Information and Software Technology 170 (2024) 107430

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Objectivity by design: The impact of AI-driven approach on employees' soft skills evaluation

Ruti Gafni [a], Itzhak Aviv [a,*], Boris Kantsepolsky [a], Sofia Sherman [a], Havana Rika [a], Yariv Itzkovich [a,b], Artem Barger [a]

[a] *The Academic College of Tel-Aviv-Yaffo, Tel Aviv-Yafo, Israel*
[b] *Kinneret College on the Sea of Galilee, Israel*

ARTICLE INFO

Keywords:
Employees analytics
Soft skills assessment
Data Science
Artificial intelligence
Machine learning

ABSTRACT

Engineers' team collaboration skills are among software development's most important success factors. Existing Artificial Intelligence practices for the engineers' soft skills assessment mainly rely on evaluations of subjective data gathered through surveys, interviews, or observations. As a result, the insights gained by these methods are biased because of the subjective data people report. To overcome the challenge of subjectivity, we offer a novel objectivity-by-design approach for continuous AI-driven team collaboration skills analytics. The method analyzes the data from workstreams gathered from data repositories like Jira. Based on the study results, we conclude that this approach enables a continuous assessment of employees' team collaboration skills, provides more accurate insights, eliminates subjective biases, and helps uncover trends and deficits on individual and team levels. Understanding and recognizing employees' strengths and weaknesses can foster an organizational culture of growth and development. An improved organizational climate is expected to result in work satisfaction, engagement, and motivation, thus positively impacting employees, businesses, and society.

1. Introduction

Soft skills are recognized as essential factors for success in today's businesses. Whether in hiring, career advancement, corporate development, or any other aspect of professional life, the study of soft skills has received much attention [15,33]. While there is an ongoing discussion in the literature about the scope and definition of soft skills, the consensus is that they incorporate elements of emotional, behavioral, and cognitive components [24]. According to predictions by the World Economic Forum [77], there will be a significant rise in demand for workers with a particular set of skills by the year 2025. These skills include analytical and critical thinking, the ability to effectively collaborate and solve complex issues, generate new ideas and approaches, the capacity for active learning, a proactive attitude and initiative, and effective decision-making. Employers across many sectors are projected to significantly value these abilities, as they are required for success in today's fast-changing and technologically advanced labor market.

In software engineering, soft skills have been acknowledged as an essential component of an employee's skill set alongside technical aptitude and knowledge. In particular, collaboration and teamwork are frequently mentioned as the most crucial for software development success [31,66]. While effective collaboration inside and across development teams affects software quality, certain aspects of teamwork, such as communication skills, are challenging to track, and their evaluation takes time.

Soft skill evaluations of employees in the software industry are traditionally conducted throughout their careers by the employees' direct managers, with some assistance from Human Resources (HR) specialists, using a series of predetermined questions in surveys, observations, or group simulations [32,81]. Traditional methods for evaluating soft skills manually are time-consuming, ineffective, and cannot be applied to large organizations [71]. Furthermore, most HR and manager evaluations are subjective, complicating the difficult by itself process [62]. Recent research reveals significant limitations in assessing and developing the social skills of software engineers. Without accurate and dependable assessments of soft skills, it can be challenging to identify the precise areas in which software engineers need to improve their skills. Accordingly, developing training programs to close the skill gaps remains problematic.

Assessing the soft skills of software engineers has proven

* Corresponding author at: The Academic College of Tel Aviv-Yaffo, Rabenu Yeruham, 2 4265934, Tel Aviv, Israel.
E-mail address: itzhakav@mta.ac.il (I. Aviv).

https://doi.org/10.1016/j.infsof.2024.107430

Available online 7 March 2024

complicated for many organizations, primarily because soft skills are difficult to measure and quantify, and assessments are frequently subjective, based on the opinions and perspectives of managers, co-workers, or other stakeholders [71]. It can lead to inconsistent or unreliable evaluations and complicates developing individualized training programs for software engineers based on their assessment and skill sets. Current methods for assessing the soft skills of software engineers rely on subjective evaluations, which can introduce bias and reduce precision. Using psychometric tests or role-playing exercises is an alternative method for measuring soft skills objectively. Still, these methodologies rely on collecting and analyzing data extracted from use-case simulations rather than actual workstreams.

In response to the gap in existing knowledge, this paper presents the results of a novel AI approach that is data science driven for objectively evaluating employees' soft skills by assessing the data collected from several real-world projects' repositories. At the core of the study is the method to assess collaboration soft skills, such as Participation, Social Impact, Responsivity, Internal Cohesion, and Communication Density within a team. This approach eliminates the subjective biases of traditional evaluation methods, enables objective assessment, and provides recommendations at personal and team levels. While our method initially assesses team-level dynamics, the granularity of the data and the sophistication of our analytical techniques allow for the accurate derivation of individual engineers' soft skills. This dual focus on group and individual assessments is crucial for a comprehensive understanding of team functionality and individual professional development. Our study's findings underscore the immense potential of AI in not just assessing but also enhancing the soft skill landscape in engineering teams. Embracing this potential can lead to significant improvements in team dynamics, process efficiencies, and ultimately, the success of engineering projects in the industry.

This paper is organized as follows. First, we review the literature on the topics of soft skills, Artificial Intelligence (AI), and Machine Learning (ML) for soft skills analytics, along with the group communication analysis methods and software platforms for team communication. Then, the research methodology is presented in section three. The results are described in section four, and the discussion in section five. Finally, we provide conclusions, address limitations, and offer directions for future research.

2. Related work

2.1. Soft skills in the software development process

The importance of soft skills in software development is no longer arguable and has been broadly discussed in industry and academic literature. Since the introduction of Agile software development methodologies, which emphasize the human-centric process, soft skills have been considered one of the factors influencing the project's outcomes [11,26]. Soft skills contribute significantly to individual learning, team performance, client relations, and business context awareness [69]. The ability of software engineers to collaborate, communicate, manage time, negotiate, solve problems, and make decisions are sometimes more critical skills than knowledge of new technologies [4,28,55]. Some employers even treat soft skills as more critical for first-job candidates than technical skills [46]. Moreover, a software project's success is affected by communication, teamwork, analysis, organization, and human relations [11,66]. While the list of soft skills focusing on the human aspects of the software development process varies, the most referred soft skills in the academic literature are communication and teamwork skills [55,66].

Some studies raise doubts regarding the ability to train soft skills in the workplace [69] and discuss the need to develop those skills during undergraduate studies [1]. However, while soft skills are difficult to teach, assessing them is even more challenging [12,66]. HR observations, peers, and managers' subjective reports constitute the primary sources of soft skill assessments. Those assessments are not synchronized with project phases, could come late in the development process, and often could not impact the project results [66,76]. As a scalable alternative, an emerging stream of research discusses the potential of assessing soft skills by applying multimodal technology using AI techniques for automatic retrieval and analysis of behavioral signals [3,34]. While AI is widely employed in fields like bioinformatics, medicine, data mining, marketing, and consumer behavior research, it is still underutilized when developing and evaluating employees' soft skills during software development.

2.2. Artificial intelligence and machine learning for soft skills assessment

Since the turn of the 21st century, AI has established a solid scientific basis and developed several applications contributing to technical and societal advancement. AI allows automated decision-making even in sectors requiring complicated decisions based on various elements and non-predetermined criteria. Since AI systems are immune to the delusion typical of human psychology, automated predictions and decisions can be cheaper, more accurate, and more impartial than those made by humans [5,30,48].

Among the prospective uses of AI is a novel Human Resource Management (HRM) approach that may foresee or prescribe future events [14,58]. In the field of HRM, AI is enticing because it implies the ability to reliably comprehend and predict human behavior within an organization, which has great appeal for managing productivity. HR analytics is characterized as a 'must have' capability for the HR profession, serves as an instrument for creating value from people, and facilitates the expansion of the HR functions' strategic influence [10]. However, human resource analytics has yet to be widely operationalized. The absence of trustworthy outcomes and the approach's measurability are two of the most significant impediments [50,58]. Existing research indicates that despite the interest above and claims regarding the applications, benefits, and impact of AI in HRM, many companies still need to experience the anticipated benefits [10]. AI-powered HR analytics enables organizations to reimagine how they manage their workforce, mainly to ensure they have a proficient workforce, such as the skills, knowledge, and experience required for organizational success [54].

A recent study on the implementation of AI-powered solutions in HRM revealed that in the area of employee skills development, AI analyses executed on data collected from the employee and managers while also taking into account the employee's job function, past learning history, and business team [10]. However, the data gained from people may create a subjective bias for several reasons. While individuals provide information about themselves, they may inadvertently or deliberately provide inaccurate or insufficient data. This may result from memory errors, social desirability bias, or misunderstanding of the query [65]. Data collectors may interpret the data based on preconceived notions or hypotheses. They may subconsciously search for patterns that support their beliefs while ignoring evidence to the contrary [40]. Participants may respond differently to extrinsic variables like survey formats, question contents, and interviewer presence. For instance, individuals may be more likely to concur with a statement if presented positively rather than negatively [73].

