6-18-2022

THE LIFE CYCLE OF DATA LABELS IN ORGANIZATIONAL LEARNING: A CASE STUDY OF THE AUTOMOTIVE INDUSTRY

Joscha Eirich
University of Bamberg, joscha.eirich@uni-bamberg.de

Diana Fischer-Pressler
University of Bamberg, diana.fischer-pressler@uni-bamberg.de

THE LIFE CYCLE OF DATA LABELS IN ORGANIZATIONAL LEARNING: A CASE STUDY OF THE AUTOMOTIVE INDUSTRY

Research Paper

Joscha Eirich, University of Bamberg, Bamberg, Germany, joscha.eirich@uni-bamberg.de

Diana Fischer-Preßler, University of Bamberg, Bamberg, Germany, diana.fischer-pressler@uni-bamberg.de

Abstract

Data labels are an integral input to develop machine learning (ML) models. In complex domains, labels represent the externalized product of complex knowledge. While prior research discussed labels typically as input of ML models, we explore their role in organizational learning (OL). Based on a case study of a German car manufacturer, we contextualize a framework of OL to the use of labels in organizations informing about organizational members who work with labels, requirements of label-based tools, label-related tasks, and impediments of label-related task performance. From our findings, we derive propositions about the role of labels in OL and outline future research opportunities. Our results inform theory about the role of labels in OL and can guide practitioners leveraging labels to create and transfer knowledge within organizations.

Keywords: Knowledge Management, Organizational Learning, Machine Learning, Human-Computer Interaction.

1 Introduction

Organizations increasingly use data-driven approaches such as machine learning (ML) methods to improve organizational processes. For example, in industrial manufacturing, ML models analyze hundreds of produced parts and predict their quality (Suschnigg et al., 2020); in engineering, ML models use multi-criteria alternatives for rotor designs and suggest the most promising ones to human experts (Cibulski et al., 2020); or, in traffic control, models based on transportation data predict traffic flows (Boukerche and Wang, 2020). All these ML approaches have one essential key ingredient in common to train and evaluate sophisticated supervised and reinforcement learning models: *labeled* input data (hereafter referred to as labels) (Afiouni, 2019) and features (Bishop, 2006).

This study focuses on labels in the context of supervised and reinforcement learning. Such labels are typically created by humans and describe data categories into which data sets are classified. A simple example of a label is the tagging of animal pictures with the corresponding label "*cat*" or "*dog*". Depending on the domain of interest, labeling can depend on expert knowledge and become a complex task. An example is a diagnosis based on CT scans (Fatima and Pasha, 2017) that only specially trained experts can perform. When an expert provides a label, the knowledge about this data is retained in the label. For instance, a label related to cells can provide knowledge about the analysis subject such as "*malaria*" or "*healthy*" cells in the image (Morang'a et al., 2020). As such, labels represent an externalized knowledge product of complex expert knowledge. Thus, organizations can leverage labels not only to create ML models but to create and transfer knowledge and support organizational learning (OL).

OL is the "dynamic process of creating new knowledge and transferring it to where it is needed and used, resulting in the creation of new knowledge for later transfer and use" (Kane and Alavi 2005, p. 796). More recently research stressed the unique capabilities of ML for OL such as that learning happens within the ML model and not only through the model (Afiouni, 2019). However, for the ML model to learn, it—in complex domains—highly depends on the knowledge of domain experts. Hence, to develop ML models, data analysts need to engage with domain experts for mutual learning processes because they need to maintain relevancy to the domain and produce knowledge independently. These processes can inform a hybrid practice where ML models and domain experts together perform knowledge work. To build models, ML developers need to understand complex domains, such as manufacturing processes (Eirich et al., 2022a) by consulting experts while domain experts need to understand the decisions of an ML model (van den Broek et al., 2021). The outcome or smallest unit of this mutual learning process are labels. The relevancy of labels is undisputed since the quality of ML models depends on the quality of labels. However, ML research thus far mostly discussed labels as a means to train ML models. We want to dive deeper and understand the role of labels in OL. In particular, we aim to answer the following research question:

How do labels affect organizational learning?

In answering the research question, we conducted a case study in collaboration with a German automotive manufacturer. We draw on the theoretical framework of OL by Argote and Miron-Spektor (2011) and contextualize it to the use of labels in organizations. In particular, we analyze how labels that are used in three visual analytics (VA) tools (Thomas and Cook, 2005) affect organizational members' tasks and knowledge retention. Our results contribute a contextualized theoretical framework, which helps to understand the role of labels in OL. Furthermore, we provide propositions on how labels affect OL and outline possible future research opportunities in this context.

2 Related Work

In the following section, we explain labeling in the context of ML. This section is followed by a description of the OL framework, which serves as an analytical tool for our analyses. Subsequently, we discuss information systems literature at the intersection of ML and OL.

2.1 Labeling

In the context of ML, a label refers to attaching a certain attribute to an instance in a data set. Examples are class labels, relevance scores, similarity judgments (Bernard et al., 2018b), or the annotation of specific data spaces in observed data, such as parts of curves or polygons in images (Wang and Hua, 2011).

Labels are of categorical or numerical nature (Weber et al., 2016). In ML models, categorical labels are more common than continuous labels (Bernard et al., 2018a) and can either be scaled nominally or ordinally (Stevens, 1946). Nominally scaled labels represent discrete units of analysis, which neither have a quantitative value nor can they be ordered (e.g., male vs. female). Ordinal scaled labels, in contrast, are discrete but ordered units of analysis, where the distance between classes is vague (e.g., fast vs. slow) (Mann and Lacke, 2013). Furthermore, categorical labels can be of a binary or multi-valued nature (Bernard et al., 2018a). While binary labels allow simple user feedback, such as "*ok*" vs. "*not ok*", multi-valued labels enable broader often more specific tagging of different classes of a data instance, such as "*dog*" vs. "*cat*" vs. "*frog*". Continuous labels are based on interval or ratio scales (Stevens, 1946). Interval scales represent ordered units of analysis with the same difference (e.g., temperature). Ratio scales are equal to interval scales but contain an absolute zero value (e.g., height) (Mann and Lacke, 2013).

In the domain of ML, a label is the result of human analysis. This analysis can vary in its complexity. In very complex domains, specific domain knowledge is required to create a label such as in labeling data in the production of highly integrated electrical engines (Eirich et al., 2021) or cell proliferation (Berg et al., 2019). Less domain knowledge is required in simple analyses such as for the classification of animals (Amershi et al., 2014) or handwritten digits (Bernard et al., 2018a).

