UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Computer Science
Software Engineering Curriculum

Karl Rapur

Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods

Master's Thesis (30 ECTS)

Supervisor: Alexander Nolte, PhD

Tartu 2024

Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods

Abstract:
Hackathons are intensive, time-bounded events where teams collaborate to solve specific problems or develop innovative artifacts. Over the past years, researchers have primarily used qualitative instruments *(e.g., observations and interviews)* to study these events. However, these instruments come with inherent limitations. Scalability issues hinder current data collection methods in the context of large hackathons, as they require significant human resource allocation. Furthermore, current methods have limited potential for live feedback, which is necessary for hackathon organizers.

This study demonstrates a combination of qualitative and quantitative data collection instruments to study collaboration dynamics. A case study was conducted during a hackathon with multiple teams, where data was gathered by combining conventional methods with smart badges and sensors. The study uncovered varied collaboration dynamics influenced by experienced participants, constellation changes, language barriers, goal clarity, adaptability, and mentorship. Badge-collected data provided insights into speaker transition logs, speech patterns, transcriptions, and spatial awareness.

As contemporary research has primarily relied on qualitative data collection instruments with inherent limitations, this study showcased the possibilities of using sensors and badges to fill this gap. Therefore, it lays the groundwork for future work to leverage such technology to overcome limitations, study hackathons on a larger scale, and provide live feedback to organizers.

Keywords:
hackathon, data analysis, collaboration dynamics, innovation, MMLA Platform

CERCS: P170 - Computer science, numerical analysis, systems, control

Häkatonide koostöö dünaamika analüüs: Kvantitatiivsete ja kvalitatiivsete meetodite ühendamine

Lühikokkuvõte:
Häkatonid on intensiivsed, ajaliselt piiratud üritused, kus meeskonnad teevad koostööd spetsiifiliste probleemide lahendamiseks või uuenduslike tehiste väljatöötamiseks. Viimastel aastatel on teadlased nende ürituste uurimiseks peamiselt kasutanud kvalitatiivseid meetodeid *(nt. vaatlusi ja intervjuusid)*. Nendega kaasnevad teatavad piirangud. Esiteks need meetodid ei skaleeru, kuna nõuavad märkimisväärse inimressursi eraldamist. See takistab suurte häkatonide uurimist. Lisaks ei võimalda kasutatavad meetodid reaalajas saada tagasisidet, mis on vajalik häkatonide korraldajatele.

See uuring demonstreerib kvalitatiivsete ja kvantitatiivsete metoodikate ühendamist koostöödünaamika uurimiseks. Juhtumiuuring teostati mitme meeskonnaga häkatonil, kus andmeid koguti kvalitatiivsete metoodikatega ning nutikate rinnamärkide ja anduritega. Uuring näitas mitmeid koostöödünaamika tegureid, milleks olid kogenud osalejad, meeskonna muutused, keelebarjäärid, eesmärgi selgus, kohanemisvõime ja mentorite tugi. Rinnamärkide kogutud andmed pakkusid teavet kõnelejate järjekorra üle, kõne mustrite, kõne transkriptsioonide ja ruumilise teadlikkuse kohta.

Kuna kaasaegne teadustöö on peamiselt toetunud kvalitatiivsetele andmekogumisinstrumentidele, näitab see uuring erinevaid võimalusi nende vahendite piirangute vähendamiseks, kasutades andureid ja rinnamärke. Ühtlasi paneb see aluse tulevastele teadustöödele, mis saavad kasutada seda tehnoloogiat piirangute ületamiseks ja sedasi uurida suuremaid häkatone ning pakkuda korraldajatele reaalajas tagasisidet.

Võtmesõnad:
häkaton, andmeanalüüs, koostöö dünaamika, innovatsioon, MMLA tehnoloogia

CERCS: P170 - Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria)

Contents

1 Introduction

Hackathons have grown in popularity in recent years [1], uniting individuals from diverse backgrounds and skill sets. Each hackathon is distinct, driven by its unique purpose and objectives. In the context of this thesis, any time-constrained collaborative event where individuals work collectively towards a shared goal is defined as a hackathon [2]. Consequently, the thesis encompasses various forms of such events, including idea hackathons, corporate hackathons, makeathons, game jams, and others under one umbrella term – hackathons.

