Work Systems in the Indian Information Technology (IT) Industry Delivering Artificial Intelligence (AI) Solutions and the Challenges of Work from Home

Vinay Reddy Venumuddala[1] ⬥ · **Rajalaxmi Kamath**[1]

Abstract
Our study is based on a workplace ethnography conducted between Jan-May 2020 in an AI research lab of an Indian service-based IT organization, whose operations shifted from co-located work to work from home (WFH) owing to the recent pandemic. The field notes of the ethnographer, working as a full-time intern in a running AI project within this lab, is the basis for the qualitative data for this study. We discuss the socio-technical aspects and the specific challenges of distributed team-working due to the WFH norms facing such emerging research units, which are rapidly diffusing across the IT industry in the offshoring context, particularly in India. We rely on work system theory as a map to bring out key findings from our ethnographic observations. The findings point to the importance of having workflows compatible with the specific work roles in such emerging work systems – particularly for the beginner roles in the AI space. Our study contributes to the IS literature by depicting the challenges of distributed teams in a relatively novel setting emerging in offshoring contexts like the Indian IT sector, and suggests implications for managers handling AI projects and tackling employee-focused Human Resource practices in such settings.

Keywords Artificial Intelligence · Work Systems Theory · Ethnography · Work from Home

1 Introduction

Indian Information Technology (IT) and Business Process Management (BPM) service providers have gained expertise in the global IT offshoring work, owing to labour cost arbitrage, standards followed in ensuring process quality, and their capability in managing projects in distributed work environments (Carmel & Agarwal, 2008; Zahedi et al., 2016). A large pool of such service providers or vendors with the highest process standard ratings are known for executing cost-optimized yet high-quality IT-BPM projects for global clients across industry verticals (Dossani & Kenney, 2007; Jalote & Natarajan, 2019; Ramasubbu et al., 2008). Standard work processes have helped these vendors effectively manage distributed software development and BPM projects and

develop and sustain long-term relationships with their clients (Rajkumar & Mani, 2001). In recent years, established IT service providers in India have started to garner expertise in emerging technologies, such as Artificial Intelligence (AI), primarily to address the digital transformation needs of their clients (Fersht & Snowdon, 2016; NASSCOM, 2018b). Clients across industries are increasingly realizing the need to manage an unprecedented amount of data that is being generated from their processes and are looking for AI-based solutions to enhance their competitive advantage by optimizing their business processes (Chen et al., 2012; Liu, 2014; Wang et al., 2016). Such digital transformation needs are pushing them to seek AI solutions from their long-standing vendors, predominantly from the Indian IT industry (Fersht & Snowdon, 2016; NASSCOM, 2018b). Many established Indian IT firms have constituted in-house research labs to address these requirements. For example, Indian IT companies such as TCS, Infosys, Wipro, HCL have all set up AI research labs in the last five years and are pitching AI solutions to cost-effectively augment the service capabilities of their mainstream IT-BPM projects (NASSCOM, 2018a).

✉ Vinay Reddy Venumuddala
vinay.venumuddala17@iimb.ac.in

Rajalaxmi Kamath
rajalaxmik@iimb.ac.in

1 Indian Institute of Management Bangalore, Bengaluru, India

In this study, we discuss the functioning of one such AI research lab whose members were co-located together in an established service-based IT firm and highlight the peculiar project management challenges faced by this unit after the sudden imposition of Work from Home (WFH) norms due to the recent COVID-19 pandemic. This AI research unit, in addition to carrying out purely research-oriented projects, was building AI solutions for clients bundled with the organization's mainstream software and BPM solutions. The WFH imposition was an unplanned occurrence, resulting in all of their AI augmented projects being carried out in a remote/distributed team-mode. This context of our study is unique compared to the project management contexts widely discussed in the Information Systems literature around distributed teams (Herbsleb, 2007; Herbsleb & Mockus, 2003). Existing studies predominantly deal with globally distributed software development (Cataldo et al., 2007; Espinosa et al., 2007; Niazi, et al., 2016a, 2016b), and a majority of them engage with the collaboration challenges faced in the offshoring context, where the work is distributed between two sites – the client, and the vendor (Niazi, et al., 2016a, 2016b; Zahedi et al., 2016). These widely discussed distributed projects operate within the rubric of software development where the task division between members situated across sites is defined and regulated through standard process methodologies and associated practices (Grinter et al., 1999; Parnas, 1972; Strode, 2016). However, the tasks and activities in AI projects are not modular and far more entangled, thereby limiting the possibility of a clear division of tasks between different team members, relative to software development (Amershi et al., 2019; Sculley et al., 2015). Therefore, co-location of members seems to be a natural option to the emerging AI research units of Indian IT organizations, given the complexity of task interdependencies in AI projects and the near absence of standard workflows or project management strategies that are needed for tackling such a complexity. Our study is motivated by the uniqueness of this research context (emerging AI research units within service-based IT organizations) and the challenges in having remote/distributed teams, given the nature of projects that they deal with (AI projects). More importantly, given the absence of modularity in such projects, what were the problems of remote working faced by these emerging units—for whom co-location seems to be a natural option? The unplanned WFH disruption gave us an opportunity to explore these problems.

The disruption came in the form of a lockdown imposed in the month of March-2020 to contain the pandemic, which necessitated a mandatory WFH for all the employees in this industry, including these research units. News articles reported the looming lay-offs if WFH continued, especially with regard to the research units pursuing AI and data sciences (Bhalerao, 2020). It was felt that employees working in such units would be the first to be laid off during this period (Chawla, 2020). They argued that clients may prefer just the descriptive analytics without any AI capabilities over AI-based predictive or prescriptive analytics, citing that the past data could now become irrelevant for prediction in the difficult times ahead in the future (Camm & Davenport, 2020). Since we were carrying out a workplace ethnographic study between Jan-May 2020 in an AI research lab of an established service-based IT organization situated in Bengaluru, India, it was a serendipitous opportunity for us to explore the problems of remote working faced by members of this lab. As part of this ethnography, one of the researchers (the ethnographer) worked full-time in an existing AI-based project handled by the members of this unit. Because of a nationwide lockdown that was announced in the month of March-2020, the ethnographer worked for close to 3 months in office and over a month from home.

In this study, we address the above discussed research problem by attempting to answer the following specific research questions,

1. What is the socio-technical context of such newly emerging AI research units in service-based IT organizations?
2. In what way did the pandemic-imposed-WFH, leading to an unplanned remote/distributed team-working, disrupt the coordination in such units?

In probing these questions, we adopt the work system theory (WST) proposed by Alter (2013) as a theoretical frame, as it serves a map-like role in allowing us to delineate the components that make up any socio-technical system. WST is a practitioner-oriented theory and provides a perspective for understanding systems within organizations as encompassing both, the technical aspects and the non-technical social and business aspects, which are of interest to managers (Alter, 2013). It also lets us discuss how such a system delivers products or services addressing client requirements while operating within the overarching organizational context – its environment, strategies, and infrastructure. WST was also neatly aligned with a major task that was assigned to the ethnographer during his work in the AI research unit. This was to document and present to the team, the otherwise unstructured activities taken up during the execution of previous AI projects of this unit. This exercise was intended to help the senior members of this unit accommodate useful insights in their current efforts at building AI-compatible workflows for future client projects. Observations and reflections related to this exercise predominantly revolved around the key components that were highlighted by Alter (2013) while defining a work system. Therefore, this motivated us to adopt WST as a lens to interpret our ethnographic observations to highlight our study's key findings.

Broadly, the findings from this study, 1) inform the socio-technical components of emerging work systems undertaking AI projects within service-based IT organizations, 2) reveal the importance of informal communication channels between participants, particularly the beginner roles in the AI space, in the absence of suitable process workflows, and 3) highlight the importance of AI compatible workflows on top of existing collaboration platforms to solve the coordination disruption that has arisen due to the unplanned WFH. Our study contributes to the IS literature by discussing challenges of distributed teams in a relatively novel setting, that of AI research units newly emerging in the offshore outsourcing contexts like the Indian IT sector. The practical implications of our study relate to the management of such emerging technology projects, and to the employee-focused human resource practices in such settings. Before arriving at our findings, in section-2, we provide some background to our study's context, AI research units in service-based IT organizations, and discuss our theoretical frame, the work system theory. We contextualise our study by providing a brief overview of the Indian IT sector and its transition into the emerging technology space, and discuss about AI project workflows and distributed software development discussed in the extant literature. In section-3, we introduce our empirical setting and discuss our methodology. In section-4, we illustrate our ethnographic observations about the operations of the AI research unit and the coordination disruption challenges it faced, owing the pandemic-induced WFH. Here we cover in detail the nature of AI projects, activities that are undertaken in a typical AI project, the organizational context and its influence on this research unit, and the impact of WFH on the operations of this unit. In section-5, we adopt a work system framework to map out the socio-technical aspects of this AI research unit and discuss the difference in the nature of coordination between participants as they worked in the office vis-à-vis during WFH. In section-6, we summarize the answers to our aforementioned research questions, provide theoretical and practical implications, and discuss the limitations of our study and suggest future work. Section-7 concludes.

2 Background and Theoretical Framework

### 2.1 Emerging Technologies and the Indian IT Sector

Solutions around emerging technologies like AI and Cloud are becoming increasingly important to organizations across industry verticals. AI embedded within traditional information management systems is expected to have profound impact on 'human decision making' within organizations (Duan et al., 2019). Increase in volume and variety of data coupled with reduction in computation costs, are together driving organizations to extract actionable insights from data by adopting AI or Big Data analytics solutions (Abbasi et al., 2016; Chen et al., 2012). Analysing such high volumes and variety of data will require cloud services to manage the data storage and service requirements (Buytendijk, 2014). Particularly in the offshoring context, owing to the scope of these new technologies, clients are looking to their long-standing vendors or service-providers from developing nations like India for providing solutions (Fersht & Snowdon, 2016).

The Indian IT sector which is mainly known for delivering IT-BPM services, has evolved into a significant player within the global technology sourcing business, catering to clients across industry verticals around the globe. Today, both the IT and BPM segments of the industry are finding it imperative to garner expertise around emerging technologies such as Artificial Intelligence (AI) and Cloud (NASSCOM, 2018b, 2019). Their clients are seeking digital transformation driven growth sensing an urgent need to modernize their legacy systems, reduce their software maintenance costs and most importantly make their software customizable for the needs of their end users. Small and medium businesses are increasingly realizing the advantages of offshoring not only their non-core business processes but also turning to new service offerings such as AI to add value even to their core business processes (NASSCOM, 2017). While BPM contracts still remain predominantly labour based, clients however are increasingly moving towards process transformation through automation and therefore expecting Indian IT service-providers to provide effective BPM solutions augmented with emerging technologies such as AI (Fersht & Snowdon, 2016). Following Fig. 1 summarizes this evolution of the Indian IT sector.

With emerging technologies such as AI and Cloud playing a crucial role across industries, requirement for work roles around these new technologies are now being incorporated into service-based Indian IT firms alongside their traditional software work roles. While partnerships with niche players in emerging technologies is one strategy for large Indian IT organizations to gain lead into emerging technologies, in-house capability development through requisite reorganization and resource balancing measures is nevertheless the primary strategy adopted by a significant many (NASSCOM, 2017). Over the past five years, AI research labs have been established by the Indian IT giants such as TCS, Wipro, Infosys, HCL, and several others, signifying a rapid diffusion of such emerging research units within the Indian IT industry which is a well-known offshoring destination for clients (NASSCOM, 2018a). Under the umbrella of one IT organization therefore new roles such as data scientists, cloud architects, data engineers, and cloud engineers work along with traditional project management roles such as business analysts and project managers, as well

Fig. 1 Evolution of the Indian
IT sector
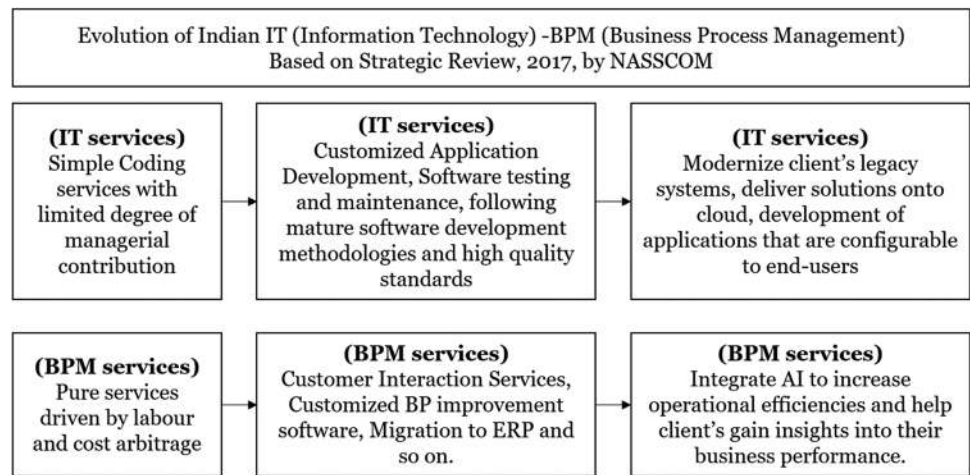
as traditional software roles such as software engineer, software architect and so on (NASSCOM, 2019).