2.2 A Framework of Organizational Learning

OL focuses on the *dynamic* processes through which knowledge is created and consumed within organizations (Vera and Crossan, 2003). It can broadly be divided into two forms: exploitation and exploration (Kane and Alavi, 2005). The former refers to incremental learning, which focuses on the diffusion, reuse, and refinement of existing knowledge (Larsson et al., 1998; March, 1991; Smith and Zeithaml, 1996), whereas the latter involves the development of new or the replacement of existing knowledge in organizations (Abernathy, 1978; March, 1991; Pentland, 1995).

In line with the notion of exploration and exploitation of knowledge in organizations, Argote and Miron-Spektor (2011) provide a conceptualization of OL (see Figure 1). The key elements comprise task performance and experience, knowledge, and the active context; the latter depends on the latent organizational context. The framework provides an analytical representation of learning in organizations in a cyclic relationship between the three key elements, in which task performance and experience create or transform knowledge through interaction with the context. While the environmental context represents elements outside an organization (e.g., suppliers or customers), the latent context represents elements within the organization (e.g., a culture of trust in technologies). Although the latent context does not "*act*" actively, it affects learning through its influence on the active context. In particular, within the active context organizational members (e.g., employees) and tools (e.g., ML-based systems) perform actions, that is, tasks related to their jobs.

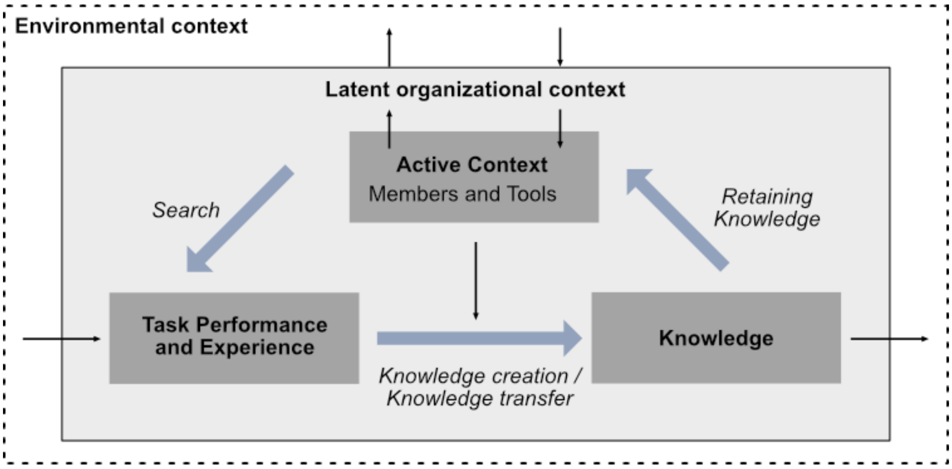

Figure 1. A framework of organizational learning based on Argote and Miron-Spektor (2011, p. 1125)

The process in which the active context affects task performance and experience is defined by Argote and Miron-Spektor as search. Knowledge is acquired through new experience (knowledge creation) or the sharing of existing experience within the organization (knowledge transfer) (Kane et al., 2005). This knowledge can either be of tacit or explicit nature (Nonaka and Takeuchi, 1995). While tacit knowledge is bound to the individual and difficult to communicate, explicit knowledge can be readily accessed by members or tools. Knowledge is retained through the active context, where organizational members or tools use it to act. Organizational members and tools, hence, represent knowledge repositories (Alavi and Leidner, 2001) and generate new experiences by performing new tasks.

2.3 ML in the Context of Organizational Learning

Research on the use of ML models in OL stresses that ML shifts the focus from the human learner to the machine. In contrast to traditional knowledge management tools (e.g., expert systems), ML technologies not only support human learning but can learn autonomously or learn in form of a hybrid practice where humans and algorithms inform each other (Sturm et al., 2021; van den Broek et al., 2021). In particular, in supervised ML models, human knowledge is crucial to inform human-in-the-loop practices, where domain experts provide inputs for algorithms, for training or debugging models, and for making sense of the results. These supervised and interactive learning approaches not only allow integrating human domain knowledge in the design of the ML model but require a huge amount of labeled data provided by human experts.

However, human-generated labels by experts can be error-prone. In knowledge-intensive domains characterized by uncertain or ambiguous knowledge work such as in generating diagnosis outputs, expert-generated labels may rely on the opinion of a single person and lack external validation. Regarding labeling tasks in such domains, the difficulty of identifying false negatives may be exacerbated because the outcome of the diagnosis may only be validated long-term, taking months or years (Lebovitz et al., 2021). Hence, especially in situations where an objective ground truth is missing, it is crucial to diligently monitor and assess the quality of labels and the ML output to avoid damage due to wrong decisions based on incorrect predictions of ML models.

In fact, a key challenge of a hybrid human-machine learning process that builds on feedback is identifying and understanding the gap between the actual and reference output. Any supervised and reinforcement ML model, hence, requires human auditing when being built, which, in turn, requires a reference measure (i.e., the ground truth provided by the domain expert) against which the output is compared. Therefore, data analysts and domain experts need to closely collaborate to alter and audit the ML model and improve its performance over time (Grønsund and Aanestad, 2020).

While research in this area focused on the development and adoption of ML models for organizational learning and its intricacies (Grønsund and Aanestad, 2020; van den Broek et al., 2021), we know little about the role of labels in this process. In response to this gap, this study investigates how labels influence OL in the context of a car manufacturer, revealing challenges in building and use of labels as well as the effects of labels (and related tools) on the key elements of OL.

3 Methodology

A case study is an empirical research approach to analyze a specific phenomenon within its environment. We chose a case study approach to study the role of labels for OL because case study designs are useful for studying context-rich sociotechnical systems (Yin, 2014), when the status quo in research is not yet well developed, when the examination of context and dynamics are important (Darke et al., 1998), and when the investigated process is difficult to observe from an outside perspective. We followed a single-case study approach (Yin, 2014). In particular, we investigated the case of a car manufacturer who introduced three ML-based VA systems building on labels. The case is situated in the context of a transformation project that many organizations currently pursue to change into a digital company.

3.1 Context and Scope

The organization in our case study is a leading German automotive manufacturer (hereafter referred to as “AutoCorp”). Like many of its competitors, AutoCorp is currently investing considerable resources in the digitization of its manufacturing processes. Therefore, dedicated analytics and engineering teams work collaboratively on the development of sophisticated VA tools to improve car production.

The two expert groups and the analyzed VA tools (see Table 1) represent our units of analysis. We refer to the two expert groups as *domain experts* and *data analysts*:

Domain Experts are responsible for specific manufacturing parts (e.g., rotor shafts of electrical engines) and have a deep understanding of a part’s physical properties (e.g., sound propagation inside electrical engines). These experts are responsible for the development of test benches, which are dedicated stations for automated part testing. Since test benches have to test a wide spectrum of a part’s properties, such as the electrical behavior of an engine as well as its mechanical attributes, domain experts also have deep knowledge about the part and its properties. For instance, two of our study participants are experts on the propagation of sound inside electrical engines. They know how an engine performs under specific conditions, for instance, in a test bench (e.g., by analyzing its magnetic field) or an electrical vehicle (e.g., by observing its acoustic signature during driving). Due to their specific expertise, domain experts are the only members of AutoCorp, who can create labels.