Effective teamwork demands setting aside differences and aligning efforts toward a common objective [3, 4]. The aim of this research is to combine qualitative and quantitative methodologies and show the potential use of technological means in this research field. By analyzing the gathered data, the thesis aims to explore collaboration dynamics and external factors, which include interactions with mentors and participating teams.

Hackathons are usually short but intensive events [2], where motivation and workload could fluctuate depending on the drawbacks or successes teams face. Therefore, it is important not to generalize the whole event but to look at the situations in the order they happen. Each situation may introduce shifts in roles, task divisions, communication, etc., affecting teams' effectiveness and productivity. For that, it is crucial to research collaboration dynamics throughout the hackathon, and therefore, the following research question is presented:

> ***RQ1:*** *How do collaboration dynamics evolve throughout a hackathon?*

By following observed events and situations in chronological order, these events can be compared with timestamped badge-collected data. This helps to show the influence of such technological instruments for future studies.

Most hackathons are surrounded by supportive people who share their domain knowledge and help teams to overcome challenges. However, considering the diversity of people involved, mentoring approaches can vary widely. Some researchers are trying to introduce standard guidelines. For instance, Nolte et al. have developed mentoring guidelines that give different mentoring strategies and setups [2]. The following thesis details mentors' role in collaboration dynamics and effectiveness while answering the following research question:

> ***RQ2:*** *How do external actors, such as mentors and other teams, influence collaboration dynamics?*

Contemporary research has mainly relied on qualitative methods, such as observations and interviews, with inherent limitations. Falk et al. proposed six potential research fields in their paper [5], one of which is scale-up. This has not been done primarily due to the

challenges current methods pose. Studying a significant event involving more than 50+ participants often necessitates a large research team, potentially introducing individual biases and subjective data, thus falling short of providing a comprehensive understanding of participant behavior throughout the event. Furthermore, current qualitative instruments lack the potential to give real-time feedback to hackathon organizers. To overcome these limitations, different research fields have started using Multimodal Learning Analytics (MMLA) [6], which is helpful in examining complex learning and collaboration dynamics for group work [7]. Therefore, this study was partially done in cooperation with a research group from the University of Copenhagen, who developed wearable intelligent badges. As a result, the primary objective of this thesis is to demonstrate the utility of wearable sensors by addressing the following question:

> ***RQ3:*** *Which aspects of collaboration can be observed using on-person wearable sensors?*

The research employs a blend of qualitative and quantitative data collection instruments. Qualitative methods encompass observations and interviews, while quantitative techniques involve online surveys and wearable MMLA [6, 7] badges with various sensors. These sensors facilitate the capture of diverse data points, including speech duration and spatial positioning within the room [7]. Furthermore, these instruments hold the potential to enhance scalability and provide near real-time feedback to hackathon organizers. The research also incorporates video and audio recordings to establish a robust ground truth.

Various methodologies are used to analyze collected data in the subsequent data analysis phase. Notably, the research leverages affinity diagrams [8] and statement encodings for qualitative analysis. These insights serve as the foundation upon which conclusions are formulated. They, in turn, are complemented and compared with data collected by the sensors.

This study provides a groundwork for future researchers interested in investigating team collaboration dynamics on a larger scale through technological means. Using sensors allows for concurrently studying multiple teams, reducing the reliance on numerous observers. Furthermore, it allows near real-time feedback for organizers [7]. Consequently, it fosters a more comprehensive and objective evaluation, enriching the understanding of collaboration dynamics at the grassroots level.

2 Background

As hackathons have gained popularity in recent years [1], it is essential to understand the background of such events and how they are defined. The following sections briefly cover hackathons' origins (Section 2.1) and define a collaboration (Section 2.2). The last section (Section 2.3) explores related work before this research, exploring the limitations and gaps this study will overcome.

2.1 History of Hackathons

Hackathons have been around for over a decade, but the initial mention of a hackathon is vague. While Nolte et al. believe hackathons emerged in the mid-2000s [9, 10], Briscoe and Mulligan show that the term appeared on June 4th, 1999, when security expert Niels Provos used the word hackathon to describe an OpenBSD event held in Calgary, Canada [11, 12]. Ten developers teamed up to overcome legal problems caused by export regulation of cryptographic software. Interestingly, just two weeks later, Sun Microsystems referred to an event at the JavaOne conference (held from June 15 to June 19) as a hackathon [11, 12]. Its goal was to write a program in Java for the new Palm V using the infrared port to communicate with other Palm users. The naming choice by Sun remains shrouded in uncertainty, leaving open the question of whether the OpenBSD event inspired them or if it was a coincidental use of the term [11]. This historical ambiguity underscores hackathons' complex and evolving nature.