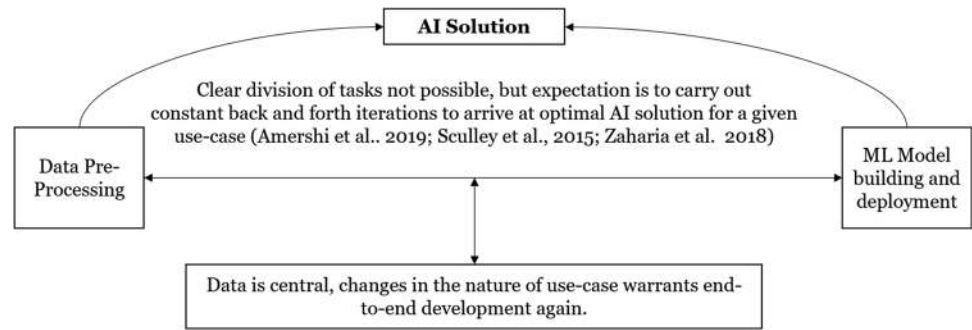## 2.2 Ideal–Typical Workflows in AI and Software Development

We study one such research unit specifically set up to work on AI projects within an established IT services organization. To understand the work carried out by members of this unit, we believe it is essential first to describe the ideal–typical AI or Machine learning (ML) project workflow, the critical work components, and their inter-relationship. We then contrast it with the workflows in traditional software development.

According to Amershi et al. (2019), 'data' which powers Machine Learning (ML) models are a central part of AI solutions. The core activities of AI/ML workflow, excluding the production deployment, typically involve two major work components, 1. Data pre-processing work, and 2. Model development work (Xin et al., 2018). The first component relies on pre-processing of structured/unstructured data in the form of cleaning, labeling, and extracting relevant features or variables. This component subsequently feeds into a second work component, ML model training and evaluation, before staging both the pre-processing and model components for production deployment (Amershi et al., 2019; Jordan & Mitchell, 2015; Sculley et al., 2015). In today's world, unstructured data in the form of text, videos, images, and speech is humongous. Therefore, pre-processing such data and extracting features from a diversity of such data forms is becoming increasingly important. For any given AI use case, pre-processing data for extracting features is intricately linked to the subsequent ML model training and evaluation. Irrespective of who handles which stage within the workflow, constant back and forth iterations experimenting with different pre-processed features and ML models are necessary for building use-case-optimal AI solutions.

According to Sculley et al. (2015), the activities within the overall ML workflow (mainly data pre-processing and model training-cum-evaluation components) are so entangled that 'changing anything changes everything.' This not only makes it difficult, but times counter-productive to use the AI modules developed for a particular use-case for another use-case, even if similar. Even if use-cases are somewhat similar, reusing the AI modules (constituted by data pre-processing components and ML models with their trained parameter settings) developed on one use-case for others result in a drop in accuracy. Interdependence between data pre-processing and ML model building, therefore makes it difficult to customize and reuse AI modules. Furthermore, the absence of strict abstraction boundaries between different ML work components forecloses the possibility of modular development, therefore warranting a constant iterative interactions between work-roles handing different stages of the ML workflow (Zaharia et al., 2018). Following Fig. 2 depicts the optimal AI solution strategy based on recent literature we discussed in this section.

Traditional software development differs from AI solution development mainly because of modularity, customizability and reusability of software programs or modules (Amershi et al., 2019; de Souza et al., 2004). Modularity requires that a software system is built as interconnected modules each of which are independently developed and communicate with other modules through application programming interfaces (APIs). According to de Souza et al. (2004), minimizing the dependency between modules by hiding implementation details of one module from another – called 'information hiding' – is one the foundational principles of software engineering. Another important aspect of software development is the practice of customization and reuse of modules. According to Amershi et al. (2019), modules that form a software system are the programs or codes in the form of functions, algorithms, libraries or other sub-modules. These modules once fully developed are available for customization

Fig. 2 Optimal AI solution
strategy

AI Solution

Clear division of tasks not possible, but expectation is to carry out
constant back and forth iterations to arrive at optimal AI solution for a given
use-case (Amershi et al.. 2019; Sculley et al., 2015; Zaharia et al. 2018)

Data Pre-Processing

ML Model
building and
deployment

Data is central, changes in the nature of use-case warrants end-to-end development again.

and reuse across several software applications. Insofar as the APIs connecting different modules remain the same, their management in terms of development and version control can proceed independent of the overall software system. These aspects of software development also enable clear division of tasks between various work roles in different stages of a typical software workflow (Dennis et al., 2018; Esbensen & Bjørn, 2014; Yilmaz et al., 2012). Business analysts and project managers arrive at high-level specifications or architecture diagrams capturing the requirements elicited from clients. These high-level specs are prepared keeping in mind modularity of tasks and customization and reuse of existing software programs. Software or technical architects refine these high-level specs by defining the functionality of and interconnections between different modules at a much finer level. The resulting modules are then independently developed by different software engineers, over standard programming platforms, following the module specific functionality and interface specifications. Independently developed modules are then integrated and tested to deliver an entire software system (Dennis et al., 2018). Clear separation of modular tasks among work roles and reusability makes it easier to have remote/distributed teams functioning across geographies and time-zones in these projects (Herbsleb & Mockus, 2003).

AI projects still rely on software development to a great extent and traditional software workflows still have their place in the context of AI projects. But given the nature of AI projects, there is a need to build customizable platforms that address the specificities of AI solution development compared to that of software development. ML flow is one such generic platform that attempts to track and manage dependencies between the pre-processing and model building work components. It facilitates the reproduction of results, given that the team can experiment with a myriad number of datasets, models, and tuning parameters. It can also support collaboration between AI and SW teams who interact during production deployment (Zaharia et al., 2018). Another more recent framework is the ModelOps developed at IBM, which serves as a platform for end-to-end development and management of AI-based applications, particularly

while working in the cloud (Hummer et al., 2019). Nevertheless, these platforms are still emerging and often are specific to the use-cases being tackled by the organizations that develop them (Amershi et al., 2019). The standardization of workflow management compatible with the peculiarities of AI project implementation is still a work in progress (Ng, 2018).

2.3 WFH and Distributed Teams

The Indian IT sector, situated at a pivotal juncture in terms of transitioning to emerging technology space, is hit by an unplanned disruption in the form of work from home (WFH), triggered by the recent COVID-19 pandemic. Though the usual software services and emerging technology projects continued to be served with employees working from home (Mathur, 2020), visible concerns around security and infrastructure limitations made it harder for Indian IT-BPM firms to facilitate their employees working from home (Economic Times, 2020). Despite these concerns, there are indications that WFH is now being seen as a new normal within the Indian IT industry (Kajarekar, 2020).

WFH as a norm brings with it a slew of changes to the culture of work within organizations. The transition of co-located teams into WFH also carries the usual implications pertaining to coordination disruption between team members, widely discussed in the IS literature around distributed teams, particularly on distributed software development. Literature on global software development speaks about some common ways in which coordination between team members gets disrupted when the members are geographically separated and interact virtually through online channels (Espinosa et al., 2007; Herbsleb, 2007). Herbsleb (2007) suggests three major ways in which coordination gets disrupted in distributed software teams – 1) Much less communication and less effective communication, 2) Lack of awareness, and 3) Incompatibilities. There is a restricted flow of information between members as they could be geographically, temporally, and socio-culturally separated (Holmstrom et al., 2006), leading to establishing connections with fewer members and less effectively (Olson & Olson, 2000). People working from

different sites share a little context about their respective work with one another, as there is a near-complete lack of informal or 'water cooler' conversations. This leads to a lack of awareness among team members about each other's work compared to co-located teams (Espinosa et al., 2007; Herbsleb & Mockus, 2003). Incompatible and conflicting work styles due to the difference in work habits such as technical tool choices or prior training and backgrounds of team members make it difficult for members to coordinate in general, specifically when the workloads are heavy (Herbsleb & Mockus, 2003).

Literature on distributed software development also discusses ways to mitigate the disadvantages of such coordination disruptions effectively. An important strategy is to optimally divide and allocate tasks between team members and ensure clarity regarding the roles and responsibilities of different team members (Nidhra et al., 2013; Orlikowski, 2002). In the context of optimal task allocation, considerations of modular design and development are crucial (Parnas, 1972). High-level architecture specifications also form an important basis for collaboration by providing greater awareness about the project to all the participants, and reduce ambiguity about the tasks that are divided between different software developers (Whitehead, 2007). Once tasks are optimally divided, collaboration technologies such as messaging tools, social networks, organizational wikis can facilitate communication channels between members in virtual mode and contribute to enhanced coordination (de Vreede et al., 2016). Instant messaging tools are particularly important in the context of improving informal communication in distributed teams, where members can chat with one another after sensing their availability status – e.g., available, busy, or in-a-meeting (Herbsleb & Mockus, 2003). Standard defined work processes also improve the communication and coordination between members distributed across sites, particularly facilitating knowledge-sharing between clients and vendors in the offshore software development contexts (Khan et al., 2009; Petersen & Wohlin, 2009).

While there is a good amount of literature on distributed teams in the context of software development, but as far as we know, the coordination disruptions around remote teams executing AI projects are not discussed in the extant literature. These disruptions are unique in comparison to those found in the context of distributed software development because as we discussed above, the tasks in AI projects are entangled in complex ways. Absence of modularity makes it less easy to divide and distribute them optimally across members, compared to software development. Further, overarching work flows for executing such projects are still emerging (Amershi et al., 2019; Sculley et al., 2015). Our study mainly focuses on this understudied context. We take the opportunity to analyse the WFH imposed challenges faced by the members of an emerging AI research

unit (whose participants were otherwise co-located) from the viewpoint of problems faced by distributed teams in such AI projects, where task division between participants is not well-defined.

2.4 Work System Theory

"A work system is a system in which human participants and/or machines perform work (processes and activities) using information, technology, and other resources to produce specific products/services for specific internal and/or external customers." (Alter, 2013)

Looking at systems merely as technical artifacts or as configurations of hardware and software that users use has been a dominant perspective in IS literature. Work System Theory (WST) proposed by Alter (2013) provides an alternative perspective to these techno-centric assumptions by providing frameworks and methods for business professionals to analyze systems in organizations – irrespective of whether or not they are related to IT. By making a shift from a purely technical system to a work system as the unit of analysis, work system theory offers a way to focus on the socio-technical aspects of such systems within organizations. For example, adopting work system theory, Marjanovic and Murthy (2016) were able to draw not only technological but also insights pertaining to organizational environment and strategies and their evolving relationships that shaped a service organization's transition from carrying out product-centric services to customer-centric services. Even before the comprehensive work systems theory was discussed in Alter (2013), previous works of Alter (Alter, 1999, 2001) inspired the adoption of the work system as the 'central and focal point of analysis for studying information systems' (Patnayakuni & Ruppel, 2010).

Work system theory offers frameworks that provide an understanding of both, static (work system framework) and dynamic (work system life cycle model) views of socio-technical systems within organizations. Work system framework represents the relatively stable 'form, function and environment' of a work system as comprising nine components divided into three major types. The first type comprises components such as processes or activities, participants, information, and technologies, which fall completely inside the work system. The second type is comprised of customers and products/services. Although they fall outside the work system, they are also required to be included within the work system because a) customers participate in the processes and activities of the system, and b) products/services intended for customers are also shaped within the work system. The third type is comprised of components such as organizational environment (e.g., organization's cultural, competitive,

Table 1 Work System Components following Alter (2013)

Component	Description
Processes and Activities	Any work system comprises activities or process-steps (if activities are structured) which can be viewed and described from a performance perspective, i.e., how the activities are actually performed within the work system
Participants	People who perform work within the work system—workforce who engage with IS development activities and those who manage development and also participate otherwise
Information	Formal (e.g., data) or informal (e.g., verbal commitments) informational entities that are exchanged by participants as they perform activities or process-steps
Technologies	Technologies include tools (hardware-software configurations) used by participants while performing different activities within the work system
Products/services	Comprise the products or services produced for the benefit and use of their customers
Customers	Customers are the recipients of products/services produced by the work system. Analysis of any work system should consider who they are, what they want and how they are going to use the provided products/services
Environment	Environment includes 'organizational, cultural, competitive, technical, regulatory, and demographic' factors that affect work system performance
Infrastructure	Resources (technical, informational, or human) that are managed outside the work system but are useful to the operation of work system
Strategies	Work system embedded in an overarching organization depends on its own strategies, and is influenced by the strategies of its parent department and the organization itself. Alignment of work system strategies with department and organizational strategies is important for the work system to function effectively

regulatory, and demographic environment within which the work system operates), infrastructure (relevant human, information, and technical resources provisioned by the organization to the work system) and strategies (alignment of enterprise strategy with work system strategy). These organizational components are largely outside the work system but directly influence the inner components of the work system. As a map, this framework allows one to identify any work system by its participants who, while performing their tasks, coordinate amongst themselves through information and technologies under structured processes or unstructured activities. It also allows for considering the interactions of this work system with clients or customers while delivering products or services satisfying their requirements. And lastly it helps to delineate the work system operations under the overarching organizational setup made up of its environment, infrastructure, and strategies. Table 1 shown below provides a brief description of these components based on the work of Alter (2013). The second framework of work system theory is the 'work system life cycle' model. This model offers a dynamic view of the work system that captures an 'iterative process through which work systems evolve… via a combination of planned change and emergent change… through adaptations, bricolage, and workarounds' (Alter, 2013, p. 13).

