AMCIS 2022 Proceedings

SIG ED - IS in Education, IS Curriculum, Education and Teaching Cases

Aug 10th, 12:00 AM

Are AI Opportunities Discovered or Created? Investigating Data Resourcing Using a No-code AI Platform in an Educational Context

Leif T. Sundberg
Umeå University, leif.sundberg@umu.se

Jonny Holmström
Umeå University, jonny.holmstrom@umu.se

Are AI Opportunities Discovered or Created? Investigating Data Resourcing Using a No-code AI Platform in an Educational Context

Completed Research

Leif Sundberg
Umeå University
leif.sundberg@umu.se

Jonny Holmström
Umeå University
jonny.holmstrom@umu.se

Abstract

With recent advances in artificial intelligence, machine learning has been identified as particularly useful for organizations seeking to create value from data resources. However, this usefulness presupposes the existence of data with sufficient structure and quality to train machine learning models. Thus, in this study, we address the research question: How does data resourcing contribute to enabling AI opportunities? We address this question by investigating an example of data resourcing in a master's level artificial intelligence (AI) course at Umeå University, Sweden. This is an empirical, case-based educational setting where students engaged in data resourcing using a no-code AI platform. Our results involve provide insights regarding constructs associated with two phases of data resourcing: data work practices, and data interpretation. Collectively, these constructs provide a framework for data resourcing that is the main contribution of this paper, together with insights into the benefits of using no-code AI tools in educational settings. Our framework reveals the nature of data resourcing as a creative process where iterative algorithmic mediation, data sensemaking and contextualization enables AI opportunities.

Keywords

Artificial intelligence, machine learning, deep learning, higher education, information systems, no-code

Introduction

The last decade has witnessed significant advances in artificial intelligence (AI), such as machine learning (ML) (Jordan and Mitchell, 2015) and deep learning (DL) (LeCun et al., 2015) systems. Massive increases in processing power of digital technology and available data have set the stage for increases in its use in many contexts (Dwivedi et al. 2021). Consequently, organizations are increasingly deploying AI technologies that can parse through large amounts of data, acquire skills and knowledge, and operate autonomously (Simsek et al., 2019).

ML has been identified as particularly useful in organizational contexts (Sturm et al., 2021). Leavitt et al. (2020) define it as "a broad subset of artificial intelligence, wherein a computer program applies algorithms and statistical models to construct complex patterns of inference within data". ML is associated with a distinctive promise to ease manipulations of descriptions of the world, and thus their transfer across contexts in ways that were not previously possible. However, realization of this promise heavily depends on preliminary data work involving important choices about what is included in datasets, what is left out, and what is considered representative or relevant for specific contexts (Lebovitz et al., 2021; Mikalsen and Monteiro, 2021; Vaast and Pinsonneault, 2021).

In this paper, we refer to the potential of using AI/ML to create value in organizational contexts as "AI opportunities". We coin the term data resourcing to refer to the active and creative data work required to leverage ML in organizational contexts and exploit ML platforms' growing algorithmic capabilities to reveal intricate patterns in highly complex datasets. Data resourcing involves active sensemaking by data scientists in situations characterized by high levels of uncertainty (Aaltonen et al., 2021; Mikalsen and Monteiro, 2021). Against this backdrop, we pose the following research question: How does data resourcing contribute to enabling AI opportunities? To address it, we present an example of data resourcing in the context of a masters' level AI course at Umeå University, Sweden. The study is based on qualitative data collected through interactions with and observation of students, and we employ the conceptualization of 'data objects' as a methodological strategy to unpack data resourcing practices (Alaimo and Kallinikos, 2022), which is further elaborated in the following section.

ML has been identified as a particularly promising tool in educational contexts (Luan and Tsai, 2021; Kayhan, 2022), despite challenges in leveraging ML associated with limited training in data science, systems, and the technicalities of coding and modeling. We argue that data resourcing requires the 'labeling' and 'packaging' of datasets for their successful use and reuse in educational contexts. Hence, we suggest that AI opportunities are not 'discovered' in datasets but rather 'created' as datasets are collected, cleansed, augmented, discarded, disassembled and re-assembled.

