Author(s): Nguyen-Duc, Anh; Sundbø, Ingrid; Nascimento, Elizamary; Conte, Tayana; Ahmed, Iftekhar; Abrahamsson, Pekka

Title: A Multiple Case Study of Artificial Intelligent System Development in Industry

Year: 2020

Version: Accepted version (Final draft)

Please cite the original version:

Nguyen-Duc, A., Sundbø, I., Nascimento, E., Conte, T., Ahmed, I., & Abrahamsson, P. (2020). A Multiple Case Study of Artificial Intelligent System Development in Industry. In EASE '20 : Proceedings of the 24th International Conference on Evaluation and Assessment in Software Engineering (pp. 1-10). ACM. https://doi.org/10.1145/3383219.3383220

A Multiple Case Study of Artificial Intelligent System Development in Industry

Anh Nguyen-Duc
School of Business, Department of Business and IT
University of South Eastern Norway
Anh.Nguyen.duc@usn.no

Ingrid Sundbø
School of Business, Department of Business and IT
University of South Eastern Norway
Ingrid.Sundbo@usn.no

Elizamary Nascimento
Institute of Computing
Universidade Federal do Amazonas
Manaus, Brazil
elizamary.souza@icomp.ufam.edu.br

Tayana Conte
Institute of Computing
Universidade Federal do Amazonas
Manaus, Brazil
tayana@icomp.ufam.edu.br

Iftekhar Ahmed
Department of Informatics
University of California, Irvine
USA
iftekha@uci.edu

Pekka Abrahamsson
Faculty of Information Technology
University of Jyväskylä
Jyväskylä, Finland
pekka.abrahamsson@jyu.fi

ABSTRACT

There is a rapidly increasing amount of Artificial Intelligence (AI) systems developed in recent years, with much expectation on its capacity of innovation and business value generation. However, the promised value of AI systems in specific business contexts might not be understood, and further integrated into the development processes. We wanted to understand how software engineering processes and practices can be applied to develop AI systems in a fast-faced, business-driven manner. As the first step, we explored contextual factors of AI development and the connections between AI developments to business opportunities. We conducted 12 semi-structured interviews in seven companies in Brazil, Norway and Southeast Asia. Our investigation revealed different types of AI systems and different AI development approaches. However, it is common that business opportunities involving with AI systems are not validated and there is lack of business-driven metrics that guide the development of AI systems. The findings have implications for future research on business-driven AI development and supporting tools and practices.

CCS CONCEPTS

• Software and its engineering • Software creation and management • Software development process management • Software development methods

EASE 2020, April 15–17, 2020, Trondheim, Norway

ACM ISBN 978-1-4503-7731-7/20/04…$15.00
https://doi.org/10.1145/3383219.3383220

KEYWORDS

Artificial Intelligent system, Software Development, Software Engineering, Business opportunity, Pivot, AI business pattern, SEMAT

ACM Reference format:

Anh Nguyen-Duc, Ingrid Sundbø, Elizamary Nascimento, Tayana Conte, Iftekhar Ahmed and Pekka Abrahamsson. 2020. A Multiple Case Study of Artificial Intelligent System Development in Industry. In *Evaluation and Assessment in Software Engineering (EASE'20), April 15–17, 2020, Trondheim, Norway*. ACM, NewYork, NY, USA, 10 pages.

https://doi.org/10.1145/3383219.3383220

1 Introduction

Among the emerging applications of software technologies, Artificial Intelligence (AI) development is among the fastest-growing track. AI system involves the interpretation of external data, learning and performing intelligent tasks [15]. According to the market research firm Tractica, the global AI software market will rapidly grow in the coming years, with revenues increasing from c.a 9.5 billion U.S. dollars in 2018 to an expected 118.6 billion U.S. dollars by 2025 [1]. Major technology companies such as Apple, Google, and Amazon are prominently featuring AI in their product launches and acquiring AI-based startups. In 2019, the adoption of AI has tripled in 12 months, which could make it the fastest paradigm shift in technology history [2]. Startups labeled as being in AI to attract 15% to 50% more funding than other technology firms [3].

The alignment of companies' business needs and requirements is the key factor driving innovation and development of software systems over the last decade. While following the trends of using

AI could increase the attractiveness of the startups, among investors, managers and entrepreneurs, many are still unsure how AI systems are developed, applied and generate business value. Moreover, a recent report found that 40% of European firms that are classified as an "*AI startup*" doesn't exploit the field of study in any tangible way for their business [4]. While business objectives should derive the requirements and design of the AI systems, the evaluation of the model against these objectives should also be systematically done. So it is of utmost importance to understand how AI development can be done in a way that business objectives can be defined, integrated seamlessly.

The development of AI systems adds an additional complexity to the overall software development, and currently, there is an emerging research area that investigates the adoption of Software Engineering principles, approaches and tools in the development of AI systems [5-9, 12]. The context, including both technical and business factors, where the AI development takes places is an important part of understanding the special characteristics of AI development. There could be special stakeholders, competence, interactions in AI development from which the AI-specific practices and tools can be derived.

This research objective are two-fold (1) contextualizing AI system development to understand business and process factors and (2) characterizing common patterns where business objectives are captured during AI system development. Research shows AI specialized roles of experts and best practices in developing AI systems [6, 7]. However, it is not clear how these roles interact with each other and how they facilitate the connection between business and technical development. We derived two research questions:

RQ1: What characterizes the context of AI system development?
RQ2: How business opportunities are captured and validated during the development of an AI system?

The paper is organized into six sections. Section 2 is a description of the background and related work. In section 3, we introduce the research methodology adopted for conducting the study. Section 4 presents our findings and Section 5 discusses these results. Finally, Section 6 concludes the paper.

2 Related work

2.1 Artificial Intelligence Systems

Computer science defines AI research as the study of "*intelligent agents*": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals [15]. The term Artificial Intelligence is used to describe machines/computers that mimic "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving" [18]. We define AI systems as a software system that includes AI modules as a sole or an important part. AI presents an opportunity to make a prediction and other types of learning tasks that have been comparatively expensive abundant and cheap [19]. As we focus on the business and engineering aspects of intelligent agents, we would consider the general area of Artificial Intelligence, instead of its subset Machine Learning.