Data analysts develop data analysis methods, such as the training of sophisticated ML or statistical models. Their work relies on labels as they need categorical or numerical labels from domain experts as input for developing ML models or label-based tools. Thus, the more labels are available, the easier it is to develop ML models and label-based tools. Data analysts are also responsible for the acquisition and processing of data. These tasks can range from connecting sensor equipment that records measurements to machines to querying datasets from existing organizational databases.

To provide domain experts with a means to improve the manufacturing process, data analysts are currently involved in developing VA tools. Such tools aim to support the analysis of large and often unstructured datasets using visual representations and abstractions of the data (Thomas and Cook, 2005). Visualizing the data with specifically designed visualization interfaces helps experts to identify patterns and derive decisions. Thus, VA tools help in making better, data-driven decisions and—by storing labels and visualizing the data—retaining knowledge in organizational knowledge bases.

VA is closely related to the discipline of *business intelligence* (BI). Similar to VA, BI uses predefined methods (e.g., data aggregation or filtering) to gather, analyze, transform, and abstract data into new information via visualization interfaces to inform business decisions (Shollo and Galliers, 2016). Nonetheless, VA and BI differ in the following aspects: VA tools are tailormade visual interfaces between human experts and ML models to support highly explorative analyses. Examples are the analysis of speech data with self-organizing maps (Sacha et al., 2018) or root cause analyses of manufacturing errors with causality graphs (Eirich et al., 2022a). Most VA tools are designed to store data in the form of formalized user feedback, which can range from simple text inputs to more abstract forms such as graphs, to continuously develop a VA tool's underlying model. In contrast, traditional BI tools focus on the visualization of data and do not store expert feedback in dedicated databases.

Tool Name	*Description*	*Context and aim*	*Role of Labels*
RfX (Random Forest Explorer)	Visualizes the decision-making process of a random forest.	The interpretation of a random forest requires the analysis of model properties and can typically only be performed by data analysts. *RfX* allows domain experts to interactively explore the properties of a random forest and thus create new knowledge by understanding its decision-making process.	Uses labels to train a random forest classifier.
IRVINE (Interactive Labeling)	Facilitates the analysis of acoustic signatures of electrical engines.	Acoustic signatures of electrical engines contain a complex data structure, which only a few domain experts can analyze. *IRVINE* facilitates the analysis of acoustic data and helps to externalize and share domain knowledge on acoustic data in the form of labels.	Stores labels as a result of an expert's analysis.
ManEx (Manufacturing Explorer)	Allows analyzing sensor data from parts across the manufacturing process.	Manufacturing data is scattered across multiple data sources along the manufacturing process. *ManEx* helps domain experts to compare measurements from defective parts to error-free parts and thus detect the root cause of errors.	Uses labels to compare measurements from erroneous parts to error-free ones.

Table 1. Summary of analyzed label-based tools at AutoCorp

In our study, we explored the use of three tailormade, in-house VA tools (see Table 1). We hereafter refer to them as *label-based tools* because the VA tools are all interactive visual digital interfaces that rely on the input from experts to their underlying statistical models to *create, leverage,* and/or *store* knowledge in the form of labels. Hence, these VA tools do not only depend on domain knowledge but can also formalize that knowledge with labels through their visual interface.

3.2 Data Collection and Coding

We conducted the case study over sixteen months. During this period, the first author was deeply involved in the development of all label-based tools. All study participants are current users of the label-based tools. To develop and evaluate each tool, we collected data from the following sources:

(1) *Semi-structured interviews*: After the roll-out of each label-based tool, we interviewed 15 participants—seven data analysts (DA-1 to DA-7) and eight domain experts (DE-8 to DE-15) of AutoCorp. All interviewees were involved in the development of the tools and are current users of each label-based tool. They are all male, were on average 34.2 years old, had a mean working experience in AutoCorps of 9.7 years, and did not use label-based tools prior to the study. Each interview took on average 33 minutes.

The interview guideline was developed based on our theoretical lens, i.e., the effect of labels on OL. First, considering the dimension active context from the original OL framework (Figure 1), we asked questions on how participants define labels, the roles of the employees, and how they use the VA-tools RfX (Eirich et al., 2022b), IRVINE (Eirich et al., 2021) or ManEx (Eirich et al., 2022a) in terms of labels. Second, in the dimension task performance and experience, we asked, for instance, about the tasks that are supported by labels or which impediments exist in labeling tasks for organizational members and tools. Third, in the knowledge dimension, questions pertained to the storage of labels in organizational databases.

(2) *Direct observations:* The first author investigated how organizational members use the three label-based tools (RfX, IRVINE, ManEx) following Someren et al.'s (1994) Think Aloud Method. The researcher asked domain experts and data analysts to express their thoughts verbally when using label-based tools (Someren et al., 1994). The observations helped us to understand how organizational members (e.g., domain experts from different organizational units) share and create knowledge in the form of labels. Furthermore, the observations helped to contextualize label-based tasks.

(3) *Documentations of design and development iterations*: Each label-based tool, introduced in section 3.1, was designed and developed in several iterations. During each iteration, the on-site researcher documented numerous types of information. This information included the tasks users perform while using the tools or VA requirements used as design targets for each tool. This data helped to abstract high-level system requirements, which *label-based tools* need to address. The data also informed the definition of general tasks, which members and tools perform when using label-based systems.

Following our theoretical framework, we *coded* the data along the dimensions of the OL process. In particular, we went through a constant comparison analysis (Boeije, 2002). In so doing, we grouped the data into themes and gave names to each theme to describe their content. Next, we categorized these themes into the dimensions of the OL framework. We went through this process three times to refine the coding of the data.

4 Case Study Results

The focus of this analysis is on how labels affect each dimension of the OL framework and how they circulate along its three dimensions (active context, task performance and experience, and knowledge). We summarized our findings in Figure 2, which shows our adapted version of the original OL framework (Argote and Miron-Spektor, 2011).

Additionally, we asked experts how they define labels. From the answers, we derived the following definition:

> *A label is a categorical or numerical result of a complex data analysis performed by domain experts to describe a product's quality. It is, hence, the representation and aggregation of expert knowledge.*

Although this definition might not apply to contexts, in which expert knowledge is unnecessary, it summarizes the views of AutoCorp's experts and thus, will be used in the subsequent sections. In the active context (Section 4.1), labels affect both *members* and *tools* to perform the tasks *provide* and *decide* (Section 4.2). When performing these tasks new *tacit* and *explicit* knowledge (Section 4.3) is created.

