Knowledge Capture for the Design of a Technology Assessment Tool

Daniela Oliveira1a and Kimiz Dalkir2b
1Independent Researcher
2School of Information Studies, McGill, Montreal, Canada

Keywords: Knowledge Management, Artificial Intelligence, Technology Assessment, Artificial Intelligence Assessment, Applied Knowledge Management, Artificial Intelligence Documentation.

Abstract: The design of technology assessment tools is an important Knowledge Management endeavour. Technology assessment is a subject where consensus is far from being achieved. Any project intended to create a technology assessment tool is expected to generate a lot of discussion or criticism. Among the most critical kinds of technology, Artificial Intelligence (AI) is a highly polemic kind of technology. Its impacts are important and multidisciplinary. Moreover, the technology evolves quickly and so do the attitudes toward that technology. Therefore, business owners intending to produce an AI assessment tool should expect extensive discussion of different points of view, but also support the continuation of the discussion throughout time and with different stakeholders. Surprisingly, technology assessment tools developed by business owners have been particularly neglected in the coalescent discussion about AI documentation, not to mention the support to create those tools. To foster a continuous innovation flow, business owners should pay particular attention to how discussions are captured. This paper explores the foundations of knowledge management initiatives to support the design of an artificial intelligence assessment tool at the business owner, in a process that supports continuous discussion and innovation. This article also suggests project aspects and supporting document structure.

1 INTRODUCTION

Making people exchange ideas, validate knowledge, and create new knowledge together are some of the challenges from the Knowledge Management field. The field is also interested in helping experts produce consumable documents and, for that, going through the process of deciding which information should be retained and which should be left out. A challenge that does not spare documentation efforts around Artificial Intelligence (AI). “Determining what information to include and how to collect that information is not a simple task”, argued Richards et al. (2020,p. 1), while designing a document structure intended to support reporting about AI services.

Documentation challenges involve identifying what knowledge is mature enough to be written down (such as a new methodology that has been tried out enough times to have an article written about it) and articulating knowledge that has not been yet expressed (such as the acknowledgement of different roles and expectations in a new process).

Documentation is particularly challenging when it involves knowledge about the design phase. Design is a project phase where several ideas are articulated. In this perspective, Design is also a knowledge process. In knowledge processes, some ideas are retained, others are not (McElroy, 2011). Design is a process very rich in terms of information about the end product. In this phase, the values surrounding the project take shape and the motivation behind the project defines itself. This definition phenomenon is maybe more visible when the discussions touch different knowledge domains, as it happens to be the case for AI evaluation tools.

AI academicians and practitioners are in the beginning of some sort of consensus regarding what to evaluate in AI systems, and when and by whom. The design of an AI evaluation tool is therefore a good candidate for heated discussions involving the interaction of concepts from different fields and
having different impact levels on different people at different stages of the AI lifecycle. This study intends to help the creation of a support structure to foster the design of a technology assessment tool. It can be particularly helpful for business owners to develop AI assessment tools.

Documentation in the design phase may capture knowledge at its state-of-the-art portrait at that time. It may capture motivation definition. The capture of the state of the knowledge at the design phase might help the end product evolve throughout time, as that knowledge also evolves, because dependencies on outdated knowledge can be more quickly identified. The capture of the motivation behind the project, in addition to increasing its transparency levels, might also help the end product evolve, as this product acceptability increases, for instance. In this sense, documentation in the design phase might help awarding the end product a continuous innovation flow, where incremental developments have their barriers lowered.

In the AI ecosystem, technology and approaches evolve quickly and so does the acceptability of the resulting products. The field is the perfect candidate for the adoption of documentation facilitating a continuous innovation flow.

2 CONVEYING KNOWLEDGE ABOUT AI

Richards et al. (2020) argue that the diversity of information needs that different stakeholders might have makes it impossible that one single document addresses all needs in a consumable format, even within the same domain or organization.

For example, Mitchell et al. (2019) have suggested a documentation paradigm to describe a machine learning model. In the short documents produced according to this documentation paradigm, named “model cards”, performing characteristics of the model should be conveyed so that potential users can understand the systematic impacts of the model before its deployment. Information such as type of model, intended use cases, attributes for which model performance may vary, measures of model performance, as well as the motivation behind chosen performance metrics, group definitions, and other relevant factors should be included. Mitchell et al. (2019) state the tool intends to help stakeholders to compare candidate models, understand the limits of each model and better decide on which model is more suitable for a given situation. In practice, the definition of stakeholders in this case is somehow limited. While the tool should “aid policy makers and regulators on questions to ask of a model, and known benchmarks around the suitability of a model in a given setting” ([p.2]), the target audience is developers, particularly those interested in including the model in a larger technological solution. For Richards et al. (2020), if the documentation is to be useful, it has to be tailored to their target audience and to the use this target audience is to make of the product. Indeed, while in Mitchell et al. (2019) there is a concern regarding the length of the document (the models cards should be “short”), in Richards et al. (2020), the perspective of reporting is changed to suit the needs of developers that would include models in a larger technological solution; instead of reporting characteristics of the model, characteristics of the AI services, that could include many models, are reported.