While conventional AI systems require significant resources to install and use, we take an interest in how a "no-code" AI platform can be used in an educational context. These platforms make AI accessible and affordable and guide users through the process of developing and deploying AI models, without the need to learn all about the intricacies associated with complex algorithms (Lins et al., 2021).

In the remaining sections of the paper we: position our study theoretically and describe the educational setting; describe the materials and methods used; present the results; discuss the results; and finally offer concluding remarks.

Theoretical background: the construction of data resources

Previous literature highlights the need for organizations to manage their data (e.g., Gillenson, 1985; Wang and Strong, 1996). Organizations that can find ways to structure and analyze data can gain competitive advantages and enjoy rapid scale-ups (Huang et al. 2017). As mentioned by Janssen et al. (2020), organizational data governance often includes consideration that data are often combined from various sources, which may not necessarily be subject to control by the organization itself.

Moreover, with the increasing generation of digital data (Hilbert and Lopéz, 2011), the roles of data are shifting from specific, siloed administrative, managerial and technical applications, to pervasive resources and means through which organizations develop, and exploit, myriad kinds of knowledge (Alaimo and Kallinikos, 2022). In diverse contexts, datasets are not merely important assets, but increasingly the foundations, pivots and mediators of organizational processes and evolution. Hence, organizations are becoming increasingly immersed in management of, and by, data (Alaimo and Kallinikos, 2021). Implications of the broad diffusion of data, or data objects, include the reconfiguration and rearrangement of organizational environments around data and algorithms, as well as increasing intertwining of organizations' identities with the technologies they deploy (Alaimo and Kallinikos, 2021). In addition, while increasing digitalization has generated a new breed of organizations that inherently rely on algorithmic operational models, traditional sectors are also increasingly being challenged by intensified use of data and algorithms (Iansiti and Lakhani, 2020).

As noted by Vial (2019), the etymology of the word data (plural of datum) stems from Latin and means 'given'. In line with this author (see also Iliadis and Russo, 2016), we argue against the view of data as a neutral, raw material (something given). On the contrary, we align with a view of data as constructed, not given as natural representations of facts (Drucker, 2011). Due to the digital nature of these data, they are at least partially determined by the technological conditions in which they are constructed (Dougherty, 2007). Moreover, through algorithmic work, data are made sense of and woven into work practices (Mikalsen and Monteiro, 2021). By shifting from a view of algorithms as rigid formulas to acknowledging that they are a part of our culture where they are given meaning and included in certain practices, their organizational roles and impacts can be addressed empirically (Seaver, 2017). To do so, it is necessary to analyze algorithms "within those systems that give them meaning and animate them" (Dourish, 2016). For example, algorithms are deployed in contexts of the professional cultures of computer scientists and machine learning experts. They also affect, operationalize and are formulated in relation to various phenomena, which requires alteration of data for algorithmic processing. Hence, as noted by Dourish (2016), algorithms and data are intertwined and co-evolve, mutually influencing each other. For example, the increasing adoption of algorithms in organizations is spurring new forms of work and occupations, such as algorithmic curation, and brokerage (Kellogg et al. 2020; Waardenburg et al. 2021). These observations are consistent with long-established understanding of technology as a social phenomenon that may alter organizational and professional processes (Barley, 1986). When algorithms' roles and impacts extend beyond the technical domains in which they are created they cannot be considered mere computational tools, but as assemblages entangled and engaged in networks of people and actions (Glaser et al. 2021). In a broader perspective, data become "sociomaterial forces of social ordering in the algorithmic age" as they contribute to the formation of epistemic architectures that enable and constrain information (Flyverbom and Murray, 2018, pp. 10).

This background prompted interest in the processes through which data are constructed for use as resources for training deep learning models, which we explore here using information acquired in a case study of a higher education AI course, described below.