Hackathons started as programming events that allowed people to cooperate on particular problems [12]. According to Rys's literature review [13], the concept of time-bounded events that last 24, 36, 48, or 72 hours was later introduced. She found that most researchers speak of 24 and 48-hour events, but a 24-hour-long hackathon is the dominant choice. With the rise of virtual hackathons, the significance of time constraints has transformed, as many of these events now extend over several weeks with conventional 9-to-5 schedules [11, 14].

2.2 Defining Collaboration

Keyton [3] defines collaboration as an interaction in which individuals, or team or organizational members, work together to reach a common shared goal, activity, or production. Most often, collaboration occurs in teams or group collaborators, although some collaborations comprise hundreds of participants. Furthermore, he labels an interaction as collaborative when it has at least two interacting parties, a shared goal or mutual goal achievement, interdependence among parties, and the interaction occurring over time [3]. According to this definition, hackathons are collaborative events as individuals come together to work as a group towards a common goal while often lasting over multiple hours.

Evaluating the success or effectiveness of collaboration requires preset measurements [15] and, therefore, is difficult to do in a hackathon setting. Luckily, some publications exist that define the fundamentals for effective collaboration. Scoular et al. have published an ACER collaboration framework [4], which gives guidelines for effective teamwork. According to them, effective group collaboration consists of three key strands. The first strand focuses on building shared understanding through communication, resource pooling, and negotiation of roles. The second strand highlights collective contribution by emphasizing active participation, acknowledgment of others' contributions, and engagement with roles and responsibilities. The third strand, dedicated to regulating group dynamics, involves ensuring constructive contributions, resolving differences, maintaining shared understanding, and adapting behavior for the group's benefit.

2.3 Related Work

The first research publications were published at the start of the last decade. Falk's and Halskov's joint research shows that the first publications about hackathons were released in 2013 [1]. The number of publications has grown cumulatively yearly, as shown in Table 1. The same fact is supported by another Falk's publication about hackathon's future research [5]. Falk's publication [1] shows how different qualitative and quantitative methods could be used to improve future research with hackathons. This Master's thesis implies some of the principles brought out by Falk. She proposes developing quantitative tools to measure and research hackathons on a bigger scale and with different perspectives is important.

Table 1. Hackathons research over the years [1]

	Research *with* hackathons Hackathons as example	**Research *with* hackathons** Hackathons as means	**Research *on* hackathons**
2007	No publications	1	No publications
2008	No publications	No publications	No publications
2009	1	1	No publications
2010	No publications	1	No publications
2011	1	2	No publications
2012	4	3	No publications
2013	11	7	1
2014	13	25	4
2015	12	24	11
2016	45	29	12
2017	33	39	16
2018	47	40	2

Research of collaboration is not something new. It has been done with virtual, in-place, and hybrid hackathons. For instance, Mendes et al. [16] investigated virtual collaboration aspects in four hackathons during the COVID-19 pandemic. They focused on social and technical constraints on leadership, collaboration, and communication. They used semi-structured interviews ranging from 15 to 90 minutes for data collection. Many teams said they felt somewhat isolated and missed external communication. In-depth collaboration forms, mentorship, and leadership were not investigated. Tillo's Master's thesis backed this research paper [17], where she primarily focused on virtual collaboration forms. She investigated two different virtual hackathons with 18 research participants in total. Interviews were used as the primary data collection instrument.

Although widely used, interviews are often backed with other qualitative instruments, such as observations. For example, Trainer et al. [18] used observations and semi-structured interviews to investigate collaboration tradeoffs in three hackathons. Similar instruments were used by Pe-Than et al. [19], who investigated participants' motivators, goal settings, coordination, and results in corporate hackathons. They heavily relied on interviews, shown by using pre-interviews, post-interviews after the hackathon, and post-interviews four months later. None of these papers used any form of quantitative tools.

