Pilot View and Transversal Aspects view differs slightly from the content

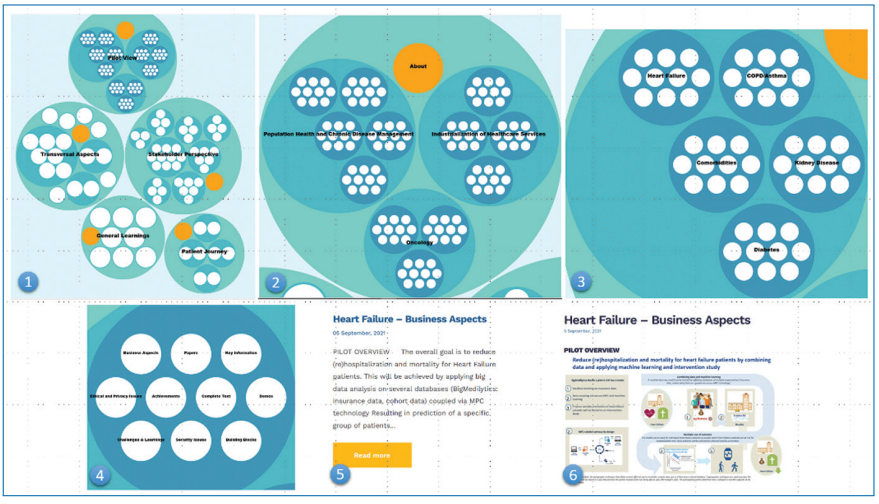

Figure 24.2. Sample Interaction – From the Top Layer (1) to Content (6).

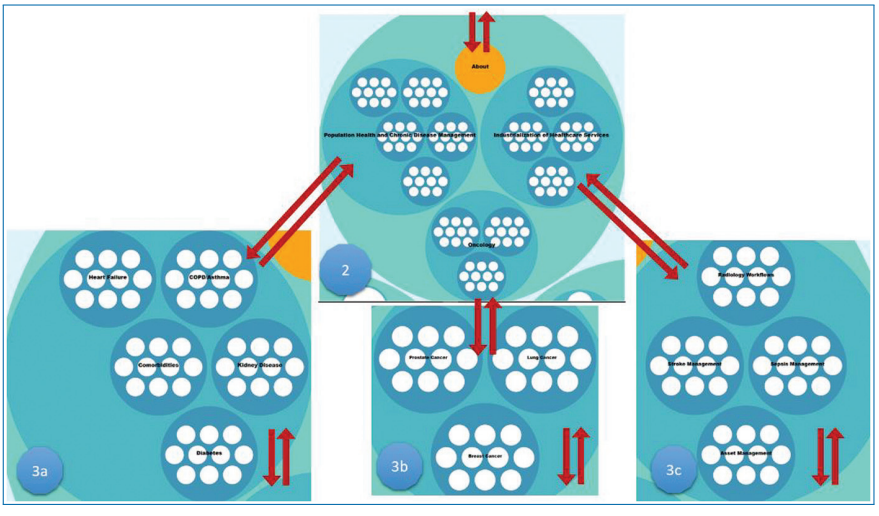

Figure 24.3. Possible forward/backward (zooming) interaction between layers.

structure in the template document (see Figures 24.4 and 24.5). This change was mainly done to balance the number of sections presented but was also necessary to make certain sections more easily to find, for example, Ethical and Privacy Issues. As, by splitting the study documents into separate content units (see Figure 24.2: No. 5 and 6.), the connection with the study itself is no longer visible, each content document is labelled with the study name in its title.

The intention of the Stakeholder Perspective is to address stakeholders in a targeted manner and provide each stakeholder with four to his/her group relevant information. We included the following stakeholder groups: Clinical Staff, Data