Educational setting

As noted by Holmström et al. (2011), rapid technological developments create challenges for maintaining up-to-date curricula for educating professionals who will work in environments with high levels of technology. These authors highlight several important issues regarding IS teaching, including the importance of ensuring that the students acquire practically relevant skills, through use of appropriate pedagogical approaches, and generic types of knowledge. As AI is being increasingly adopted in diverse domains (Dwivedi et al. 2021), most, if not all, professionals will engage with or be affected by intelligent systems in their careers. However, as already mentioned, AI in the form of algorithms is also associated with skills stemming from computer science and engineering. Professionals rooted in other disciplines face challenges here, not because they have nothing to contribute to AI, but because of a lack of fundamental knowledge of how, for example, a machine learning system works. A potential remedy for this is to use 'lightweight' AI in the form of AI service platforms (Geske et al. 2021; Lins et al. 2021), which are easy to use and install (as they are cloud-based), and have graphical interfaces that help users to train ML models. Here we present an approach for using such a system, the Peltarion (2022) 'no-code' deep learning AI platform, in a higher education environment at the Department of Informatics, Umeå University, Sweden. This AI solution enables non-data scientists to upload data, train and evaluate a ML model that can be deployed via an application programming interface (API). The platform guides users via a graphical interface together with suggestions regarding problem types, workflows, pre-trained models and iterative improvements. The platform was used in an 'AI for business' course at Umeå University, to give the students hands-on experience in training ML models, based on a case-based task that the students had to solve. In line with the course curriculum (Umeå University, 2022), the learning objectives of the exercise was to "Account for and explain the role of AI in organizational value creation", by giving the students first-hand experience of training ML models. The educational approach is further described in the following section.

Materials and Methods

To address the research question posed above, we have subjected empirical material drawn from the case study of use of the no-code AI platform in the educational setting described above to thematic analysis (Braun and Clarke 2012; Clarke and Braun 2014). Following Mathiassen and Purao (2002), a group-based project approach was applied in the course, inviting the students to engage in development of ways of working and participating in communicative activities regarding 'real-life' problems. As noted by Leidner and Jarvenpaa (1995), such approaches provide opportunities for students to understand the 'messiness' professionals face in industry, acknowledging the social situatedness of these contexts, and acknowledge that the problems students will face are "unstructured, ambiguous, and immune to purely technical solutions" (Holmström et al. 2011).

The case presented to the students described a fictive organization, 'WeldCorp', specialized in welding, seeking to expand and acquire customers in additional geographical markets while retaining and automating quality measures. To assist the company, we invited the students to develop ways to use ML as a tool to assess welding points based, at least initially, on a limited dataset of 157 images of good and bad examples. The task given to the students involved the following questions: 1. Describe and motivate your choices regarding the data processing, problem selection, and model training in the Peltarion platform. 2. Describe how you evaluated the predictions of your model. Are they accurate enough to use live for WeldCorp? Why / why not? 3. Discuss: What could be done by WeldCorp to improve the results of the model? How would they implement this type of solution in their business?

We applied the principles of instructions framework advocated by Merrill (2007, 2013) in the educational setting, which incorporates five principles: problem-centered learning, activation, demonstration, application, and integration. These principles, together with our operationalization of them, are summarized in Table 1. The framework provides an integrated, multi-strand strategy for teaching students how to solve real-world problems, or complete complex real-world tasks.