Some research papers show the use of both qualitative and quantitative instruments. Nolte et al. [9] investigated the mentorship role in a hackathon, where they primarily focused on a scientific community. For data collection, they used interviews, observations, and before-after surveys. In addition, they had access to documents produced by participants. A similar methodology was used by Leemet [20] in his thesis, where he

investigated innovation in corporate hackathons. Observations, interviews, and surveys were used as data collection instruments.

Some researchers have used data analysis to dig deep into collaboration forms. For instance, Schulten et al. [21] thoroughly analyzed online communication channels like Slack. They extracted the participants' chats in timestamped order. They compared them with timely essential events, such as checkpoints in which participants had to give an overview to the committee about done work. This method applies and scales well for virtual hackathons but cannot be used with hybrid or in-place hackathons.

In conclusion, the existing literature primarily explores collaboration, communication, and teamwork within small to medium-sized hackathons. This research primarily relies on conventional qualitative methodologies such as interviews and observations. Quantitative investigations are comparatively rare, and there is a notable absence of evidence indicating the integration of technological means, such as sensors, to expand and scale the scope of research in this domain. Furthermore, a gap exists in the literature regarding comprehensive comparisons between qualitative and quantitative research, as most studies tend to concentrate exclusively on one or the other.

3 Methodology

A mixed-method case study was conducted to address the research questions proposed in the introduction. Observations and interviews were used to answer RQ1 and RQ2, which require an exploratory approach. Previous researchers have used these methods to study collaboration [9, 18, 19], and they have been proven to be effective. To answer research question RQ3 and come across the limitations the beforementioned methods pose, an MMLA Platform [7] was utilized. Furthermore, quantitative data collection was extended with surveys to find supportive and aggregatable data for the questions RQ1 and RQ2.

The MMLA Platform [7] was created by an interdisciplinary group between the Center for Digital Education and Human-Centered Computing from the University of Copenhagen. Their innovative system allows tracking people's movement and communication via wearable intelligent badges. All the privacy concerns related to video and audio badges were taken into consideration[1]. Qualitative data collection instruments were aligned with these badges to provide comparable datasets.

The methodology chapter is organized into three key sections, each addressing distinct facets of the hackathon research methodology. Section 3.1 outlines the research setting, including the selection of the hackathon, the selection of the participants, and a data collection timeline. In Section 3.2, the used data collection instruments are discussed, encompassing various methods such as observation, interviews, surveys, and technological means. Finally, Section 3.3 describes the analysis methods, showing the strategies employed to analyze the collected data and derive findings. These sections provide a comprehensive overview of the methodological approach guiding this thesis, ensuring transparency and rigor in the research process.

3.1 Research Setting

This research was partnered with Garage48 under the Creative Impact Research Centre Europe[2], which imposed some constraints on hackathon selection and introduced strict deadlines. These aspects are covered in the subsequent sections.

3.1.1 Hackathon Selection

Hackathons are organized in Estonia by different companies and communities. They happen throughout the year, but there is no specific database from where researchers could gather information about upcoming hackathons like Devpost [14] or MLH [22]. Some are closed to the public as the specific corporation organizes them for internal innovation [20]. Furthermore, some hackathons are hybrid or purely remote, which

[1] Research Ethics Committee of The University of Tartu protocol number: 381/T-13

[2] `https://creativeimpact.eu/en/`

is unsuitable for this research, where participants are given intelligent badges to wear. Considering these aspects and the fact that this research partnered with Garage48, the choices for hackathons were limited.

Garage48 [23] has organised hackathons and other open-innovation events since 2010. The idea was initiated by Estonian entrepreneurs behind successful startups such as Pipedrive, Fortumo, Weekdone, Mooncascade, and Taxify. Within those 13 years, they have organized over 250 hackathons in 70 different cities and 50 countries all across the globe. According to them, their main goal is 'To enrich the startup ecosystems all across the globe from a grassroots level to bring the utmost impact to the startup world on a lean budget in a limited amount of time' [23]. Garage48 is located in the Telliskivi Creative City in Palo Alto Club, Tallinn.