People analytics may promise to provide an accurate and objective picture of employees' performance based on the premise that digital data can correctly reflect the reality of individuals' characteristics, experiences, and talents [30,75]. Nonetheless, algorithmic conclusions may sometimes be erroneous or biased, repeating and creating new human prejudices. Moreover, even if fair and accurate, computerized evaluations of persons may harm those individuals [14,30,56]. Consequently, AI produces significant risks, including unemployment, inequality, discrimination, social isolation, monitoring, and manipulation [9,30,48]. The uncertainties associated with the newness, complexity, and breadth of the technologies' personal and societal consequences underlined the urgent need for a regulatory framework to guarantee that AI systems are safe and comply with current laws on fundamental rights and societal values.

To address employees' privacy concerns, limit the misuse of employee personal information, and standardize how employee data is utilized, businesses apply legislation such as the General Data Protection Regulation (GDPR) [41,57]. Although AI is not explicitly referenced in the GDPR, its concepts nonetheless apply. AI-based systems leverage collecting and analyzing vast data about persons and their social relationships. When designing and implementing AI systems, conventional data security concepts such as purpose restriction, data minimization, handling sensitive data, and limiting automated choices should be considered [48]. However, the GDPR gives scant direction on how to accomplish this objective. In addition, the GDPR places the burden on determining how to manage risk and develop appropriate solutions for controllers, which may be difficult and expensive [41]. Consequently, the effective implementation of GDPR to AI applications highly depends on data protection agencies and other competent authorities' advice to controllers and data subjects [35,48].

Many AI systems depend on ML approaches for detecting data connections and developing predictive models, using algorithms for processing collective data such as hidden Markov models [53], social network analysis [79], and process mining [72]. Using ML, different types of personal data may be utilized to study, predict, and affect human behavior. Scholars and practitioners have recently begun utilizing advanced automated ML approaches for recruitment and applicant selection [10]. ML algorithms extract patterns from social media activity, resumes, and interview reactions. For example, the Support Vector Machine (SVM) type classification approach has been applied in pattern identification, face detection, and spam filtering [44,61]. ML algorithms can also help gain insights from group interactions that result in high-quality outcomes in conflict resolution and cooperation in collaborative environments [42,72,74].

2.3. The use of software management platforms for communication among teams

Software developers use collaborative environments incorporating comprehensive data repositories that host all communications among the team, including development assignments, new features, bug fixes, task progress updates, comments to discuss progress, and additional information. Asana, Jira, and Trello are popular product and project management software tools containing data repositories that allow teams to store and organize various information related to their projects. Jira facilitates tracking team progress, managing backlogs, and planning sprints [59]. It serves as a collaborative problem-solving discussion platform and can be used to conduct group communication analysis, detect patterns of contributor interaction, and improve soft skills. Jira's issue-tracking system allows team members to create, assign, and comment on tasks. It enables them to engage in ongoing discussions and provide relevant arguments and responses related to the subject matter [43]. Additionally, Jira's customizable dashboards and reporting tools enable real-time monitoring of project progress, ensuring that team members are aware of each other's work and can provide timely feedback [39].

2.4. Collaborative problem solving and group communication analysis

Learning Sciences and Psychology have long studied Collaborative Problem Solving (CPS) as a social-cognitive process in which people work together to achieve a common objective. CPS is emphasized as one of the critical skills of the 21st century that should be studied, taught, and evaluated according to AI-based skills evaluation [78]. As an example, in several recent studies, deep neural networks and other supervised ML algorithms were utilized to predict the various stages of student collaboration automatically [20,51]. Although most of the research was done with students, the results may be extended to employees at business organizations. Recent publications shed light on how people can learn via CPS by externalizing ideas, negotiating to reach an agreement, and engaging in a worthwhile task that serves society [38,60,80]. Given the collaborative partnerships' dynamic and complex nature, investigating how different CPS abilities may improve or worsen collaboration results is crucial.

Group communication analysis (GCA) is one of the core subskills of CPS [27]. GCA builds on cohesion and contingency studies to define individual CPS processes using automated computational linguistics and the study of sequential interactions. GCA reveals people's coherent behaviors in CPS circumstances by integrating various socio-cognitive factors, including engagement, internal coherence, responsivity, social impact, newness, and communication density [19]. Five metrics of GCA measures [18] for data classification, categorization, and clustering, as presented in Table 1.

Participation is a critical component of successful team collaboration in software engineering projects. It involves the extent to which team members actively engage in discussions, contribute ideas, and assume responsibilities [13,23]. High levels of participation and engagement enhance a team's communication, problem-solving, and productivity. Team members who actively participate in team meetings, discussions, and other collaborative activities are more likely to be involved in the team's success and contribute to accomplishing the team's objectives. In turn, the indication of limited participation may impede cooperation and inhibit progress toward team objectives. Research has shown that effective participation within a team can lead to improved communication, problem-solving, and overall project outcomes [47]. In the context of Jira, the tool's features, such as issue tracking, commenting, and notifications, can help foster and maintain high levels of participation among team members [39]. Software engineers' willingness to actively contribute to group communication is crucial for effective teamwork. Studies have highlighted the importance of open and transparent communication in software engineering projects, especially using agile development methodologies [47]. Engineers actively contributing to discussions and providing ideas can facilitate decision-making, improve overall team dynamics, and improve project outcomes [23]. Engineers can use various communication channels of the Jira platform, such as commenting on issues, sharing project updates, and participating in sprint reviews, to actively contribute to group communication [43].

The employee's social impact refers to the impact of the employee's activities on society or the team [23]. In a team environment, each

Table 1
Group Communication Analysis Measures description [adopted from Dowell et al. [18]].

GCA Measure	Psychological and Discursive Process	Description
Participation	Engagement	An employee's willingness to actively contribute to group communication by providing ideas on an ongoing topic.
Social Impact	Productive communication	An employee's willingness to spark debate by providing ideas that engage other participants to continue the discussion and evolve the proposed viewpoint.
Responsivity	Uptake and transactivity	An employee's willingness to respond coherently to the conversation subject.
Internal ohesion	Monitoring and reflecting	An employee's willingness to keep a continuous discourse while presenting their views.
Communication Density	Consistent communication	An employee's willingness to continue a debate by delivering arguments and replies that are consistent with the subject of the important discussion.

employee's activities may add to the work's societal influence. For instance, employees committed to sustainability and environmental preservation may endeavor to lower the team's carbon footprint or advocate for eco-friendly methods. The high levels of commitment to social impact within the group may indicate that teamwork is productive and successful. A team's commitment to having a sound social effect may provide new views, ideas, and motivation to a team's work, which can assist the team in achieving its objectives and positively influencing society or the community. In software engineering, team members with high levels of social impact contribute positively to the team's dynamics and the project's overall success. These team members often bring fresh perspectives and innovative ideas and foster a sense of shared purpose [70]. Jira platform features such as issue tracking, commenting, and notifications can promote transparency and encourage team members to contribute to discussions, thus enabling social impact within the team [39]. Team members can use features like sharing updates, participating in sprint reviews, providing new viewpoints, and stimulating debate [43]. Besides, Jira's communication channels, such as comments on issues, real-time messaging, and video conferencing integrations, can spark debate and lead to meaningful discussions [39]. The willingness to engage in discussions is essential to effective teamwork, as it promotes critical thinking, innovation, and shared understanding among team members [23]. Jira facilitates these aspects by providing a range of features that promote transparency, communication, and active engagement in discussions. Software engineers who actively participate in discussions and offer ideas that challenge conventional thinking contribute to improved problem-solving and decision-making processes [47].

Responsivity refers to the ability of team members to react promptly, efficiently, and coherently to requests, comments, and changes, contributing to the team's overall effectiveness [52]. High responsiveness indicates an employee's ability to solve problems, provide the necessary information, and make timely decisions. Responsiveness is crucial in a team context because it enables the team to move ahead and achieve progress on tasks and projects. Team members can rapidly make choices and handle issues, thus enhancing efficiency and effectiveness. In software engineering, responsiveness is vital for addressing issues, providing information, and making timely decisions that help advance tasks and projects. Jira's features can promote responsivity among software engineering team members. For instance, Jira's issue tracking system allows for comments, updates, and the assignment of tasks, ensuring that team members are aware of the ongoing discussion and their responsibilities [43]. Jira's customizable notifications and integrations with communication tools like Slack or Microsoft Teams can help keep team members informed and engaged, promoting timely responses to ongoing conversations [39]. By promoting effective communication, Jira can help team members respond coherently to the conversation subject and contribute to the overall success of their projects.