Principle	**Description**	**Our operationalization**
Problem-centered learning	Humans learn better when they are solving problems, so learning is promoted when learners acquire skill in contexts of real-world problems.	The students were presented with a case of a welding company, WeldCorp, seeking to expand and scale up its business while improving quality control. To help these efforts they were encouraged to apply ML to differentiate between good and bad weld points.
Activation	Learning is promoted when learners activate existing knowledge and skills as a foundation for a new skill. An important step here is to start where the learner is. Activation requires learning activities that stimulate the development of mental models and schemes that can help learners to incorporate new knowledge or skill into their existing knowledge framework.	Since the students had different educational backgrounds (business and administration, computer science, and behavioral science), we chose to use the Peltarion no-code deep learning platform. This enabled them to incorporate previous skills and work during the course, even if they lacked previous experience of data science.
Demonstration	Learning is promoted when learners observe a demonstration of the skill to be learned, e.g., by exposure to examples of good and bad practice.	We showed the students several examples of ways to train ML models via the platform. Students were encouraged to take tutorials and experiment with different types of data (e.g. tabular, text, images) and problems that can be accessed through the platform.
Application	Learning is promoted when learners apply new skills they have acquired to solve problems. Applying new knowledge or skills to real world problems is treated as almost essential for effective learning.	The students were divided into two groups and each student was given access to an enterprise account in the deep learning platform where they were encouraged to engage with different types of data and activities in addition to the group assignment.
Integration	Learning is promoted when learners reflect on, discuss, and defend knowledge or skill they have acquired. The effectiveness of a course is enhanced when learners are provided with opportunities to discuss and reflect on what they have learned in order to revise, synthesize, recombine and modify their new knowledge or skills.	Students were encouraged to reflect on their learning during the final seminar, in a survey, and in the course evaluation. During the final seminar they were also expected to learn from each other by preparing questions for the other group.

Table 1. Principles of the educational approach

The course module described in this paper consists of a workshop, a Q&A session, supervising sessions, and a final seminar. Its content is further outlined here. First, features of the AI platform were demonstrated by showing the students how to solve various ML tasks using tabular, text, and image data. After this, the students were divided into two groups, representing 'organizations' in the platform. This division was in line with our educational approach as we encouraged the groups to reflect on their experiences by discussing their approaches to solving the problem at hand at the end of the course. The division also enrichened our empirical material as we could identify differences and similarities between the approaches undertaken by the groups.

Nineteen students attended the course (14 male and five female), with educational backgrounds including bachelor's degrees in computer science, business and administration, or behavioral science. The empirical materials used in the study presented here, as summarized in Table 2, stem from interactions with the students, the Peltarion AI deep learning platform, and teachers' reflections. These materials allowed us to contribute both to research on AI, and present interesting findings from the educational setting.

Materials	**Source(s)**
Students' feedback and course evaluations	E-mails, notes taken during the course, written evaluations and feedback from students.
Students' written assignments and presentations	Two written group reports, and two presentations during a final seminar.
Datasets, models and deployments created by the students	The Peltarion (2022) AI platform
Observations	Teachers' experiences and reflections during and after the course

Table 2. Materials

We subjected the materials to thematic analysis (Braun and Clarke 2012; Clarke and Braun 2014) by inductively coding the students' activities during the module. More specifically, we coded activities associated with data resourcing in our empirical setting mentioned and observed in the materials, then aggregated them into themes to inform a process model (Langley, 1999). This provided a framework describing their data resourcing practices using ML tools, and thus foundations for answering our research question.

Results: Using machine learning in an educational context

To prepare the students for the assignment, we demonstrated various examples of ML problems, and their possible solutions using the Peltarion platform. Examples of such problems included image classifications and image similarity searches, as well as other types of ML problems, such as text-classification with natural language processing models, (e.g., BERT). At the end of this session, the students were divided into two groups and assigned the task of helping WeldCorp to assess welds using the set of 157 images of good and bad welds. They were then given enterprise accounts providing access to the Peltarion platform. The teachers had access to both accounts to enable them to observe and aid the students as they uploaded data, trained and deployed ML models to solve the problem.