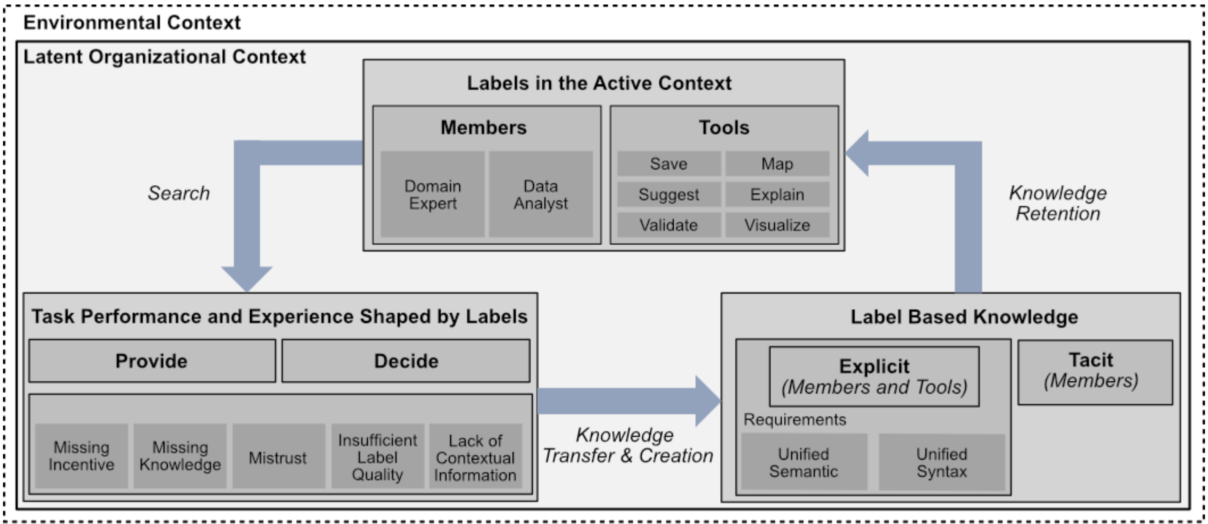

Figure 2. The organizational learning framework contextualized to labels

4.1 Active Context

As outlined in Section 3.1, members who work with labels are either *domain experts* or *data analysts*. Furthermore, the active context of the framework includes *label-based tools*, which are developed by data analysts in close collaboration with domain experts.

Labels in the active context represent a semantic, which facilitates the collaboration between domain experts and data analysts. As such, labels capture and hence, serve to share explicit knowledge. Specifically, domain experts provide and share labels with data analysts. Data analysts can use these labels to develop label-based tools or to better comprehend the problems that domain experts are facing. Label-based tools also allow domain experts to increase their analytical reasoning and understanding because they provide an interactive means to compare complex data structures, such as the production quality of different manufacturing stations. Thus, label-based tools allow members to build on each other's knowledge and facilitate OL. Hence, we propose:

P1a: *Labels representing explicit knowledge provide a semantic for organizational members which positively affects the comprehension of knowledge-intensive domains and thus positively impacts OL.*

Based on the experts' opinion of the three label-based tools, we derived the six high-level requirements *save, map, suggest, explain, validate*, and *visualize* that these tools should address.

Save: Label-based tools should store labels in organizational databases. This is necessary to make labels readily accessible for analyses by other organizational members within the same or different domains. IRVINE, for instance, allows saving labels from analyses of acoustic data from electrical engines. These labels are also relevant for domain experts outside the domain of acoustics. Engineers, who are responsible for the development of novel generations of electrical engines, can use stored labels to reflect on the design of electrical engines and include this externalized knowledge in future part generations.

Map: Label-based tools should provide information about how labels are stored in different organizational databases. For example, at each step of the manufacturing process, measurements for engines are recorded and stored either locally or scattered across organizational cloud infrastructures. Domain experts reported that it is "*extremely important to harmonize measurements and labels from different steps to understand the relation between them across the manufacturing process*" (DE-8). In this regard, DE-10 explained why he uses the tool ManEx: "*Fully comprehending the relation of measurements and groups of labeled engines is crucial to find the root cause of an error as early as possible in the manufacturing process.*"

Suggest: Tools should make recommendations about analysis objects (e.g., electrical engines) that are most suitable for labeling. Especially when members analyze large numbers of data, this requirement is necessary to reduce the number of analysis objects from the plethora of available options. IRVINE addresses this requirement by suggesting groups of most relevant engines for an in-depth analysis using ML predictions reducing the time domain experts need to find and analyze particularly relevant engines by over 30%. ManEx addresses this requirement by guiding domain experts and data analysts to groups of engines with the most deviating measurements along the manufacturing process, which helps members to easily locate and label engines with previously unnoticed errors.

Explain: Since label-based tools often address high-cognition tasks, they also should provide additional information about their decision-making process (Rudin, 2019). This helps organizational members to better understand a tool's decision, for instance, a prediction about the quality of a part. DA-2, for example, argued that label-based tools "*which use ML models should provide data so I can understand the reason why it predicted a specific error for a particular engine.*" The explanation of a model depends on the label type. For example, RfX visualizes predictions of multiple decision trees in a random forest for categorical labels as a node-link diagram. In addition, IRVINE outlines regions of interest in time series data of acoustic signatures.

Validate: Label-based tools should provide a quality measure for labels. This is important to ensure consistent and standardized labeling processes, which, in turn, allow tools to continually learn from labels. ManEx, for instance, addresses this requirement by providing a validation mechanic where domain experts can evaluate how different measurements influence groups of labeled engines. Furthermore, RfX shows different quality metrics for ML models results, which help to evaluate whether more labels are needed to train models or whether the ML model achieves a sufficient prediction accuracy. DE-14 argued to establish even more validation mechanisms, for example, that "*labels must be validated by multiple experts*" or that "*tools should include wrong predictions as an additional safety check, which need to be detected by an expert.*"

Visualize: Label-based tools should include an interactive visualization interface (Thomas and Cook, 2005). A visualization interface is necessary to relate labels to associated data by giving them a visual context. This visualization should enable data analysts and domain experts to comprehend the data or to analyze patterns, trends, or errors in large datasets easily. Although all three tools introduced in the case description include visualization interfaces, this requirement is undisputed for all new tools, which use labels. DE-8 explains why visualization is so important: "*If I cannot see the model predictions embedded in a visualization interface, I find it hard to make a decision based on the prediction.*"

Search: In the original OL framework (see Figure 1), *search* is the sub-process that connects the active context with task performance and experience. To perform a task, members and tools need to effectively locate labels and "*need to know where to find them*" (DA-3). The easier the locatibility of labels is, the easier it is for members and tools to perform a label-related task. Data analysts need to know which type of label exists in which database and how to write queries for each database to retrieve labels. Domain experts need to know the location of specific file storages to search for labels efficiently.