Employee internal cohesion within team cooperation refers to the extent to which team members work well together, get along, and have a feeling of oneness inside the team [23]. The team's connection or link allows them to cooperate toward a shared objective. A team with high internal cohesiveness may boost communication, problem-solving, and production. Team members with good ties and mutual trust are more inclined to communicate openly and honestly, which may aid in dispute resolution and enhance team performance. For software engineering teams, internal cohesion is a vital aspect of collaboration. It refers to the extent to which team members work well together, get along with one another, and maintain a sense of unity within the team [70]. Strong internal cohesion is characterized by team members willing to engage in continuous discourse, openly and honestly present their views, and maintain a supportive and respectful atmosphere [52]. Jira's features help promote internal cohesion by facilitating communication and collaboration within software engineering teams. For instance, Jira's issue-tracking system allows team members to create, assign, and comment on tasks, enabling them to engage in continuous discourse and share their perspectives [43]. Jira's customizable dashboards and reporting tools provide real-time insights into the project's progress, promoting transparency and a shared understanding of the team's objectives [39]. Furthermore, Jira's integrations with communication tools, such as Slack or Microsoft Teams, can help foster internal cohesion by streamlining communication channels and enabling team members to collaborate more effectively [39]. These tools enable more informal and spontaneous conversations, which can contribute to building trust and rapport among team members [70].

Employees' communication density refers to the frequency and efficacy of communication between team members. It includes the quality, quantity, and communication routes [23]. Team members who communicate regularly and effectively can remain abreast of the group's progress and resolve difficulties expeditiously. This ensures that team members work toward a unified objective and that tasks are accomplished on time. In software engineering, communication density refers to the frequency and efficacy of communication between team members encompassing the quality, quantity, and communication channels [52]. Effective communication is critical for the success of software engineering teams because it ensures that team members work toward a unified objective and that tasks are accomplished on time. High communication density within a team can lead to enhanced cooperation, productivity, and problem-solving capabilities, while poor communication can hinder teamwork and impede progress toward team objectives [70]. Jira's features promote communication density within software engineering teams by providing a centralized platform for managing tasks, tracking issues, and facilitating collaboration [39].GCA uses cross-correlation measures at various lag values to ascertain how closely social participation and cognitive levels defined by latent semantic analysis interactions of individual coworkers are connected. To analyze the skills of software developers, previous studies have used the GCA framework [19] and a computational linguistics approach to group communication analysis [8]. This approach identifies socio-cognitive positions within multiparty interactions and extracts socio-cognitive functions from group communication using linguistic characteristics. The researchers used a combination of lexical, syntactic, and cohesion-based characteristics to create a model that classifies the responsibilities of individuals in group discussions. The analysis was applied to various datasets, including online conversations and face-to-face interactions, to assess the generalizability of the method. The results of the study show that the computational linguistics approach effectively identifies socio-cognitive roles in group interactions. The model accurately detects roles across multiple datasets, which contributes to understanding group dynamics and has potential applications in fields such as education, teamwork, and social science research. By providing a dependable and scalable method for analyzing group communication, the approach can aid researchers and practitioners in gaining a deeper understanding of and facilitating effective group interactions.

3. Methodology

In this study, the GCA assessment method is implemented using AI-based data analysis to overcome the subjective biases of traditional evaluation methods and enable objective assessment and recommendations at personal and team levels. For the purpose of the study, we adopt five metrics of GCA measures [18] for data classification, categorization, and clustering, as presented in Table 1. The study is based on the analyses of three different software teams working on the Lynx Foundation software project using the Jira management platform. The method was empirically evaluated using LSA application algorithms for GCA soft skill categorization from three open-source projects. JIRA was chosen for this study because it provides a centralized repository for managing and reporting project-related stuff [59]. Utilizing the Jira platform as the primary data source enabled the collection of a vast amount of information whose analysis sheds light on the soft skills of individual team members and the posture of the entire software team. The initial phase of the research was dedicated to collecting data from Jira. The second phase involved developing and implementing a protocol for data processing and semantic analysis that was used for data analytics. The outcome of this step is the GCA measures of participation, responsivity, internal cohesion, and communication density. In the last step, we clustered contributors into 5 clusters according to the GCA measures. In the following subsections, we provide additional details on each one of the study stages.

3.1. Data collection

The data set was obtained from the JIRA platform of three prominent and popular open-source projects: Moodle, Apache Zookeeper, and Hyperledge. The data acquired from JIRA included ticket descriptions, comments, assignees, status, and more, which were used for analysis and reporting purposes. In addition to the ticket information in JIRA, comments were extracted to capture dialogues about individual issues and were used for calculating GCA measures. Moreover, user stories, sprints, and distributed tasks were tracked in JIRA to account for all cooperation performed during a project's lifecycle and delivery criteria.

3.2. Data processing and semantic analysis

A content analytics protocol was followed in the data science protocol used for Jira data analytics [8]. The foremost step involved data preparation, where Jira ticket information was processed to be ready for analysis. Each ticket is a conversation topic, with descriptions and comments converted into vectors representing word occurrences. Comments from single-ticket discussions were isolated, and tickets with minimal dialog were filtered out, leaving only those with at least ten substantial comments. We exported relevant data from Jira, such as ticket information, comments, labels, and timestamps, and pre-processed the data by cleaning, filtering, and transforming it into a suitable format for analysis. Next, we pre-processed the text data by removing any stop words, punctuation, or irrelevant characters. We normalized the text by converting it to lowercase and performed stemming or lemmatization to reduce the words to their base form. This process resulted in a word-by-document matrix that shows term co-occurrence patterns within Jira tickets, which is crucial for subsequent stages. Each ticket is considered a discussion content and we consider the ticket's description and its following comments as documents. Whenever the user opens a ticket or comments in response, we consider such a contribution as a single document for the sake of LSA computation. Therefore, each row in the matrix represents a single document.

In the next phase, Term-Frequency Inverse-Document-Frequency (TF-IDF) [36] was used to measure the significance of each term in the document corpus. This method balances the frequency of a term in a document against its inverse frequency across all documents. TF-IDF assigns weights to the words in the dataset and identifies the most important terms and phrases by considering both the frequency of the term in a particular document and its inverse frequency across the entire dataset. Next, the model applies Log Entropy Transformation (LET) [22] to further refine the term weighting by considering the overall distribution of terms in the dataset. The TF-IDF scores are based on the distribution of term occurrences across documents in Jira. This approach reduces the impact of common terms and emphasizes more meaningful ones. As the next step, Latent Semantic Analysis (LSA) is performed to uncover hidden semantic connections between words and documents in the dataset [22]. LSA enhances search and retrieval performance and identifies underlying patterns and concepts. LSA employs singular value decomposition (SVD) to reduce the dimensionality of the term-document matrix, making it easier to uncover relationships between terms and documents. Following LSA, Topic Modeling (TM) is used as a probabilistic modeling approach to identify latent topics within the document corpus and their associated word probabilities within the Jira tickets. TM groups Jira data into distinct topics to identify common themes and patterns within the dataset. Next, Principal Component Analysis (PCA) is employed to reduce the feature space's dimensionality and pinpoint the principal components explaining data variance [2]. PCA reduces the dimensionality of data and visualizes the relationships between various topics and trends in Jira. We kept components that explain 80 % variance in the data. Finally, the K-means technique [68] is utilized to cluster GCA metrics. K-means groups n employees into $k = 5$ clusters based on their similarity in the GCA measures which we calculated as follows.

To determine the GCA Participation measure, the model analyzes the frequency and content of each team member's comments, mainly focusing on their contributions to new ideas or solutions to ongoing topics. Additionally, the model calculates metrics reflecting their levels of engagement, such as the number of tickets they actively participate in, the number of original ideas they propose, and their involvement in decision-making processes. The model assesses team members' willingness to contribute, allowing the identification of patterns, proactive individuals, and potential areas for improvement in team dynamics and communication. To calculate the GCA Social Impact measure, the model analyzes the content and context of each team member's comments, mainly focusing on their contributions that generate further discussions, challenge existing viewpoints, or inspire new perspectives, considering the number of replies or follow-up comments generated by their initial input to indicate their ability to engage others in the conversation. To compute the GCA Responsivity measure, the model analyzes the relevance and coherence of each team member's comments concerning the ongoing discussion topic. This can be done by measuring the semantic similarity between a team member's comments and the main subject or by evaluating the degree to which their input advances the conversation meaningfully. To calculate the GCA Internal Cohesion measure, the model analyzes the consistency and flow of each team member's comments within the context of the ongoing discussion. This is performed by examining the logical progression of ideas in their input, the frequency of their contributions, and the extent to which they address previous comments or build upon existing ideas. To evaluate the GCA Communication Density measure, the model analyzes the quality and relevance of each team member's comments with the ongoing discussion topic. This is done by evaluating the logical coherence of their arguments, the extent to which they address counterpoints or opposing views, and the degree to which their input advances the conversation meaningfully.