A week after receiving the initial dataset with a limited number of images of good and bad welds, a Q&A session was held with the student groups. During this session, the teacher who led it observed that most of the questions concerned the dataset. Illustrative queries from the students concerned the quality of the supplied dataset, tentative workarounds, and image formats. After the Q&A session we observed how the students engaged in **data collection** and uploaded larger datasets with various images to the platform. They trained several models, and iteratively fine-tuned the settings in the platform. Students used several sources here, including social media, Google image search, and Kaggle. Both groups chose to label their images in a binary fashion as 'good' or 'bad'. To create consensus necessary for creating "ground truths", one of the groups formalized the **data labeling** process in their report with a 'weld quality framework', but both groups adhered that they were not experts in the welding domain.

One of the groups ran into problems and received error messages while training their models due to inappropriate image properties. They solved the problems by manually adjusting the images in photoshop using a batch function that resized and reduced the bit depth from 32 to 24. As some of the images the students collected came in batches the students engaged in **data cleaning**, by removing unwanted features or 'noise' as they called them in their report. Examples of noise included irrelevant features such as screws, multiple welds etc. Here, the students also mentioned the importance of retaining a certain degree of complexity in the images since reality is rarely noise-free. One of the two groups strongly engaged in **data augmentation** as they extended their dataset 4-to 5-fold by manipulating the images through zooming, cutting and rotating them.

Both groups ended up using a pre-trained model (EfficientNetB0) to solve an image classification problem (single label) in the Peltarion platform. Each group formed training, validation and test sets containing 80, 10 and 10% of their full datasets (images), which is common practice and a default option in the platform. The students refined the output of the model in two ways. First they iteratively adjusted settings in the platform, such as increasing the training rate (with careful monitoring of the variances of accuracy of the predictions among the three datasets generated by splitting to avoid overtraining the model). Secondly, as particularly strongly emphasized by one of the groups, the students strove to ensure that included data were contextually relevant, and suitable for WeldCorp's purposes. This was done after they received output from the ML model via receiver operating characteristic (ROC) curves and confusion matrices (see Figure 1) and could assess whether certain types of images were incorrectly classified, identify potential biases in the data, and signs of model overtraining. Examples mentioned during the final seminar were images of painted welds, which would not be relevant in the industrial **context** they imagined.

Figure 1. Screenshot of students' work in the AI platform

We noticed in the students' course evaluations and written feedback that they heavily emphasized the role of data in their learning, as illustrated by the following two quotations:

"I've obtained practical knowledge and experience of the impact of data. And I've even seen the impact of flaws in the dataset first-hand. Thus, I think this was an optimal learning method considering our (and my) educational background." – student evaluation.

“[I’ve learnt] that data matters! The choice, generating and cleansing of data are crucial.” – student evaluation.

During the discussions in the final seminar, the students also proposed ideas for operationalizing their work in a live setting, such as using automated cameras to feed data on welding points for evaluation by the deep learning model. Based on our materials, we generated constructs in the form of distinct activities that the students conducted during their assignments. These constructs, together with exemplars are described in Table 3.

Construct	**Description**	**Exemplars - source**
Data collection	Continuously searching for and adding to a dataset and monitoring numbers of data points of each category (e.g. ‘good and ‘bad’ welds).	“We initially had few images of bad welds. This was solved by including a dataset containing images of bad welds from Kaggle” - student during presentation.
Data labeling	Generating consensus on the use of categories for labeling, and criteria (ground truths) associated with each category.	“This [data labeling] was a big issue in our group. We tried several times to create consensus regarding what made good/bad welds, and we tried to have one person doing a cleanup. But since none of us had any real experience in the area, it was really hard and it took a lot of debate before we could proceed.” - student evaluation
Data cleaning	Removing unwanted features from the data, while retaining the ML model’s ability to account for a certain ‘messiness’.	“The original data set was cleaned of excessive noise (images containing screws, horns, overview pictures or multiple welds in the same image) and duplicates.” - student report
Data augmentation	Enhancing and extending an existing dataset by altering data.	“By augmenting the data by cropping, zooming, flipping and rotating, the original dataset was multiplied 4- to 5-fold.” - student report
Algorithmic mediation	The feedback and constraints stemming from the output of the ML model.	"If we look at some of the images that were classified incorrectly by our model, it quickly becomes clear why the error has occurred." - student report.
Data sensemaking	Interpreting the ML model’s output to ensure relevance for the context.	“Images of painted weld joints are not relevant for WeldCorp” - student during presentation