Label-based tools, which address the high-level requirements, can increase the knowledge of domain experts and data analysts by supporting locating and analyzing interesting data instances. For instance, label-based tools that suggest new data instances for labeling reduce the overall amount of data domain experts must analyze to a subset of particularly interesting data. These subsets may represent engine parts with highly anomalous sensor signals. Such signals, in turn, are particularly helpful in detecting and understanding new types of possible errors in the manufacturing process and thus foster the creation of new knowledge. Hence, we propose:

P1b: *Label-based tools that save, map, suggest, explain, validate, or visualize labels positively affect OL through their property to easily locate and analyze data instances of interest for human analyses.*

4.2 Task Performance and Experience

In the dimension *task performance and experience*, we surfaced the two label-related tasks *provide* and *decide* that label-based tools and members frequently perform.

In the task **provide**, domain experts create labels by a manual analysis that can be supported by label-based tools, such as IRVINE. An example of a manual labeling task is the analysis of time-series measurements, in which a label is created by assigning a property to the measurement. For example, DE-10 explained how he manually provides labels: "*When the time-series surpasses a certain threshold, I assign an error to the data*". Furthermore, providing labels is supported with designated labeling systems such as IRVINE. IRVINE facilitates the labeling of multiple engines by clustering engines with similar data patterns and presenting them to users for an in-depth analysis.

In the task **decide**, members and tools process labels for a specific analysis of manufacturing data and make a decision based upon this analysis. Domain experts perform analyses with labels, for instance, "*to find the exact location of an error in a labeled part*" (DE-11) or "*to understand the cause of an error*" (DE-8). Labels support decision-making tasks by "*providing a direction that points towards a decision*" (DA-6). Domain experts, for instance, need to decide on "*the adaptation of an existing testing procedure*" (DE-12). In contrast, data analysts, for example, decide based on "*a model recommendation, which electric vehicle is best suited for an in-depth analysis*" (DA-7). The tool ManEx automatically decides when to notify domain experts. For instance, in case of an unexpectedly high number of errors in the manufacturing process, the model will alert an expert about the errors.

Labels shape task performance and experience when the respective tasks either rely on labels or are strongly influenced by labels. For example, labels can be used to improve label-based tools to make better suggestions of new relevant data instances for labeling. During the process of providing labels, domain experts gain a better understanding of the data they analyze and thus, create new knowledge, which can be externalized in the form of labels. In terms of the task "decide", labels can improve decision-making and thus result in new knowledge. For example, label-based tools can proactively inform domain experts when produced parts contain a large number of anomalies in their sensor data. Domain experts then need to decide whether to send these parts to reworking stations or not. By taking this label-based decision, domain experts gain a better understanding of sensor data in the manufacturing process resulting in new tacit knowledge. Hence, we propose:

P2a: *Label-based tools support domain experts in providing labels and thus have a positive impact on the storing and sharing of explicit knowledge in organizations.*

P2b: *Label-based tools positively affect decision-making capabilities and thus the creation of new knowledge.*

In addition, we identified the five impediments *missing incentives*, *missing knowledge*, *mistrust*, *insufficient labels quality*, and *lack of contextual information*, which can hinder task performance:

Missing incentives is a lack of extrinsic or intrinsic motivation of a domain expert to provide a label. DE-13, for instance, pointed out that he does "*not have enough time to manually analyze hundreds of engines and provide a label for every single one*." Missing incentives, however, negatively affect the work of data analysts, who rely on labels. To tackle this impediment, data analysts try to run analyses with few labels only. However, these analyses do not provide good results as DA-4 explains: "*With few labels, I can only run simple statistical models, which give little insights about the cause of an error*."

Missing knowledge is a lack of understanding the domain (e.g., the exact functionality of a manufacturing station) by a data analyst or domain expert to perform the tasks provide and decide. By missing knowledge, we mean a lack of knowledge about specific domain problems and not the use of label-based tools. DA-4 explained how missing domain knowledge slows down his work: "*I first have to fully comprehend the problem domain and understand how to create labels, which is a very tedious process*." In addition, DE-13 found it "*particularly difficult to create new labels or use existing labels if the problem to solve is new and I do not have the necessary knowledge to understand it*."

Mistrust is a suspicious attitude of data analysts or domain experts when collaborating with other members or working with label-based tools (Thiebes et al., 2021). Trust often relies on personal relationships between members who create and use labels. DE-12 explained how trust in relationships affects his work with labels: "*I have more trust in a label if it is created by a person I know than in a label from a person I do not know. If I do not know the person, I always crosscheck the label myself.*" Trust also relies on a member's perception of a domain expert's or data analyst's expertise. A data analyst's ability to explain the statistical results of an analysis to a domain expert is an example of that. In this context, DA-5 explained: "*If I cannot respond to the question of why I believe that the specific anomaly is significant, the person for whom I have to perform the analysis does not trust my analysis.*" Also, a lack of trust in *label-based tools* can harm task performance as DE-13 noted: "*I need some kind of information on the trustworthiness of the model's prediction, for example, a confidence interval. Lacking this information, I find it difficult to make an informed decision based on a prediction.*"

Insufficient label quality refers to a label that captures false information about a part. Examples are *false positives*, such as an error-free part labeled as erroneous, or *false negatives,* such as an erroneous part labeled as error-free. Analyses based on false labels can result in incorrect analyses results of data analysts and domain experts or false predictions by label-based tools. Decisions based on wrong results can have severe negative effects on the work of members and the quality of parts. DA-4 provided the following example: "*Someone presented an ML model and all thought that it performed well. However, after a while, the model did make wrong predictions. It took us a lot of time to identify the reason, which was that it was trained with low-quality labels.*"

Lack of contextual information refers to the missing but necessary information to interpret a label correctly. Examples of such information are the "*location where the label was produced*" (DA-1), "*who provided the label*" (DE-12), "*the timestamp of a label*" (DE-10), or missing/incomplete data (Gashi et al., 2021) as described by DA-7 as a *"lack of relevant sensor data for a label.*" If labels do not include relevant context information, members have difficulties in fully comprehending the meaning of a label as DE-15 stated: "*I need to have additional information, for example, on which day the label was provided, to know what it means.*"

Knowledge creation and transfer connects task performance and knowledge. While performing the two label-related tasks (provide and decide), members gain new experiences and learn to perform tasks more efficiently. For instance, by discussing individual decision trees visualized by RfX, we observed that members started to develop a shared meaning for a random forest. A shared meaning of this ML model, in turn, enhances the collaboration between domain experts and data analysts to improve RfX. P1 provides another example of how he creates new knowledge with labels: "*The more I work with labels, the better I comprehend the manufacturing process and gain new knowledge about it.*" Labels also help to transfer knowledge between members and tools; DE-15 stated: "*Labels help to exchange information since I can pass labels to my colleagues so they can use them for their analyses*."