4. Results

In accordance with the GCA framework, the data science protocol was applied to a Jira dataset of software teams to cluster, identified and provide the following socially engaged roles for participants: Followers, Influential Actors, Lurkers, and Socially Detached Drivers. Each employee was categorized based on his or her performance across the GCA measures. These clusters were derived and identified by the GCA measures as follows. Drivers rated among the highest in all GCA measures except internal cohesion. Followers are characterized by low participation but high internal cohesion, responsiveness, and social impact. Influential Actors provide slightly more than the average, they evaluated reasonably positively in terms of societal impact, responsiveness, and newness. Lurkers have some of the lowest values Among all GCA measures. And finally, Socially Detached have the highest involvement rate, as well as high communication density and newness scores. They do, however, score poorly on responsivity, implying that their discourse is more in response to themselves than to others because they have a more robust internal cohesion than overall responsivity.

The experiment was conducted on three distinct software projects, and the results revealed that all three projects displayed comparable clustering tendencies. This consistency across projects suggests that the identified social behavior categories are not project-specific but reflect general patterns in employee collaboration and group dynamics. The clustering results for Hyperledger Fabric, Moodle, and Apache Zookeeper, are presented in Fig. 1 (ordered from left to right).

Due to the nature of software development work, which requires communication, problem-solving, and coordinated efforts among team members, we determine that there are likely universal collaboration patterns and group dynamics that apply to all software teams. Based on these patterns, the K-means clustering algorithm may identify similar tendencies. Moreover, software development teams adhere to comparable organizational structures and Agile methodologies, such as Scrum. These structures can result in similar team dynamics and collaboration styles, which also may be reflected in the K-means clustering outcomes. In addition, the software development discipline has a shared professional culture, which may influence how team members collaborate and interact. This shared culture may contribute to comparable collaboration and behavior patterns among software development teams.

Moreover, software development teams typically include members with specialized skill sets and duties, such as programmers, testers, and project managers. As team members must collaborate to accomplish shared objectives, combining these roles within a team can result in similar collaboration patterns and group dynamics. Group behavior tends to manifest certain similarities regardless of the specific domain or industry. When applying the K-means clustering algorithm to various software teams, the behaviors, such as communication styles, problem-solving approaches, and interpersonal dynamics, may result in similar clustering tendencies.

The method used for validating the cluster results by discussing them with the managers of the clustered people can be described as expert validation or qualitative validation [37]. Expert validation involves seeking feedback from knowledgeable individuals - in this case, the managers, who deeply understand the people and their roles in group collaboration. In our study, team managers were conferred to validate the accuracy of the AI clustering by evaluating the model's output. This approach helped to assess the accuracy and relevance of the clustering results by comparing them with the insights and observations of these experts [17]. The managers' evaluation of the clustering results revealed high accuracy, indicating that the model captured the underlying patterns of intrapersonal socio-cognitive parameters within team interactions.

5. Discussion

The AI model presented in this paper has provided a way to objectively evaluate employee soft skills, group dynamics, and problem-solving skills derived from the collected data, also known as the Data Science approach. In the following paragraphs, we discuss how the soft skills evaluation may affect the organization's individuals, teams, and organizations and recommend how to improve them over time.

5.1. Individual level impact

Our methodology, while initially focused on team dynamics, inherently captures individual contributions and interactions that collectively form the group's communication pattern. The AI-driven approach analyzes individual contributions within the group context by examining metrics like participation frequency, nature of communication, and responsiveness. By quantifying these aspects, we can discern each engineer's role and interaction style within the group, thereby shedding light on individual soft skills such as teamwork, communication, and adaptability.

The method considers the context in which individual interactions occur. For instance, an engineer's ability to influence group decisions, or their responsiveness to queries, is directly indicative of leadership and collaborative skills. By analyzing these interactions within the group's dynamics, we gain insights into individual competencies.

Through K-means clustering, individuals are categorized based on their communication patterns and contributions. This clustering not only helps in understanding group dynamics but also highlights individual skill sets, as each cluster can be indicative of certain soft skill traits.

We correlate individual behaviors and contributions with overall group performance. By understanding how individual actions impact group outcomes, we can infer specific soft skills. For example, a team's success in meeting project milestones can be partially attributed to individuals' problem-solving and time-management skills. Our approach allows for continuous monitoring of group and individual performance. This ongoing assessment enables us to track the development of individual skills over time, providing a dynamic rather than static view of each engineer's abilities. The alignment of our AI-driven assessments with expert evaluations further substantiates the accuracy of individual skill assessments. Experts can validate the inferred individual skills by comparing them with their observations and interactions.

Using collaborative problem-solving skills assessment may have twofold implications for individuals. On the one hand, it can increase employees' awareness of their teamwork performance and contribution to the shared aim. On the other hand, such monitoring may make employees defensive, resulting in an unnatural change in their behavior and undue stress [45]. Prior research [29] found that voluntary knowledge sharing, such as participation and contribution to user's

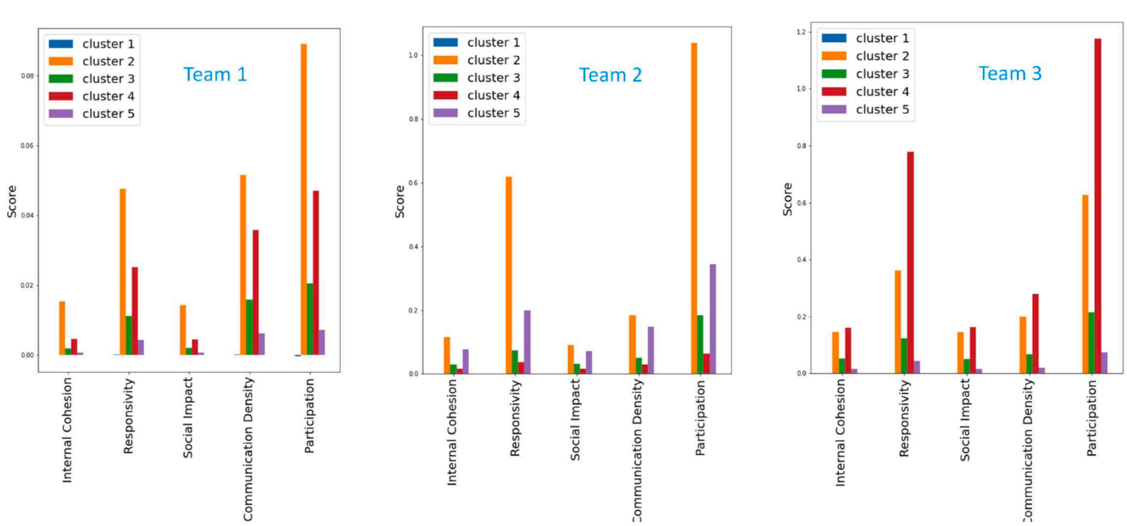

Fig. 1. GCA Assessment Results of Three Teams.

generated content, is triggered by extrinsic motivation like reciprocity, prestige, or awards (either tangible or virtual), especially if they are aware of the possible acknowledge and reward [25]. Thus, employees may deceive the system by contributing irrelevant or ineffective communication to get their extrinsic reward, causing the tool to score their performance higher.

This research demonstrated that methods of statistical learning implemented on operational data could clean the noise and minimize subjective bias. The results reveal that continuous ML-based clustering enables real-time monitoring of employees' soft skills. Understanding of the current solution's incompleteness must be communicated to everyone using this assessment tool. One of the benefits of the Data Science-based solution is its objectivity, as it is based solely on data analysis, with no prejudices or emotional or other ties. However, it also has a flaw in that the tool doesn't know what caused the employee to behave in a specific way. The metadata, which can shed some light on employee performance, is not included in the data analysis. Furthermore, after the tool input is received, there is currently no possibility to provide feedback to calibrate outcomes based on metadata that is not part of the tool's algorithm decision-making.

In the discussion of the effectiveness of our AI-driven approach for assessing soft skills in software engineering teams, a critical aspect is its comparison with other prevalent AI methodologies in terms of reliability and trust. Our approach stands distinct in its utilization of data derived from real-world project management tools like Jira. This choice of data source is pivotal, as it relates directly to work performance and team interactions, thus providing a robust and relevant dataset for assessing soft skills. In contrast, other AI methods may rely on data from surveys, psychometric tests, or observational data, which, while useful, can be subjective and influenced by personal biases, potentially affecting their reliability. In terms of analytical techniques, our method employs advanced NLP techniques like LSA and TF-IDF, coupled with K-means clustering. These sophisticated techniques enable a deep and nuanced analysis of communication patterns, enhancing our understanding of soft skills. This is in comparison to other AI methods that often resort to simpler NLP or basic machine learning models, which, although effective, might not capture the complexity of soft skills as comprehensively.