Table 3. Data resourcing constructs

Concluding Discussion

By aggregating the constructs presented in Table 3 we decompose an integrative ML process model of data resourcing as a combination of schemes of actions (Figure 2). In combination with earlier views of data as a resource (e.g. Aaltonen et al., 2021; Mikalsen and Monteiro, 2021) this provides a cohesive framework, with the following elements:

- Data work practices: the activities of collecting, labeling, cleaning and augmenting data that form a training dataset.
- Data interpretation: following initial training, outputs of the ML model provide feedback that stimulate interpretation of the outputs due to the black-boxed nature of the neural network. These interpretations subsequently lead to contextualized knowledge that informs the continued process of constructing data resources.

The elements of our framework are illustrated in Figure 2. The building blocks identified from literature together with the analysis of our case allowed us to zoom into the dynamics of data resourcing in the ML context. In addition to the constructs relating to the preparation of data and establishment of ground truths (c.f. Lebovitz et al., 2021) our model highlights the importance of data interpretation as the students actively "translated" the output of their trained model (c.f. Waardenburg et al., 2021). This practice, enabled by algorithmic mediation and data sensemaking, generated important contextualized knowledge, vital to create value from AI, and to reduce biases in data. Thus, the results demonstrate ways in which *data resourcing* involves active, creative data work that is essential for leveraging ML in organizational contexts, due to associated complexities and the need for sensemaking in situations with typically high levels of uncertainty (Aaltonen et al., 2021; Mikalsen and Monteiro, 2021). They also illustrate how data was made sense of and integrated into work practices (Mikaelsen and Monteiro, 2022) while being intertwined with deep learning algorithms (Dourish, 2016) and subject to interpretative practices (Waardenburg et al. 2021). Echoing Waardenburg et al. (2021), we see these interpretations as important in the context of ML, as algorithmic outputs are subject to translations by human actors. We see these translations as especially important in environments where data may be subject to societal bias associated with factors such as gender and race.