Like other software systems [17], AI systems can be classified based on their business objectives [20]. They can be used to facilitate business activity without a direct contribution to core business value [17, 20]. For instance, AI software is built or bought to automate form filling in a startup providing Customer Relationship Management solutions. As a mediator, AI can be used to connect startups to their clients, such as using chatbots to reduce customer service costs. As a bearing object, AI is used as a part of the company's infrastructures and products. And as ubiquitous, a company's value creation entirely relies on software technologies. For instance, AI companies provide advanced image recognition systems for visual searches. Another company detects personalized patterns of a patient's health condition and can find leading indicators of potential health problems.

Lwkatare et al. [12] describe five stages of adopting AI systems for commercial use. At the first level, companies can experiment and prototype an AI solution. Secondly, they can have a non-critical deployment of AI modules. Thirdly, they can have a critical deployment of the modules. Fourthly, it is a cascading deployment of AI modules and lastly, the adoption of autonomous AI modules.

2.2 Engineering aspects of Artificial Intelligence systems

Recent studies in SE provide several frameworks capturing common steps in the process of developing an AI system [7, 8, 13]. As shown in Figure 1, the authors described a workflow of nine steps employed in Microsoft [8]. In the beginning, model builders need to decide predictive goals of the AI model, input data, expected output, and algorithms. After that input data is collected, processed (data cleaning and data labeling). The next step is feature engineering, including all activities that are performed to extract and select informative features for the AI models. After that, the model is trained, evaluated with tested datasets. When the validation is done, the model can be deployed and monitored in a real-world context.

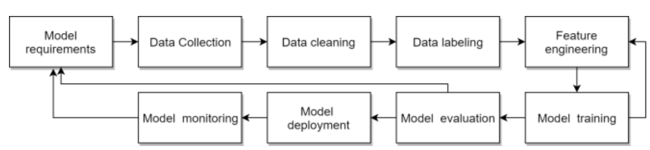

Figure 1: Nine-step workflow of AI development at Microsoft [8]

In the context of small companies, Nascimento et al. [7] describe four common steps in building AI systems (as seen in Figure 2). At the problem understanding, a team needs to understand the problem and define the goals and business metrics. At the data handling, the team define on required data and data collection strategies. At the model building, the team create the model, perform training and testing. Eventually, at the model monitoring, the AI model in the production environment is continuously monitored.

From a data management perspective, Polyzois et al. [13] organize challenges of developing an AI system into data understanding, data validation and cleaning and data preparation. Raj et al. describe AI challenges into (1) data collection, (2) data exploration, (3) data preprocessing, (4) dataset preparation, (5) data testing, (6) deployment and (7) post-deployment [14].

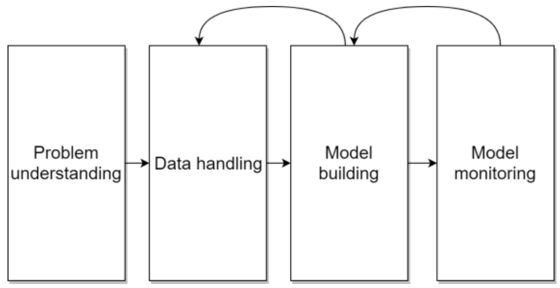

Figure 2: Common activities to build AI systems in small companies [7]

2.3 Challenges of developing Artificial Intelligence systems

The engineering challenges for developing and operating AI systems in real-world commercial settings are addressed in SE literature [6, 7, 8, 10, 12, 14, 16, 21, 22]. These challenges can be organized into seven categories: (1) Requirement, (2) Data management, (3) Model design and implementation, (4) Model configuration, (5) Model testing, (6) Evaluation and Deployment and (7) Processes and Practice. For instance, SE literature reveals requirement-type problems of identifying business metrics [7, 12], proving added value to customers [6, 10], setting goals for developing AI development [12] and transferring the problem definition into specifications [21].

In AI systems, data is the foundation for algorithmic and computational activities [6, 7, 8, 10], and there are currently many challenges relating to manage, store and process data [6, 7, 8, 10, 16].

While there are known technical challenges in building models, configuration and evaluation, there is fewer number of problems about processes and practices. Some challenges are revealed, i.e. undefined processes [7], communication challenges [16], team competences [10] and effort estimation [28]. However, these challenges are claimed in a general context, without actionable insights.

Different from the focuses of existing work, we investigated the business-technical aspect of AI systems, how the business objectives are connected to the development and integration of AI modules into a larger software system.

2.4 SEMAT ESSENCE

SEMAT (Software Engineering Method and Theory) is an initiative launched by Ivar Jacobson and colleagues [24], aiming at addressing major challenges of SE. SEMAT Essence is defined as a language that captures the essential elements (alphas) of software engineering, those that are integral to all software engineering methods. Alphas can be in certain states and are categorized in areas of concern. As shown in Figure 3, seven alphas are mentioned in the SEMAT kernel, which are Opportunity, Stakeholders, Requirements, Software, Work Team and Way of Working.

Due to our previous experience [31, 32], SEMAT Essence offers a flexible yet powerful theoretical framework to capture software engineering practices and processes. We used the SEMAT alpha, as shown in Figure 3 to design our interview protocol and later, to guide the data analysis and coding steps.

As shown in Figure 3, the SEMAT Kernel consists of seven alpha in customer, solution and endeavour categories [24], details as below:

- Opportunity: is a chance of doing something, i.e. building or enhancing software to meet a goal and to generate value
- Stakeholder: individuals, organizations, or groups that have an interest in the development and operation of the software system
- Requirement: indicates what the software system must do, but do not express how it must be done.
- Software system: includes functionality, quality and extensibility.
- Teams with a concern of the collaboration of diverse competencies and skills to achieve team effectiveness
- Work with a concern of the number of tasks to implement the requirements. Work needs to be prepared, coordinated and tracked by teams.
- Way of working is about practices, processes, tactics, tools, methods, strategies that are used by the teams to accomplish work.