The aforementioned impediments can hinder the successful execution of the tasks *provide* and *decide*, and hence the creation and transfer of knowledge. For example, missing incentives or knowledge to provide labels result in fewer labels provided by domain experts. Hence, less explicit knowledge captured in labels circulates inside the organization. Furthermore, mistrust, insufficient label quality, and a lack of contextual information may result in the negative attitude of members towards label-based tools (Chatzimparmpas et al., 2020) and thus, negatively affect knowledge transfer and creation through labels. Hence, we propose:

P2c: *Missing incentives and knowledge to provide labels, mistrust in tools, insufficient label quality, and a lack of contextual information negatively affect knowledge transfer and creation.*

4.3 Label-based Knowledge

For the category knowledge, we surfaced the two knowledge types, *tacit* and *explicit knowledge,* in the context of creating and using labels:

Tacit knowledge is created in the interaction between members or between members and label-based tools. DE-8, for instance, creates knowledge by sharing experiences or discussing labels with other experts: "*Labels help me in discussions, where my colleagues and I compare error-free parts to erroneous parts and collaboratively find the causes of an error.*" Members can also interpret the representation of labels and associated data, for example, by using the tools RfX, IRVINE, or ManEx and thus continuously learn from them.

Explicit knowledge is captured in labels (Bernard et al., 2018a), which can be stored and accessed by organizational members or tools. For instance, when a domain expert creates a label with IRVINE it is saved in an organizational database. Labels can then be used to develop ML models as DE-9 noted: "*When I run the same analyses with new labels, I can evaluate the quality of my analysis process over time and gain new interesting insights.*" Furthermore, domain experts, data analysts, or label-based tools can combine labels with different data sources into more systematic and comprehensive sets of explicit knowledge (Wang et al., 2009).

Both, a *unified semantic* and *unified syntax* are important requirements for storing labels and the storage of explicit knowledge.

Unified semantic refers to the description of labels, as well as their properties and relations by defining a set of abstract concepts and categories that represent and relate labels. This is necessary so that labels are interpreted the same way by different members of AutoCorp. DE-15 provided one approach to build a unified semantic: "*different domain experts should decide on a common metric, which fits best to create labels.*" Furthermore, DA-7 pointed out that "*we need to define a common ontology, to understand how labels are embedded in the manufacturing process.*" A unified semantic is also important for the interpretation of labels in label-based tools because "*different tools, which all process labels with a common data processing routine would make it easier to compare them*" (DE-15).

Unified syntax refers to a common structure and data format for storing and sharing labels within organizations. In particular, labels should be available in "*numerical interchangeable data formats*" (DE-9) and be "*readable by machines*" (DE-7). A unified syntax also includes that labels are "*compatible with different kinds of data sources*" (DE-7) allowing the combination of labels with multiple datasets.

Knowledge retention connects knowledge with the active context. To continually retain knowledge in the form of labels and to transfer it from the short-term to long-term organizational memory, labels should be "*easily accessible to all individuals, who need to work with the labels*" (DA-1). Hence, labels need to be stored in a central database. Nevertheless, labels that contain sensitive organizational knowledge "*should only be available to specific members who have to work with these labels*" (DE-12). Sensitive knowledge can pertain to labels about prototypical parts, which the organization wants to keep secret. In the case of sensitive data, label databases should contain only restricted user access.

That being said, for the efficient retention of label-based knowledge in an organization, labels must be stored uniformly with a unified semantic and syntax. IS researchers have stressed that the standardization of knowledge is crucial to improve organizational knowledge retention (Förderer et al., 2014; Hsiao, 2008). To implement a unified semantic, standardization methods for labels such as ontologies (Alvarez-Coello and Gomez, 2021) or knowledge graphs (Ehrlinger and Wöß, 2016) can be used. To make labels readily accessible across organizational units, a common syntax, for example, compatible data types of labels are necessary. Hence, we propose:

P3a: *A unified semantic for labels positively affects the retention of tacit and explicit knowledge.*

P3b: *A unified syntax for labels positively affects the retention of tacit and explicit knowledge.*

5 Discussion

To answer our research question and to gain a better understanding of how labels influence the process of OL, we adapted Argote and Miron-Spektor's (2011) theoretical OL framework (see Figure 2). Drawing on our results, we propose a set of research questions summarized in Table 2.

Since labels present explicit knowledge (Wang et al., 2009), they play an important role in OL and complement well-established explicit but more generic knowledge products such as images or symbols (Peltokorpi et al., 2007). In contrast to such generic knowledge products, labels and related tools explicitly shape domain experts' routine work processes and help to better understand certain domains. Hence, sharing labels in organizations can have a positive effect on organizational learning since members can use them to understand knowledge-intensive domains. In our study, we analyzed the role of labels inside a single organizational unit. Since the original process model of OL (Kane et al., 2005) is also about OL across organizational units and between organizations, our first RQ proposes to understand how labels can improve knowledge creation and transfer in these broader contexts (RQ1).

Propositions	*Future research opportunities (RQ)*
1a and 1b	*RQ1*: How can labels efficiently be used across different organizational units and between organizations? *RQ2*: How must label-based tools be designed to improve knowledge creation and transfer across different organizational units and between organizations?
2a, 2b, and 2c	*RQ3*: How can organizations design incentives to foster label-based tasks and roles? *RQ4:* How can organizations overcome impediments related to labels that hinder effective task performance? *RQ5*: Is a data labeler a new position that organizations need to incorporate? And if so, what are the tasks and what knowledge is required of a labeler?
3a and 3b	*RQ6*: How must labels be modeled in organizational ontologies to support organizational knowledge management? *RQ7*: How can ontologies support the development of label-based tools?

Table 2. Future research opportunities

A further relevant future research direction is the design and development of label-based tools to support the efficient sharing of labels inside and between organizations (RQ2). In so doing, we believe that the two stakeholder groups—domain experts and data analysts—are crucial. For instance, data analysts should not only consider what data to include when building label-based tools but also how domain experts can provide labels during the development of label-based tools. Furthermore, our high-level requirements of label-based tools can guide researchers and practitioners with the design of similar decision support systems, such as the presented VA tools or similar systems, such as BI.

In addition, we surfaced impediments, which hinder the execution of the tasks provide and decide. One impediment refers to missing incentives to provide labels. Hence, we suggest investigating how organizations could design incentives to foster label-based tasks and roles to overcome such impediments (*RQ3*). One strategy to increase the engagement of domain experts in these tasks could be to apply gamification (Khakpour and Colomo-Palacios, 2020) or rewards for labeling activities similar to rewarding inventions or patents (Giarratana et al., 2018).