For the purpose of this study, we rely only on the work system framework – the static view of a work system (See Fig. 3 for a diagrammatic view). Through this framework, we try to understand the operation and components of a newly emerging work system embedded within a service-based IT organization and is delivering AI solutions to the organization's long-standing clients. We are particularly interested in understanding the challenges this emerging work system faces, specifically the coordination disruption between its participants, owing to the pandemic-induced WFH. Given the limited duration of our ethnographic study, it will be premature for us to advance a dynamic view of this work system. However, our findings hint at the workarounds that participants of this work system drove to sustain its operations within the overarching organizational context.

3 Methodology

3.1 Empirical Context

The empirical context of our study is a service-based IT organization situated in Bengaluru, India, where one of us joined as an intern in the month of January 2020 in its recently formed research unit for undertaking AI projects. Being in this industry for over 25 years, this organization is well-known as a vendor or service provider for IT-BPM services to clients across industry verticals situated across the globe. Recently, to catch up with the clients' emerging technology needs, it gathered members from traditional software teams, recruited employees for new work roles, and started an in-house research unit. This unit deals primarily with solutions around AI and Cloud. Since clients are increasingly moving from on-premise servers to cloud environments in anticipation of reduced software and data management costs, this organization has evolved itself into developing customizable cloud-based software applications. However, a major focus of this team has been to build customized AI solutions for

Fig. 3 **The Work System Framework.** (Diagram is prepared following the original depicted in Alter (2013))
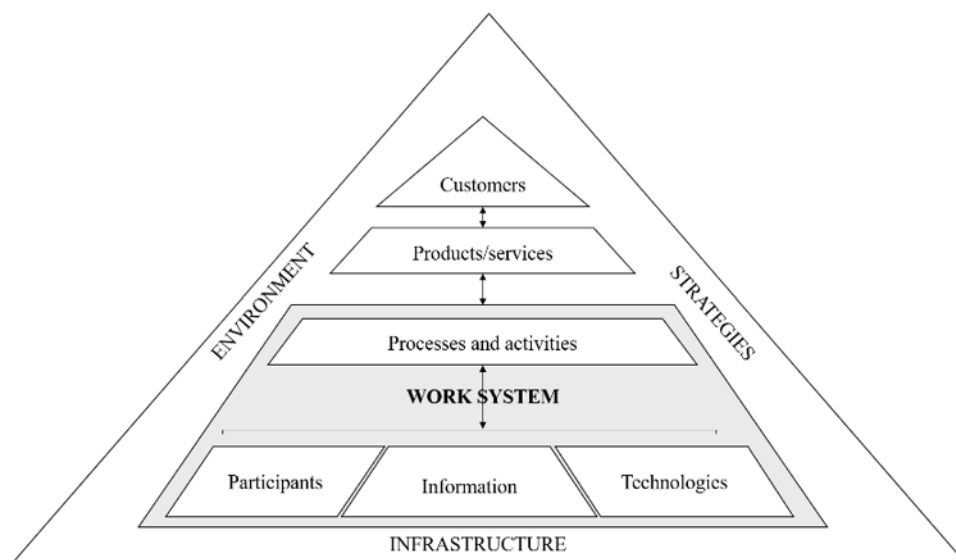
its clients seeking business process automation, and in the process, it became quite successful in accumulating proprietary AI solutions as IPs over the past five years. While client-centric AI projects constituted a major portion of IPs, research projects utilizing off-the-shelf data are another source. Its solution stack constituted by such IPs has enabled this organization to not only offer turnkey services around AI for its future clients but also to deploy them over cloud platforms such as AWS (Amazon Web Services), Azure, and GCP (Google Cloud Platform), allowing their generic usage in the online marketplace. The ethnographer worked in this team as a full-time resource in a running AI project until May 2020. Owing to the pandemic that caused lock-down and WFH close to the end of March, the ethnographer was engaged with this project work for about three months in office and slightly over a month from home.

3.2 Ethnography at Workplace

Our primary research objective was to draw insights about different work roles and their work interactions in the broader context of social mobility of Engineers within the Indian IT industry. This topic was timely as this industry is transitioning towards developing solutions around emerging technologies such as AI and Cloud, and therefore was of particular interest to us. Since there is insufficient secondary data available on these subjects, we decided to undertake primary research. We opted for participant-observation-based ethnography as we believed it would provide us a way to uncover the complexities of work and work interactions as it happened in-situ within one organization over a significant period of time. Our choice of ethnography as a methodology is motivated by the

works of Barley and Kunda (2001). According to Barley (1996), work in organizational theory has gone into the background, and many studies either simply acknowledge the complexity of work or just gloss over the 'issue of how work might be changing.' However, according to him, pushing the complexities of work into the background may not help in the understanding of any social phenomenon occurring within organizations. In today's world, the nature of work is intricately linked to organizational structures, and according to Barley and Kunda (2001), the latter have to adapt to changing nature of work to avoid the risk of 'becoming misaligned with the activities they organize' (p.2). Therefore, it is argued that bringing work back into studying organizations is imperative, especially in today's context where technologies are rapidly evolving, and markets are constantly expanding, affecting organizations both from within and from outside. In this regard, they argue for the importance of workplace ethnographies as a suitable methodology. While providing a native perspective about in-situ work practices and processes, ethnographies also motivate researchers to relate their findings with a more general understanding of organizations embedded in similar environments and comprising similar work contexts (Barley, 1996; Barley & Kunda, 2001).

Taking into account our research objectives and the ethnographer's prior qualifications (industry experience and knowledge of machine learning from doctoral coursework), the head of the research unit assigned three specific tasks to the ethnographer. (1) to work as a full-time intern in a running Natural Language Processing (NLP) based project, (2) to understand, identify and document a common workflow based on the team members' experience of working across multiple projects over the past five years, and (3) to periodically report research findings to

members of this unit. The first two tasks were considered as a value-addition that the ethnographer would bring to the team, and the last task was to help the ethnographer triangulate and validate his/her findings. The first two of these tasks allowed the ethnographer to gain a considerable understanding of the operations of this AI research unit. Observations and reflections which the ethnographer gathered while conducting these tasks, both while working from the office and from home after the lockdown, form the basis for this study. It also motivated us to consider work system theory as a theoretical frame given its focus on participants and their coordination through information and technologies while performing work to produce products/services for the clients. Working with team members in a project, joining them for breakfast, lunch, and coffee, engaging in impromptu conversations near cubicles, attending team meetings, knowledge transfer sessions, birthday celebrations, employee farewells, and many more events laid the basis for the ethnographer to build a strong rapport with the team members and allowed him/her to become a member of the team.

3.3 Data and Analysis

Our participant-observation-based ethnography provided us a way to uncover the complexities of work and work interactions as they happened in-situ within this one organization for a period of four months. Working as a full-time intern on a running AI project with other team members helped the ethnographer internalize the intricacies and complexities of work practices, work processes, and how different work roles interacted during the project and beyond. Further, the assigned task of conducting unstructured interviews with some of the team members to understand the project work flows allowed the researchers to get a concrete sense of the overarching work processes within which AI project works were carried out. Working from the office for about three months and from home for over a month allowed the ethnographer to document differences in the nature of work-role interactions during the project execution, in-office vis-à-vis at home. The contrasting observations between work interactions at the office and from home in the context of this project were the prime motivation to carry out this particular study and informed the emic perspective of this ethnography. Engaging with the members or participants of this unit allowed the ethnographer to understand the rationale for AI projects and of different work roles, the nature of the interaction between these participants, their prior educational and occupational backgrounds, and their future aspirations.

The ethnographer's own detailed notes based on the embedded understanding of the day's work, relations with co-workers, observations about the work place and

unstructured interviews (UI) and impromptu conversations (IC) with the team members formed the source of qualitative data for our study. For example, for the task assigned to the ethnographer by the team lead to document the AI project workflows by talking to the team members required the ethnographer to conduct interviews with different members of the team. The team lead formally approved taking unstructured interviews concerning their previous AI project workflows, but there was a strict prohibition to setting up meetings for such interviews to record and transcribe them. The use of devices to record conversations was against the organizational policies, and given the work schedules, the team members did not welcome formal and lengthy discussions enquiring about the project workflows. The Ethnographer, therefore, had to undertake such interviews in multiple short-duration conversations with the participants. Occasions were also sought when the team members were having their breakfast, coffee, or lunch and the ethnographer made it a point to join them. Whenever there was an opportunity, the ethnographer took notes capturing the key points during such interviews or otherwise registered such points or recalled quotes (in italics) after completing the conversation in the form of memory markers on the mobile phone. The usual impromptu conversations with team members daily, like in project status meetings, casual discussions near cubicles, during lunch, breakfast, or coffee, were registered by the ethnographer solely in the latter form. Notepad within the ethnographer's mobile phone helped in capturing these memory markers as and when he/she got the time to be physically away from the team members. At the end of each workday, the ethnographer made detailed field notes, based on these short memory markers and any other key points by recollecting and paraphrasing the unstructured interviews and conversations These were shared with the research collaborator and mentor, daily. Gaps in understanding any particular observation or participant's viewpoint were filled in during subsequent conversations with the team members. The ethnographer had to summarize his/her research findings to the team lead and members every week, and towards the end, a formal presentation was also organized to provide the team with a summary of field notes. These presentations gave scope for listening to the queries raised by the team members and also gave opportunity to ask the members questions that could correct and refine our understanding of the overall functioning of this unit.

Given the ethnographer's full-time engagement with the team for a project, there was a risk of him/her getting bogged down by the project deadlines, losing the necessary perspective for drawing generalizable insights about the industry beyond this organization's context, or even take a more 'distanced' view of the ongoings. The advantage of participant observation is that it brings rare insights through proximity,

Table 2 Coding our field notes—An example

Sample Paragraph from Field notes:

"It's difficult to show weekly status updates [during the prototype stage in AI projects] to the clients' team. They [the client, whose BPO work is also taken up by this organization] are interested in business insights from their data [using AI] in a relatively non-technical fashion from day one. To show such non-tech outputs, we need to extract insights from the data and in the beginning it will be difficult because there won't be any proper logic [for such extraction] without an understanding about clients' data. In the last few weeks, it's still ok as we get some understanding about the relevant features and insights to extract from data by then. But they [the client's team] want this at the beginning itself, also expect us to give a statement of work before we started. That's why we do some preliminary work to understand how the client's data is currently processed by the BPO work-force. They evaluate our work based on this statement." Paraphrased from discussion with a data scientist

Coding and summarization-in-brief, done by the mentor to guide the ethnographer's observations during subsequent visits:

		Emic perspective	Etic perspective	Questions/Pointers
Structure	**Technological aspects**	It is difficult to plan, structure, and phase-out the execution of AI projects during the prototype stage	AI projects call for greater interconnectedness between the business and technology, and their raison-detre is to be able to make smart business decisions	What are the technical and non-technical outputs that the participants talk about? What business insights are the clients interested in?
	Organizational aspects	Translating technical output in a non-technical fashion is critical for ensuring communication flow between vendor and the client	AI project work happens at the cusp of emerging technologies, business process outsourcing, and consultancy work	If billability is related to the work hours and work statements, what kind of problems are posed for billing AI work—since much of it still is unstructured in comparison to SW/BPO work?
Agency	**Participants' viewpoint**	The data scientist wants to build models that work optimally for the client's use case. But for that per se, she needs to have an understanding of the client's business and data	Data scientists need to be good at both technical, and business aspects of the AI projects	How do the data scientists look at AI projects which augment BPO work relative to their own research projects?

but this could also be a cause for distortion. It was here that the role of the other research collaborator, the mentor outside of the organisation, with whom the field notes were shared daily, helped. Each day's field notes were coded and independently summarized by the mentor (See Table 2, for an example) and questions and pointers to guide the ethnographer's observations during subsequent visits were made. These fields notes were classified into categories covering the meta-lens of structure and agency that was the theoretical bed-rock of this ethnography. Structure in this AI unit was captured by the technological aspects and organizational aspects and agency was captured by the participants' viewpoints. The coding was done with an aim to provide a prism for further observations and theoretical refinement. Given the nature of our ethnography, we chose to achieve reliability by continually analysing our data in a systematic fashion given the structure-agency lens. Finer coding for example, within the technological structures – workflows and work-processes, were done after the completion of the project. The later coding was also determined by Alter's WST that was the framework adopted to map the workflows. We also found it useful to summarize along the emic and etic perspectives, suggesting questions and pointers for the ethnographer to pay attention to, during the subsequent field work. The etic perspective was naturally provided by the mentor and from outside the context of this organization. By interpreting and reflecting over the field notes daily, he/she helped the ethnographer connect field observations beyond this one organization and enabled the ethnographer to sharpen his/her focus on the broad research objectives. There were weekly discussions between them about the happenings in the office and about changes happening in the Indian IT industry, and the intricacies of work in emerging technology projects. In addition, the weekly work-summary at the workplace with the team provided a means for triangulation of the insights, and gave scope for listening to the queries raised by the team members. It provided an opportunity to ask members questions that clarified the understanding about the overall functioning of this unit.