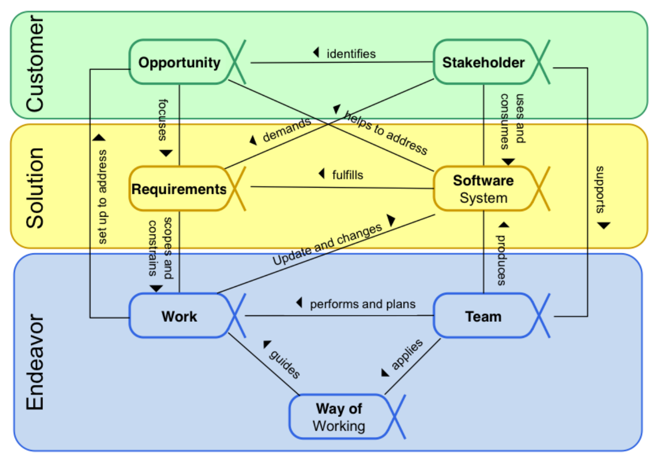

Figure 3: Alphas of the SEMAT Kernel [24]

3 Research Methodologies

Empirical AI development is an emerging research topic with limited existing knowledge [6, 12, 13]. As we would like to investigate the state-of-the-art AI integration concerns and practices in real software systems, an exploratory study is a suitable approach [25, 26]. The challenges and practices identified are based on our interpretations of the experiences of experts who implement AI systems in a real-time scenario with real-world datasets. This research approach is appropriate as it facilitates the exploration of the real-life challenges in its context through a variety of lenses [25]. A multiple exploratory study enables researchers to explore differences within and between cases. A large number of cases can help to triangulate findings, which was particularly important since we obtained only one, or at most, three interviews per case.

The data collection process is performed in two research teams, Team A (the first, second and the last author) and team B (the third and fourth author). With Team A, the first and second authors of the paper explored our professional network to search for a convenient yet suitable case. The data collection process started by exploring companies doing AI in our professional network. We collected a list of companies that develop AI systems. After that, invitations are sent to contact people from the companies. We included a wide range of stakeholders, mainly AI engineers, but also software engineers, project managers, business owners, who have in-depth understanding of the context and detail of the AI systems. Interviews were conducted with practitioners between September and November 2019.

Table 1: Interview questions relevant to this study

Q1.	Describe about the AI project you are involving
Q2.	What is your role in the AI project?
Q3.	Briefly describe other stakeholders in the project or in your team
Q4.	How do you interact with these people?
Q5.	What is the position of the project in a larger product?
Q6.	What are the business objectives of building the AI system?
Q7.	How do we know the AI system satisfies initial business objective?
Q8.	Are there any business-driven metrics applied to guide the development of the AI system?
Q9.	Architectural overview of AI systems?
Q10.	Activities involved in the development of the systems?

With Team B, the other authors of the paper investigated three companies, located in Manaus, Brazil. Two companies develop retail solution-oriented AI systems, such as credit prediction, smart stock, insurance recommendations. The third company develops AI systems to support the supervision of electronic tax documents for Amazonas State Government. In these companies, all contacted professionals had prior experience in developing non-AI systems (3-20 years) and AI systems (2-6 years). Most of the professionals we interviewed were data scientists or programmers. Interviews were conducted with practitioners between February and April 2019, with some follow-up questions in November 2019. In total, there are 13 interviews included in this study.

3.1 Data collection

Semi-structured interviews were used to acquire qualitative data. Based on the research objective of exploring state-of-practice of AI development, an interview protocol with 20 questions categorized into five sections was formulated. The first section focuses on the background of the interviewee and the project. The second section focuses on the human factors in the team. The third section addresses the alignment of business objectives to AI systems. The fourth section concerns the technical integration of AI into a large software system. Finally, the fifth section wrap-ups the interview. The interview guide was reviewed by the authors, with several modifications before finalizing into a stable version. All interviews were conducted both face-to-face and remotely via video conference. Each interview lasted from 45 to 60 minutes. All the interviews were recorded with the permission of respondents and were transcribed later for analysis. Main relevant questions to this study is shown in Table 1.

3.2 Data analysis

We used thematic analysis to analyze the data, a technique for identifying, analyzing, and reporting standards (or themes) found in qualitative data [29]. First of all, interview records were transcribed by Team A and Team B separately. Notes taken during interviews helped to focus on relevant parts of the interview transcripts. We also use a data extraction form based on the SEMAT ESSENCE framework (Figure 3) to ensure consistent extraction of information across cases. Relevant text that is within the topics of SEMAT alphas or captures the relationship among them is coded according to open coding [30]. In this work, we synthesized themes that are relevant to SEMAT concepts: (1) Opportunities, (2) Requirements, (3) Stakeholder, (4) Team, (5) System, (6) Work and (7) Way of Working.

3.3 Case description

We collected seven cases about developing and operating Artificial Intelligence systems, anonymized with alias Case 1 to Case 7.

Case 1 is an SME company, headquartered in Oslo, Norway. The company has 70-100 employees and growing. The project that was interviewed is failure detection in Smart Grid. The idea was that a drone flying along the grid lines and sending pictures of the line back to the center, where AI components will run and detect the defective grid components.

Case 2 is an emerging startup providing digital services for telecommunication and banking systems, headquartered in South East Asia. The company has c.a 20 employees. Their key product is eKYC (Know Your Customer), providing a solution to verify their clients remotely. The eKYC system scans identifiable documents (passports, ID cards, etc) and performs the verification of the clients' profile images to detect faked documents. The verified documents, together with scanned data will be integrated into customer's workflow and databases. The AI services has a high traffic, c.a. more than one million requests per month.