In line with prior research in complex and uncertain domains where high uncertainty and lack of validity of labels can lead to wrong predictions in ML models (Lebovith et al. 2021), we found that a lack of domain knowledge can affect the quality of ML models. Organizations need to develop strategies to overcome such issues, for instance, by flagging labels that are uncertain as well as continuously monitoring, assessing, and developing the ML model over time *(RQ4)*.

Concerning the task *provide*, a new organizational role similar to established organizational roles, which deals with data-related tasks (Crisan et al., 2020) could be introduced. Crisan et al. (2020) summarized existing roles, such as the "*data engineer*", who is mainly responsible for the development of data gathering and processing tasks. In addition to the roles defined by Crisan et al. (2020), organizations may consider introducing the "*data labeler*" (*RQ5*). This role could include organizational members, who have not only knowledge about a specific domain such as electrical engineering but are also responsible for providing high-quality labels for label-based tools as part of their job descriptions.

One way to represent knowledge in organizations should be in a machine-interpretable form. In the field of artificial intelligence, researchers investigated the role of knowledge representation with ontologies (Studer, 2007) to provide ML models with conceptual abstractions for particular domains of interest (Gonçalves et al., 2019; Harispe et al., 2014; Smaili et al., 2019). Future research may focus on the role of labels as knowledge products in the design of ontologies for organizational knowledge representation (RQ6). When a unified semantic and syntax for labels in organizational knowledge bases are established, another future research avenue is to understand how these affect the creation of label-based tools (RQ7). In particular, we assume that access to labels structured by a unified semantic and compatible syntax will foster the development of label-based tools by reducing the need to cross-check labels with domain experts or to convert and harmonize labels into a consistent data format.

6 Limitations

We acknowledge that our study has limitations. First, the study is based on qualitative data drawn from analyzing the use of three label-based tools, participant observations, and interviews with experts who work all in the same automotive company. Hence, our results are limited to the context of our study and the results may have limited validity in other domains. Nevertheless, since labels are such a basic data ingredient for any ML model, also other domains may reveal similar patterns regarding the circulation of labels between the three dimensions of the OL framework.

Second, our adapted OL framework does not address the latent organizational context nor the environmental context of the original framework by Argote and Miron-Spektor (2011). This was because in AutoCorp the role of labels is relatively new and label-based tools are not yet institutionalized. Further research can build on our propositions and research opportunities to investigate how labels affect OL.

7 Conclusion

In this paper, we analyze the impact of labels on OL. In so doing, we conducted a case study in collaboration with a German car manufacturer to understand how labels can support OL. The result is an adapted framework of OL focusing on the role of labels in OL that informs about organizational members who work with labels, requirements of label-based tools, label-related tasks and impediments of task performance, and how all these affect OL within an organization. Furthermore, we outline seven propositions about the role of labels regarding OL and suggest possible label-related future research directions. Practitioners can use the framework to understand labels along different stages of OL and to improve learning and knowledge management within their organizations.

References

Abernathy, W.J., 1978. The productivity dilemma: roadblock to innovation in the automobile industry. Johns Hopkins University Press, Baltimore.

Afiouni, R., 2019. Organizational Learning in the Rise of Machine Learning, in: International Conference on Information Systems.

Alavi, M., Leidner, D.E., 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly 25, 107-136.

Alvarez-Coello, D., Gomez, J.M., 2021. Ontology-Based Integration of Vehicle-Related Data, in: IEEE International Conference on Semantic Computing.

Amershi, S., Cakmak, M., Knox, W.B., Kulesza, T., 2014. Power to the People: The Role of Humans in Interactive Machine Learning. AIMag 35, 105–120.

Argote, L., Miron-Spektor, E., 2011. Organizational Learning: From Experience to Knowledge. Organization Science 22, 1123–1137.

Berg, S., Kutra, D., Kroeger, T., Straehle, C., Kausler, B., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Köthe, U., Hamprecht, F., Kreshuk, A., 2019. ilastik: interactive machine learning for (bio)image analysis. Nature Methods 16, 1–7.

Bernard, J., Hutter, M., Zeppelzauer, M., Fellner, D., Sedlmair, M., 2018a. Comparing Visual-Interactive Labeling with Active Learning: An Experimental Study. IEEE Transactions on Visualization and Computer Graphics, 298-308.

Bernard, J., Zeppelzauer, M., Lehmann, M., Müller, M., Sedlmair, M., 2018b. Towards User-Centered Active Learning Algorithms. Computer Graphics Forum 37, 121–132.

Bishop, C.M., 2006. Pattern recognition and machine learning, Information science and statistics. Springer, New York.

Boeije, H., 2002. A Purposeful Approach to the Constant Comparative Method in the Analysis of Qualitative Interviews. Quality and Quantity 36, 391–409.

Boukerche, A., Wang, J., 2020. Machine Learning-based traffic prediction models for Intelligent Transportation Systems. Computer Networks 181, 107530.

Chatzimparmpas, A., Martins, R.M., Jusufi, I., Kucher, K., Rossi, F., Kerren, A., 2020. The State of the Art in Enhancing Trust in Machine Learning Models with the Use of Visualizations. Computer Graphics Forum 39, 713–756.

Cibulski, L., Mitterhofer, H., May, T., Kohlhammer, J., 2020. PAVED: Pareto Front Visualization for Engineering Design. Computer Graphics Forum 39, 405–416.

Crisan, A., Fiore-Gartland, B., Tory, M., 2020. Passing the Data Baton: A Retrospective Analysis on Data Science Work and Workers, in: IEEE Visualization in Data Science.

Darke, P., Shanks, G., Broadbent, M., 1998. Successfully completing case study research: combining rigour, relevance and pragmatism. Information Systems Journal 8, 273–289.

Ehrlinger, L., Wöß, W., 2016. Towards a Definition of Knowledge Graphs, in: SEMANTiCS.

Eirich, J., Bonart, J., Jäckle, D., Sedlmair, M., Schmid, U., Fischbach, K., Schreck, T., Bernard, J., 2021. IRVINE: A Design Study on Analyzing Correlation Patterns of Electrical Engines. IEEE Transactions on Visualization and Computer Graphics, 11-21.

Eirich, J., Koutroulis, G., Mutlu, B., Jackle, D., Kern, R., Schreck, T., Bernard, J., 2022a. ManEx: The Visual Analysis of Measurements for the Assessment of Errors in Electrical Engines. IEEE Comput. Grap. Appl. 1–18.

Eirich, J., Münch, M., Jäckle, D., Sedlmair, M., Bonart, J., Schreck, T., 2022. RfX: A Design Study for the Interactive Exploration of a Random Forest to Enhance Testing Procedures for Electrical Engines. Computer Graphics Forum 1–18.