The resulting field notes were thus a synthesis of emic and etic perspectives of a two-member research team. We feel that multi-member research teams work best in such non-traditional short/medium duration ethnographies. Over four months, observations during interactions at work (either in office or from home), essence of unstructured interviews, impromptu conversations, and internalized experiences constituted our ethnographic field notes that ran into 114 single-spaced pages that amounted to close to 80,000 words. The following Table 3 summarizes the overview of respondents (anonymized) and the nature of data from portions of our field notes relating to AI projects and work from office/home, which became relevant for this study.

4 AI research unit and challenges from WFH disruption

In this section, we discuss our ethnographic observations regarding the socio-technical components of this AI research unit functioning within a service-based IT organization and the effects that the unplanned WFH cast on this unit, specifically in terms of disrupting the coordination between its participants. In particular, we describe 1) the nature of AI projects undertaken by the participants of this unit, 2) The major activities or stages of a typical AI project and specifically about the activity where the ethnographer worked hands-on with other participants, 3) the organizational context, underpinned by client-vendor relationship, within which this unit operated, and 4) the influence of WFH on this unit and its participants.

4.1 Nature of AI projects

Like most service-based IT companies in India, this organization is known for delivering Software (SW) and Business Process Outsourcing (BPO) services to clients across industry verticals such as banking and finance, logistics, healthcare, and several others. Offshore development centers (ODCs) are a significant part of this organization undertaking BPO related work, which involves maintenance of client's business processes through manual and software-based interventions. With an increasing need to digitalize and extract business insights from unstructured process data such as process logs or text documents, clients now expect AI-augmented solutions from this organization mainly to reduce costs. This is where the role of the AI research unit comes in. According to the business analyst of the AI research unit, clients are expecting an increase in per-worker productivity from the existing BPO teams so that it could lower the expected costs of services. This became important to clients as they are now witnessing a tremendous increase in the need for digitalization of their process documents, and they see a potential in the use of AI—particularly natural language processing (NLP) – to augment and hasten the manual data-entry work. The BPO workforce is usually responsible for manual or simple rule-based extraction of fields from the unstructured process data – typically text documents. Therefore, supporting BPO work with some level of automation through AI is considered crucial for this organization to gain a competitive advantage over other similar service providers.

The AI research unit within this organization builds IPs through its research projects involving the development of AI solutions for generic use-cases relying on off-the-shelf data sets. However, a major source of projects

Table 3 Overview of respondents and nature of data

Work Role(s)	Number	Data
Data Engineers	2	*The ethnographer worked with these two engineers on the on-going AI projects which gave him/ her an opportunity to converse with them on a daily basis during work. The emphasis of the conversations was around the intricacies of the current project, their work interests, and their hobbies*
Software Engineers from SW team	2	*The ethnographer got introduced to these SW engineers through the two data engineers. There were six different occasions, on the ethnographer interacted with these two to understand the nature of deployment in AI projects*
Project Manager	1	*The ethnographer was formally assigned to work under this manager. This gave the ethnographer regular opportunities to interact during work or otherwise. Summarizing our research insights, gathering feedback, and conversations around other happenings at work were the key topics of our discussion almost every day*
Cloud Architect	1	*While at work, the ethnographer accompanied the cloud architect and Project Manager to the office cafeteria for breakfast, lunch, and coffee several times. This gave the ethnographer a chance to be a part of several impromptu discussions that they had around the nature of augmented AI projects, client-vendor relationships, and other aspects related to work in this research unit*
Senior data scientist	1	*The ethnographer had interviews regarding project workflows with this senior data scientist split across three different occasions. On reaching the office early, the ethnographer was able to catch this data-scientist every morning for coffee. Discussions were primarily around the intricacies of AI project work, the challenges they face, and corrective steps they are taking to mitigate such challenges*
Junior data scientist	1	*The ethnographer had an hour-long unstructured interview with this data scientist regarding the project workflow and also regularly sought help regarding any programming-related bottlenecks that the ethnographer faced during the daily project work. The emphasis of conversations was primarily around the differences between traditional SW and AI projects*
Team Lead	1	*The Team lead was responsible for taking regular feedback on our findings. Project status meetings were the key occasions where the ethnographer got the opportunity to listen to the conversations initiated by the team lead. Such conversations were primarily concerning the progress of the work currently performed by the two data engineers and the ethnographer. The team lead also enquired about the insights we gathered from our ethnography and gave comments and posed queries that helped in refining our understanding of the complexities of work taken up by this unit*
In-house cloud/SW Engineers	2	*These two were the in-house cloud/SW engineers who occupied cubicles adjacent to the ethnographer's cubicle. This allowed the ethnographer to participate in conversations with them almost daily while in the office. Due to a sustained relationship with these members, the ethnographer was able to have informal conversations on a few occasions during WFH. Most conversations were around the intricacies of SW engineering work in AI projects and particularly the component integration and deployment aspects*
Senior Business Analyst	1	*The project manager and business analyst conducted weekly status meetings with the two data engineers and the ethnographer to track the progress of the current project. This gave the ethnographer the opportunity to witness conversations that the business analyst initiated. The ethnographer also had unstructured interviews about project workflows, scattered on two different occasions, and impromptu conversations during breakfast and lunch on several occasions around the BPO work that formed the basis for AI-augmented projects in this unit*

While in office, the ethnographer was allocated a cubicle situated in the midst of the team, which had about 30 members. The following table summarizes an overview of about 12 employees with whom the ethnographer had, a) multiple informal conversations during lunch, breakfast, and coffee, b) conversations that engaged the ethnographer near his/her cubicle, c) conversations that happened around others cubicles in which researcher got involved, d) unstructured interviews about project workflows and e) conversations that happened both in office as well as during WFH with the team members who were part of the running AI project in which the ethnographer was also involved. Beyond the members of this AI research unit, the ethnographer also interacted several times in the office with two software engineers from a SW team stationed on the same floor and worked closely during the deployment of several previous AI solutions developed by this unit.

for this unit comes in the form of AI augmentation use-cases that clients require on top of mainstream IT-BPM projects which this organization delivers. These projects often require the members of this unit to solve novel AI use-cases that again help this unit build IPs. IP development relying on client-centric projects is critical for this team to sustain itself within the organization. According to the team lead,

"IP development is not the cup of tea for all players, you need deep pockets to develop IP, because if you take our team with 35 odd people, it's a cost centre for the company. Even after investing in the IP … if we do not show regular results in terms of the IP development for [our] clients being signed up… I am sure the management might have lost patience with us and said that ok you guys are not building enough good quality IPs so let us all put you in consulting."

Within AI, this team has now gained expertise in the domains of Natural Language Processing (NLP), computer vision, graph-based machine learning, speech recognition, and has also recently ventured into quantum machine learning. Almost all client-centric projects taken up by the AI team are related to NLP and are augmented in nature, where AI augments software or manual work taken up by the BPO workforce. Not only this team but similar research teams emerging across many established service-based IT organizations in the country are also catching up to these kind of AI projects. Since a majority of such projects currently deal with processing information from unstructured text data coming from clients, the domain of NLP is dominating the landscape of AI projects, especially in the context of the Indian IT industry (NASSCOM, 2018a).

4.2 Activities in a typical AI project undertaken by this unit

As discussed above, augmented AI solutions supporting the mainstream IT-BPM projects, particularly the BPO work, are the kind of AI projects that this AI research unit predominantly undertakes. The participants involved in these AI projects are the team lead, business analysts, project managers, data scientists, data engineers, and software/cloud engineers. Team lead, business analyst, project manager are predominantly project management roles, and the remaining are technical or execution-related work roles.

4.2.1 Stages of a typical AI project

The ethnographer's task of documenting project workflows enabled us to understand the major activities undertaken by the participants of this unit while implementing a typical AI project. There are three important stages, 1) Project Initiation, requiring the unit to market their AI expertise to clients 2) Use-case identification, which requires identification of AI-augmentation use-cases in existing IT-BPM projects, and may or may not follow an intermediate proof of concept (POC) stage, depending on the novelty of use-cases, and 3) Productionizing stage, which involves execution and deployment of the resulting AI solution.

Stage I: *Project Initiation* The first stage is called *'project initiation'* which is driven by the sales team within this organization. During the regular client-vendor meetings around the organization's mainstream IT-BPM projects, the sales team of this organization often also introduces clients to the possible AI expertise their AI research unit could offer. In this regard, the sales team discusses the IPs developed by the AI unit and does a high-level overview of their functionality with clients. If clients see potential in integrating AI into mainstream projects, the sales team involves senior members of the AI research unit during the client discussions – mainly the team lead and the business analyst. The team lead requires the business analyst to attend these meetings to map out potential AI use-cases corresponding to the mainstream IT-BPM project(s). These meetings occur in the virtual mode since the client teams, the sales team, and the members of the AI research unit are geographically separated. These meetings are typically spread out over a week, and the business analyst gets ample time to discuss with his other teammates from the AI research unit while preparing any preliminary plan proposals. The business analyst often sits with data scientists and discusses with them while preparing these preliminary plan presentations. Such presentations typically capture 1) Previous project use-cases or generic research IPs that might be of interest to the client, 2) High-level architecture or functionality maps to address the potential AI use cases of the client, 3) Suggested integration with the client's mainstream IT-BPM project, 4) Tentative figures capturing the total cost of ownership if such projects are taken up. As more meetings happen with the clients, the business analyst consolidates his understanding of the client's business processes and potential AI use-cases, based on inputs gathered from the clients, their in-house BPO teams, and the data scientists. Based on this information, he/she prepares or updates high-level architecture or functionality diagrams detailing the scheme for stitching existing IPs and AI components from previous projects and how they plan to integrate it with the client's mainstream IT-BPM project.

Coordination between the team-lead, business analyst, and data scientists, typically happens face-to-face given the way they are located in the office. While team lead has a closed cubicle, it is situated in a location facing the open cubicles where all the participants – including business analysts and the data scientists sit. Given his position higher up the ladder, the team lead reserves the authority to call upon business analyst at any time and ask for the status regarding the plans for the prospective AI project. If the client approves these plans, then the next activity is the use-case identification.

Stage II: *Use-case identification* Once the project initiation phase is complete, the next stage often takes place at the client's site. This is called the *'use-case identification'* stage.

This stage aims to map client requirements for business process improvement into AI use cases. The business analyst, project manager, and data scientist are involved during this stage along with the client's business team. Data scientist brings her expertise in terms of assessing the applicability of various ML-based solutions to client's requirements. The business analyst and project manager jointly mediate communication between the data scientist and the client's business team and ensure that the use cases are mapped for client requirements in a workshop mode. At the end of this stage, the business analyst, along with the project manager, refine and prepare high-level architecture diagrams for each of the client use-cases, capturing strategies for reusing existing IPs, previous project components, and wherever required the need to develop new ones. In the latter case, data scientists are engaged further beyond this stage to showcase a prototype AI solution to the client, which the team calls 'proof of concept' (POC) for their new ideas. In this stage, a senior data scientist along with junior data scientists, experiment together on preliminary client data, if available, or any other off-the-shelf data closer to the structure of client data for developing prototype solutions. Off-the-shelf data is often present as open-source benchmark datasets or any non-confidential data in use from the previous client projects taken up by the team. Programming environments around Python and R provide packages that serve as platforms to work on the two main components – pre-processing and model building – for this prototype AI solution. This POC stage is applicable only where the AI project requires some novel strategies that the team hasn't worked on before and engages their costliest resources – the data scientists. In such projects, the components developed during POC, and high-level architecture diagrams will be the basis for customization/reuse. Work taken up during the subsequent stage is called *'productionizing'*.