Case 3 is an AI startup software and product company. The company has its headquarter in Oslo, Norway, a large office in Houston, Texas and some in Palo Alto, California. Case 3 was founded in mid-2015, focus on asset-heavy industry, and offers business applications and enabling technology to companies. Case 3 is experiencing their success, and they currently have 105 employees (35 of them holds a PhD, who are working with AI models). They offer scalable products for asset administration, failure prediction, performance monitoring and optimization, deep learning for reading industrial drawing and converting it into a virtual twin. They also offer platforms for data pipeline deploying a model and edge-computing.

Case 4 is a multinational telecommunication company headquartered in Oslo, Norway. The studied case is the Artificial Intelligence research lab of the company, which locates in Trondheim, Norway. The lab investigates various applications of AI in telecommunication, i.e. detecting abnormal sensor signals, information leak, etc. The most recent project is Artificial Intelligence for Air Quality Prediction system. The project involves around 20 people from Case 4, a local university and a municipality. The project aims at building a platform that predicts real-time air quality using signals from environmental sensors attached to fixed locations or buses going around the city.

Case 5 is a startup that develops AI systems for debt renegotiation and online sales using a self-service chatbot, employee turnover prediction, and customer experience mapping across retail and hospital groups. Its headquarter is in Manaus, Amazonas, Brazil. The company was founded seven years ago, and it has been working in AI solutions development for four years. They provide AI solutions to a major retailer that has 50 years in the marketplace and more than 1.3 million registered customers. Case 5 has experts in data scientists and other necessary competencies to develop and sell products.

Case 6 is a startup that started its activities in late 2018 and has been developing AI solutions focused on retail credit solutions. Its main products are AI systems for retail credit risk assessment and smart stock. Its headquarter is in Manaus, Amazonas, Brazil, and it also has an office in São Paulo, Brazil. Case 6 has experts among its collaborators, such as data scientists and a product manager specialized in the retail market. The product manager role is responsible for negotiating the product requirements that will be developed and offered to the customer. The product manager also validates the model results and presents it to the customer, since he understands the problem and knows what the AI product can be meaningful in the customer's businesses. At the time of the study, it had a team of five data scientists and one product manager who worked on these projects.

Case 7 is an agency that works directly for the State government with its headquarter in Manaus, Amazonas, Brazil. It is composed of a team of IT professionals who work with software development for more than 20 years. It started the development of AI projects three years ago, with two professionals that are data scientists. The agency develops AI software systems to support the supervision of documents for the State Government.

4 Results

This section devotes to answering our RQs, the contextual factors in AI system development (Section 4.1) and the patterns that business opportunities are captured and validated (Section 4.2).

4.1 RQ1: What characterizes the context of AI system development?

The characteristics of AI system developments observed in our seven companies are presented according to seven SEMAT alphas, as below. For each alpha, we have a summary table to summarize the cross-case observations.

Opportunities: according to the SEMAT framework, there are many levels where opportunities present to stakeholders, i.e. being identified, validated or value are established [27]. In all cases, opportunities are identified at the business level, either CEOs or Head of Development. The reasons for building AI systems can be explorative, demonstrate the innovation capacities of the companies, search for funding, and investments. In Case 2, and Case 3 the opportunities were early addressed due to the demands from big customers.

> "*The opportunity is clear! These companies need to apply our solution to complete the task they supposed to do a long time ago… there was a similar solution provider working with them, but they fail to deliver. So the chance came to us …"* (Case 2)

In most cases, the value of desired AI systems is not directly from the real production objective. As can be seen from Table 2, all seven cases have their business opportunities identified. However, there are only four cases having their business value validated in the early stages and three cases having proofs of value establishment via operation and measurement.

Table 2: Extends that opportunities exposed to our cases

Opportunity level	Cases
Identified	1234567
Validated	-23-5-7
Value established	-23-5--

Observation 1: Opportunities with AI are often identified early in a project, however, fewer of them are validated

Stakeholders: in all of our cases, AI projects often involve multiple organizations. Common stakeholders during the development of AI systems are data providers, business use-case owners, system developers, sales and marketing staffs and end users. *Data providers:* provide access to a dataset to train and build AI models. In many cases (Case 2, 3, 5, 6, 7), the data providers are customer organizations who own the data. In Case 1, data was provided by a partner company in the project. In Case 4, data was co-created by multiple stakeholders.

> "*…the training we do separately … it comes from a different source, we have some customers flying helicopters and they have the images. When becoming our partner, they send us the images..."* (Case 1)

Business case owners: are also early customers (Case 2, 3 and 5) is an enabling factor for AI development, establishing opportunities and deriving concrete requirements. Early customers are often large organizations who own specific problems that can be automated or optimized with the adoption of AI technologies. *Sales & marketing* has an important role in providing input for the AI team, being a bridge between the the AI team and customers (as shown in Table 3). In many cases, the interaction between sale teams and AI team is very close to capture and validate business opportunities. *System developers*: AI is developed in the context of larger software systems, that are developed, tested, maintained by a team of software developers. Eventually, *end users* in our cases, actual users are clearly defined from the beginning of the AI projects.

Table 3: Stakeholder type

Stakeholders	**Cases**
Data providers	-23-56-
Business use-case Owner	-23-5--
End Users	-23-567
Software developer	1234567
Sale & Marketing	123456-

Requirements: AI systems must capture the identified opportunities and cover detailed specifications that are sufficient to infer model development. The requirements are also expected to associate with quantifiable criteria to evaluate the completeness of these requirements at the end of the projects. In our cases, there often lack criteria for evaluating if the AI models is successfully achieved. Instead, the expectation of the functional performance of the AI systems evolves over time. Another important part of the systems is non-functional requirements, i.e explainability [28]. However, we find this concern is application domain-specific. As shown in Table 4, in all of our cases, there are no specific requirement for explaining the AI models or their prediction outcomes. There are, however, requirements regarding privacy and security of data that is owned by customer organizations in five cases.

Table 4: Different types of requirements in our cases

Types of requirements	**Cases**
Functional req.	1234567
Model accuracy	1234567
Explainability	-------
Privacy and security	-23-567
Business or heuristic metric	-23-56-

Often when the customer has a business or heuristic metric already employed to solve the problem, the team typically uses this metric as a baseline to guide the development of the ML model. But when the customer does not have a defined business metric, the first model developed by the team can become a baseline, as we found in Cases 5 and Case 6.

