Challenges of Trustable AI and Added-Value on Health
B. Séroussi et al. (Eds.)

doi:10.3233/SHTI220387

Scaling AI Projects for Radiology – Causes and Consequences

Gunnar ELLINGSEN[a,1], Line SILSAND[b], Gro Hilde SEVERINSEN[b] and Line LINSTAD[b]
[a] *UIT – The Arctic University of Norway, Tromsø*
[b] *Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø*

Abstract. Artificial intelligence (AI) for radiology has the potential to handle an ever-increasing volume of imaging examinations. However, the implementation of AI for clinical practice has not lived up to expectations. We suggest that a key problem with AI projects in radiology is that high expectations associated with new and unproven AI technology tend to scale the projects in ways that challenge their anchoring in local practice and their initial purpose of serving local needs. Empirically, we focus on the procurement of an AI solution for radiology practice at a large health trust in Norway where it was intended that AI technology would be used to process the screening of images more effectively. Theoretically, we draw on the information infrastructure literature, which is concerned with scaling innovative technologies from local settings, with a limited number of users, to broad-use contexts with many users.

Keywords. Artificial intelligence, radiology, socio-technical, scaling

1. Introduction

AI for radiology has the potential to handle an ever-increasing volume of imaging examinations and thus be a countermeasure against the lack of human radiological resources [1]. AI can support radiologists in many of their core responsibilities, such as scheduling examinations, prioritizing images according to severity, and interpreting images [2]. However, the implementation of AI in clinical practice has not lived up to expectations, as only a few solutions have made it into actual use [2-4]. Existing studies have attributed this mainly to limitations in the technology [5] while organizational issues have basically been ignored [3]. In this paper, we suggest that a key problem with AI projects in radiology is that high expectations associated with new and unproven AI technology tend to scale the projects in ways that challenge their anchoring in local practice and their initial purpose of serving local needs. We don't necessarily attribute this to the project per se, but rather see the scaling as the effect of a combination of internal and external factors. We ask the following research question: What causes scaling of AI projects in radiology, how does it happen, and what are the consequences? Empirically, we focus on the procurement of an AI solution for radiology practice at a large health trust (denoted Health Trust N) in Norway. The trust is one of the largest in

[1] Corresponding author, Gunnar Ellingsen, UIT – The Arctic University of Norway, research group SOF NORD, 9038 Tromsø, Norway; E-mail: gunnar.elllingsen@uit.no

Norway, with about 9,800 employees and is responsible for providing specialist health services for about 500,000 people. The motivation for Health Trust N's ambitions of implementing AI solutions is caused by a steady increase in labor-intensive imaging examinations, estimated at 5–10% per year. Therefore, the trust intended to use AI technology to process the screening of images more effectively. Theoretically, this paper draws on the socio-technical literature, meaning that technology must always be understood in the organizational context in which it is implemented and used. Specifically, we draw on the information infrastructure literature, which is concerned with scaling innovative technologies from local settings, with a limited number of users, to broad-use contexts with many users [6]. This concern is sometimes referred to as the "bootstrap problem." In this regard, it is crucial to understand the organizational factors that drive the scaling process and thereby assess how sustainable the scaling process is.

2. Methods

This study is based on an interpretive research approach [7,8], in which reality is socially constructed among participants. We aim to see the causes for scaling the AI project from the viewpoints of different stakeholders while also considering the broader context. Data are collected through approximately hour-long semi-structured interviews with 20 informants, whereof five are radiologists, seven are hospital managers and eight are managers in collaborating organizations. The data analysis is based on a hermeneutic approach, wherein we understand a complex whole from preconceptions about the meanings of its parts and their interrelationships [7]. We discussed the empirical data thoroughly within the research project, as well as with members of the radiology project, to get a balanced picture of the process. In the interviews, we also looked for recurrent themes and patterns, which we explored further in the interviews that followed. All interviews were audio-recorded and transcribed for analysis.