Stage III: *Productionizing* Within the productionizing stage, data engineers execute the AI project and software/cloud engineers deploy the resulting solution onto cloud or on-premise client servers. Data scientists only play an informal and advisory role in this stage. Clients typically share a preliminary data sample comprising variety, but a very low volume, owing to confidentiality reasons, and assume that the AI team will build pre-processing and ML model components based on it. They expose their data fully to these components only after software/cloud engineers create appropriate wrappers and user interfaces that can maintain the anonymity of client data. Data engineers are primarily engaged in the pre-processing aspect of the AI project and rely on Python and R programming environments for building new pre-processing programs or customizing existing ones. Software/cloud engineers rely on platforms such as flask and Kubernetes to build web-based environments for

clients to access AI/ML components and to manage the overall solution deployment in the cloud. One of the reasons why the team lead advocates for customization and reuse of their previous project components is because the client data this unit works with is rarely representative. Given the absence of fully representative data, the pre-processing programs and ML models (previously developed by data-scientists as part of the POC stage or from the previously developed IP cores) go online independently. Only when these components are jointly exposed to the full volume and variety of client's data will the true performance of the AI model and the need for carrying out re-training with the same or different models, etc.., be known. Such independent development of the two main components of the AI project results in sub-optimal accuracy, as against an iterative exercise involving joint experimentation with the two components (Sculley et al., 2015). But, given that the BPO workforce gets to manually validate the results, the achieved accuracy levels are good enough to augment their work and suffice meeting the AI project objectives.

4.2.2 The Ethnographer's engagement during the productionizing stage

Productionizing is the stage that takes up a major portion of time and effort from the execution engineers of this team, i.e., Data and SW/cloud engineers. The Ethnographer worked in one AI project in its productionizing stage, along with two data engineers, a business analyst, a project manager, and the team lead. His/her role was closer to the data engineer's role, and therefore participation in project-related meetings became necessary. Weekly status meetings were organized by the business analyst and project manager with the data engineers to oversee their progress. These were precursors to the monthly status meetings that were headed by the team lead, where he/she evaluated the overall project progress. During the weekly status meetings, the business analyst evaluated the work of data engineers along a high-level framework comprising the strategy for customization/reuse from previous project components. In these meetings, statements by the business analyst, for example, *"don't do everything at once, take up one task after the other… refer to what we did in our previous project, look at the files that were prepared."* (Essence of a conversation during a weekly meeting) directed at data-engineers reminded them of the importance of customization/reuse. According to the business analyst, this minimizes the cost of the development and maintenance of AI solutions. The project manager managed the tasks and responsibilities assigned between the data engineers. Unlike the business analyst, the project manager engaged in matters related to the time allocation of data engineers in the project. As the data engineers

worked on multiple projects at a time, their challenges of multi-teaming – working simultaneously in multiple projects – are discussed periodically during the status meetings. For the technical aspects of work, they were directed to the business analyst or any of the senior or junior data scientists.

Data scientists believe that during the productionizing of client-centric AI projects, a high-level experimentation strategy – the crux of AI solutions, is often missing. Since this stage of the project takes the longest duration, data scientists, who are among the costliest human resources of this unit, are formally excluded from participating in this stage. They only undertake an informal and advisory role. Although in an advisory role, they make conscious efforts to informally mentor data engineers in their execution-related work. Data engineers informally interact with data scientists asking them about the appropriate model building strategies that could go with the kind of client data they foresee during the eventual deployment of this project. Data-scientists would suggest multiple strategies for feature extraction in their pre-processing programs so that they may work with a diverse set of ML models during deployment. This was important because it provided some leverage for experimentation in arriving at a more optimal and accurate ML model when faced with the full volume and variety of client data during deployment.

4.3 The organizational context of this AI research unit

The AI research unit ultimately had to operate within an overarching organizational context – that of a service-based IT organization. Below we discuss the influence of the client-vendor relationship that underpins this organization on the functioning of this AI research unit and some additional flexibility given to this unit to function with relatively greater independence compared to other units of this organization.

4.3.1 Influence of client-vendor relationship on AI Project Workflows

"we have carved out our own way … and developed a sweet spot to deliver [AI solutions] … [it] seems to be working well for us… it may be different in other companies where the types of clients they deal with are different… we see [that] we are aligned to the overall scheme of things, that is the organizational objectives, that is where I speak from, which are typically in terms of the financial metrics, the social metrics, and the employee client metrics" – Team Lead

The AI research unit is embedded within a service-based IT organization that is heavily influenced by the clients in

terms of its organizational values, objectives, and prospects. Although developing novel IPs around AI is a stated objective for the AI research unit, it is nevertheless influenced significantly by the overarching organizational context.

A predominant influence of this client-vendor relationship over the AI research unit comes in the form of IT-BPM project workflows and practices used in existing projects, like Agile and Scrum,[1] which end up determining the flow of activities within the AI projects. As one of the junior data scientists said, *'in service-delivery companies where research wings on AI and data analytics are setup, they still may try to follow similar kind [similar to traditional software] of agile methodologies (especially) if the team leaders are from classical software.'* Data-scientists express a concern that on many occasions during the project, there is a dominance of traditional software workflows. For example, the breakup of tasks, customization and reuse strategies prescribed by the business analyst, etc., are closer to the tenets of software development but are in stark contrast with the ideal–typical expectation of AI projects. In many AI projects, scrum managers are also allocated to oversee the weekly progress of AI projects not just during the final stages of productionizing but sometimes also during the time of building prototype solutions.

"[it was] total 8-week project [divided] into 4 sprints [within an] agile setup… scrum master will keep track of what is happening… for each sprint we will decide milestones, they will check if milestone is complete, or not. … when we were 2 weeks into, we realized they can't share their data with us… they realized it's not easy… [after] 4 months …team changed… scrum master changed… other team members here changed… problem is with new team we have to give them similar amount of time to come to pace" – Junior data scientist about one of their previous projects in productionizing stage.

According to data scientists, tracking weekly status is difficult in AI projects where there is a lot of experimentation. They have to do it because it is convenient for the clients to track progress for billing-related requirements. To make the most of such periodic status evaluation exercises, this team had to devise off-the-cuff approaches. Each week they incrementally showed descriptive statistics about different features and their correlations as extracted from a preliminary client data-corpus. Looking at these, clients provided inputs about features that were relevant and which were not, and this helped the data-scientists to conceive of appropriate strategies for building prototype pre-processing programs for

---

[1] https://en.wikipedia.org/wiki/Agile_software_development, last accessed 3rd February, 2021.

feature extraction and relevant ML models. Another reason as to why clients are more in agreement with the traditional flows, is because they also do not wish to share their full volume and variety of data which is essential for experimentation in AI. As we mentioned earlier, only after the entire AI solution gets deployed and appropriate UIs provided to clients, they prefer to expose their confidential data to the AI models.

Realizing the inefficiencies associated with the current workflows, the data-scientists are actively working to streamline such workflows and make them compatible with the ideal-type AI/ML projects, as followed by many product-based MNCs such as Google, IBM, Microsoft and so on. They are working to incorporate an open-source platform called ML-Flow, a recently developed software (Zaharia et al., 2018), and customize it for the specific kind of AI projects that are done here. One of the junior data-scientists explained about this as,

> *"To make AI solution delivery faster we need to integrate development and deployment… This can be achieved by building tools [like ML Flow] that automate many intermediate tasks that hitherto were meant to serve as hand-shake between different stages of the AI/ML work. [With this] we can iterate [or interface] between various components of the project much faster, and therefore build, test and deploy simultaneously, instead of going step by step, that is to first develop solution completely and then deploy. This is a work in progress." –Junior-data scientist*

However, it may take time to enforce these new workflows as there is inertia associated with existing workflows given that they evolved with the decades' long client-vendor interactions, which happened around the mainstream IT-BPM projects, and are relatively settled into the operations of this organization.

4.3.2 Additional flexibility given to the AI research unit

Being a research team, there is a lot more flexibility for the employees of this unit relative to other units in terms of work timings and access to resources. For them, the start and end times in office are flexible so long as the time spent in office is at least 7 h. This is in contrast to other teams in this organization, where the timings are rigid. IT staff who help employees with setting up their laptops, installing software and so on, admin staff who help employees with setting up their office infrastructure, and providing other infrastructure for enabling team meetings and discussions; and teams in ODCs—all these teams are expected to work in strict time-shifts in office as stipulated by their managers. The members of this unit also get access to open internet access which other teams do not. Additional resources such as discussion

boards, dedicated meeting rooms, etc.., are again in contrast to the facilities allowed for other teams. All the members of this research unit are seated near to each other within one floor of a building. Their office space comprising cubicles and the pathways between them are designed in such a way that individuals can freely move around and talk to one another irrespective of how far he/she is located from others within the floor. One can just get up and see if someone is available in their cubicle or not. Many employees chat across cubicles, and their conversations are heard across the hall, but no one is worried about the impression that they might create about not being busy. It seems to be a commonly accepted norm. Even the managers do not wish to control the friendly banter that employees make during their work hours.

More importantly, the office space facilitates informal collaboration between different project work roles that are not easy to establish formally. For instance, two data engineers involved in the current project, complemented each other's work while in the office. One among them was more interested in the business understanding of the project from client's point of view, and about the theoretical underpinnings of their existing ML models developed for previous projects. He used to interact with business analyst informally during their short coffee breaks and inquire about it. Before one of the project status meetings, he also requested the latter to provide him some opportunity to be present during client meetings. This person also consciously sought occasions to talk to data-scientists about the theoretical background of pertinent ML models that could be applicable for this project. He also used to involve his fellow data engineer in such discussions. The second data engineer was more interested in hands-on programming work for building pre-processing programs. He was more concerned about getting an understanding of the process of cloud deployment. He used to take his fellow mate for discussions with software/cloud engineers and architects to clarify queries about best ways to program pre-processing modules and to gain general understanding of how they usually integrate and deploy AI modules into the cloud. Since both of them mostly worked together, despite one person's initiative, both the team members used to benefit from such interactions. Office spaces were thus the key in facilitating such collaborations and work interactions.

Informal discussions between data engineers and data scientists were also possible in the office as they are seated close to each other. Given the latter's engagement with building advanced AI/ML models as part of IP development for this team and their efforts at streamlining machine learning workflows, it was not easy to get a dedicated time slot for discussion with them. Their skype (formal chat platform for employees) status used to be one of 'do-not-disturb', 'busy' or, 'in-a-meeting' all the time. To give an instance, for the

very first meeting around executing the current AI project, the ethnographer and two data engineers met a data scientist during lunch and talked to him about the project. We also asked him for his free time when we could have a detailed discussion about this. Since he was busy, he called sometime later in the evening and explained on a whiteboard in a discussion room about how he developed ML models around similar use-case for a previous project. After an hour, he quickly realized that there is another meeting that he was required to attend, and left abruptly. But over the next couple of days, the data engineers were still able to ask him their doubts and get clarifications, whenever they found him free in office.

In summary, the overarching service-based organization casts a predominant influence over the operations of this AI research unit. It imparted client-centric values and influenced project workflows, thereby shaping the task allocation between work roles and the nature of information exchanged between members. However, at the same time, the organization has provided some flexibility for this team to operate as a research wing. It provided access to resources such as open internet access, which allowed this team access to open-source tools, packages, and platforms that were critical to the implementation of AI projects at any stage. Greater flexibility in access to infrastructure, in terms of flexible work timings, collaboration conducive office spaces, and other provisions were also provided.

4.4 Mandatory WFH and its effects

Starting with a nationwide lock-down imposed due to the pandemic, WFH became an expectation that the Indian IT industry had to deal with. The Influence of this mandatory WFH was also felt in this organization and its AI research unit.

> "Less or no scope for informal knowledge sharing like in office. Every call becomes a client call where one reports status and other evaluates status. Persuading supervisors through effective communication becomes most important." – Ethnographer's reflections from attending project status meetings online, Field Notes.

In the context of the AI research unit, this unplanned shift to WFH had a unique set of repercussions. It primarily affected the dynamics of interaction between team members working on the running AI project, which was currently in the productionizing stage. The formal status meetings with project managers and the team lead happened online but still followed similar style of interactions. However, the data engineers now lost much of those informal pathways to gain the implicit knowledge from experienced data scientists, which used to be possible while in office.

Another conspicuous observation we could identify based on the online project status meetings was the mode of communication. We noticed that the two data engineers who used to work together while preparing their status presentations in the office failed to do that when the meetings shifted online. Each one of them independently prepared detailed presentations indicating their share of work done in the project. Hitherto in office, especially during the status meetings with the business analysts and managers, it was normal for the two data engineers to go onto the board intermittently while presenting their progress of work in the project. Sometimes when their presentations weren't ready, they used to present their status on the board, extempore. However, with work from home, we noticed that in every weekly status meeting and a monthly meeting, they came with detailed presentations describing what they had done over that particular week or month and what will they do in the future. They are also seen to adhere more to the high-level customization/reuse framework set by the business analyst and religiously attempt to align their presentations about tasks performed along the expectations set by such a framework. We also noticed that clarity of presentation became extremely important during these online meetings. For example, the two data engineers during their presentations in the office were able to complement each other while presenting as they worked on it or discussed it with each other in the office. However, during the online meeting, we noticed that this did not happen. Since each of them prepared their presentations independently and didn't get a chance to informally discuss, like they would have done in the office, we noticed that while one of them was presenting, the other was silent.