Garage48 and Estonian Refugee Council, in partnership with the British Council in Estonia, Swedbank Estonia, and BPW, organized an idea garage-type hackathon titled *"Empowering Women Idea Garage"* on 9-10 September 2023 at their headquarters [24]. It was a part of the more extended entrepreneurship program where Ukrainian refugees were helped to start or integrate their businesses into Estonia. They have been organizing that program since 2019, including hackathons in the Donbas area, Ukraine, before the full-scale war. The hackathon aimed to support refugees forced to leave their home country to rebuild their lives in Estonia through entrepreneurship. Professional mentors from different fields helped the participants by guiding their idea development. Although this hackathon focused on Ukranian women, other nationalities and genders were also welcomed.

This idea garage was a medium-sized two-day hackathon with roughly 50+ participants, who all participated in place. The goal was to develop a business idea and value proposition via teamwork sessions and the guidance of mentors. The event featured three workshops on the first day and two on the second. This format was a good fit for this research, as teams had to communicate and discuss aspects of the idea, making an ideal setting for voice badge data collection and analysis. Pauses between workshops allowed researchers to change the environment, switch and recharge badges, and fine-tune system parameters. Furthermore, the event provided many participants, guaranteeing a diverse pool to select candidates for the research. Lastly, the timing of the hackathon aligned seamlessly with the research timeline, making it a suitable choice for the study's objectives.

3.1.2 Participant Selection

The selected hackathon presented challenges in participant selection, mainly as most participants were refugees, raising unique privacy concerns. Additionally, the research had to adapt to a foreign language by involving more researchers who could speak Russian. The minimum number of required teams was two. The one to be researched with technological means and one without. Ideally, the number of teams would be four,

considering the ample availability of observers. English-speaking teams were preferred whenever possible. Participation in this research was restricted to individuals aged 18 and above to avoid parental consent issues. Given the multiple objectives and intricate privacy concerns, the task of introducing the research was assigned to the organizers.

The Garage48 organizers introduced the research concept and its objectives to participants during the initial briefings held in the days leading up to the hackathon. At the start of the hackathon, participants were given the Russian or English consent forms at the registration table shown in Appendix I. These forms followed the guidelines of the Ethics Committee and Runeson et al. [25]. In addition, next to the table was a Ukrainian/Russian-speaking researcher who answered questions related to research, privacy, etc. Participation in the research was voluntary, and individuals did not receive compensation. Furthermore, participants were allowed to opt out of the research at any given time. In total, 35 individuals gave consent to participate.

Due to the high number of consents, teams were allowed to form naturally, meaning no team was formed artificially by researchers. Later, four teams were selected as shown in Table 2, each with the consent of all its members. Two teams were Russian/Ukrainian-speaking, one spoke Persian/English/Russian, and the fourth was English/Russian-speaking. The team sizes were 4, 6, 3, and 2 members. Every team was assigned a dedicated researcher who closely observed their team dynamics during the hackathon and conducted post-event interviews. A team of four, further referred to as Team A, was chosen to be researched with technological means, as they had optimal team size and were not multilingual. In total, 15 participants were selected for this research.

Table 2. Selected teams for the research

Code	Team Language	Team Size	Study Type
A	Russian, Ukrainian	4	Quantitative, Qualitative
B	Russian, Ukrainian	3 to 6	Qualitative
C	Persian, English, Russian	3	Qualitative
D	English, Russian	2	Qualitative

To ensure the anonymity of each participant, this thesis introduces the coding system outlined in Table 2. Participants affiliated with Team A are denoted as A1, A2, A3, and A4, with a similar denotation for members of the remaining three teams and mentors.

3.1.3 Timeline Overview

Before the hackathon, all data collection methods were developed, discussed with the ethics board, and tested on test subjects. The research was introduced to the participants during the program's introductory sessions that took place in batches before the hackathon. The event itself took place from the 9th to the 10th of September, 2023, shown in Figure 1.

During the hackathon, data was collected with observations and utilizing the MMLA Platform developed by a research team from the University of Copenhagen. They also organized the data collection with this technology.

Participants were interviewed within one week after the hackathon, particularly in week 37, shown in Figure 1. Each interview was scheduled between the participant and interviewer at a time convenient to them. After the interview, participants were given a link to a short online survey to be completed within the same week. Once all the data was collected, it was translated, and then the analysis started.