A crucial factor in our approach is the focus on objective metrics derived from actual work-related activities and communications. This emphasis significantly reduces subjective bias and increases trust in the results. Conversely, AI methods based on self-reported data or observational analysis might exhibit a higher degree of subjectivity, potentially diminishing the trustworthiness of the assessment. Another distinguishing feature of our approach is its capability for continuous and dynamic assessment. This allows for ongoing monitoring of soft skills over time, providing a dynamic view of an individual's or team's development. Many other AI methods, however, might offer a more static perspective based on one-time assessments or periodic evaluations, which may not accurately reflect changes over time. The approach also aligns well with expert evaluations, as detailed in the manuscript, enhancing trust in our AI-driven method. This alignment is not always evident in other AI methodologies, where the degree of trust can vary, especially in methods that heavily rely on algorithmic predictions without expert validation. Furthermore, by using well-defined metrics and clear analytical techniques, our method likely offers greater transparency, a crucial factor for trust. In comparison, methods using more opaque or complex models, like deep learning, can struggle with interpretability, which might affect trust, particularly among non-technical stakeholders. Our method is designed to analyze work-related data while respecting privacy, a key factor in building trust. This is in contrast to some AI approaches that use more invasive data, like emotion recognition from video feeds, which might raise ethical concerns and affect their trustworthiness.

Applying ML to measure workers' soft skills may favor their skill development. By giving an objective and reliable evaluation of workers' soft skills, an assessment based on ML may assist in identifying particular areas in which employees need to develop. This insight may be utilized to create individualized training regimens targeting these regions, resulting in more efficient and successful training. In addition, using ML to measure soft skills may give near real-time feedback to workers, allowing them to better comprehend their performance. Employees may see the consequences of their work and understand how they contribute to the organization's objectives as an acknowledgement or reward, which can enhance motivation and engagement [29].

Evaluating soft skills based on ML also enables firms to work on improving individuals over time. This may aid in identifying skill development trends and provide training resources where needed. Based on understanding the individual's strengths and shortcomings, organizations may develop individualized training programs that target particular areas for growth. However, it is essential to stress that the influence of ML-based soft skills evaluation on the development of employees' abilities relies on the data quality, the accuracy of the models, and the execution of the assessment. In addition, one must ensure that the evaluation process matches the organization's objectives and personnel demands. Finally, it is crucial to remember that soft skills are complex and multifaceted. To receive a complete presentation of an individual's soft skills, it is necessary to use numerous data sources and evaluation methodologies.

5.2. Team level impact

Using a Data Science-based solution may lead to prioritizing communication tools (e.g., JIRA) as the recommended collaboration mode on a team level. For example, conversations may be held primarily through JIRA comments rather than in frontal or remote meetings [66]. Such a shift in communication style has both advantages and disadvantages. On the one hand, the team receives a well-documented, traceable history of conversations and decision-making. This makes maintenance and retrospective activities more accessible and more efficient. However, the social components of communication are being impacted, resulting in a less personal touch, which might influence teamwork.

On the other hand, verbal or face-to-face communication is more efficient as a process, with the ability to clarify issues quickly, challenge assumptions, and reach decisions quickly. Face-to-face communication enables updating the tool with the final decision without flooding it with intermediate and unnecessary clarifications. Another consideration is that the Data Science-based solution provides employees with an indirect and anonymous method to affect peer evaluation. Once employees understand how evaluation is done and what influences their score, they can project the same onto others and potentially use the tool in unintended ways. Hence, it is recommended that the method be considered as a complementary assessment.

In addition, using AI-based solutions to analyze soft skills may positively affect team equilibrium. By providing an objective and reliable evaluation of team members' soft skills, an assessment based on AI may assist in discovering particular areas in which team members thrive and those in which they need to be improved. This knowledge may be utilized to establish balanced teams whose strengths and weaknesses complement one another, resulting in more efficient and successful team interactions. Team members may see the outcomes of their work and know how they contribute to the team's objectives, which can enhance their motivation and engagement. Finally, recognizing and correcting any difficulties related to communication, collaboration, and additional soft skills may contribute to a more pleasant and productive work environment by enhancing the overall team dynamics. The influence of AI-based soft skills evaluation on team balance relies on the quality of the data, the precision of the models, and the execution of the assessment. One should verify that the evaluation process matches the organizational objectives and that team requirements are also essential. Additionally, it is necessary to recognize that the team's dynamic and culture are significant in its performance and should be considered carefully.

5.3. Organizational level impact

Recognizing an employee's strengths and limitations, anticipating, and preventing employee defection, and detecting changes in employee behavior or team interaction can signal problems and help the organization with in-time intervention. These are just a few notable benefits of using this technology for the organization. However, such oversight might be considered overly intrusive, stressing the employees, and making them defensive. Changing the culture of the organization is necessary to overcome this obstacle. Employees are expected to appreciate the importance of soft skills through personal and professional development.

Continuous evaluation of soft skills may assist in enhancing a company's environment by fostering a culture of growth and development. Additionally, it may help employers monitor their staff's development over time. This may aid in identifying skill development trends and areas where more resources or assistance may be required. Evaluation of soft skills may also help firms make better people management and development choices. By offering consistent feedback and chances for growth, employees may feel that their talents and contributions are appreciated and that progression prospects exist within the organization. This may result in improved work satisfaction, engagement, and motivation.

Ongoing evaluation of soft skills may help firms make better choices about people management and development. With a thorough awareness of their workers' strengths and shortcomings, employers may make more informed judgments about promotions, transfers, and other career advancement options. Continuous evaluation of soft skills enables firms to track the growth of employees' soft skills over time. By frequently measuring workers' soft skills, such as communication, cooperation, and problem-solving, employers may monitor employee growth and uncover trends of skill acquisition. Besides, continuous evaluation of soft skills may also allow firms to discover skill gaps across the workforce, which can aid in determining training and development priorities.

Finally, utilizing soft skills assessments may have a favorable effect on employee retention. Continuous evaluation of soft skills might help prevent voluntary resignations by offering frequent feedback and development chances. This may make workers feel appreciated and invested in their work, leading to greater job satisfaction and a more profound commitment to the firm. In addition, detecting and fixing skill weaknesses early on helps avoid them from becoming severe difficulties that may drive an employee to quit the organization. This may increase employee happiness and engagement, which can result in a rise in staff retention rates. Nevertheless, it is crucial to emphasize that the effect of continuous evaluation of soft skills will rely on the quality of the assessment technique, the execution of the assessment, and the organization's subsequent actions. Furthermore, it is essential to recognize that other elements, such as remuneration, perks, and general work satisfaction, influence employee retention.

In considering the broader context of our study's findings, particularly in the realm of engineering team collaboration, it is imperative to acknowledge the role of organizational culture, a factor prominently identified in the literature. Organizational culture, with its unique set of values, norms, and practices, has a profound impact on team collaboration and interactions. This influence extends to the realm of our AI-driven approach to assessing soft skills within engineering teams. The integration of organizational culture into the assessment of soft skills is not just a theoretical expansion but a practical necessity. Organizational culture shapes the ways in which team members communicate, collaborate, and resolve conflicts. These cultural elements, often subtle and ingrained, can significantly sway the manifestation of soft skills within a team. For instance, a culture that fosters open communication and collaborative problem-solving will naturally encourage the display of these soft skills among its members. Conversely, a culture that prioritizes individual achievement and competitiveness might yield a different set of soft skill dynamics. The implications of this for our approach are multifaceted. Firstly, the ability of our ML model to predict the degree of soft skills must be contextualized within the framework of the prevailing organizational culture. The model, trained on data reflective of certain cultural norms, may exhibit biases or limitations when applied to a different cultural setting. This suggests a potential variability in the predictive accuracy of the ML model depending on the organizational culture within which it is deployed. Moreover, this cultural aspect underscores the need for our ML approach to be adaptive and sensitive to the nuances of different organizational cultures. Incorporating cultural variables or adapting the model to recognize and adjust to different cultural contexts could enhance its predictive accuracy and applicability. This adaptation could be achieved through the analysis of additional data sources that capture cultural indicators or by implementing methodologies that account for cultural differences in team dynamics.