It is therefore reasonable to expect that the documentation support needed for reporting characteristics of an AI model or an AI service should be different from the documentation support needed for assessing the suitability of that AI model or service. This assessment should evaluate the alignment of that model or service with other criteria, for example, the policies and practices in an organization.

No documentation approach seems to cover the whole machine learning cycle, neither to address all the needs of all audiences (Garbin & Marques, 2022; Laato et al., 2021).

2.1 Technology Assessment

Technology assessment involves, but is not limited to, approaches and tools that allow:

- The evaluation of the suitability of a technological solution to a particular situation or business need;
- As mentioned, the alignment of the technological solution with the policies and practices of the organization;
- The evaluation of positive and negative, intended and unintended, current and future impact on the situation, people, environment and other technological solutions;
- The comparison of one technological solution to another;
- In a larger spectrum, the comparison among approaches or technologies.
2.2 Artificial Intelligence Applications Assessment

AI applications may have high positive impact, but they also have a great risk of generating negative impact. AI applications may “violate privacy, discriminate, avoid accountability, manipulate and misinform public opinion, and be used for surveillance” (Janssen et al., 2020), among other risks.

Every professional involved in the AI cycle has a responsibility towards increasing AI transparency and limiting its potential to cause harm, be they AI producers, regulators or executive board members and managers of organizations of organizations making use of AI (Laato et al., 2021). On one hand, it is necessary for AI producers to better report the characteristics, uses and limitations of AI models and services. On the other, well-designed and well-suited governance approaches to AI are necessary, to define and monitor its potential negative implications and limit those implications with effective and timely responses to incidents.

AI risk assessments are necessary at many levels. At the level of application domains and at the institutional level (Winfield & Jirotka, 2018) and at the level of individual systems (Janssen et al., 2020; Winfield & Jirotka, 2018). Even if the AI solution in question has an explainable AI approach, what to explain and how to explain it might differ from one domain to another and from one organization to another (Laato et al., 2021). In addition, solutions containing explainable AI modules still must be monitored, as “blindly trusting findings from any usability research in the XAI field would be counterproductive due to the novelty and formative state of the research area” (Laato et al., 2021, p. 20)

In each domain or organization, regulations, culture and then policies, principles and procedures are mechanisms for the establishment of thresholds of acceptable behavior, mechanisms that both influence and are influenced by societal expectations, norms and values (Janssen et al., 2020). How can these mechanisms be used for the creation of AI technology solutions assessment tools? An approach and a tool example from Knowledge Management research and practice follows.

3 KNOWLEDGE MANAGEMENT

Knowledge Management concerns all questions regarding the acquisition, the development, the sharing, the exploitation, and protection of knowledge (Dalkir, 2011).

Applied Knowledge Management is about the development and tailoring of initiatives and tools from Knowledge Management regarding a particular field or activity.

3.1 Applied Knowledge Management and Technology Assessment

The creation of approaches and tools is some ways no different from other creativity endeavour in a business environment. The ideas to be generated need motivation, expertise and creative skills at their origin and are required to be “appropriate, useful and actionable” (Amabile, 1998, p. 79). Creative work is often expected to be developed in groups (Hennessey & Amabile, 2010) and work involving knowledge is often linked to connectedness (Nahapiet, 2009). It is possible that collaborative work increases the flexibility and robustness of the solution. In any case, it includes different perspectives and has the potential to increase buy-in (Oliveira, 2022).

Applied Knowledge Management can support this process by removing possible roadblocks and otherwise creating conducive conditions so that better goals can be attained more quickly, in addition to providing individuals, groups and organizations with a positive experience.

3.2 Knowledge Management and Artificial Intelligence

Knowledge management around the evaluation of technological solutions using artificial intelligence tends to raise challenges that may not be raised in the evaluation of other technologies. Some of these challenges are:

- The multidisciplinarity of fields required for the evaluation. Portraits of Artificial intelligence solutions have raised social, economic, technological, linguistic, ethical, legal, management and philosophical issues, to name only some;
- The global nature of collaboration: research from academia and from companies around the world are mutually influenced by new developments in the field;
- The field is still in its early stages.