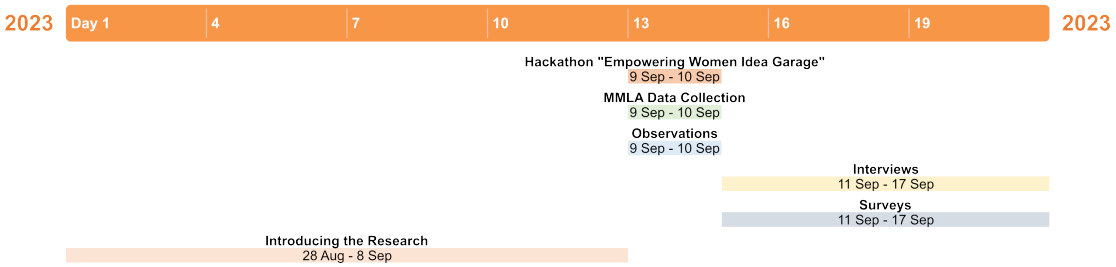

Figure 1. Timeline of research activities of the study

3.2 Data Collection Instruments

This research facilitated multiple data collection instruments. The primary instruments were observations and intelligent badges from the mBox system [7] to achieve the essential dataset for answering proposed research questions. Furthermore, interviews and surveys were also utilized to gather data from individual perspectives and served as supplementary datasets in the event of complete system failure. To mitigate the risks of privacy and data leaks, all the data were collected offline, anonymized with encoding introduced in Section 3.1.2, and later stored in the University of Tartu's servers. The following sections cover all the data collection instruments used in this research.

3.2.1 Observation

Observations have shown their importance throughout the hackathon research of the past years, as explored in Section 2.3. This study employs observational methods to address research questions and provide ground evidence for quantitative data. Furthermore, traditional observation notes are enhanced with timestamps, improving their synchronization with data collected through sensors.

Every team was assigned an observer who could understand their primary language. All observers were previously briefed and given observation guidelines shown in Appendix II. To understand collaboration dynamics (RQ1) and external influence over it (RQ2), these guidelines covered topics partially derived from previous works and ACER

collaboration framework [4], which could be compared or complimented by possible findings from quantitative datasets required by RQ3. The topics are as follows:

- **Task Division & Prioritization** assesses the processes, strategies, and considerations involved in decision-making and work distribution.
- **Team Constellation** helps to understand when teams break up and join.
- **Leadership** assesses the leader's presence and its role.
- **Collaboration & Communication** examines how people work together towards a common goal and how they share information.
- **Challanges** highlight potential problems in team dynamics, for instance, disagreements, confusion, trust issues, lack of skill, etc.
- **Mentorship Influence** assesses the mentor's role and influence over team dynamics and outcomes.
- **External Influence** helps to understand other teams' and individuals' influence over team dynamics and outcomes.
- **Hackathon Setting** evaluates the feasibility of facility and time navigation.

Observation notes needed to follow a chronological order to address the progressive part of RQ1 and be comparable with MMLA data RQ3. Timestamping tools were employed to achieve this. Since the observers operated on different systems, distinct methods had to be suggested. The preferred approach for Windows involved Notepad with the built-in timestamp shortcut F5. On Mac, the recommendation was to utilize the Automator script[3], compatible with any text editor. Similarly, a BASH script[4] activated by a specific hotkey was proposed for Linux. Subsequently, the observation notes, enriched with timestamps, enabled easy comparison with data collected with sensors.

3.2.2 Interview

While observations cover collaboration dynamics, interviews give insight into individual perspectives. Over the past, different types of interviews have been used. For instance, Nolte et al. used pre-event, post-event, and post-post-event interviews to study the outcomes of a corporate hackathon [10]. This research only facilitated post-event interviews, which were semi-structured, meaning that interviewers could adapt and

[3] https://explog.in/notes/inserttimestamp.html
[4] https://softhints.com/linux-mint-hotkey-paste-insert-timestamp/

modify the open-ended questions and their order based on their observations and the interviewee's responses as explained by Runeson [25].

Before the hackathon, a test subject from the *"Defence for Ukraine 2023"* hackathon organized by Garage48 was chosen to conduct the test interview. This was aimed to validate the questions and conformance to the expected interview length. Interviewers were given the interview guidelines (see Appendix III), which covered the following topics:

- **Team Collaboration** assesses the idea stage before the hackathon, task division (e.g., *"How were tasks assigned among team members?"*), decision makings and challenges faced during the hackathon (e.g., *"Were there any challenges where teamwork could have been improved?"*). These aspects gather the participants' personal view of team dynamics necessary to answer RQ1.

