

Developing or Undermining Expertise in the Shadow of AI?

Marjolaine Rostain, Warwick Business School, Marjolaine.Rostain@wbs.ac.uk

Hila Lifshitz-Assaf, Warwick Business School, Hila.Lifshitz-Assaf@wbs.ac.uk

Nowadays, with the development of intelligent technologies, more and more of our decisions are taken based on algorithms' output. For example, traders may decide their investment based on financial models (Beunza and Stark, 2012), art expert may rely on algorithms to evaluate artist (Sachs, 2020), lawyers may rely on machine learning to support their decision (Klutzn and Mulligan, 2019) or engineers may base their prediction from weather forecast algorithms (Barley, 2015).

However, as AI users are, in most cases, different from AI developers, the algorithm behind AI is opaque for most of them. This opacity in use is problematic when there is a conflict in diagnosis between human and AI. For example, when AI, in health care, diagnoses a tumor when the doctor doesn't, it may be problematic as patient care are highly different. Recent studies (e.g Lebovitz, Lifshitz-Assaf, and Levina, 2022; Anthony, 2021) have highlighted the possibility to question those algorithms in order to understand the underlying assumptions of those *black-box* and to validate or refute the AI diagnosis. While we know that there is the possibility to question AI, it is still unclear how those questioning practices are enacting in practices and what are the consequences on workers' knowledge. Some may argue that AI will automate workers and decrease their ability to learn and develop new knowledge whereas others will argue, on the contrary, that AI will help workers to develop new knowledge. In this paper we ask "How AI is enacting in practices and What are the consequences on workers' knowledge?"

Based on an ethnography study of doctor doing colonoscopy with the help of AI, we found two distinct paths that doctors use to decide which diagnosis is the right one. In one path, they engage in a questioning process that requires to describe and discuss their actual medical knowledge with senior doctors. In the other path, doctor do not engage in questioning practices and integrate AI diagnosis in their decision in a blinded manner. AI outcomes in the questioning group strengthens doctors' knowledge and help to develop confidence in their own diagnosis whereas AI outcomes in the blinded group impede their own judgement and do not provide opportunities to develop new knowledge.

Research Setting

The setting is a public hospital in France, in a gastro-enterology department, where doctors do endoscopy procedure to detect and eventually remove polyps. The department is a leading department in Europe with well-known experts in endoscopy. It is composed of three professors, nine physicians, five assistants (equivalent of a two-year post-doctoral position; they are young doctors who have just defended their thesis), about ten residents (changing every 6 months) and several medical students (who are not engaged in endoscopy procedure).

To do the endoscopy, the patient is asleep in an operating room and in this room, there are at least three medical staffs: the gastroenterologist, the nurse who helps the doctor during the procedure and, the anesthetist nurse who cares about the patient's sleep. The standard procedure consists of inserting inside the patient, a tube with a camera at the end. On a screen, in front of the doctor, he sees what is happening inside the digestive tract. At the same time, the doctor is going up and down inside the digestive tract, he is looking for polyps. He needs to look at each area of the digestive tract, from the rectum to the intestine, to be sure not to forget any space. This standard procedure is called the "white light" as the light used to see inside the digestive tract is a normal light.

Before AI arrives, they had two other technologies that help them to better detect and characterize polyps. The first one the "blue" is a procedure consisting of inserting a blue liquid inside the colon. This procedure helps to better see the relief between the normal and the abnormal mucosa (polyp). While this procedure significantly increases the detection and characterization of polyps, it also increases the duration and the cost of the endoscopy. For this reason, the procedure is not automatically used by doctors but rather used at discretion. Another technology called "NBI" is also widely used to help doctor characterizing polyp. NBI (Narrow Band Imaging) option is just a button to press on the endoscopy itself which provides a new visual as it does not use the whole spectrum of white light but only certain colors. When using the NBI option, the polyp can be better characterized but at the same time these colors are really tiring for doctors, and they do not use it all the time.