Knowledge management initiatives supporting the development of technology evaluation approaches and tools for technological solutions involving artificial intelligence must then take into account...
collaborative work among professionals with a plurality of backgrounds and a high level of knowledge acquisition and development.

3.3 Knowledge Development and Innovation

Development, sharing and exploitation of knowledge are processes strongly related. Knowledge development is associated with innovation, as the creation of new knowledge has the potential to propel the organization into new venues. While the development of knowledge or of new ideas can be done individually, more and more frequently this process is undertaken in groups (Carrier & Gélinas, 2011; Fisher & Amabile, 2008). Knowledge sharing is then a process that influences knowledge development. Among other reasons, knowers might share developed knowledge in order to validate this knowledge (Mokyr, 2000), a process that also occurs with knowledge acquired by an individual outside the organization. Knowledge validation is necessary for the subsequent application of this knowledge. Once the knowledge has been embedded in processes, services or products, it can be said to be exploited. In the case of the evaluation of technological solutions involving Artificial intelligence, knowledge surrounding artificial intelligence, technology evaluation and relating themes must be sought outside the organization or developed internally and then validated. These processes might occur before or during the process of design of an actual technology evaluation approach or tool.

3.3.1 Supporting Knowledge Acquisition

Knowledge from outside the organization can be acquired through a structured organizational initiative, but it can also enter the organization through an employee that acquired that knowledge on their own (Shoham & Hasgall, 2005). This employee may act as a sponsor of this knowledge and advocate its integration into the organizational knowledge.

There are many initiatives that can support knowledge acquisition. Direct support can include providing access to academic resources or training and allowing employees time to explore those resources. The knowledge acquisition process can also be supported indirectly through the organizational endorsement of the whole knowledge management cycle, particularly development, sharing and exploitation stages. If employees are allowed and encouraged to validate knowledge externally acquired, they will feel also encouraged to seek future acquisition of knowledge.

One important element when validating knowledge that was acquired outside the organization is to acknowledge its provenance. Provenance holds a symbolic weight that might be useful when advocating for the acquired knowledge. This symbolic weight might indicate, among other aspects, maturity of scholarship, interdisciplinary points of view, importance of the subject or the practical potential of the knowledge in question. It is therefore interesting to include provenance in knowledge management tools designed to support the validation of acquired knowledge.

3.3.2 Supporting Knowledge Development

Knowledge developed inside the organization might combine acquired knowledge with previously internally developed knowledge. Therefore, knowledge management tools supporting knowledge development should include the elaboration of the new idea or statement and the possibility of mentioning the provenance of both externally acquired and internally developed knowledge.

3.3.3 Supporting Knowledge Application

Once a particular knowledge claim has been validated (McElroy, 2011), it is then time to validate the application of this same knowledge claim. Statements explaining the application of knowledge tend to be prescriptive and respect practical constraints. They are therefore different in nature from the statements describing the knowledge at their origins, which can be more abstract or general.

In the design of an evaluation tool, the presentation of the knowledge acquisition or development statement beside the knowledge application statement allows the reader to understand the reasoning behind the application statement and imply the organizational constraints that were considered along with the knowledge acquired or developed. It is the knowledge about knowledge, or metaknowledge, helping the understanding of the knowledge itself.

The promotion of the understanding of the design process, beyond the end result, is one of the most important elements in supporting the creation of evaluation approaches and tools of technological solutions involving artificial intelligence.

As the technology evolves and more of its impacts and possible mitigation solutions are known, it is important to facilitate the identification of which areas of the evaluation tool are less current or
applicable and should be revised. Therefore, while the understanding of the design process as a whole is to be supported, the understanding of the design process of each application statement is equally important. It is also important to consider that the team charged with the revision of the evaluation tool might be considerably different from the team that developed the tool. Facilitating the granular understanding of the design process supports the updating process of the evaluation tool in general and the increasing diversity in the revision team in particular.

The support to the granular understanding of the design process reduces the need for the revision team to understand all the aspects of the project. This facilitates the collaboration with experts in different domains and reduces the pressure on the tacit knowledge of the team members.

3.3.4 Supporting Explicit and Tacit Kinds of Knowledge

The term “tacit knowledge” refers to knowledge that has not been expressed in any kind of language (Nonaka & Takeuchi, 1995). “Explicit knowledge” addresses the knowledge that has been articulated to the extent that it can be understood without needing direct access to the holder of that knowledge. If a continuum of the media holding knowledge is considered, human minds would be on one extreme whereas documents would be on the other (Oliveira et al., 2021).