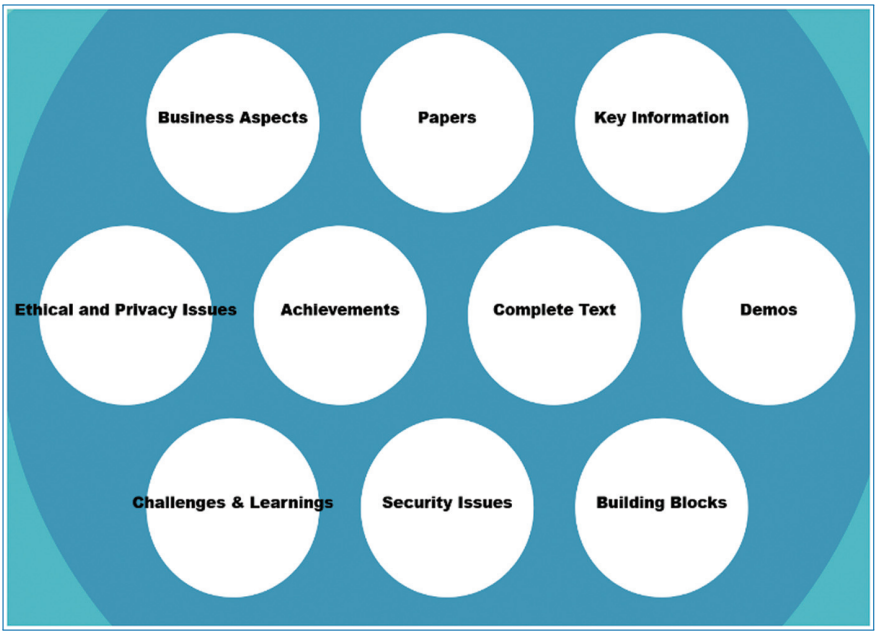

Figure 24.4. Pilot view: structure of contents.

Scientists, Health Insurance, Hospital Decision Makers, Hospital IT/Equipment, Patients, Policymakers, and Privacy Officers. For example, for stakeholders from Hospital IT/Equipment information regarding system architectures, access control, data processing, and learnings were considered relevant. In contrast, stakeholders from Privacy Officers are shown the findings with respect to trustworthy AI, privacy measures, and data protection.

To illustrate the Patient Journey, we constructed a realistic but fictitious medical case, as shown in Figure 24.6, in which a patient goes through the following four stages: General Practitioner (GP), at the Hospital, at Home, and at Follow-Up GP. The Patient Journey view of the Interactive BigMedilytics Website relates relevant studies and BigMedilytics technology to nodes in the sample patient journey. For example, for the At Home stage, the Interactive BigMedilytics Website lists all studies that use technology that supports telemedicine or remote patient monitoring, e.g., the Kidney study (Chapter 9 – eHealth and Telemedicine for Risk Prediction and Monitoring in Kidney Transplantation Recipients) that uses a web app for telemedicine and remote patient monitoring or the Diabetes study (Chapter 10 – Remote Monitoring to Improve Gestational Diabetes Care) that develops an integrated self-monitoring and self-management system for Gestational Diabetes Mellitus (GDM).

The General Learnings view differs from the other views/perspectives. While the other perspectives mostly focus on the presentation of the scientific results from the

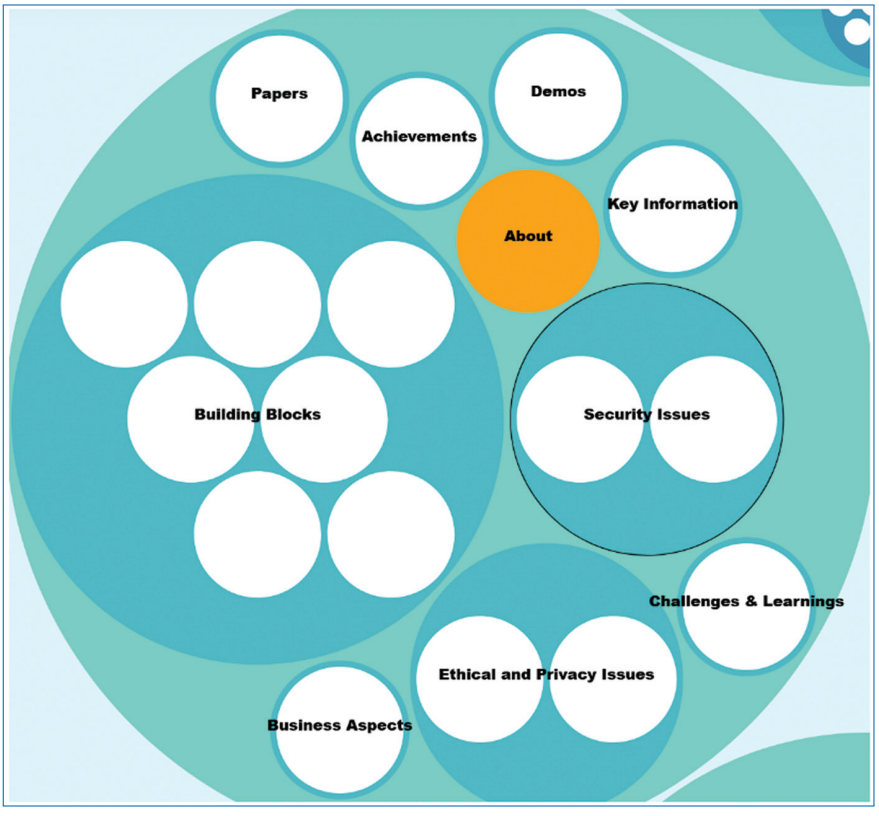

Figure 24.5. Transversal aspects: structure of contents.