> *"When the customer has a credit policy, that is, uses some algorithm that predicts the customer score, which is what we give to the person who has credit. What we do is compare our score with the customer score, which is what the customer has already defined." (Case 6)*

However, it is also common that the customers not having pre-defined business metric (Case 1, 4, 7).

Observation 2: There is no specific non-functional requirement addressing explainability of AI models in our cases

Teams: the AI development team might include outsourcing partners. This happens in Case 1:

> "*we hired another company to build the platform for us. We think about using in-house resource to do that in the future … when we collected the data, we build the model, at the same time, they build the platform*" (Case 1)

In all cases, teams are specialized in four common roles. *AI scientist* has a broad knowledge of ML, knows how to train model, create models and how to use statistics, and can advise multiple pipelines. *Software developers* are able to maintain the whole AI pipeline, automate the deployment, wrap the prediction outcomes into services. *Data engineers* are ones who have domain knowledge, understand the characteristics of data and perform data management activities, i.e. data cleaning, data pre-processing. *Project coordinators* coordinate the activities of all team members, manage the communication flow between internal and external partners. In addition, there is also a *Product manager* who is responsible for the product that will be developed and offered to the customer. The product manager also validates the model results and then presents it to the customer because he or she knows the market and knows what to do. As shown in Table 5, in most of our cases, there are common team roles in AI system development.

Table 5: Common team roles in AI development

Common team role	**Cases**
AI scientist	1234567
Data engineer	1234567
Software developer	1234---
Project coordinator	12345-7

Systems: Table 6 presents three scenarios for the development of AI systems, which are (1) AI as a prototype, (2) AI as a component and (3) AI as a stand-alone system. AI system is a prototype (Case 1, Case 4) and developed in a separate process with the main value propositions of the company or the development team. The outcomes of AI systems might or might not be input for real-world use cases. At this level, the concern of retraining or redeployment of models is irrelevant:

> *"In any case, do you re-train the model? No, I do not think we have done that … we may have some ideas about retraining and then redeploy the model, but I am not sure it is really in place now if the mechanism has been there yet"* (Case 1)

AI as a component is the most popular case (Case 1, 3, 5, 6, and 7): AI is developed as a component in a newly built system or an existing one. In this case, AI components is an architectural part of the whole software system, sharing cross-cutting functional and non-functional concerns.

> *"we build a microservice, the whole classification is the service, which is independent of the platform"* (Case 1)

AI as a stand-alone system (Case 2): AI is developed to provide a service that solely bases on the outcomes of AI models. Beyond the scope of AI systems, there is work required to connect the outcome of the Ai systems and the existing business workflow.

Table 6: Architectural positions of AI systems

Architectural position	Cases
AI as a prototype	1--45--
AI as a component	1-3-567
AI as a standalone system	-2-----

Observation 3: AI can be developed in different engineering contexts that makes software engineering a relevant matter

Work: as a possible unit of a task, artifacts or work in AI systems include not only model, data, but also infrastructure needs to build and deploy the models. In most of our cases, AI models are not newly developed, but adapt from existing ones:

> "*... we get the same model with known X model, and then we modify the output ... for the input images we do not need to modify much, but for the output we need because the components are different, like the Y components, number of classes, the characteristics might be different*" (Case 1)
>
> "*AI algorithm is developed for 50 years. Nowadays, when there is a new algorithm, after a few weeks, you can find their source code in Google pages ... So it does not make sense to think about algorithm development*" (Case 3)
>
> "*Normally, a ML model replaces the previous one…*" (Case 5)

An important aspect of the model is to extract and select informative features. In Case 3, 5 and 6, it is mentioned the construction of a model of models that combines input and output from different models.

Table 7: Types of work in AI systems development

Work type	Cases
AI models	1234567
Combined AI models	--356--
Digital data	1234567
Non-digital data	-23----
Free data	1234567
Paid data	---4---

Data is an essential type of work to deal with in AI system development. As shown in Table 7, our cases present a various type of data as inputs for AI models, including both physical data (data from oil and gas industry, scanned old documents, data stored in old tapes) and digital data (signals from environmental sensors, images, etc.), both structure and unstructured data, both free data (i.e. satellite images gathered from internet) and paid data. In Case 2, besides a set of real data, the company also includes a mechanism to generate simulate data to deal with the problem of insufficient seed data.

Way of working: AI team can perform their work in different ways, as shown in Table 8. As AI work involves research and experiment, the natural approach is explorative and ad-hoc manner. As a part of system development, however, the AI development might also be iterative, with regular Sprint meetings and deliveries. This can be seen from Case 1, Case 3, Case 4, Case 5, and Case 6.

Table 8: Way of working in AI system development

Development approach	Cases
Research-based	-2-4--7
Waterfall-like	-------
Scrum	1-3456-
Inter-organization coordination	123456-

Besides, as AI projects often involve more than one organization, the way of working also includes cross-team collaboration mechanisms. This is observed in six out of seven cases. In Case 4, one of the AI scientists also plays the role of a coordinator with partners who provide data and build the platform. In Case 5, the PO also plays the role of AI project coordinator, provides and makes available the dataset to the team and validates model results. In Case 7, one of the AI scientists also plays the role of a technical leader for the AI project.

Observation 4: AI development often involve external data providers and significant inter-organization coordination efforts

4.2 RQ2: How business value can be defined and evaluated for the development of an AI system?

Our cases present two common patterns, as shown in Figure 4 and Figure 5. In the scope of this work, our analysis focuses on the customer dimensions, which are about Opportunity, Stakeholder, Requirement and System. In Pattern A, a business use-case owner is the central part of the model, enables the opportunity captured and established. Pattern A is observed clearly in Case 2, 3, and 5. The business goals are quantified into specific metrics that specify both functional and non-functional requirements of AI systems. AI systems are typically deployed as a part of a bespoken enterprise-level software system. The development and operation of AI systems are integrated into the whole software system, guided by business metrics.