Fatima, M., Pasha, M., 2017. Survey of Machine Learning Algorithms for Disease Diagnostic. JILSA 9, 1–16.

Förderer, J., Schütz, S., Kude, T., 2014. Add-on Solution Success: A Configurational View on Knowledge Sharing in Digital Platforms, in: International Conference on Information Systems.

Gashi, M., Ofner, P., Ennsbrunner, H., Thalmann, S., 2021. Dealing with missing usage data in defect prediction: A case study of a welding supplier. Computers in Industry 132, 103505.

Giarratana, M.S., Mariani, M., Weller, I., 2018. Rewards for Patents and Inventor Behaviors in Industrial Research and Development. AMJ 61, 264–292.

Gonçalves, R.S., Kamdar, M.R., Musen, M.A., 2019. Aligning Biomedical Metadata with Ontologies Using Clustering and Embeddings, in: Hitzler, P., Fernández, M., Janowicz, K., Zaveri, A., Gray, A.J.G., Lopez, V., Haller, A., Hammar, K. (Eds.), The Semantic Web. Springer International Publishing, Cham, pp. 146–161.

Grønsund, T., Aanestad, M., 2020. Augmenting the algorithm: Emerging human-in-the-loop work configurations. The Journal of Strategic Information Systems 29, 101614.

Harispe, S., Ranwez, S., Janaqi, S., Montmain, J., 2014. The semantic measures library and toolkit: fast computation of semantic similarity and relatedness using biomedical ontologies. Bioinformatics 30, 740–742.

Hsiao, R., 2008. Knowledge Sharing in a Global Professional Service Firm. MIS Q. Executive 7.

Kane, A.A., Argote, L., Levine, J.M., 2005. Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. Organizational Behavior and Human Decision Processes 96, 56–71.

Kane, G., Alavi, M., 2005. Information Technology and Organizational Learning: An Investigation of Exploration and Exploitation Processes, in: Organization Science.

Khakpour, A., Colomo-Palacios, R., 2020. Convergence of Gamification and Machine Learning: A Systematic Literature Review. Technology, Knowledge and Learning.

Larsson, R., Bengtsson, L., Henriksson, K., Sparks, J., 1998. The Interorganizational Learning Dilemma: Collective Knowledge Development in Strategic Alliances. Organization Science 9, 285–305.

Lebovitz, S., Levina, N., Lifshitz-Assa, H., 2021. Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What. MIS Quarterly 45, 1501–1526.

Mann, P.S., Lacke, C.J., 2013. Introductory statistics, 8. ed., international student version. ed. Wiley, Hoboken, NJ.

March, J.G., 1991. Exploration and Exploitation in Organizational Learning. Organization Science 2, 71–87.

Morang'a, C.M., Amenga–Etego, L., Bah, S.Y., Appiah, V., Amuzu, D.S.Y., Amoako, N., Abugri, J., Oduro, A.R., Cunnington, A.J., Awandare, G.A., Otto, T.D., 2020. Machine learning approaches classify clinical malaria outcomes based on haematological parameters. BMC Med 18, 375.

Nonaka, I., Takeuchi, H., 1995. The knowledge-creating company: how Japanese companies create the dynamics of innovation. Oxford University Press, New York.

Peltokorpi, V., Nonaka, I., Kodama, M., 2007. NTT DoCoMo's Launch of I-Mode in the Japanese Mobile Phone Market: A Knowledge Creation Perspective. Journal of Management Studies 44, 50–72.

Pentland, B.T., 1995. Information systems and organizational learning: The social epistemology of organizational knowledge systems. Accounting, Management and Information Technologies 5, 1–21.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1, 206–215.

Sacha, D., Kraus, M., Bernard, J., Behrisch, M., Schreck, T., Asano, Y., Keim, D.A., 2018. SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance. IEEE Transactions on Visualization and Computer Graphics 24, 120–130.

Shollo, A., Galliers, R.D., 2016. Towards an understanding of the role of business intelligence systems in organisational knowing: The role of BI systems in organisational knowing. Information Systems Journal 26, 339–367.

Smaili, F.Z., Gao, X., Hoehndorf, R., 2019. OPA2Vec: combining formal and informal content of biomedical ontologies to improve similarity-based prediction. Bioinformatics 35, 2133–2140.

Smith, A., Zeithaml, C., 1996. Garbage Cans and Advancing Hypercompetition: The Creation and Exploitation of New Capabilities and Strategic Flexibility in Two Regional Bell Operating Companies. Organization Science 7, 388–399.

Someren, M.W. van, Barnard, Y.F., Sandberg, J.A.C., 1994. The think aloud method: a practical guide to modelling cognitive processes. Academic Press, London.

Stevens, S.S., 1946. On the Theory of Scales of Measurement. Science 103, 677–680.

Studer, R., 2007. Semantic Web Services: concepts, technologies, and applications, Semantic Web Services. Springer, Berlin.

Sturm, T., Gerlach, J., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., Nan, N., Buxmann, P., 2021. Coordinating Human and Machine Learning for Effective Organizational Learning. MIS Quarterly 45, 1581–1602.

Suschnigg, J., Mutlu, B., Fuchs, A., Sabol, V., Thalmann, S., Schreck, T., 2020. Exploration of Anomalies in Cyclic Multivariate Industrial Time Series Data for Condition Monitoring, in: International Workshop on Big Data Visual Exploration and Analytics.

Thiebes, S., Lins, S., Sunyaev, A., 2021. Trustworthy artificial intelligence. Electron Markets 31, 447–464.

Thomas, J.J., Cook, K.A. (Eds.), 2005. Illuminating the path: the research and development agenda for visual analytics. IEEE Computer Soc, Los Alamitos, Calif.

van den Broek, E., Sergeeva, A., Huysman Vrije, M., 2021. When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. MISQ 45, 1557–1580.

Vera, D., Crossan, M., 2003. Organizational learning and knowledge management: toward an integrative framework. Handbook of Organizational Learning and Knowledge Management 122–141.
Wang, M., Hua, X.-S., 2011. Active Learning in Multimedia Annotation and Retrieval: A Survey. ACM Transactions on Intellilgent Systems and Technologies 2.
Wang, X., Jeong, D.H., Dou, W., Lee, S.-W., Ribarsky, W., Chang, R., 2009. Knowledge Assisted Visualization: Defining and Applying Knowledge Conversion Processes to a Visual Analytics System. Computer Graphics 33, 616–623.
Weber, N., Waechter, M., Amend, S.C., Guthe, S., Goesele, M., 2016. Rapid, Detail-Preserving Image Downscaling. SIGGRAPH Asia 6.
Yin, R.K., 2014. Case study research: design and methods, Fifth edition. ed. SAGE, Los Angeles.