- **Leadership** helps characterize the team's leader presence and its role by asking questions like *"How would you characterize the leadership within your team?"*. These questions deepen an understanding of the collaboration aspects addressing the RQ1 and providing qualitative material for validating the use of sensors by answering RQ3.

- **Individual** category explores personal contribution (e.g., *"What do you feel you contributed? What was your role?"*), personal goals (e.g., *How did your personal goals or objectives align with your teams?*) and voice of speech (e.g., *"Did you ever feel like you were not being listened to or your opinions weren't considered? Could you bring out any specific instance of this?"*). These questions assess individual participation in team dynamics and address RQ1 and RQ3.

- **Mentorship** helps to interpret the mentor's role (e.g., *"How would you describe the role of mentors during the hackathon?"*) and importance (e.g., *"Could you share a specific instance where a mentor played a crucial role in helping you overcome a challenge or providing valuable insights?"*) from the participant's perspective to explore RQ2.

- **Other Teams** questions outline the interactions with other teams (e.g., *"Did you assist other teams? How did you help them?"*) to assess the external influence aspects in RQ2.

- **Hackathon Setting** evaluates the general aspects of the hackathon, such as time navigation (e.g., *"Was there a workshop where you felt you had too much or too little time? How did it influence the teamwork?"*), facility (e.g., *"How did the facility help your team's collaboration? Was everything easily accessible, or did you feel lost?"*), etc. This category did not address specific research questions

but provided complementary information about the setting for the research and organizer.

In most cases, interviews were conducted by the same researchers who observed the team. For participants C2 and D2, a different interviewer had to be employed due to the language barrier. In total, 12 interviews were conducted in this research. They all lasted around 10-25 minutes, even though the test interview took 42 minutes.

3.2.3 Survey

A survey is a secondary data collection instrument for this research, providing quantitative data necessary to answer RQ1 and RQ2. Similarly to interviews, it will complement the observations by shedding light on individual perspectives. Questionnaires have been a valuable, in some cases - primary data collection instrument to many studies [9, 10, 16, 20].

The survey uses pre-defined scales adopted from Reining's article on process and outcome satisfaction [26], Sawyer's scales for goal clarity [27], and satisfaction scales from Filippova's publication [28]. These scales were previously tested and validated, demonstrating excellent internal consistency, as assessed by Cronbach's alpha as a statistical measure. For example, Nolte et al. utilized the same scales in their studies [9, 10]. Recognizing that there is no one-size-fits-all solution for hackathon research, some existing scales were modified, and new scales (e.g., Hackathon Setting) were developed. This process followed Nemoto's guidelines for creating Likert-Scale questionnaires [29]. All scales used a five-point rating system: Strongly disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Strongly agree. To ensure clarity and adherence to time constraints, the modified survey underwent testing on two pilot subjects.

The final survey (see Appendix IV) was administered through the University of Tartu's LimeSurvery platform[5]. It consisted of the following categories:

- **Teamwork** assesses the individual perspective on team processes, individual goal clarity, and teamwork.
- **Leadership** indicates the leader's presence and its effectiveness.
- **Individual** category explored individual learning and skill utilization during the hackathon.
- **External Actors** examines participant's views on mentors and other teams.
- **Hackathon Setting** provides essential feedback to the organizer and gives a general insight about the setting.

[5] `https://survey.ut.ee/`

3.2.4 Multimodal Learning Analytics Platform – mBox

Studying hackathons with traditional methods, such as observations and interviews, is problematic as they occupy a large group of researchers simultaneously, introduce bias, and do not scale well [1, 5]. Furthermore, they do not allow on-the-fly analysis of collaboration, which could be valuable insights for hackathon organizers. An interdisciplinary group between the Center for Digital Education and the Human-Centered Computing group from the University of Copenhagen tackled this problem by developing an open system incorporating diverse data sensors that can connect to a system for capturing a range of modalities. As this system is still in the prototype phase, this thesis paper will try to examine the use of such sensors by answering RQ3.