6. Conclusions

Although Data Science applications in the HR sector are still in their infancy, they necessitate cooperation between HR professionals and Data Science specialists for intricate organizational procedures involving hundreds or thousands of individuals. Based on the study's results, the conclusion is that Data Science is an acceptable and relevant approach for automatically evaluating employees' soft skills. Moreover, we argue that the demonstrated approach allows us to claim to introduce the "objectivity by design" term concerning the chosen design of the assessment methodology based on the objective data sources.

However, it is vital to note that while using Data Science-based soft skills assessment as a decision-making tool for HR and leadership teams, HR professionals must carefully evaluate the algorithm findings. Algorithms cannot make autonomous judgments, such as selecting whether to promote an employee or establishing an employee's development program based on AI-based soft skills evaluations. The conceptual and practical insights offered in this study are anticipated to contribute to advancing Data Science-based HR and leadership management.

Using AI to assess soft skills may have beneficial and harmful effects on organizations. An AI-based soft skills assessment may give a more objective and accurate evaluation of employee abilities, resulting in improved knowledge of each employee's strengths and weaknesses. This may aid in identifying personnel improvement areas, resulting in more efficient and successful training and development programs. In addition, providing employees with objective feedback may help them comprehend their performance and growth, thus boosting their motivation and engagement. Conversely, AI-based assessments of soft skills might have a potentially detrimental effect if not applied appropriately. It may lead to employee alienation and distrust, resulting in poor morale and higher turnover. In addition, AI-based assessments of soft skills may not be able to completely capture the complexity of human behavior, resulting in erroneous or incomplete employee evaluations, which may lead to poor personnel management and development choices. Considering potentially positive and negative impacts, one should understand that the influence of AI-based soft skills evaluation will rely on the data quality, the models' precision, and the execution of the assessment. In addition to ensuring that the evaluation technique matches the organization's objectives and employees' requirements, it is essential to evaluate any negative consequences and make efforts to prevent them.

In aligning with the conclusions of our study, it is pertinent to extend our discussion to include an additional significant outcome: the potential of our AI-driven approach to identify specific soft skill shortcomings in highly technically skilled engineers and the opportunity to work with them for improvement. In the industry, it is a recognized phenomenon that team cohesion and process efficiency often evolve organically over time and experience. However, this natural progression can be significantly augmented by intentionally focusing on strengthening the soft skills of individual team members, especially those who excel technically but may lag in areas such as communication, teamwork, or leadership.

The advantage of our AI-driven method lies in its ability to provide precise and objective insights into these softer aspects of team dynamics. By identifying specific areas where an engineer might need development, our approach opens the door to targeted interventions and tailored skill enhancement programs. This is particularly valuable in the industry, where the balance between technical prowess and soft skills is crucial for overall team performance and project success. For instance, an engineer who is exceptional in technical problem-solving but struggles with effective communication could benefit from tailored training programs focused on communication strategies. Similarly, an engineer who excels in individual tasks but shows limitations in teamwork could be supported through collaborative workshops or mentorship programs aimed at enhancing team participation and interpersonal skills.

The potential of this approach extends beyond individual development; it has implications for team dynamics and organizational productivity as well. By proactively working to enhance the soft skills of essential engineers, organizations can foster a more cohesive, adaptable, and effective workforce. This proactive stance contrasts with the traditional reactive approach and represents a paradigm shift in how soft skills are valued and developed in technical fields. Incorporating AI-based assessments to identify and address soft skill gaps ensures a data-driven, objective approach to personal and professional development. This strategy not only benefits individual engineers by rounding out their skill sets but also contributes to the creation of well-balanced teams capable of handling complex, multifaceted challenges in today's dynamic work environments.

Limitations and future research

This study's limitations follow the AI explainability, ethics, and privacy concerns. In April 2021, after years of preliminary work, the European Commission [21] released the AI regulatory framework, commonly known as the Harmonized Rules on AI or the AI Act as follows.

a. Ensuring that AI systems are safe and respect existing laws on fundamental rights and Union values.
b. Ensuring legal certainty to facilitate investment and innovation in AI.
c. Enhancing governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems.
d. Facilitating the development of a single market for lawful, safe, and trustworthy AI applications and preventing market failures.

The AI Act creates a legislative framework and regulatory strategy to manage risks and issues associated with AI. Although the legislative framework does not restrict technical advancements, it specifies the principle-based standards that AI systems must meet. In addition, the framework incorporates flexible methods to facilitate future technological progress and the emergence of new applications. Among the manifest accomplishments of the AI Act are the laws governing the creation, distribution, and use of AI systems. The AI Act defines and distinguishes illegal and high-risk AI activities. According to the AI Act, an AI system is defined as software produced using one or more of the following methods and approaches capable of producing outputs such as content, forecasts, suggestions, or judgments that affect the environments in which they interact.

The AI Act lists a variety of AI systems in employment, workers management, and access to self-employment as high-risk AI systems. Specifically, the AI Act lists AI systems for recruitment and AI systems "intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation, and monitoring and evaluating the performance and behavior of persons in such relationships" ([21], Annex III) and specifies the requirements for these systems. The criteria for high-risk AI systems include risk management systems, data governance, technical documentation, record keeping, transparency and information given to users, human supervision, precision, robustness, and cybersecurity.

Our AI system meets the AI Act criteria for high-risk AI systems. Thus, it is crucial that future research address AI explainability concerns. Given the inherent difficulty and challenges associated with explaining the outputs of an AI system utilizing classical ML algorithms, we propose an alternative approach by employing advanced tools derived from Quantum Cognition. Quantum Cognition is a framework rooted in Quantum Computing and Quantum Probability (QP), specifically designed for the evaluation of soft skills. QP offers numerous advantages over the limited outcomes provided by classical ML algorithms that rely on classical probability [64]. Traditionally, cognitive studies have relied on classical probability theory and principles derived from classical mechanics. However, recent investigations have uncovered that certain experimental data on human cognition, such as violations of the sure-thing principle, conjunction fallacies, disjunction fallacies, and order effects, cannot be adequately explored using classical theory [63,67]. Furthermore, QP operates within a unique vector space known as Hilbert space, where all computations and results can be reverse-engineered as a linear combination of the orthogonal basis vectors spanning the specific Hilbert space. This capability might enable one to provide explanations for the calculation process and the outcomes generated by our prediction model, which will overcome the explainability challenge. Additionally, in order to enhance the reliability and accuracy of our measurement model beyond measures taken in the current research, it would be advantageous for future studies to conduct a comprehensive comparison between our measurement model and traditional survey measurement methods. Such an approach can help to identify the strengths and limitations of our proposed model and provide additional insights for further improvements. Future research should also examine the impact of organizational culture on teams. Previous studies have shown that organizational culture influences teams in various ways (for example, [6,7,16]). When studying teams, it's important to consider these cultural differences. However, it's equally important to recognize that AI algorithms may have biases towards certain cultures. Therefore, any AI implementation must follow global standards, as discussed earlier in this section and emphasized by Leavy et al. [49]. In future research, exploring ways to reduce cultural biases in AI implementation and comparing different approaches would be an interesting avenue for research.

Focusing on organizational culture and its influence on soft skills also opens additional avenues for future research. Exploring how different organizational cultures impact the expression and assessment of soft skills can provide deeper insights and lead to more refined ML models. This exploration could ultimately lead to the development of AI tools that are not only technically proficient but also culturally intelligent, capable of operating effectively across diverse organizational landscapes. The influence of organizational culture is a critical factor in the effective assessment of soft skills in engineering teams. Recognizing and addressing this influence is vital for the development of ML approaches that are robust, accurate, and culturally aware, thereby enhancing their utility and effectiveness in real-world applications.

While our approach has some limitations, we believe that it holds tremendous potential for advancing the field of team studies. By enabling more accurate measurements, this approach can help researchers gain a deeper understanding of how teams function and what factors contribute to their success or failure. In our view, this new approach can have a significant impact on the field of team studies, and we are excited to see how it can be further developed and refined in the future. We believe that this method can help researchers better understand the complexities of team dynamics and ultimately lead to more effective team-building strategies and better outcomes for teams in various contexts.