In 2020, the hospital received a new AI tool for helping doctors in their detection of polyps. The AI is based on a neural network and has been trained to recognize polyps. The material used and the procedure is the same as the usual procedure. The only difference resides in the appearance on the screen. In the operating room where there is AI, polyps appear highlighted with a green square whereas in operating room where there is not AI, polyps are not highlighted. In the operating room where there is AI detection, the doctor has nothing to do to add AI diagnosis, it appears instantaneously on the screen.

In 2022, the hospital has received another AI tool for helping doctors in the detection but also in the characterization of polyps. This AI is located in another operating room than the one that does only the detection. The difference with AI detection is that once the polyps is detected and highlighted with a green square, it is also written on the top if it, if it is an adenoma (malign polyp) or not. This is crucial for doctors as they could not remove all the polyps due to time and money concerns.

This setting is highly interesting to study the enactment of questioning practices as AI is used as the primary technology to detect and characterize polyps and when there is a conflict between doctor's diagnosis and AI diagnosis, they can use other technologies to validate or refute AI diagnosis.

Data Collection

This study is based on ethnographic methods combined with interviews. The ethnography consists of being in the operating room (without AI, with the AI detection tool and, with the AI detection and characterization tool) between 2 and 3 days a week, observing doctors and nurses in doing the endoscopy. Each day, a doctor, a nurse and an anesthetist's nurse are affected to a specific operating room by the administrative manager. Between 4 and 6 patients can be examined during that day.

The ethnography consists of being in the operating room, taking field notes about all the exchange, gesture and actions that the doctors and the nurses are enacting. The ethnography started in January 2022 and will continue until June 2023. After each day of observation, the fields notes are transcribed in a word document and a memo is written. In October 2022, 154 patients have been observed. In addition to observation, interviews were also conducted with doctors and nurses. Initially, interviews were done outside of operating time, asking doctors and nurses question about their practices and their use of AI. After 5 interviews we ended up with no clear answer. The doctors themselves, seems to be unable to explain why they are using some technologies and no other, how and why they validate or refute AI diagnosis. For this reason, we decided to interview them in a different manner. After each colonoscopy we have asked doctor to reflect on their behavior and explain why they have taken this decision.

Data Analysis

We started the analysis in a really open manner (Locke, Feldman, & Golden-Biddle, 2020) by rereading our field notes looking for similarities and differences between the different technologies used using a constant comparative grounded approach (Glaser and Strauss, 1967). We soon realized that there was a huge difference in the enactment of AI. This difference was not only dependent on doctor's expertise but also on the type and location of the polyps the doctor needs to look at. We decided to use as our unit of analysis each polyp detection. We went back to our data and started coding line by line each polyp detection. For each polyp detection, we coded the interaction, behaviors, gestures of doctors and nurses. From this first coding, two groups of doctors emerged: those who engage in a questioning process that requires to describe and discuss their actual medical knowledge with senior doctors and those who do not engage in discussion process and integrate AI diagnosis in their decision in a blinded manner. AI outcomes in the questioning group strengthens doctors' knowledge and develop confidence in their own diagnosis whereas AI outcomes in the blinded group decrease workers' confidence in their own diagnosis and do not provide opportunities for those workers to increase their knowledge.

Findings

When doctors disagree with AI diagnosis, we found that they either engage in a questioning process or engage in a blinded agreement process.