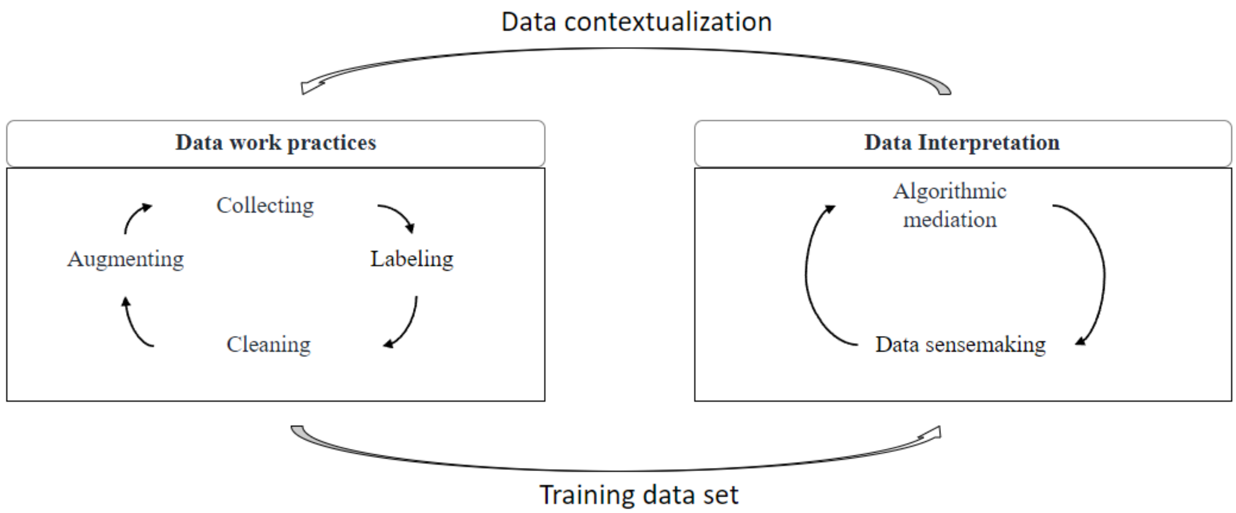

Figure 2. Data resourcing for machine learning framework

As the no-code approach enabled students to engage in collective data work the selected empirical setting provided an ideal opportunity to address our research question, *How does data resourcing contribute to enabling AI opportunities?* Dwelling in the empirical context of a master's level AI course at Umeå University, where students were encouraged to use the AI tool to solve a weld assessment problem, enabled us to construct a framework of activities (data work practices and data interpretation) related to data resourcing. The results contribute to insights into the process of the generation of data as resources for use in ML and deep learning, and demonstrate the benefits of using a no-code AI solution in higher education contexts. Our framework also shows that data resourcing is an active, creative process where algorithmic feedback mediates contextualized knowledge, which enables creation of AI opportunities.

Acknowledgements

This work was supported by the Kempe Foundation, grant no. JCK-2024.

REFERENCES

Aaltonen, A., Alaimo, C., & Kallinikos, J. 2021. "The making of data commodities: Data analytics as an embedded process," *Journal of Management Information Systems*.

Alaimo, C., & Kallinikos, J. 2022. "Organizations decentered: Data objects, technology and knowledge," *Organization Science*.

Alaimo, C., & Kallinikos, J. 2021. "Managing by data: Algorithmic categories and organizing," *Organization Studies 42*(9), 1385-1407.

Barley, S. R. 1986. "Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments," *Administrative Science Quarterly*, 78-108.

Braun, V., & Clarke, V. 2012. "Thematic analysis," In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological (pp. 57–71). American Psychological Association.

Clarke, V., & Braun, V. 2014. *Thematic analysis*. In Encyclopedia of Critical Psychology (pp. 1947-1952). Springer, New York, NY.

Dougherty, D. 2007. "Trapped in the 20th century? Why models of organizational learning, knowledge and capabilities do not fit bio-pharmaceuticals, and what to do about that," *Management Learning*, 38(3), 265-270.

Dourish, P. 2016. "Algorithms and their others: Algorithmic culture in context," *Big Data & Society*, *3*(2).

Drucker, J. 2011. "Humanities approaches to graphical display," *Digital Humanities Quarterly*, 5(1), 1-21.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. 2021. "Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *International Journal of Information Management*, 57.

Flyverbom, M., & Murray, J. 2018. "Datastructuring—Organizing and curating digital traces into action," *Big Data & Society*, 5(2).

Geske, F., Hofmann, P., Lämmermann, L., Schlatt, V., & Urbach, N. 2021. "Gateways to Artificial Intelligence: Developing a taxonomy for AI Service platforms," *European Conference on Information Systems (ECIS)*.

Gillenson, M. L. 1985. "Trends in data administration," *MIS Quarterly*, 317-325.

Glaser, V. L., Pollock, N., & D'Adderio, L. 2021. "The biography of an algorithm: Performing algorithmic technologies in organizations," *Organization Theory*, 2(2).

Hilbert, M., & López, P. 2011. "The world's technological capacity to store, communicate, and compute information," *Science*, 332(6025), 60-65.

Holmström, J., Sandberg, J., & Mathiassen, L. 2011. "Educating reflective practitioners: The design of an IT Management Masters Program," *Americas Conference on Information Systems (AMCIS)*.

Huang, J., Henfridsson, O., Liu, M. J., & Newell, S. 2017. "Growing on steroids: Rapidly scaling the user base of digital ventures through digital innovation," *MIS Quarterly*, 41(1).

Iansiti, M., & Lakhani, K. R. 2020. Competing in the age of AI: Strategy and leadership when algorithms and networks run the world. Harvard Business Press.