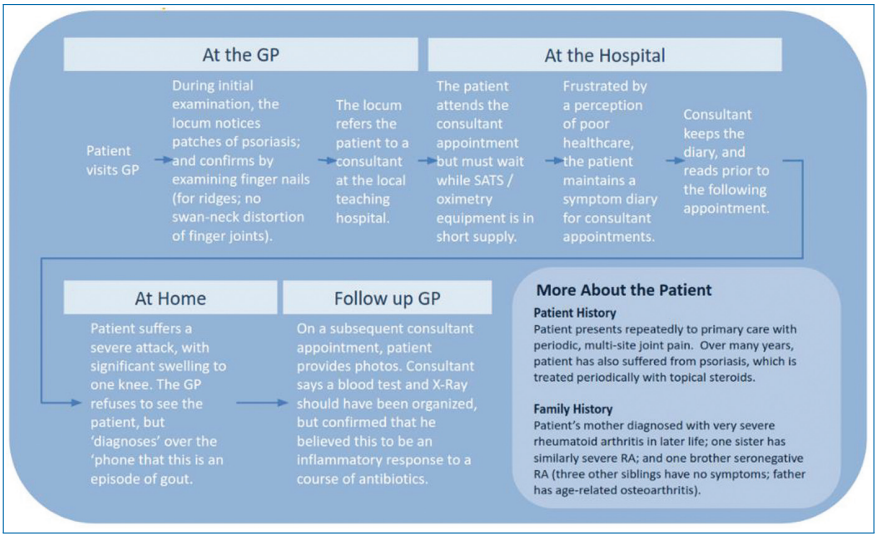

Figure 24.6. Patient journey.

BigMedilytics project and include practical aspects on how to carry out an AI and Big Data technology-based healthcare project, the General Learnings view mainly is a collection of pitfalls to avoid, and lessons learned gathered throughout the realization of the 12 study projects and that may be valuable knowledge for future projects in the area. The General Learnings that are presented on the website are included in Chapter 27 (General Learnings From the Horizon 2020 Project BigMedilytics).

24.4 Conclusion

BigMedilytics resulted in more the 50 scientific publications that demonstrate the disruptive potential of applying Big Data technology in combination with AI in healthcare projects (see Section II, Section III, and Section IV). But BigMedilytics with 36 partners and 12 studies covering several medical and healthcare domains (Population Health and Chronic Disease Management, Oncology, and Industrialization of healthcare services) due to its size also provided a unique opportunity to present a more holistic view and thereby support the implementation of forthcoming projects in the healthcare sector.

In this chapter, we have presented the structure and content of the Interactive BigMedilytics Website: a collection of scientific findings, best practices, and lessons learned from the BigMedilytics project. The Blueprint presents both facts about each of the 12 studies carried out in BigMedilytics, as well as information targeted at specific stakeholder groups, reveals general learnings from planning and executing the studies, and finally relates technology used and findings from the studies to stages in a sample patient journey.

To be able to compare the methods and technologies of different studies, we first asked each study team to describe their study in a document with a pre-defined structure. So, for the website, this information could be split into smaller, comparable chunks for Pilot View and Transversal Aspects. General Learnings from BigMedilytics (see also Chapter 27 – General Learnings From the Horizon 2020 Project BigMedilytics) have been collected partly based on these documents and manually extended by observations made by consortium members. To provide information that is relevant for a specific stakeholder group in Stakeholder Perspective, the content collected has been filtered, so only information is shown that is relevant for the specific group of stakeholders. The Patient Journey stands out a little bit as a realistic but fictitious story about a patient's journey provides the background and defines stages (At the GP, At the Hospital, At Home, and At Follow-Up GP) in the journey and exemplifies illness and treatment of the patient. The four stages are used on the website to explain how technology implemented in the studies can transform healthcare.