> "WFH demands that individuals spend a lot more time before having to present anything in a meeting. A lot more work needs to be done in preparing the presentation so that it can be shared, read out and articulated properly so that the others could understand. Further it happened that the two data engineers were asked to present their respective works, and it became difficult for them to segregate and present. While in the meeting room last time they were able to take turns and jointly present, here it was a bit difficult for them to do that, so they divided the same presentation between the two and presented." - Reflections from attending project status meetings online, Field Notes.

The overall work in the office had allowed these data engineers to gain knowledge from their teammates about other projects and also about other kinds of work beyond AI, such as integrating AI modules into existing software and their deployment in the cloud. The researcher was part of many such informal discussions that these data engineers had with their lunch mates who were also their college friends. Conversations between them always started with a common question, 'what are you currently working on?'

Fig. 4 Depiction of AI research unit through Work System Framework
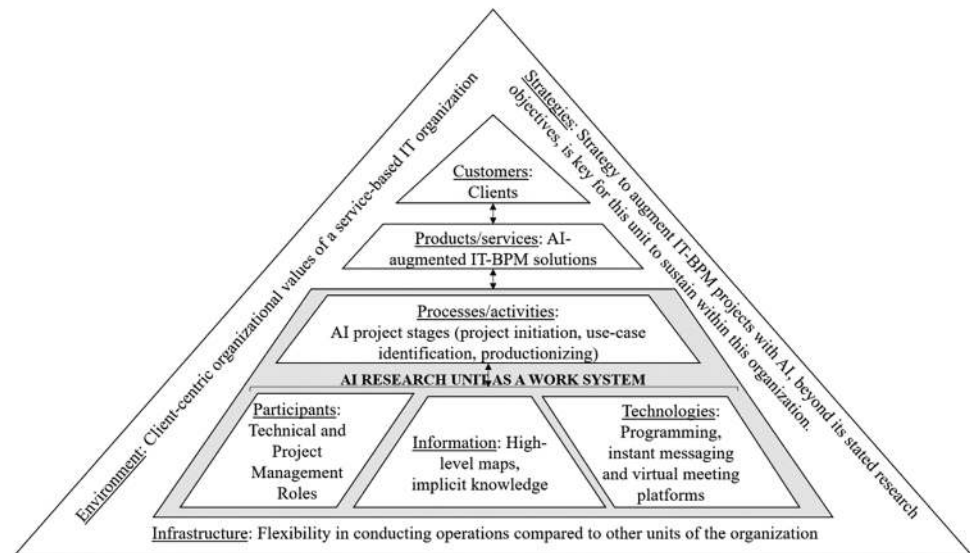
posed on a daily basis. Unknowingly, in these conversations, a good amount of talk about the organization, the projects they are working on, and the projects others were working on got passed around. However, for the data engineers, all these informal channels are now disrupted due to WFH.

5 Findings – Through the Work System Framework

Following Fig. 4 diagrammatically represents the key socio-technical components associated with the work system – the AI research unit of a service-based IT organization. At the core of this work system are participants such as the team lead, business analysts, project managers, data scientists, data engineers, and software/cloud engineers. These participants coordinated amongst themselves through certain kinds of information, like a) the illustration of AI use-cases associated with client projects as high-level functionality maps or architecture diagrams by the business analyst, b) presentations by the data engineers about their work in status meetings, c) data shared between clients and the research unit, d) implicit details about compatible pre-processing and ML model building activities shared by data-scientists, and so on. In addition to the office spaces, which allowed for dynamic and instantaneous interaction between members of this unit, there are technologies like programming environments or platforms (such as Jupyter notebook, Spyder, RStudio environments to work on Python and R programming tasks), and tools which enable instant messaging and virtual online meetings (e.g., Skype). This work system is an emerging one and is characterized by activities or processes that are still under development. Although the unit has carved out its own AI project stages, the activities undertaken in

each of these stages are largely ad-hoc, and there is a significant influence of the mainstream IT-BPM project workflows in shaping the flow of activities – particularly during the POC and productionizing stages. The end products or services delivered were in the form of AI-augmented IT-BPM solutions for the organization's long-standing clients. The client-centric organizational values and objectives clearly influenced the operation of this work system, and this is clear from a) the nature of AI projects undertaken and b) the influence of client-convenient software workflows on AI project activities. However, the parent organization made efforts to ensure some sort of flexibility in operations of this research unit by providing them additional infrastructure for conducting research activities that are typically unavailable to other organizational units.

5.1 Coordination disruption due to unplanned WFH

Following Fig. 5 diagrammatically captures the coordination between participants of this work system in office, and through technologies and information under the context of still-emerging work processes (Within the links depicted in the diagram, **T** indicates pertinent technologies or media that support coordination, and **I** indicates the nature of information exchanged between participants). In this emerging work system, the work processes or execution workflows are not yet compatible with the nature of work expected in AI projects – for example, the current process workflows, specifically during the productionizing stage, do not formally provide avenues for iterative experimentation between participants handling pre-processing and model building work which is essential to arrive at an optimal AI solution. Present workflows, which are closer to software project management, significantly influence the coordination between
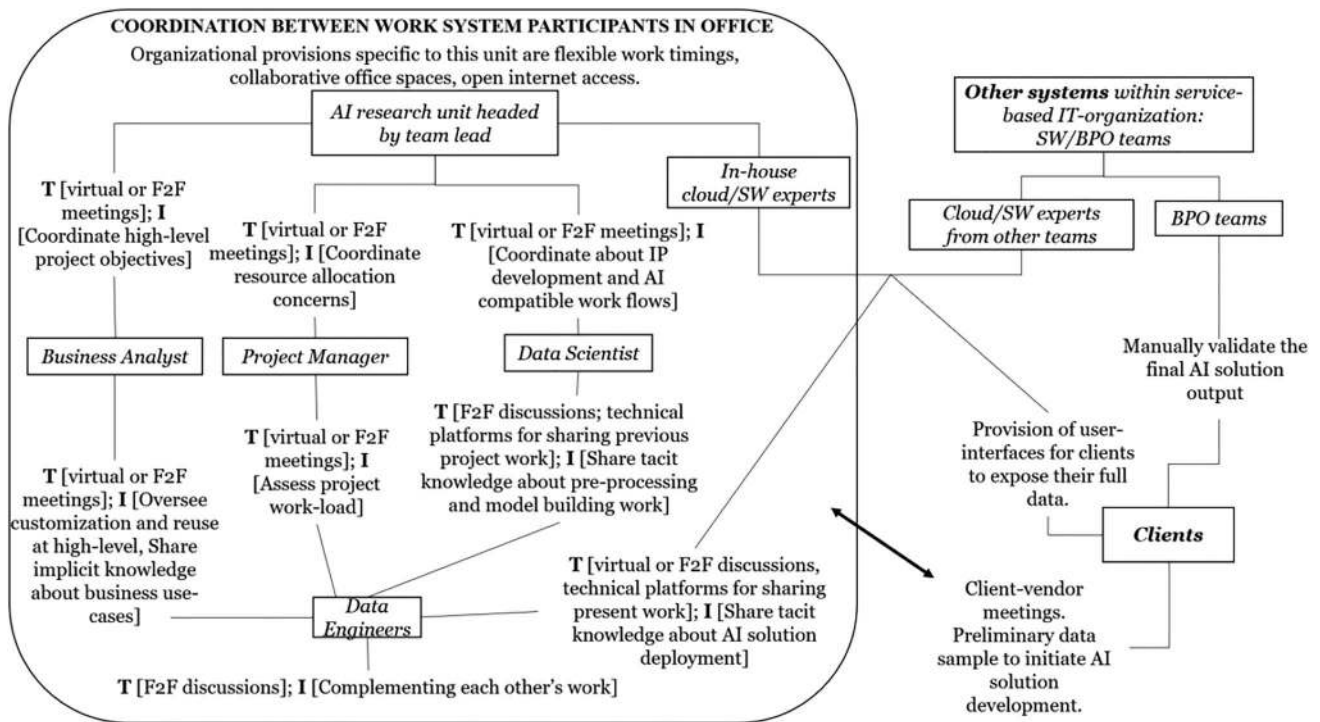
COORDINATION BETWEEN WORK SYSTEM PARTICIPANTS IN OFFICE
Organizational provisions specific to this unit are flexible work timings, collaborative office spaces, open internet access.

Fig. 5 Coordination between participants of the work system in the office

management roles, and between them and the beginner technical roles (e.g., between team lead, business analyst, and, project manager; or between them and the data engineers). The available coordination enabling technologies or media, in the form of virtual or face-to-face (F2F) meetings, only helped these participants exchange very high-level status updates. There was a conspicuous absence of technical platforms that could enable collaboration, between data engineers and data scientists, centering around iterative experimentation for building use-case optimal AI solutions. Although there was greater emphasis on customization and reuse from previous projects, the unstructured work activities did not allow this unit to build internal wikis that can help quickly access previous work and experiment with it. The usual data-sharing platforms like One-Drive were used for sharing previous project files between participants, particularly among the technical roles – data scientists, data engineers, and cloud/SW engineers.

Informal channels of communication in the office had been critical for the data engineers as they gave them a better context about their work. This context came from data scientists in the form of ideas about the best ways to perform pre-processing tasks keeping in mind the subsequent project activities, from cloud/SW engineers in the form of ideas about how different work components get integrated and deployed over the cloud, and from their co-data-engineers in the form of interest and capabilities that complemented each

other's work. These informal communications, which gave context awareness to data engineers, were not a one-time affair and were quite frequent while in office. Virtual platforms were used more for pinging each other to ask about their availability for lunch, coffee, or a casual discussion. Pre-set virtual meetings for sharing such tacit knowledge were rare. Given the tentative nature of the process workflows, such context informing communication was especially important for the data engineers – who were the beginner roles handling AI projects – and formally had very few avenues to understand where their work feeds into the overall solution.

Unfortunately, WFH norms that resulted in a distributed team environment impacted these informal work-flow channels. It made this knowledge exchange, which otherwise improved context-awareness for data engineers, especially difficult. Data engineers were central to the execution of AI projects, but a reduced scope for informal communications gave them little context about their work beyond the high-level assignments stipulated by the business analyst. For participants engaged in monitoring and overseeing roles such as business analyst and project manager, for data scientists who are engaged in purely research driven activities such as IP development, for cloud/SW experts where division of tasks are relatively clear, the impact of WFH was relatively less severe. Following Fig. 6 highlights the key coordination aspects that were disrupted due to the unplanned WFH,
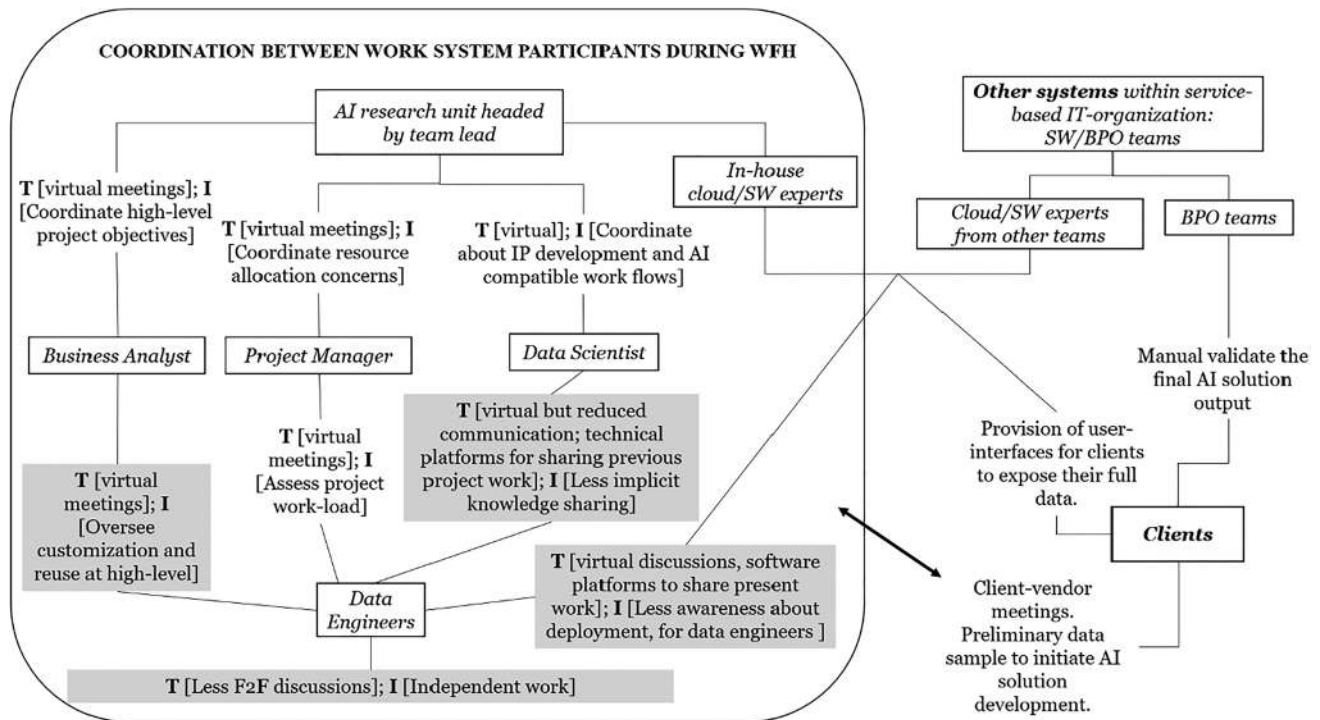
Fig. 6 Coordination between participants of the work system during WFH

particularly concerning the work carried out by the data engineers (the shaded boxes) – who were the beginner roles in the context of AI as an emerging technology.