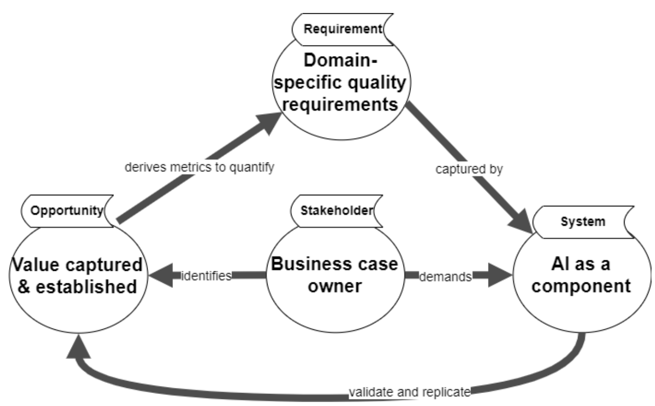

Figure 4: Pattern A - business value captured and established

An example of this pattern is Case 2. The business use-case owner is a large telecommunication company who needs to adopt a new online customer verification approach in its current business workflow. A top-down push from management level triggers the adoption of the digital solution, in which AI modules help to verify the authenticity of client's profile photos.

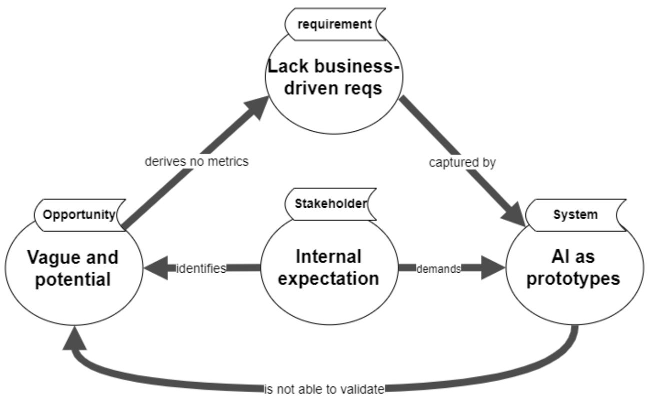

Figure 5: Pattern B - business experiments with AI prototypes

The primary stakeholder is business experts that are a part of the current workflow, who directly use the outputs from the digital solutions. The two major non-function requirements in Case 2 is (1) model accuracy and (2) data security. The model needs to scan and check more than one million documents every month. Hence, the difference between the accuracy of 98% and 99% is 10.000 wrongly classified documents. This has a direct consequence on the amount of manual fix in the workflow. Another requirement is to keep clients' data confidentiality. This is ensured by offering various ways to deliver the model that Case 2 did not need to store the customer data.

Another example of this scenario is Case 5. The business use-case owner is a large retail company that needs to take a new approach to insurance recommendation in the current store presale system. This presale system triggers an API that queries the ML model to return a prediction for the presale system on the seller's screen. The main stakeholders are store sellers, as they are part of the current workflow and deal directly with product sales. The two main non-functional requirements of Case 5 are (1) model accuracy and (2) data consistency. The model needs to verify the history of more than 1.3 million customers, specifically customers who have previously purchased insurance. The prediction system must be precise in offering insurance to customers who have previously purchased. Another requirement is to ensure that model training and test data are consistent with the current database.

Observation 5: Use-case owner enables the validation of business opportunities via the implementation and testing of business-driven requirements.

In Pattern B, business opportunities are identified and maybe somehow validated. There is a lack of a clear business case in the beginning. The ideas of AI systems mainly originates from internal sources. The requirements are derived and evolved overtime without concrete business-level metrics. The development of AI systems involves multiple MVPs. This also causes a problem of "feature creep" that AI teams try to do too many things at once and underestimating the effort needed to prepare the data. Examples of this pattern are Case 1, 4, 6 and 7. The original requirements for the AI system in Case 1 are from internal managers. As mentioned in Case 1:

> "*…at the beginning, we did not have a clear goal with what we are going to do… try to have the power provider to detect the components automatically … it is kind of a spontaneous idea…*" (Case 1)

Case 1 mentions the AI project is mainly explorative and they actively search for funding. Many demonstrations occur during the constructions of the AI model. However, they are to demonstrate useful scenarios where AI models can be useful. There is no business metric used during these demonstrations:

> *"…do you know if there is some metric to measure the success of the project? I am not sure. We have not come to an end …KPI, for example, when we have more funding, when we have more customers, this may be a bit vague, I think."* (Case 1)

There are several potential customers providing inputs to the AI systems throughout the whole project duration. In one case, the input requires a different application domain with a different type of input data. The team ends up with modifying their MVPs and chasing an actual business use case. However, it is also mentioned by the interviewee that the application domain is still new and underexplored, hence, the problem is not yet clearly defined, and available data is also scarce.

Other examples of this scenario are Case 6. The interviewee from Case 6 mentions that the AI project was initially exploratory, and they sought to verify that the data submitted by the customer could respond to the problem they expect to solve:

> "*Sometimes, the customer does not understand what he wants to solve, and he does not know the limitations. So, first, we try to understand the problem and then think of alternatives on how to solve the customer problem. We also offer new solutions to solve other problems. From the analysis of the initial problem, and that sort of thing happened in the smart inventory project.*" (Case 6)

Table 9: Comparison among two patterns

Element	Pattern A	Pattern B
Opportunity		
Value identified	235	14567
Value captured and established	235	
Requirements		
Metric-driven requirements	2356	5
Scenario-driven requirements	2356	147
Stakeholder		
Business case owner	235	14567
Internal project manager	235	
AI systems		
AI as a prototype		145
AI as a component	35	167
AI as a stand-alone	2	

Case 5 is an interesting observation, where both patterns are present. The case was initiated by an external customer, but the opportunity is captured internally without actual validation. The company had several prototypes that implement customers' ideas. Pivots occurred when technical feasibility is not achieved.