To sum up, with the Interactive BigMedilytics Website, we created a unique knowledge source containing facts and findings from the implementation of the 12 Big Data and AI healthcare study projects. Beyond that, cross-cutting topics that concern all project phases, such as ethical and privacy issues and general learnings, are included as well, so future projects in the same domain can benefit and learn from experiences from BigMedilytics and avoid common pitfalls. The content has been prepared in such a way that items of interest can be accessed quickly and easily compared between studies and persons from different stakeholder groups can find relevant information. To achieve this, a modern website design has been selected and implemented using an existing JavaScript library module. While the goal of the BigMedilytics Website mainly has been to provide a holistic view on findings and lessons learned from the BigMedilytics project, the chosen approach can also be used as a template for other lighthouse projects in healthcare or other domains.

The Interactive BigMedilytics Website can still be accessed online (https://www.bigmedilytics.eu/blueprint/). The main parts of the content of the Interactive BigMedilytics Website are presented in this book. Newly updated information regarding a selection of BigMedilytics study projects can be found in Sections II–IV, ethical and privacy issues are discussed in Section I, and general learnings are presented in Chapter 27 (General Learnings From the Horizon 2020 Project BigMedilytics).

Acknowledgement

We would like to thank the team from Inclivia that implemented the interactive webpage.

References

[1] Leopold JA, Maron BA, Loscalzo J. The Application of Big Data to Cardiovascular Disease: Paths to Precision Medicine. Journal of Clinical Investigation 130, Nr. 1. January 2020. 29–38. https://doi.org/10.1172/JCI129203.

[2] Dai H, Younis A, Kong JD, Puce L, Jabbour G, Yuan H, Bragazzi NL. Big Data in Cardiology: State-of-Art and Future Prospects. Frontiers in Cardiovascular Medicine 9. April 2022. 844296. https://doi.org/10.3389/fcvm.2022.844296.

[3] Thongprayoon C, Kaewput W, Kovvuru K, Hansrivijit P, Kanduri SR, Bathini T, Chewcharat A, Leeaphorn N, Gonzalez-Suarez ML, Cheungpasitporn W. Promises of Big Data and Artificial Intelligence in Nephrology

and Transplantation. Journal of Clinical Medicine 9, Nr. 4. April 2020. 1107. https://doi.org/10.3390/jcm9041107.

[4] Zeng XX, Liu J, Ma L, Fu P. Big Data Research in Chronic Kidney Disease. Chinese Medical Journal 131, Nr. 22. November 2018. 2647–50. https://doi.org/10.4103/0366-6999.245275.

[5] Rossetto, AM, Zhou W. Deep Learning for Categorization of Lung Cancer CT Images. IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). 2017. 272–73. Philadelphia, PA, USA. https://doi.org/10.1109/CHASE.2017.98.

[6] Rumbold JMM, O'Kane M, Philip N, Pierscionek BK. Big Data and Diabetes: The Applications of Big Data for Diabetes Care Now and in the Future. Diabetic Medicine 37, Nr. 2. 2020. 187–93. https://doi.org/10.1111/dme.14044.

[7] Price WN, Cohen IG. Privacy in the Age of Medical Big Data. Nature Medicine 25, Nr. 1. January 2019. 37–43. https://doi.org/10.1038/s41591-018-0272-7.

[8] Ferretti A, Ienca M, Sheehan M, Blasimme A, Dove ES, Farsides B, Friesen P, *et al.* Ethics Review of Big Data Research: What Should Stay and What Should Be Reformed? BMC Medical Ethics 22, Nr. 1. December 2021. 51. https://doi.org/10.1186/s12910-021-00616-4.

[9] Roski J, Bo-Linn GW, Andrews TA. Creating Value In Health Care Through Big Data: Opportunities And Policy Implications. Health Affairs 33, Nr. 7. July 2014. 1115–22. https://doi.org/10.1377/hlthaff.2014.0147.

[10] Bakker L, Aarts J, Uyl-de Groot C, Redekop W. Economic Evaluations of Big Data Analytics for Clinical Decision-Making: A Scoping Review. Journal of the American Medical Informatics Association 27, Nr. 9. September 2020. 1466–75. https://doi.org/10.1093/jamia/ocaa102.

[11] Agrawal R, Sudhakaran P. Big Data in Digital Healthcare: Lessons Learnt and Recommendations for General Practice. Heredity 124, Nr. 4. April 2020. 525–34. https://doi.org/10.1038/s41437-020-0303-2.

[12] Hansen MM, Miron-Shatz T, Lau AYS, Paton C. Big Data in Science and Healthcare: A Review of Recent Literature and Perspectives: Contribution of the IMIA Social Media Working Group. Yearbook of Medical Informatics 23, Nr. 01. August 2014. 21–26. https://doi.org/10.15265/IY-2014-0004.

[13] Nielsen J. Designing Web Usability: The Practice of Simplicity. New Riders Publishing, USA. 1999.