These findings indicate that the usual collaboration-related technologies that are in place within this work system were not sufficient to mitigate the unplanned disruption due to the pandemic-induced WFH. We speculate that the beginner roles who are relatively isolated in terms of understanding their work context, such as the data engineers, are more affected, primarily because of the incompatible processes or workflows within which they operate. Such a situation could arise in any remote working context dealing with emerging technologies where the work practices do not gel well with the traditional software project management workflows. The need is to have compatible workflows that formalize the collaboration between the execution level engineers, such as the data engineers and cloud/SW engineers, and also makes the output of this collaboration transparent for the data-scientists to check and advise. Platforms like ML flow and ModelOps are poised to facilitate 1) tracking the data and algorithmic (rule-based programs or ML-based) dependencies between members handling pre-processing and ML model building work, 2) help the teams experimenting with different data-sets, models, software packages, and tuning parameters, to reproduce their results, and 3) facilitate the collaboration of AI teams with software teams especially during production deployment by automatically scheduling updates to the

pre-processing programs and ML models whenever there are triggers owing to data drift or for any other use-case specific reasons (Hummer et al., 2019; Zaharia et al., 2018). If such workflows are effectively realized, informal interactions between data engineers working on pre-processing components and data-scientists with their implicit knowledge from advanced ML model building can easily and formally collaborate. The same can happen between the data engineers and cloud engineers or between any two data engineers. However, given that the AI projects are relatively new even for clients and when they still have data sharing concerns, the implementation of these AI-compatible workflows may take time to see the light of the day.

6 Discussion

The above findings interpreted through the work system framework broadly tries to address the questions we started with. With regard to the first question, which relates to the socio-technical context within which AI research units of service-based IT organizations operate, our findings show an overwhelming dominance of the client-service provider relationship over such emerging work systems. Our study shows that the organizational strategies central to sustaining client-vendor relationships significantly influenced the activities undertaken by the AI research unit while executing

AI projects. The nature of AI projects undertaken by this unit was driven primarily by the client requirements specific to the offshoring context. Most of these projects were around AI augmentation of the mainstream IT-BPM outsourcing projects, where clients see the potential of AI in achieving process optimization and an overall cost reduction. Participants in the AI research unit occupied traditional project management roles, software development roles, and AI-specific roles such as data engineers and data scientists. Although these participants coordinated through programming platforms and environments, instant messaging, and virtual meeting tools, their informal communication channels within the office were critical to their functions, given the relatively under-developed and incompatible process workflows. In the context of this study, such informal channels were particularly important for the beginner roles in the AI space – the data engineers. Co-located work, therefore, seemed a natural option for these emerging work systems around AI operating within a service-based IT organizational context.

The second question concerns the nature of WFH disruption that resulted in remote/distributed team-working of this unit. Our findings showed that this affected the informal workflow channels, which were otherwise providing a rich context given the relative absence of any formal processes. This disruption was particularly critical for the beginner roles in the AI space – the data engineers, because the usual collaboration-related technologies widely used to facilitate communication in distributed teams were insufficient to address their peculiar challenges. There were no such collaboration tools or platforms supporting iterative experimentation for building use-case optimal AI solutions. High-level architecture diagrams which typically facilitate a model or artifact centric collaboration in traditional software development (Herbsleb, 2007; Whitehead, 2007), were less compatible for AI projects, owing to a lack of clear task division and modularity in such projects. The available unstructured and informal workflow channels were important for the data engineers to plan and execute work on their own. Therefore, for these emerging work systems there is an immediate need to reconsider the efficacy of extant processes or workflows and how the work of different participants fit into them. Our research findings indicate the need to pay attention to the extant processes or workflows particularly in the context of emerging technology projects, where the flow of work is characteristically different from traditional IT-BPM projects. We speculate that for these newly emerging work systems in an offshoring context, addressing the context-awareness problem for different participants by building compatible workflows is critical.

6.1 Theoretical implications

Our study complements existing literature on distributed teams within the IS literature. Extant literature on this subject focuses on the coordination challenges observed in the distributed software development contexts and technological tools to mitigate coordination problems between distributed teams in such contexts (de Vreede et al., 2016; Herbsleb, 2007). Much of this literature engages with the distributed teams in an offshoring context by paying attention to the challenges and solutions for ensuring effective coordination between client teams and the vendor teams (Zahedi et al., 2016). Our study complements this literature by illustrating the coordination disruption challenges in a relatively new work system that is now emerging in the offshoring context. These emerging work systems are working on projects around new technologies like AI but are not fully independent from the organizational context of the traditional service-based IT organizations. Such work systems where the work and work roles are relatively new, see co-location as a natural option, given the absence of any standard workflows or ways of optimally dividing tasks between spatially distributed participants (Herbsleb & Mockus, 2003). Given the unique context of these work systems, our study illustrates their key socio-technical components and their operations and highlights the peculiarity of coordination disruption that happens when the co-located participants of these emerging work systems are suddenly required to work remotely. Our findings point to the extent of organizational influence over these emerging work systems, particularly how the organization's extant workflows influence the operations of such work systems. It also points to the need to define compatible workflows and adopt necessary platforms to address the context-awareness problems of different participants – particularly of the beginner roles in emerging technology space, where division of tasks is difficult relative to software development projects.

This study also speaks to the importance of ethnography as a methodology to bring out the richness of contexts in which the coordination between distributed teams manifest. It is widely accepted in the IS domain that qualitative methodologies such as ethnography and case studies offer strong tools to bring out the contexts pertaining to coordination within distributed teams when compared to other positivist methods such as controlled experiments (Runeson & Höst, 2009; Zahedi et al., 2016). In this regard, we felt that the choice of our theoretical frame, Alter's work system theory, helped us map and interpret our ethnographic observations in a structured manner around the context surrounding the newly emerging work systems. Through this, we believe that our study was partly successful in bringing out the illustrative benefits of adopting theoretical frames such as the work system theory to explain ethnographic observations.

Although we were unable to discuss the dynamic view of the work system, long-duration ethnography studies can uncover rich insights about how work systems evolve or stabilize. Despite being a relatively medium-duration ethnography, our findings suggested some adaptations and workarounds that the participants carried out while adjusting to customers' expectations, overarching organizational environment, and strategies. In this regard, we believe that the findings from this study could help business professionals and academicians to gain insights into the future prospects of these new emerging technology work systems that are rapidly diffusing in an offshoring context.

6.2 Practical implications

The practical implications of this study relate to a) the individual/employee level factors often leveraged by the human resource management practitioners, and b) the management of emerging technology projects within offshore-outsourcing contexts, which is a relatively new and evolving landscape within the Indian IT sector.

Findings from our study pointed to the disruption of key context-relevant communication channels for data engineers who occupied the beginner roles in the AI space, thereby putting them at distinct disadvantage in navigating the new emerging technology work. In this unit, these engineers also came with industrial engineering backgrounds and were keen to enrich their expertise around data analytics and a substantive understanding of clients' business use-cases in AI, as they aspired to become data scientists. Despite the absence of AI-compatible project workflows, the co-located work environment offered them sufficient context that could potentially help them realize such an aspiration. Nevertheless, these data engineers expressed a crucial gap in understanding the over-all context of their work, given their modularised and stylised work processes. The following paraphrased quote from one of the data engineers clarifies this gap as stemming from an absence of clarity with regard to the new AI roles within this industry.

"I have applied to several roles before in [the] AI/ ML [space]. JDs [Job descriptions] of their work roles, like the data engineers, data scientists and others are not very clearly defined. They talk about the need to understand client business use-cases, their data, identify relevant AI use-cases, and build models. But reality is very different. Often some people without hands-on expertise tend to promise wonders to clients around AI and up their expectations. To meet their expectations given the limited time and resources [like sufficient volume and variety of data], the data engineers, who work hands-on, often end up building too many rule-based programs that cannot really be

called machine learning. We have some expectations about the role when we join, but rarely they match [with these expectations]." – Data Engineer

Our study suggests that the lack of AI-compatible project workflows could be one possible reason for this absence of role clarity and the dissatisfaction of the data engineers. It is widely acknowledged that role clarity is an important constituent of mature process workflows that ensure distributed software development work (Herbsleb & Mockus, 2003; Ramasubbu et al., 2008; Yilmaz et al., 2012). The disillusionment of data engineers in terms of mismatched expectations vs. reality can get further exacerbated in distributed environments unless compatible workflows and clarity of roles are established. Failing to do this can reduce the relatively new AI work within the Indian software services industry into a monotonous legacy software system maintenance work that is often seen as a major reason for employee attrition issues within this offshore-outsourcing-dominated industry (Agrawal et al., 2012). HR practitioners often target such attrition issues through individualistic employee-focused tools such as managerial training programs, upskilling, encouraging flexible work practices, facilitating on-site visits, and adopting fair pay and procedures. It is anticipated that these tools help employees cope with the job demands, tackle stressful environment, and navigate cultural differences, especially in globally distributed environments like the Indian software services industry (Adamovic, 2018; Agrawal et al., 2012). Our study indicates that socio-technical interventions, such as ensuring compatible project workflows, enriching context awareness, and building necessary clarity for different work roles, could be necessary prerequisites for such employee-focused management tools to work in the context of emerging technology projects like that of AI.

Findings from this study could also have practical implications for the client-facing project management roles while managing emerging technology projects in an offshoring context like the Indian IT industry. With the Indian IT industry dominated by IT-BPM service-based companies, it is very likely that the work systems around emerging technologies such as AI may be similar to the one we described in this study. The strength of the relationship between client and service provider is the key to competitive advantage in this industry, and their substantive revenue is generated from client-centric IT-BPM projects (NASSCOM, 2017; Rajkumar & Mani, 2001). Although the pandemic-induced WFH initially saw these companies struggle to facilitate their employees working from home, more and more companies are now adjusting to WFH as the new norm for this industry (Kajarekar, 2020). Owing to the pandemic, clients across verticals are realizing the importance of emerging technologies such as Cloud and AI to transform

their business processes. This change is driving several established IT companies to recruit skilled workforce in these technologies and ramp up the operations of their emerging research units (Baruah, 2020; Tavaga, 2020). Given the changing conditions witnessed after the pandemic in the context of offshoring, experts go to the extent of quoting that the next few years could be another Y2K (an earlier event that put Indian IT on the global pedestal) moment for the Indian IT industry (Roy, 2020). Despite this turnaround for the Indian IT sector, our study points to substantive unresolved challenges that the Indian IT industry cannot ignore, particularly in managing the emerging technology projects in an offshoring context. Development and standardization of emerging technology compatible workflows is particularly important as the members of such labs now increasingly have to transition from co-located work to WFH, and our study shows the importance of building such workflows for mitigating coordination disruption problems. Urgent efforts are also expected from the client-facing project management roles within these emerging work systems to engage their long-standing clients while defining, adapting, and standardizing such workflows.

6.3 Limitations and future directions

The scope of this study was to understand, a) the kind of work systems that are diffusing within the Indian IT industry, in the context of emerging technologies like AI, and b) the challenges to this work system operating as a distributed team work system owing to a sudden unplanned WFH caused by the pandemic. This high-level scope limited us in going into the finer details about how different participants used available technological platforms while working on AI-related pre-processing, model building, and deployment-related activities. Delving into the finer details could have allowed us to talk about the exact affordances offered by different technological platforms to different participants as they went about conducting project-related activities. Paying attention to these finer details could have revealed concrete findings about the possible technological interventions or workflow improvements that could be taken up to mitigate the coordination disruptions resulting from WFH in the context of distributed or remote teams in this context. Given the scope of our study, it was not possible for us to go into these finer details. Another limitation of our study is that it is specific to the Indian context, and generalizing our findings about work system characteristics, their operations, and pandemic effects etc.., may not be generalizable to other offshoring contexts outside India. The difference in the policies may also shape the prospects of such emerging work systems differently in different countries. And, this is one aspect that can be undertaken in future research.