3. Results

In 2019 Health Trust N applied for innovation funds from its regional health authority to explore the possibilities of implementing an AI solution. The Trust was granted a budget of 1.7 MNOK. Apart from this, the project was expected to use internal human resources, and, most importantly, the future key users: the radiologists from the radiology department. In early 2020, the procurement project started the first phase, preparing the procurement of a test solution. This included clarifying the clinical needs, an associated requirements specification and a strategy for evaluating the solution in clinical practice. The project concluded that AI should be used for the following examinations:

- CT (thorax for lung nodules, pulmonary embolism, and lung metastases, which can make the examination process more efficient and improve survival rates from lung cancer.
- MR caput for multiple sclerosis follow-up is an MR scan of the brain, which relieves the radiologists from a time-consuming examination.
- Conventional X-ray for skeletal X-ray and chest X-ray, which can provide faster diagnostics and sift out irrelevant findings.

The procurement process is aligned with EU regulations for public tender acquisitions and conforms to the principle of competitive dialogue. However, after starting out relatively modestly, the project has gradually scaled in ambition, purpose and outlook. In the following, we elaborate on three different ways the scaling is happening.

3.1. From a local to a national scope

First, while there are several research initiatives related to the development of algorithms for diagnostic imaging in Norway, none of these have been implemented in clinical practice. This paradox is of national concern and therefore the Norwegian Directorate of Health has established a national project named "Better Use of Artificial Intelligence" to facilitate the process towards realizing the benefits of AI in clinical practice. Based on Health Trust N's initiative, the Norwegian Directorate of Health, through its regional health authority, commissioned it to explore how the optimal procurement process of a commercial ready-to-use AI solution for clinical practice could be done most efficiently. This positioned the AI project at Health Trust N as the national pilot for such a purpose. One of the project members said: "We try as best we can to keep the regional health authority up to speed on how we are doing things, so that the next trust in line to implement this can do so better and faster."

Second, several other health trusts became interested in Health Trust N's initiative. As a result, eight of these trusts have joined the framework deal that Health Trust N is negotiating and plan to acquire the same solution if this project is successful. However, the various health trusts all have different Radiology Information Systems (RIS) and Picture Archiving Systems (PACS). This is expected to increase the complexity of integration and interoperability because these systems need to communicate with the AI solutions. The heterogeneous portfolios of RIS/PACS in the health trusts make it more urgent for the project to find a solution that will work well for everyone. One of the project members has asked rhetorically: "There is really mounting pressure in the project … what happens if we can't come up with a solution that works for everybody?" Still, the project perceives the common framework deal for several trusts as a win-win situation, and a risk worth taking because it will be beneficial for the Norwegian healthcare sector. Acquiring an AI solution as part of the same framework agreement therefore means that the trusts can put an AI solution into real use much faster than they would otherwise have been able to.

3.2. From algorithm to platform

From early on, the AI procurement project planned to procure a commercially available CE-marked solution tested in another European country to limit the need to validate the algorithms. In addition, the requirement was that the algorithms would be static, meaning that they should not be learning by themselves, that is, by continuous feedback from datasets in use. In August 2021, the project started the procurement process with five international vendors. One vendor offered single algorithms and the other four offered platform solutions. The vendors offered a variety of AI solutions, all conforming to the six different examinations. The platform vendors offered some of the same AI applications developed by third-party vendors, some apps from different third-party vendors, and some apps developed in-house. The vendor offering only single algorithms had developed these in-house. The AI solutions offered by the different vendors ranged from immature products with weak documentation, both internally and in peer-reviewed publications, to established and mature products that were well documented. The variety of maturity and documentation made it difficult for the members of the project to know which solutions to choose and which one would fit a particular clinical use in a specific organizational context. During the procurement process, the members of the project came to an understanding that choosing applications from different vendors for each of the six use cases would involve more complex technical integration and business processes for the health trust, compared to choosing a platform solution with only one point of integration and only one vendor to address. By choosing a platform vendor, the AI project gained access to several AI applications through a single point of integration to the platform residing in the cloud. If an application did not meet expectations, then it would be easy to replace it with another app from the platform menu. The platform vendor would take care of all legal, ethical and practical collaboration with third-party vendors. One of the project members expressed the expectations of such a platform like this: "We believe that there are some areas where you can reap the benefits straight away, but that the potential is very much greater in the future. And based on this we chose to go for a platform solution."