Iliadis, A., & Russo, F. 2016. "Critical data studies: An introduction," *Big Data & Society*, 3(2).

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. 2020. "Data governance: Organizing data for trustworthy Artificial Intelligence," *Government Information Quarterly*, 37(3).

Jordan, M. I., & Mitchell, T. M. 2015. "Machine learning: Trends, perspectives, and prospects," *Science, 349*(6245), 255-260.

Kayhan, V. (2022). "When to Use Machine Learning: A Course Assignment," *Communications of the Association for Information Systems*, 50.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). "Algorithms at work: The new contested terrain of control," *Academy of Management Annals*, *14*(1), 366-410.

Langley, A. (1999). "Strategies for theorizing from process data," *Academy of Management review*, *24*(4), 691-710.

Leavitt, K., Schabram, K., Hariharan, P., & Barnes, C. M. 2020. "Ghost in the Machine: On Organizational Theory in the Age of Machine Learning," *Academy of Management Review*.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). "Is AI ground truth really "true"? The dangers of training and evaluating AI tools based on experts' know-what," *Management Information Systems Quarterly*.

LeCun, Y., Bengio, Y., & Hinton, G. 2015. "Deep learning*," Nature*, *521*(7553), 436-444.

Leidner, D. E., & Jarvenpaa, S. L. 1995. "The use of information technology to enhance management school education: A theoretical view," *MIS quarterly*, 265-291.

Lins, S., Pandl, K. D., Teigeler, H., Thiebes, S., Bayer, C., & Sunyaev, A. 2021. "Artificial Intelligence as a service," *Business & Information Systems Engineering*, 63(4), 441-456.

Luan, H., & Tsai, C. C. 2021. "A review of using machine learning approaches for precision education," *Educational Technology & Society*, 24(1), 250-266.

Mathiassen, L., & Purao, S. 2002. "Educating reflective systems developers," *Information Systems Journal*, 12(2), 81-102.

Merrill, M.D. 2007. "A task-centered instructional strategy," *Journal of Research on Technology in Education*, 40(1), 5-22.

Merrill, M.D. 2013. "First principles of instruction: Identifying and designing effective, efficient and engaging instruction," Hoboken, NJ: Pfeiffer/John Wiley & Sons.

Mikalsen, M., & Monteiro, E. 2021. "Acting with inherently uncertain data: Practices of data-centric knowing," *Journal of the Association for Information Systems*, 22(6), 1715-1735.

Peltarion 2022. The Peltarion deep learning platform. www.peltarion.com, accessed April 2022.

Seaver, N. 2017. "Algorithms as culture: Some tactics for the ethnography of algorithmic systems," *Big data & society*, 4(2).

Simsek, Z., Vaara, E., Paruchuri, S., Nadkarni, S., & Shaw, J. D. 2019. "New ways of seeing big data," *Academy of Management Journal*, 62(4), 971-978.

Sturm, T., Gerlach, J. P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., ... & Buxmann, P. 2021. "Coordinating human and machine learning for effective organizational learning," *MIS Quarterly*.

Umeå University (2022). "AI for business course curriculum," https://www.umu.se/en/education/syllabus/2in408/ accessed April 2022.

Waardenburg, L., Huysman, M., & Sergeeva, A. V. (2021). "In the land of the blind, the one-eyed man is king: Knowledge brokerage in the age of learning algorithms," *Organization Science*, *33*(1), 59-82.

Vaast, E., & Pinsonneault, A. 2021. "When digital technologies enable and threaten occupational identity: The delicate balancing act of data scientists," *Management Information Systems Quarterly*, 45(3).

Wang, R., & Strong, D. 1996. "Beyond accuracy: What data quality means to data consumers," *Journal of Management Information Systems*, 12(4), 5-33.

Vial, G. 2019. "Reflections on quality requirements for digital trace data in IS research," *Decision Support Systems*, 126.