> "*It is complicated to say something at the first contact with the customer, without first analyzing the data (...) In the beginning, the customer wanted something, but when we analyzed the data provided by the customer, we concluded that it not could be done. However, we found that with the data he had, after deep analysis, we saw that we could do something else*." (Case 5)

There was the point that the model value is proof of the right business problem. The company builds a large system where the AI system is one of the key modules.

Observation 6: Internal expectation leads to the development of different AI prototypes without concrete business-driven requirements to capture evolving business opportunities.

The comparison among the two patterns is shown in Table 9, which implies that Pattern A could be a successful evolution from Pattern B with clear opportunity captured and proper requirement specifications.

5 Discussion

5.1 AI development as a business-driven, cross-organization initiative

The development of AI systems is known for its additional complexity in comparison to traditional software development due to specific AI-specific issues [16]. Efficient engineering principles and processes for AI systems is a serious concern for industry and research [8-12]. Our findings contribute to this emerging knowledge area.

At a customer and business level, our findings confirm challenges reported in previous literature i.e., lack of business metric [7, 12], convincing customers about AI value [6, 10], and transferring the problem definitions into specifications [21]. Our study demonstrates a root-cause for the lack of business-level guidelines for AI development via two patterns. We also see that there are initial intentions associated with a new AI project, but in many cases, the goals change over time to capture new business opportunities.

Lwakatare et al. [12] described their view on the evolution of AI development in a five-stage model. However, they did not mention about the transitions from one stage to another. We observed three cases that are at the lowest level "*experimentation and prototyping*" in his model. While we do not have enough cases for generalization, insights from the three cases revealed that many companies, regardless of their size and business maturity, would actually be at this stage. We strongly believe that when connecting the staged model and our AI development context factors, it would be possible for practitioners to have an overview of gaps in business, competence and practices that prevent the company to move to a higher level in the staged model.

Research has recently focused much on explainability with the ability of understanding and interpreting "black-box" AI models [23, 33]. The expectation is that an explainable model can increase the model adoption when gaining trust. Our finding reveals that explainability is not considered as a business-driven non-functional requirement, meaning that they are neither included in model specification nor driven by business objectives. While we are limited in detailed insights on different types of non-functional requirements, our thought is that the lack of adoption could be the result of challenges in defining, quantifying and testing model explainability [23].

Our findings reveal a concern that is not explored in the Software Engineering literature. AI development is often an initiative that goes beyond the boundary of a single company, which is largely due to the requirement on a large amount of data for building models. Ensuring data privacy and security, general data protection regulation (GDPR), Intellectual property (IP) are possible concerns involving cross- organizational collaboration. Future research might investigate human factors and organizational factors in the development of AI systems.

5.2 Threats to validity

Our observations were done in seven software development companies that have AI development and operation. These companies are various in terms of geographical locations, industry sectors, sizes, process maturity, and revenue. Although we saw common themes repeating across cases, and some specific topics, reaching saturation, we are aware of the limited size of our sample. Comparison to existing literature shows some similarity, however, our findings might not be applicable to all kinds of AI system development.

Regarding internal validity, our data collection protocol has been reviewed and piloted. Every interview is participated by at least two researchers, which reduces the bias and misunderstanding during interviews. Interviewees include CEO, CTO, chief scientist, leading researchers, etc., who can offer deep insights on the investigated projects. We also send a summary of our finding from interviews to the interviewee for their confirmation. The audio files and summary forms are made available online for future research.

6 Conclusions

Advancement in data technologies, machine learning algorithms and computing resources have enabled the practical development and adoption of AI in large software systems. Research has revealed a number of technical challenges when developing AI systems [11, 14, 16]. However, there is a relatively limited understanding of engineering processes and practices for AI development, especially the guideline for adopting these practices in their companies' contexts.

Our findings from seven AI developing companies in three different countries revealed different types of AI systems and different AI development approaches. In many cases, AI development is explorative and experimental. Opportunities with AI are often identified, however, not many cases have them validated and derived to a business metric that guides AI engineering activities. The journey of capturing business opportunities comes along with a series of experiments and internal validation. In this context, collaboration with business case owner is a key enabler for achieving AI-business value.

With the rapidly increasing number of AI teams across industries and application domains, our findings raise the awareness of several contextual factors that influence the successful implementation of AI systems from a business perspective. In the context that business objective, expert team and matured business domains, needed data in place, the implementation of AI systems can be straightforward. In this situation, the adoption of engineering practices and processes would enhance the productivity and quality of AI development. In other contexts, where uncertainties apply to either business domain, team, opportunities or data, there is a need for a "probe-and-sense" approach that integrates both business and technical activities to maximize the gained business value of building AI systems.

ACKNOWLEDGMENTS

In this paper, Conte's work is funded by CNPq (423149/2016-4, 311494/2017-0, and 204081/2018-1/PDE). We would like to thank all of the people who participated in piloting an early version of the interview guide and the interview participants who took time from their workdays to participate in our interviews.

REFERENCES

[1] Tractica report, 2019. Available at: https://www.tractica.com/research/artificial-intelligence-market-forecasts/

[2] MMC Ventures, 2019. The state of AI: Divergent. Available at : https://www.mmcventures.com/wp-content/uploads/2019/02/The-State-of-AI-2019-Divergence.pdf

[3] Olson, P. Forbes Staff, 2019. Available at: https://www.forbes.com/sites/parmyolson/2019/03/04/nearly-half-of-all-ai-startups-are-cashing-in-on-hype/#452f2966d022

[4] OECD, 2018. Private Equity Investment in Artificial Intelligence. OECD Going Digital Policy Note, OECD, Paris, www.oecd.org/going-digital/ai/private-equity-investment-in-artificial-intelligence.pdf..

[5] Byrne, C. 2017. Development Workflows for Data Scientists, 1st ed., O'Reilly Media, 2017, pp. 1-28

[6] Kim, M., Zimmermann, T., DeLine, R. and Begel, A. 2018. Data Scientists in Software Teams: State of the Art and Challenges. IEEE Transactions on Software Engineering. 44, 11 (Nov. 2018), 1024–1038.