Explicit knowledge can be supported through fields where the provenance of the knowledge claim can be indicated. Supporting tacit knowledge requires different strategies. One of those is a field where the name of an employee sponsoring the knowledge can be codified, as well as the name of an employee endorsing the proposition. This strategy credits employees with their efforts in the design of the evaluation tool, for instance, and carries the symbolic weight of their expertise.

4 CODIFICATION EFFORTS OF KNOWLEDGE AND METAKNOWLEDGE, AI AND INNOVATION FLOW

Documentation, in most technological solutions, focuses on the resulting tool. It is intended to accompany the tool and help client developers make good use of the tool. This documentation will usually cover only the application of the knowledge claims that have been validated in the design of the tool. The documentation would not articulate the knowledge claims nor describe the knowledge and processes involved in validating those claims. In other words, it would only present the knowledge itself, and not the metaknowledge surrounding the technological solution. After all, the aim of the documentation is to support use of the tool, not necessarily the development of the knowledge involved in designing the tool.

Transparency in AI applications require that codification efforts for the client developer go a little further, both in terms of the knowledge and of the metaknowledge surrounding the solution. In terms of the knowledge, AI documentation developed by its producer should cover how the evaluation the AI model or service went through before it was made available (Mitchell et al., 2019; Richards et al., 2020), as a part of its quality control or application limits definition. In terms of metaknowledge, the documentation should cover the motivation behind the choice of the metrics used to evaluate the AI application (Mitchell et al., 2019). There is then a need to promote codification efforts of this metaknowledge, even for the documentation intended for client developers’ use. The process supported by the documentation suggested in this communication, however, is the assessment of AI applications performed by the business owner, or the first client of AI applications.

4.1 The Business Owner’s Assessment Tool

The business owner’s assessment team should be composed of developers and members from other functions in the business working together. Some portions of the assessment would require more technical expertise, while other portions would need a broader view of the business to assess impact and alignment, but most of the assessment will need collaboration among professionals of different backgrounds. The AI application would be evaluated and compared with other applications with respect to its technical approaches, but the alignment of the AI application with the domain and the business culture and with the policies and regulations surrounding the business and the intended application environment must also be ensured.

Reaching coherent decisions in this diverse knowledge environment is a challenge, particularly if the members of the assessment team change during the assessment process. Documentation supporting the knowledge acquisition, development and
application processes would help the assessment team remain coherent throughout the assessment. Along the same lines, this documentation can be especially helpful when the organization is ready to transform the ad-hoc AI assessment process into a structured and more established one and starts the design of an AI assessment tool to guide the process.

4.2 Documentation Granularity at the Level of Knowledge Claims

When a technology assessment tool is designed, the knowledge claims surrounding the tool are validated or rejected. The validated claims will join the design process, and they will most likely be embedded in the resulting tool.

In the case of the design of an AI assessment tool, one approach would be to have the documentation cover the assessment tool as a whole. This approach gives a broad view of the assessment tool as is probably the best option for the purposes of training of new members of the assessment team or to present the work of the assessment team to other departments of the organization.

A more granular approach is however necessary when knowledge surrounding AI evolves and knowledge claims have to be revaluated. In this case, it is interesting to quickly and unambiguously identify the knowledge claims that might be affected by new knowledge. In this way, barriers for a continuous innovation flow are lowered and the updating process of the assessment tool is facilitated.

In these codification efforts aiming at a continuous innovation flow, the connection between knowledge and metaknowledge must be made clear. The documentation efforts should present which portions of the end product are connected to which knowledge claims, and the knowledge surrounding the validation of those claims. Some documentation of the social capital, “the sum of the actual and potential resources embedded within, available through, and derived from the network of relationships possessed by an individual or social unit” (Nahapiet & Ghoshal, 1998, p. 243), involved would also help understand the importance of knowledge claims. The documentation suggested in this communication is part of the codification efforts that support the evolution of the knowledge surrounding AI applications. The documentation suggested enforces codification efforts of the knowledge claims validation process, covering the knowledge management phases of acquisition, development and application and both tacit and explicit knowledge.

5 CAPTURING ASPECTS OF THE KNOWLEDGE CLAIM

An AI assessment tool encompasses a number of criteria used to identify the degree in which a particular AI application presents interesting features. Each criterion comes from a knowledge claim that was validated. The documentation suggested captures aspects of the knowledge claim validation process.

The principle is the expression of the knowledge claim.

The motivation shows the context in which the knowledge claim was acquired or developed in the organization.