CRediT authorship contribution statement

Ruti Gafni: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Itzhak Aviv:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Boris Kantsepolsky:** Conceptualization, Formal analysis, Investigation, Validation, Writing – original draft, Writing – review & editing. **Sofia Sherman:** Methodology, Project administration, Visualization. **Havana Rika:** Software, Validation. **Yariv Itzkovich:** Validation, Writing – review & editing. **Artem Barger:** Formal analysis, Methodology, Software.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

References

[1] C.L. Aasheim, S. Williams, E.S. Butler, Knowledge and skill requirements for IT graduates, J. Comp. Info. Sys. 49 (3) (2009) 48–53.
[2] H. Abdi, L.J. Williams, Principal component analysis, Computat. Stat. 2 (4) (2010) 433–459.
[3] S. Achchab, Y.K. Temsamani, Use of artificial intelligence in human resource management: application of machine learning algorithms to an intelligent recruitment system, in: In Advances in Deep Learning, Artificial Intelligence and Robotics, 2022, Springer, Cham, 2022, pp. 203–215.
[4] S.T. Acuna, N. Juristo, A.M. Moreno, Emphasizing human capabilities in software development, IEEe Softw. 23 (2) (2006) 94–101.
[5] M. Arora, A. Prakash, A. Mittal, S. Singh, HR analytics and artificial intelligence-transforming human resource management, in: International Conference on Decision Aid Sciences and Application (DASA), 2021, pp. 288–293, https://doi.org/10.1109/DASA53625.2021.9682325.
[6] N.J. Adler, Z. Aycan, Cross-cultural interaction: what we know and what we need to know, Annual Rev. Organizat. Psych. Organizat. Behav. 5 (2018) 307–333.
[7] G. Arman, C.K. Adair, Cross-cultural differences in perception of time: implications for multinational teams, Eur. J. Work Organ. Psychol. 21 (5) (2012) 657–680.
[8] I. Aviv, A. Barger, S. Pyatigorsky, Novel machine learning approach for automatic employees' soft skills assessment: group collaboration analysis case study, in: In 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS), IEEE, 2021, pp. 1–7.
[9] S. Chatterjee, R. Chaudhuri, D. Vrontis, E. Siachou, Examining the dark side of human resource analytics: an empirical investigation using the privacy calculus approach, Int J Manpow 43 (1) (2022) 52–74.
[10] S. Chowdhury, P. Dey, S. Joel-Edgar, S. Bhattacharya, O. Rodriguez-Espindola, A. Abadie, L. Truong, Unlocking the value of artificial intelligence in human resource management through AI capability framework, Human Res. Manage. Rev. 33 (1) (2023) 100899.
[11] E.L. Christensen, M. Paasivaara, Learning soft skills through distributed software development, in: In Proceedings of the *International Conference on Software and System Processes and International Conference on Global Software Engineering*, 2022, pp. 93–103.
[12] B. Cimatti, Definition, development, assessment of soft skills and their role for the quality of organizations and enterprises, Internat. J. Quality Res. 10 (1) (2016) 97.
[13] R.E. de Souza Santos, P. Ralph, A grounded theory of coordination in remote-first and hybrid software teams, in: In Pro*ceedings of the 44th International Conference on Software Engineering*, 2022, pp. 25–35.
[14] V. de Stefano, M. Wouters, AI and digital tools in workplace management and evaluation: an assessment of the EU's legal framework, Osgoode Legal Studies Res. Paper Forthcom. (2022), https://doi.org/10.2861/305539.
[15] I. Deebom, P. Ordu, Employment Opportunities in Soft Skills Acquisition of Business Education Graduates in the Covid-19 Labour Market, Internat. J. Modern Innovat. Knowl. 2 (4) (2022) 88–100.
[16] D.M. Dekker, C.G. Rutte, P.T Van den Berg, Cultural differences in the perception of critical interaction behaviors in global virtual teams, Internat. J. Intercult. Relat. 32 (5) (2008) 441–452.
[17] N.K. Denzin, Y.S. Lincoln (Eds.), The sage handbook of qualitative research, Sage publications, 2018.
[18] N.M. Dowell, Y. Lin, A. Godfrey, C. Brooks, Exploring the relationship between emergent sociocognitive roles, collaborative problem-solving skills, and outcomes: a group communication analysis, J, Learn, Analyt, 7 (1) (2020) 38–57.
[19] N.M. Dowell, T.M. Nixon, A.C. Graesser, Group communication analysis: a computational linguistics approach for detecting sociocognitive roles in multiparty interactions, Behav Res Methods 51 (3) (2019) 1007–1041.
[20] X. Du, L. Zhang, J.L. Hung, H. Li, H. Tang, Y. Xie, Understand group interaction and cognitive state in online collaborative problem solving: leveraging brain-to-brain synchrony data, Int. J. Educ. Technol. High. Educ. 19 (1) (2022) 1–21.
[21] European Commission. (2021). Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. Accessed April 26, 2023.
[22] N.E. Evangelopoulos, Latent semantic analysis, Cognitive Sci, 4 (6) (2013) 683–692.
[23] D. Falessi, N. Juristo, C. Wohlin, B. Turhan, J. Münch, A. Jedlitschka, M. Oivo, Empirical software engineering experts on the use of students and professionals in experiments, Empir. Softw. Eng. 23 (2018) 452–489.
[24] T. Feraco, D. Resnati, D. Fregonese, A. Spoto, C. Meneghetti, An integrated model of school students' academic achievement and life satisfaction. Linking soft skills, extracurricular activities, self-regulated learning, motivation, and emotions, European J. Psych. Educat. 38 (1) (2023) 109–130.
[25] R. Gafni, N. Geri, P. Bengov, Investigating the effect of tangible and virtual rewards on knowledge contribution in online communities, Online J. Appl. Knowl. Manage. (OJAKM) 2 (2) (2014) 1–11.
[26] M. Galster, A. Mitrovic, S. Malinen, J. Holland, What soft skills does the software industry really want? an exploratory study of software positions in New Zealand, in: In Proceed*ings of the 16th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, 2022, pp. 272–282.
[27] Q. Gao, S. Zhang, Z. Cai, K. Liu, N. Hui, M. Tong, Understanding student teachers' collaborative problem solving competency: insights from process data and multidimensional item response theory, Think. Skills. Creat. 45 (2022) 101097.
[28] V. Garousi, G. Giray, E. Tuzun, C. Catal, M. Felderer, Closing the gap between software engineering education and industrial needs, IEEe Softw. 37 (2) (2019) 68–77.
[29] N. Geri, R. Gafni, P. Bengov, Crowdsourcing as a business model: extrinsic motivations for knowledge sharing in user-generated content websites, J. Global Operat. Strat. Sourc. 10 (1) (2017), https://doi.org/10.1108/JGOSS-05-2016-0018.
[30] L.M. Giermindl, F. Strich, O. Christ, U. Leicht-Deobald, A. Redzepi, The dark sides of people analytics: reviewing the perils for organisations and employees, European Journal of Information Systems 31 (3) (2022) 410–435.
[31] E. Gockel Blessing, T.A. Wood, N.E. Grahovec, Case-by-Collaboration: an Adaptable Soft Skills-Based Educational Model for Health Disciplines, Internet J. Allied Health Sci. Pract. 20 (4) (2022) 12.
[32] V. Gomes, M.G. da Silva, D.C. Kowaltowski, Long-term experience of teaching life cycle assessment and circular design to future architects: a learning by doing approach in a design studio setting, Sustainability. 14 (12) (2022) 7355.
[33] V.G. Goulart, L.B. Liboni, L.O. Cezarino, Balancing skills in the digital transformation era: the future of jobs and the role of higher education, Industry Higher Educat. 36 (2) (2022) 118–127.
[34] E. Hadzhikolev, S. Hadzhikoleva, K. Yotov, M. Borisova, Automated assessment of lower and higher-order thinking skills using artificial intelligence methods, in: In International *Conference on ICT Innovations*, Springer, 2022, pp. 13–25.
[35] R.H. Hamilton, H.K. Davison, Legal and ethical challenges for HR in machine learning, Empl. Responsibil. Rights J. 34 (1) (2022) 19–39.
[36] L. Havrlant, V. Kreinovich, A simple probabilistic explanation of term frequency-inverse document frequency (tf-idf) heuristic (and variations motivated by this explanation), Int. J. Gen. Syst. 46 (1) (2017) 27–36.
[37] M. Hirose, J.W. Creswell, Applying core quality criteria of mixed methods research to an empirical study, J. Mix Methods Res. 17 (1) (2023) 12–28.
[38] H.C. Ho, Y.C. Chan, Longitudinal associations between psychological capital and problem-solving among social workers: a two-wave cross-lagged study, Health. Soc. Care. Comm. 30 (5) (2022) e2702–e2711.
[39] S. Hooda, V.M. Sood, Y. Singh, S. Dalal, M. Sood (Eds.), Agile software development: trends, challenges and applications, 2023.
[40] S.W. Hung, C.W. Chang, S.Y. Chen, Beyond a bunch of reviews: the quality and quantity of electronic word-of-mouth, Info. Manage. 60 (3) (2023) 103777.
[41] A.L. Hunkenschroer, C. Luetge, Ethics of AI-enabled recruiting and selection: a review and research agenda, J. Bus. Ethics 178 (4) (2022) 977–1007.
[42] R. Isufaj, D. Aranega-Sebastia, M. Angel-Piera, Toward conflict resolution with deep multi-agent reinforcement learning, J. Air Transport. 30 (3) (2022) 71–80.
[43] D. Jadhav, J. Kundale, S. Bhagwat, J. Joshi, A systematic review of the tools and techniques in distributed agile software development, Agile Software Develop. (2023) 161–186.
[44] T. Jain, J. Yadav, An enhanced support vector machine for face recognition in fisher subspace, in: InProceedings of Second Doctoral Symposium on Computational Intelligence, Springer, Singapore, 2022, pp. 393–407.
[45] D. Jelonek, T. Nitkiewicz, P. Koomsap, Soft skills of engineers in view of industry 4.0 challenges, in: In Co*nference Quality Production Improvement–CQPI* 2, 2020, pp. 107–116.
[46] K. Jones, L.N. Leonard, G. Lang, Desired skills for entry level IS positions: identification and assessment, J. Comp. Informat. Syst. 58 (3) (2018) 214–220.
[47] F. Kortum, Supporting the understanding of team dynamics in agile software development through computer-aided sprint feedback, Logos Verlag Berlin GmbH, 2022.