[7] Nascimento, E. de S., Ahmed, I., Oliveira, E., Palheta, M.P., Steinmacher, I. and Conte, T. 2019. Understanding Development Process of Machine Learning Systems: Challenges and Solutions. 2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (Sep. 2019), 1–6.

[8] Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B. and Zimmermann, T. 2019. Software Engineering for Machine Learning: A Case Study. Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Practice (Piscataway, NJ, USA, 2019), 291–300.

[9] Begel, A. and Zimmermann, T. 2014. Analyze This! 145 Questions for Data Scientists in Software Engineering. Proceedings of the 36th International Conference on Software Engineering (New York, NY, USA, 2014), 12–23.

[10] Huijgens, H., Spadini, D., Stevens, D., Visser, N. and van Deursen, A. 2018. Software Analytics in Continuous Delivery: A Case Study on Success Factors. *Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement* (New York, NY, USA, 2018), 25:1–25:10.

[11] Khomh, F., Adams, B., Cheng, J., Fokaefs, M. and Antoniol, G. 2018. Software Engineering for Machine-Learning Applications: The Road Ahead. *IEEE Software*. 35, 5 (Sep. 2018), 81–84.

[12] Lwakatare, L.E., Raj, A., Bosch, J., Olsson, H.H. and Crnkovic, I. 2019. A Taxonomy of Software Engineering Challenges for Machine Learning Systems: An Empirical Investigation. *Agile Processes in Software Engineering and Extreme Programming* (Cham, 2019), 227–243.

[13] Polyzotis, N., Roy, S., Whang, S.E. and Zinkevich, M. 2018. Data Lifecycle Challenges in Production Machine Learning: A Survey. *SIGMOD Rec.* 47, 2 (Dec. 2018), 17–28. DOI:https://doi.org/10.1145/3299887.3299891.

[14] Munappy, A., Bosch, J., Olsson, H.H., Arpteg, A. and Brinne, B. 2019. Data Management Challenges for Deep Learning. 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA) (Aug. 2019), 140–147.

[15] Poole, D., Mackworth, A. and Goebel, R. 1997. Computational intelligence: a logical approach. Oxford University Press, Inc.

[16] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F. and Dennison, D. 2015. Hidden Technical Debt in Machine Learning Systems. *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2* (Cambridge, MA, USA, 2015), 2503–2511.

[17] Nguyen-Duc, A., Khalid, K., Shahid Bajwa, S. and Lønnestad, T. 2019. Minimum Viable Products for Internet of Things Applications: Common Pitfalls and Practices. Future Internet. 11, 2 (Feb. 2019), 50. DOI:https://doi.org/10.3390/fi11020050.

[18] Russel, S. J., and Norvig, P. 2009. Artificial Intelligence. A Modern Approach (3rd ed.). Prentice Hall

[19] Goldfarb, A., Joshua S. G., and Avi, G. 2017. What to Expect From Artificial Intelligence. MIT Sloan Management Review.

[20] Steininger, D.M. 2019. Linking information systems and entrepreneurship: A review and agenda for IT-associated and digital entrepreneurship research. Information Systems Journal. 29, 2 (2019), 363–407. DOI:https://doi.org/10.1111/isj.12206.

[21] Ishikawa, F. and Yoshioka, N. 2019. How Do Engineers Perceive Difficulties in Engineering of Machine-Learning Systems? - Questionnaire Survey. 2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SER IP) (May 2019), 2–9.

[22] Nalchigar, S. and Yu, E. 2018. Business-driven data analytics: A conceptual modeling framework. Data & Knowledge Engineering. 117, (Sep. 2018), 359–372. DOI:https://doi.org/10.1016/j.datak.2018.04.006.

[23] Gade, K., Geyik, S.C., Kenthapadi, K., Mithal, V. and Taly, A. 2019. Explainable AI in Industry. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (New York, NY, USA, 2019), 3203–3204.

[24] Jacobson, I., Ng, P.-W., McMahon, P., Spence, I. and Lidman, S. 2012. The Essence of Software Engineering: The SEMAT Kernel. Queue. 10, 10 (Oct. 2012), 40–51. DOI:https://doi.org/10.1145/2381996.2389616..

[25] Yin, R. K. 2009**.** Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage. The Canadian Journal of Action Research 14:69–71.

[26] Runeson, P. and Höst, M. 2008. Guidelines for conducting and reporting case study research in software engineering. Empirical Software Engineering. 14, 2 (Dec. 2008), 131. DOI:https://doi.org/10.1007/s10664-008-9102-8.

[27] Jacobson, I., Lawson, H. "Bud," Ng, P.-W., McMahon, P.E. and Goedicke, M. 2019. The Essentials of Modern Software Engineering: Free the Practices from the Method Prisons!. Association for Computing Machinery and Morgan & Claypool.

[28] Belani, H., Vukovic, M., Car, Z. 2019. Requirements Engineering Challenges in Building AI-Based Complex Systems. CoRR abs/1908.11791

[29] Strauss, A., Corbin, J. 1998. Basics of Qualitative Research Techniques and Procedures for Developing Grounded Theory. 2nd edition

[30] Boyatzis, R.E. 1998. Transforming qualitative information: thematic analysis and code development. Sage Publications, Thousand Oaks

[31] Kemell, K.-K., Evensen, A., Wang, X., Risku, J., Nguyen-Duc, A. and Abrahamsson, P. 2019. A Tool-Based Approach for Essentializing Software Engineering Practices. 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA) (Aug. 2019), 51–55.

[32] Nguyen-Duc, A.., Khalid, K., Lønnestad, T., Bajwa, S.S., Wang, X. and Abrahamsson, P. 2019. How do startups develop internet-of-things systems: a multiple exploratory case study. Proceedings of the International Conference on Software and System Processes (Montreal, Quebec, Canada, May 2019), 74–83.

[33] Dam, H.K., Tran, T. and Ghose, A. 2018. Explainable Software Analytics. Proceedings of the 40th International Conference on Software Engineering: New Ideas and Emerging Results (New York, NY, USA, 2018), 53–56.