3.3. From one to several application areas

The project started out rather narrowly, focusing on solving Health Trust N's problem of the lack of radiology resources. However, increasingly, the project is exploring other possible use areas. First, using radiology-related AI in emergency rooms may deal with long waiting lists. AI may then offer preliminary results that enable the triaging of patients. Patients with a low probability of injury or fracture can then be sent home, pending a radiologist's careful examination of the images and a final assessment. The project members emphasize that such solutions have already been implemented at many health institutions in Europe. This could potentially contribute to the more efficient use of resources in the specialist health service and could streamline the operation of the emergency room. Second, AI may potentially improve the practice of referral by general practitioners (GPs) to hospitals by ensuring that the patient's need is described accurately. For example, AI can improve the quality of the referral to the X-ray department so that the correct examination is chosen for the patient's problem. Not all conditions need such advanced examinations as CT and MRI. An MRI examination is quite expensive, plus there are long waiting lists of patients who need this examination. A machine learning algorithm can support the requisition process by assessing the usefulness of certain examinations for given conditions and making this information available to the GP. Third, based on a recent inquiry by the Norwegian Cancer Registry, the AI project is considering how the platform solution can support the Registry's research project on how AI algorithms can make breast cancer screening more efficient. In the Cancer Registry's research project, they have faced quite big challenges in implementing algorithms for testing. In this regard, Health Trust N's platform solution may be a way forward, since it is supposed to be quite easy to add new algorithms.

4. Concluding Discussion

A key concern in the information infrastructure literature is how technology is implemented into real use and how this technology scales from one setting to the next.

This concern is called the “bootstrap problem” since the user community is initially non-existent or very small [6]. One strategy to deal with this is to start with a small and delimited technology, which the targeted users find useful. After the technology is established as a working solution, new user groups can be recruited, and new functionality can gradually be added to complement what already exists, and thus create more value for the users. In this way, scaling takes place gradually and always with a foothold in real use. Also, in our study, the AI project at Health Trust N started quite narrowly with limited ambition, focusing on solving Health Trust N’s problem of the lack of radiology resources. The initial idea was to explore the possibilities of implementing commercial ready-to-use AI solutions conforming to the local needs of Health Trust N. However, increasingly, the project is facing scaling challenges as a result of the involvement of many more stakeholders on different levels and expectations of new application areas. A major challenge is that the scaling happens before there is any real use, which in turn can make it difficult to maintain the interest of the original users – the radiologists [6]. While the project is no doubt trying to keep the interest of the radiologists intact, there are now many other stakeholders that expect to be listened to on both the local and the regional levels. The project must balance these interests carefully, which ultimately means that everybody must give and take, which may cause delays and unforeseen results (at least for some). To face some of the challenges associated with scaling, the project has invested in a platform solution. On the one hand, this seems wise, given that the project is in the situation where it needs to align the expectations of different stakeholders. On the other hand, a platform approach also represents a scaling from Health Trust N’s initial ambition of simply requiring a ready-made algorithm. While perhaps unintended, the consequence is that both new and old stakeholders may see potential for new application areas, which in turn may further scale the ambitions and scope. An illustration is how the Norwegian Cancer Registry sees great potential in the platform. Also, internally, the project considers new application areas in the emergency room and in the GPs’ practices. While these represent long-term prospects, they nonetheless illustrate how easily scaling may happen for the much-applauded technology.

References

[1] Oren O, Gersh, BJ, Bhatt, DL. Artificial intelligence in medical imaging: Switching from radiographic pathological data to clinically meaningful endpoints. The Lancet Digital Health. 2020; 2(9):e486–e488.

[2] Syed AB, Zoga AC. Artificial intelligence in radiology: Current technology and future directions. Seminars in Musculoskeletal Radiology. 2018; 22(5):540–545.

[3] Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: Hindering and facilitating factors. European Radiology. 2020; 30(10):5525–5532.

[4] van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. European Radiology. 2021; 31(6): 3797–3804.

[5] Sogani JAB, Dreyer K, McGinty G. Artificial intelligence in radiology: The ecosystem essential to improving patient care. Clinical Imaging. 2019; 59:8–11.

[6] Aanestad M, Jensen TB. Building nation-wide information infrastructures in healthcare through modular implementation strategies. The Journal of Strategic Information Systems. 2011; 20(2): 161–176.

[7] Klein H, Myers, M. A set of principles for conducting and evaluating interpretive field studies in information systems. MIS Quarterly. 1999; 23: 67-94.

[8] Walsham G. Interpretive case studies in IS research. Nature and method. European Journal of Information Systems. 1995; 4: 74-81.