The academic / legal references present artifacts of external knowledge, such as journal articles and proceedings, books and book chapters, law text and somehow captured (in documents, emails, audio or video files, for example) legal advice.

The mentions field provides a space for a description of the knowledge claim internal validation process: the mention of the knowledge claim in conferences, formal or informal discussions in which members of the assessment team or executive board members took part.

The previous use field offers the possibility to codify the identification of the knowledge claim in benchmarking efforts, either internal or external to the organization.

The criterion is the short sentence that is an actual part of the assessment tool. It aims to assess a particular aspect of the AI application. It is the actual expression of the knowledge claim in the assessment tool.

The application field offers a space for alternative ways to express practical aspects of the knowledge claim.

The fields proposed by and seconded by capture a little of the social capital surrounding the knowledge claim validation process, as they should present the names of assessment team members that sponsored the knowledge claim.

The decision field adds to the organizational memory as it captures the result of the knowledge claim validation process, whether it was retained, rejected or if the group has not reached a decision about it yet.
Table 1: Fields of the suggested documentation.

<table>
<thead>
<tr>
<th>Proposition sheet field</th>
<th>Values</th>
<th>KM cycle process or kind of knowledge supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle</td>
<td>The knowledge claim</td>
<td>Knowledge acquisition or development</td>
</tr>
<tr>
<td>Motivation</td>
<td>The inspiration or reasoning behind the knowledge claim</td>
<td>Knowledge acquisition, development or application</td>
</tr>
<tr>
<td>Academic / legal references</td>
<td>External knowledge artifacts supporting the knowledge claim</td>
<td>Knowledge acquisition or development</td>
</tr>
<tr>
<td>Mentions</td>
<td>Mentions of the knowledge claim in the media, conferences or internal discussions</td>
<td>Knowledge acquisition or development</td>
</tr>
<tr>
<td>Previous use</td>
<td>Presence of the knowledge claim in a benchmarked assessment tool</td>
<td>Knowledge application</td>
</tr>
<tr>
<td>Criterion</td>
<td>The knowledge claim applied to a solution</td>
<td>Knowledge application</td>
</tr>
<tr>
<td>Application</td>
<td>More context about the application of the knowledge claim</td>
<td>Knowledge application</td>
</tr>
<tr>
<td>Proposed by</td>
<td>Employee advocating for the knowledge claim</td>
<td>Tacit knowledge</td>
</tr>
<tr>
<td>Seconded by</td>
<td>Employee supporting the knowledge claim</td>
<td>Tacit knowledge</td>
</tr>
<tr>
<td>Decision</td>
<td>The result of the knowledge claim validation process</td>
<td>Knowledge development</td>
</tr>
</tbody>
</table>

6 CONCLUSIONS

Documentation helps to foster transparency. To build trust in Artificial Intelligence solutions in general, documentation is needed in many levels (Winfield & Jirotka, 2018) and in many steps of the AI cycle (Richards et al., 2020).

This communication explored the documentation to be developed by the business owner regarding the assessment of AI applications. Technological assessment, in general, is influenced by characteristics of the application domain, of the organization and of the category of the solution in question. Intellectual work is then necessary to make sure these characteristics are included in the assessment tool design. This intellectual work has to also be collaborative, as expertise from different backgrounds is necessary to evaluate technological solutions not only from a technological viewpoint, but also from the organization’s mission perspective, in addition to a management perspective.

The negotiation process of what to assess and how can be seen as a knowledge process. In this knowledge process, a knowledge claim is advocated, supported, defended, discussed, sponsored, rejected or, in some cases, just left aside until a consensus among team members can be reached.

Supporting the validation process of knowledge claims during the design of assessment tools has the benefit of providing a map of the knowledge dependencies of the end product, in this case, the technology assessment tool.

The diversity of knowledge involved in the design of a technology assessment tool already justifies the documentation support of the validation process of knowledge claims. However, the design of a technology assessment tool for an AI application also involves the consideration of multifaceted impacts, that have to be considered under the light of different disciplines, which adds to the complexity of the knowledge claim validation process. In addition, AI technology and the understanding of its impacts evolve quickly. In this scenario, being able to quickly identify knowledge dependencies on outdated knowledge helps to keep documentation updated, by triggering a new knowledge claim validation process.

For these reasons, a template to codify aspects of the knowledge claim validation process is suggested. The document provides fields that capture elements of the knowledge management steps of knowledge acquisition, development and application. It also fosters ways to capture the transformation of tacit knowledge into explicit knowledge and of the social capital involved in the knowledge claim validation process.

REFERENCES