7 Conclusion

This study explores the operations of an AI research unit of an Indian service-based IT organization and illustrates the challenges it faced owing to the recent pandemic, which required participants of this unit to shift from co-located work in the office to working as remote teams. In this regard, we discuss our ethnographic observations to bring out the socio-technical aspects of such emerging AI research units and describe how the pandemic-induced WFH has altered the nature of coordination between participants. We bring out the study's key findings by interpreting our ethnographic observations through the lens of work system framework – a key aspect of work system theory (Alter, 2013). Our findings broadly indicate the need for streamlining workflows in these emerging work systems, engaging with new technologies like AI that are rapidly diffusing across the IT industry, particularly in offshoring contexts like India. We highlight the importance of such workflows in solving the coordination disruption problems in distributed team environments, and also talk about their importance in the context of beginner-roles working in the AI space. Our study complements extant literature on distributed teams within the IS literature by discussing challenges of distributed teams in a relatively novel setting, that of AI research units newly emerging in the offshore outsourcing contexts like the Indian IT sector. The practical implications of our study relate to the management of emerging technology projects (like AI) and to the employee-focused human resource practices tackling the issues of individual workers in such contexts.

Declarations

Conflict of Interest The authors declare that there is no conflict of interest.

References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems, 17*(2), 3.

Adamovic, M. (2018). An employee-focused human resource management perspective for the management of global virtual teams. *The International Journal of Human Resource Management, 29*(14), 2159–2187.

Agrawal, N. M., Khatri, N., & Srinivasan, R. (2012). Managing growth: Human resource management challenges facing the Indian software industry. *Journal of World Business, 47*(2), 159–166.

Alter, S. (1999). A general, yet useful theory of information systems. *Communications of the Association for Information Systems, 1*(1), 13.

Alter, S. (2001). Which Life Cycle-Work System, Information System, or Software? *Communications of the Association for Information Systems, 7*(1), 17.

Alter, S. (2013). Work system theory: Overview of core concepts, extensions, and challenges for the future. *Journal of the Association for Information Systems, 14*(2). https://doi.org/10.17705/1jais.00323

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019). Software engineering for machine learning: A case study. *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, 291–300.

Barley, S. R. (1996). Technicians in the workplace: Ethnographic evidence for bringing work into organizational studies. *Administrative Science Quarterly, 41*(3), 404–441.

Barley, S. R., & Kunda, G. (2001). Bringing work back in. *Organization Science, 12*(1), 76–95.

Baruah, A. (2020). *IT sector hiring in new-age skills to pick up in 2021*. https://www.livemint.com/companies/news/it-sector-hiring-in-new-age-skills-to-pick-up-in-2021-11605686860932.html. Accessed 8 Jan 2021.

Bhalerao, S. (2020). *Nasscom warns of layoffs in IT sector; clients from these sectors hardest hit*. https://trak.in/tags/business/2020/05/25/nasscom-warns-of-mass-layoffs-in-it-sector-clients-from-these-sectors-hardest-hit/. Accessed 8 Jan 2021.

Buytendijk, F. (2014). Hype cycle for big data, 2014. *Gartner Research*. Available at: https://www.gartner.com/en/documents/2814517/hype-cycle-for-big-data-2014

Camm, J., & Davenport, T. (2020). *Data science, quarantined – MIT sloan management review*. https://sloanreview.mit.edu/article/data-science-quarantined/amp. Accessed 8 Jan 2021.

Carmel, E., & Agarwal, R. S. (2008). The Maturation of Offshore Sourcing of Information Technology Work. *MIS Quarterly Executive, 1*(2), 5.

Cataldo, M., Bass, M., Herbsleb, J. D., & Bass, L. (2007). On coordination mechanisms in global software development. *International Conference on Global Software Engineering (ICGSE 2007)*, 71–80.

Chawla, V. (2020). Are we seeing the data science bubble burst? *Analytics India Magazine*. https://analyticsindiamag.com/are-we-seeing-the-data-science-bubble-burst/. Accessed 8 Jan 2021.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188. https://doi.org/10.2307/41703503

de Souza, C. R., Redmiles, D., Cheng, L.-T., Millen, D., & Patterson, J. (2004). Sometimes you need to see through walls: A field study of application programming interfaces. *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, 63–71.

de Vreede, G.-J., Antunes, P., Vassileva, J., Gerosa, M. A., & Wu, K. (2016). Collaboration technology in teams and organizations: Introduction to the special issue. *Information Systems Frontiers, 18*(1), 1–6.

Dennis, A., Wixom, B. H., & Roth, R. M. (2018). *Systems analysis and design*. John wiley & sons.

Dossani, R., & Kenney, M. (2007). The next wave of globalization: Relocating service provision to India. *World Development, 35*(5), 772–791.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.

Economic Times. (2020). *Covid-19 impact: IT firms seek client waivers so staff can work from home*. https://economictimes.indiatimes.com/tech/ites/covid-19-impact-it-firms-seek-client-waivers-so-staff-can-work-from-home/articleshow/74642447.cms. Accessed 31 June 2020.

Esbensen, M., & Bjørn, P. (2014). Routine and standardization in global software development. *Proceedings of the 18th International Conference on Supporting Group Work*, 12–23.

Espinosa, J. A., Slaughter, S. A., Kraut, R. E., & Herbsleb, J. D. (2007). Team knowledge and coordination in geographically distributed software development. *Journal of Management Information Systems, 24*(1), 135–169.

Fersht, P., & Snowdon, J. (2016). *Making the leap from strategic to effective BPM*. National Association of Software and Service Companies. https://www.nasscom.in/knowledge-center/publications/making-leap-effective-strategic-bpm. Accessed 16 Apr 2020.

Grinter, R. E., Herbsleb, J. D., & Perry, D. E. (1999). The geography of coordination: Dealing with distance in R&D work. *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work*, 306–315.

Herbsleb, J. D. (2007). Global software engineering: The future of socio-technical coordination. *Future of Software Engineering (FOSE '07), 2007*, 188–198. https://doi.org/10.1109/FOSE.2007.11

Herbsleb, J. D., & Mockus, A. (2003). An empirical study of speed and communication in globally distributed software development. *IEEE Transactions on Software Engineering, 29*(6), 481–494.

Holmstrom, H., Conchúir, E. Ó., Agerfalk, J., & Fitzgerald, B. (2006). Global software development challenges: A case study on temporal, geographical and socio-cultural distance. *2006 IEEE International Conference on Global Software Engineering (ICGSE'06)*, 3–11.

Hummer, W., Muthusamy, V., Rausch, T., Dube, P., El Maghraoui, K., Murthi, A., & Oum, P. (2019). Modelops: Cloud-based lifecycle management for reliable and trusted ai. *2019 IEEE International Conference on Cloud Engineering (IC2E)*, 113–120.

Jalote, P., & Natarajan, P. (2019). The growth and evolution of India's software industry. *Communications of the ACM, 62*(11), 64–69.

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

Kajarekar, R. (2020). *Nasscom says 100% work from home not feasible for IT companies; Is Nasscom rooting for hybrid model?* Trak.in - Indian Business of Tech, Mobile & Startups. https://trak.in/tags/business/2020/12/05/nasscom-says-100-work-from-home-not-feasible-for-it-companies-is-nasscom-rooting-for-hybrid-model/. Accessed 8 Jan 2021.

Khan, S. U., Niazi, M., & Ahmad, R. (2009). Critical success factors for offshore software development outsourcing vendors: A systematic literature review. *2009 Fourth IEEE International Conference on Global Software Engineering*, 207–216.

Liu, Y. (2014). Big data and predictive business analytics. *The Journal of Business Forecasting, 33*(4), 40.

Marjanovic, O., & Murthy, V. (2016). From product-centric to customer-centric services in a financial institution–exploring the organizational challenges of the transition process. *Information Systems Frontiers, 18*(3), 479–497.

Mathur, V. (2020). Work from home: If the experiment works, it can continue when the world is better again. news18.com. https://www.news18.com/news/tech/work-from-home-if-the-experiment-works-it-can-continue-when-the-world-is-better-again-2561401.html

NASSCOM. (2017). *IT-BPM Strategic Review*. National Association of Software and Service Companies.

NASSCOM. (2018a). *Artificial Intelligence Primer*. National Association of Software and Service Companies.

NASSCOM. (2018b). *Talent Demand and Supply Report, AI and Big Data Analytics*. National Association of Software and Service Companies.

NASSCOM. (2019). *Uncovering the True Value of AI, Executive AI Playbook for Enterprises*. National Association of Software and Service Companies.

Ng, A. (2018). AI transformation playbook: How to lead your company into the AI era. *Landing AI*. https://landing.ai/resources/ai-transformation-playbook/

Niazi, M., Mahmood, S., Alshayeb, M., Qureshi, A. M., Faisal, K., & Cerpa, N. (2016a). Toward successful project management in global software development. *International Journal of Project Management, 34*(8), 1553–1567.

Niazi, M., Mahmood, S., Alshayeb, M., Riaz, M. R., Faisal, K., Cerpa, N., Khan, S. U., & Richardson, I. (2016b). Challenges of project management in global software development: A client-vendor analysis. *Information and Software Technology, 80*, 1–19.

Nidhra, S., Yanamadala, M., Afzal, W., & Torkar, R. (2013). Knowledge transfer challenges and mitigation strategies in global software development—A systematic literature review and industrial validation. *International Journal of Information Management, 33*(2), 333–355.

Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction, 15*(2–3), 139–178.

Orlikowski, W. J. (2002). Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science, 13*(3), 249–273.

Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. In *Pioneers and Their Contributions to Software Engineering* (pp. 479–498). Springer.

Patnayakuni, R., & Ruppel, C. P. (2010). A socio-technical approach to improving the systems development process. *Information Systems Frontiers, 12*(2), 219–234.

Petersen, K., & Wohlin, C. (2009). Context in industrial software engineering research. *2009 3rd International Symposium on Empirical Software Engineering and Measurement*, 401–404.

Rajkumar, T. M., & Mani, R. V. S. (2001). Offshore software development. *Information Systems Management, 18*(2), 63–74.

Ramasubbu, N., Mithas, S., Krishnan, M. S., & Kemerer, C. F. (2008). Work dispersion, process-based learning, and offshore software development performance. *MIS Quarterly, 32*(2), 437–458. https://doi.org/10.2307/25148847

Roy, S. (2020). *Pandemic proves to be a boon for Indian IT*. Tribuneindia News Service. https://www.tribuneindia.com/news/comment/pandemic-proves-to-be-a-boon-for-indian-it-166070. Accessed 8 Jan 2021.

Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering, 14*(2), 131–164.

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., & Dennison, D. (2015). Hidden technical debt in machine learning systems. *NIPS*, 2494–2502. https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf

Strode, D. E. (2016). A dependency taxonomy for agile software development projects. *Information Systems Frontiers, 18*(1), 23–46.

Tavaga, R. (2020). *Indian IT Stocks On Cloud 9! What Next?* Investing. Com India. https://in.investing.com/analysis/indian-it-stocks-on-cloud-9-what-next-200449643

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics, 176*, 98–110.

Whitehead, J. (2007). Collaboration in software engineering: A roadmap. *Future of Software Engineering (FOSE'07)*, 214–225.

Xin, D., Ma, L., Song, S., & Parameswaran, A. (2018). How Developers Iterate on Machine Learning Workflows–A Survey of the Applied Machine Learning Literature. *ArXiv Preprint ArXiv:1803.10311*.

Yilmaz, M., O'Connor, R. V., & Clarke, P. (2012). A systematic approach to the comparison of roles in the software development processes. *International Conference on Software Process Improvement and Capability Determination*, 198–209.

Zaharia, M., Chen, A., Davidson, A., Ghodsi, A., Hong, S. A., Konwinski, A., Murching, S., Nykodym, T., Ogilvie, P., & Parkhe, M. (2018). Accelerating the Machine Learning Lifecycle with MLflow. *IEEE Data Eng. Bull., 41*(4), 39–45.

Zahedi, M., Shahin, M., & Babar, M. A. (2016). A systematic review of knowledge sharing challenges and practices in global software development. *International Journal of Information Management, 36*(6), 995–1019.

Vinay Reddy Venumuddala completed his PhD from Public Policy area, at the Indian Institute of Management Bangalore (IIMB). His research interests lie at the intersection of Information systems, Public Policy, and Electrical and Electronics Engineering. He also presented his thesis work in collaboration with his doctoral advisor, Prof. Rajalaxmi Kamath, in Information Systems conferences such as ICIS and AMCIS.

Rajalaxmi Kamath is an Associate Professor at the Centre for Public Policy, Indian Institute of Management Bangalore (IIMB). Her current research interests primarily involve in-depth field and organisational studies around issues of microfinance and labour (formal and informal).