Questioning Group

To better illustrate how workers engage in a discussion process we provide an example. Victor, an assistant (recently graduated doctor), is doing the colonoscopy when he suddenly sees something in the screen that looks like a polyp. He does not know what it is and the AI highlights it and says « adenoma » (first stage of a malign polyp). He looks at the resident next to him and says « it's strange, it's really plane, there is a strange area there ». At that moment, Jérôme, an expert doctor arrives and asks where are they in their planning? Victor does not answer but says that he has something strange on his screen. Jérôme enters the room, looks at it and tells Victor that he has to zoom in to better characterize the polyp. Jérôme explains Victor how to zoom it with the endoscope. There he says “Now you have to analyse the pit pattern [meaning the shape, colors and size of the polyps]”. Jérôme points one area on the screen and says “Here there are a lot of large area but down it looks more like a scallop polyp (one type of adenoma)”. He asks Victor to continue to zoom in. Once Victor has zoomed in, he says “we can say that it is not inflammatory, it's not a cancer neither” and he continues “it's a hyperplastic[benign polyp], it is not a scallop polyp”. He says again to Victor that he has to zoom in to classify the polyp. Victor thanks him and Jérôme get out of the room.

In this example, the disagreement with AI diagnosis, has provided an opportunity for Victor to enter in a questioning process and to strengthen his own knowledge. As he disagreed with AI diagnosis, he started talking with the resident and then with a senior doctor. Through this discussion the senior doctor provides Victor some tips to better characterize the polyps such as zooming in but also describe what he thinks the polyp is and with this description he teaches Victor new knowledge. At the end of this episode Victor has learnt how to characterize a malign polyp and how to differentiate it with a hyperplastic one.

Blinded agreement

To illustrate the second path, « the blinded agreement » we will also take an example. Alin, a resident, is doing a colonoscopy and he is now in the rectum. He sees a polyp that he considers as a hyperplastic (benign polyp) whereas AI says it is an adenoma (potential malign polyp). He asks, Paul, an assistant to come to look at it. Paul validates Victor diagnosis and says « for me it is a hyperplastic but as AI has said adenoma, in the doute, we should remove it”. Alain removes it and sees other polyps in the rectum that he considers as hyperplastic. Victor says this is normal as in the rectum there are often several hyperplastic polyps next to each other. Again AI says that there are adenoma. Paul looks at it and says « Frankly it's a hyperplastic but we remove it anyway”.

In this example, Alain and Paul do not engage in a questioning process and just accept AI diagnosis and remove the polyp. Contrary to the previous example, they do not try to describe the polyp and to confront their own diagnosis with the one of AI but just consider that they have to remove it. This path impedes doctor's own knowledge and decrease their

own confidence as explained by Paul after the colonoscopy when he said during the interview “Frankly it’s really disturbing..., we do not know anymore with AI”

Contributions

This paper is for now at an early stage. However, we think that we could contribute to better understand how AI technologies are enacted in practice and how does it change knowledge and learning at work. We highlight in this paper that the augmentation/automation dilemma about AI may be too simplistic and we propose to analyse deeper those process of augmentation and automation. We have found that, in some cases, AI can augment the efficiency of the task whereas at the same time decrease the knowledge of the worker who do that task. In other cases, AI can augment both the efficiency of the task and the knowledge of the worker.

References

Anthony, C. (2021). When knowledge work and analytical technologies collide: The practices and consequences of black boxing algorithmic technologies. *Administrative Science Quarterly*, 66(4), 1173-1212.

Barley, W. C. (2015). Anticipatory work: How the need to represent knowledge across boundaries shapes work practices within them. *Organization Science*, 26(6), 1612-1628.

Beunza, D., & Stark, D. (2012). From dissonance to resonance: Cognitive interdependence in quantitative finance. *Economy and society*, 41(3), 383-417.

Kluttz, D., & Mulligan, D. K. (2019). Automated decision support technologies and the legal profession. *SSRN Electronic Journal*.

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126-148.

Locke, K., Feldman, M., & Golden-Biddle, K. (2020). Coding practices and iterativity: Beyond templates for analyzing qualitative data. *Organizational Research Methods*, 109442812094860.

Glaser, B., & Strauss, A. (1967). *Discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing.

Sachs, S. E. (2020). The algorithm at work? Explanation and repair in the enactment of similarity in art data. *Information, Communication & Society*, 23(11), 1689-1705.