[48] F. Lagioia, The impact of the general data protection regulation (GDPR) oN ARTIFICIAL INTELLIGence. European Parliamentary Research Service, Scientific Foresight Unit, 2020. Retrieved from, https://policycommons.net/artifacts/1336910/the-impact-of-the-general-data-protection-regulation-gdpr-on-artificial-intelligence/1944455/on, 26 Apr 2023. CID: 20.500.12592/v7bt0p.
[49] S. Leavy, G. Meaney, K. Wade, D. Greene, Mitigating gender bias in machine learning data sets, in: In Bias and Social Aspects in Search and Recommendation: *First International Workshop, BIAS 2020* 1, Springer International Publishing, Lisbon, Portugal, 2020, pp. 12–26. April 14, Proceedings.
[50] E. Ledet, K. McNulty, D. Morales, M. Shandell, How to be great at people analytics, McKinsey & Company, 2020.
[51] ... & J. Lin, T. Gao, Y. Wen, X. Yu, B. You, Y. Yin, H. Pu, Graph convolutional neural networks-based assessment of students' collaboration ability, Pract. Exper. 34 (28) (2022) e7395.
[52] J. Lindsjørn, D.I.K. Sjøberg, T. Dingsøyr, G.R. Bergersen, T. Dybå, Teamwork quality and project success in software development: a survey of agile development teams, J. Syst. Software 122 (2016) 274–286.
[53] T. Liu, M. Israel, Uncovering students' problem-solving processes in game-based learning environments, Comput. Educ. 182 (2022) 104462.
[54] A. Malik, P. Budhwar, H. Mohan, S. NR, Employee experience–the missing link for engaging employees: insights from an MNE's AI-based HR ecosystem, Hum Resour Manage 62 (1) (2023) 97–115.
[55] G. Matturro, F. Raschetti, C. Fontán, A systematic mapping study on soft skills in software engineering, J. Universal Comp. Sci. 25 (1) (2019) 16–41.
[56] A.S. Mayer, F. Strich, M. Fiedler, Unintended consequences of introducing AI systems for decision making, MIS Quart. Execut. 19 (4) (2020) 239–257. https://aisel.aisnet.org/misqe/vol19/iss4/6.
[57] S. McCartney, N. Fu, Promise versus reality: a systematic review of the ongoing debates in people analytics, J. Organizat. Effective. 9 (2) (2022) 281–311, https://doi.org/10.1108/JOEPP-01-2021-0013.
[58] D.B. Minbaeva, Building credible human capital analytics for organizational competitive advantage, Hum. Resour. Manage. 57 (3) (2018) 701–713.
[59] M. Ortu, G. Destefanis, B. Adams, A. Murgia, M. Marchesi, R. Tonelli, The Jira repository dataset: understanding social aspects of software development, in: In *Proceedings of the 11th international conference on predictive models and data analytics in software engineering*, 2015, pp. 1–4.
[60] F. Ouyang, T. Ling, P. Jiao, Development of group cognition in online collaborative problem-solving processes, J. Edu. Comp. Res. 60 (3) (2022) 599–630.
[61] B.L.E. Pamplona, B.C.A. Henao, S.A.F. Calvo, Human activity recognition using penalized support vector machines and hidden Markov models, Revista Facultad de Ingeniería Universidad de Antioquia 103 (2022) 152–163, https://doi.org/10.17533/udea.redin.20210532.
[62] H.R. Pinto, L.B. Gouveia, M. Trigo, Assessment of the evaluation of technical and soft skills in a professional training course: pilot study to analyze the efficiency of the instruments and results. In *Perspectives and Trends in Education and Technology*, Springer, Singapore, 2023, pp. 165–173.
[63] H. Rika, I. Aviv, R. Weitzfeld, Unleashing the potentials of quantum probability theory for customer experience analytics, Big. Data Cogn. Comput. 6 (4) (2022) 135.
[64] H. Rika, I. Aviv, A. Bertha, Novel data science approach for emotion analytics: from machine learning to quantum cognition, in: IEEE World Conference on Applied Intelligence and Computing (AIC 2022), 2022.
[65] R. Schwartz, A. Vassilev, K. Greene, L. Perine, A. Burt, P. Hall, Towards a standard for identifying and managing bias in artificial intelligence, NIST Special Publicat. 1270 (2022) 1–77.
[66] S. Sherman, B. Kantsepolsky, I. Roth, Teamwork skills: the case of a practical course, Issues Informat. Syst. 23 (2) (2022), https://doi.org/10.48009/2_iis_2022_102.
[67] Q. Song, W. Wang, W. Fu, Y. Sun, D. Wang, Z. Gao, Research on quantum cognition in autonomous driving, Sci Rep 12 (1) (2022) 300.
[68] D. Steinley, K-means clustering: a half-century synthesis, British J. Mathemat. Statistical Psych. 59 (1) (2006) 1–34.
[69] M. Stevens, R. Norman, Industry expectations of soft skills in IT graduates: a regional survey, in: In Proceedings of the Australasian Computer Science Week Multiconference, 2016, pp. 1–9.
[70] D. Strode, T. Dingsøyr, Y. Lindsjorn, A teamwork effectiveness model for agile software development, Empir. Softw. Eng. 27 (2) (2022) 56.
[71] K. Styron, Forward-Looking Practices to Improve the Soft Skills of Software Engineers, Business Manage, Res, Applicat 2 (1) (2023) 1–36.
[72] C. Sun, V.J. Shute, A.E. Stewart, Q. Beck-White, C.R. Reinhardt, G. Zhou, N. Duran, S.K D'Mello, The relationship between collaborative problem solving behaviors and solution outcomes in a game-based learning environment, Comput. Human Behav. 128 (2022) 107120, https://doi.org/10.1016/j.chb.2021.107120.
[73] O. Sykes, D. Shaw, B. Webb, Research design and methods for international planning studies. In International Planning Studies: An Introduction, Springer Nature Singapore, Singapore, 2023, pp. 85–112.
[74] H. Tang, M. Dai, S. Yang, X. Du, J.L. Hung, H. Li, Using multimodal analytics to systemically investigate online collaborative problem-solving, Distance Education 43 (2) (2022) 290–317.
[75] A. Tursunbayeva, C. Pagliari, S. Di Lauro, G. Antonelli, The ethics of people analytics: risks, opportunities and recommendations, Personnel Rev. 51 (3) (2022) 900–921.
[76] A. Volkov, Y. Rishko, Y. Kostyukhin, E. Sidorova, D. Boboshko, D. Savinova, V. Ershova, Using digital tools to teach soft skill-oriented subjects to university students during the COVID-19 pandemic, Educ. Sci. (Basel) 12 (5) (2022) 335.
[77] World Economic Forum, The future of jobs report 2020, 2020. Retrieved from Geneva. 2020 Oct. Accessed 26 April 2023.
[78] J. Yuan, Y. Xiao, H. Liu, Assessment of collaborative problem solving based on process stream data: a new paradigm for extracting indicators and modeling dyad data, Front Psychol. 10 (2019) 369.
[79] S. Zhang, Q. Gao, M. Sun, Z. Cai, H. Li, Y. Tang, Q. Liu, Understanding student teachers' collaborative problem solving: insights from an epistemic network analysis (ENA), Comput. Educ 183 (2022) 104485.
[80] S. Zhang, Y. Cao, M.C. Chan, M.E. Wan, A comparison of meaning negotiation during collaborative problem solving in mathematics between students in China and Australia, ZDM–Mathemat. Educat. 54 (2) (2022) 287–302.
[81] W. Zhou, A.J. Hill, The spillover effects of parental verbal conflict on classmates' cognitive and noncognitive outcomes, Eco.n Inq. 61 (2) (2023) 342–363.