Their developed Open MMLA Platform mBox [7] architecture revolves around multi-faceted badges for on-person data collection and specialized base stations for data processing and synchronization, as shown in Figure 2. This system employs four different types of badges. Vision Badges are the most comprehensive ones as they are equipped with Nicla Vision Arduino boards[6] and AprilTags [30], facilitating onboard AprilTag detection that aids in constructing a participant network graph. Regular Badges, also called passive badges, feature a basic AprilTag to provide essential data on location and orientation. Voice Badges, utilizing Nicla Vision, are designed to stream audio data to the audio base station for speaker and speech recognition. This study did not use RFID Badges shown in Figure 2. Physical badges used in this study are shown in Figure 3 part A.

[6] https://www.arduino.cc/pro/hardware-product-nicla-vision/

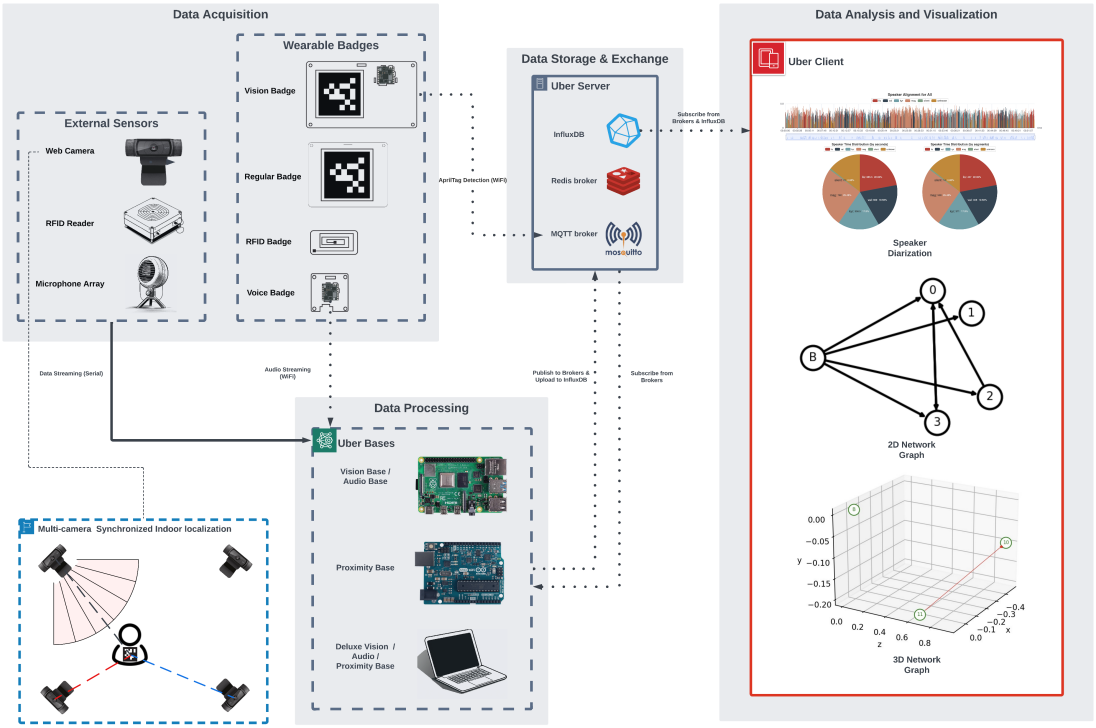

Figure 2. Visualization of mBox architecture [7]

According to the mBox description [7], the mentioned badges and external sensors (e.g., web cameras, microphone arrays, etc.) collect and transmit data to base stations, depicted as Uber Bases in Figure 2. These hubs are specialized units designed for handling specific data types. The Vision Base Station, powered by either a Raspberry Pi 4 or high-performance computers, processes visual data captured from web cameras or visual badges to detect badge-to-badge spatial relationships. Similarly, the Audio Base Station, also powered by a Raspberry Pi 4 or high-performance computers, processes audio data recorded by Voice Badges or the Jabra conference microphone array speaker. This base station synchronizes the audio and identifies the most dominant speaker at any given moment. These base stations communicate over WiFi and USB interfaces. They use the Message Queuing Telemetry Transport (MQTT) protocol for standardized communication and utilize Redis as a message broker. Furthermore, collected data and visualizations could be accessed in real-time in the system's dashboard, as shown in Figure 3 part B, provided by the Uber Client.

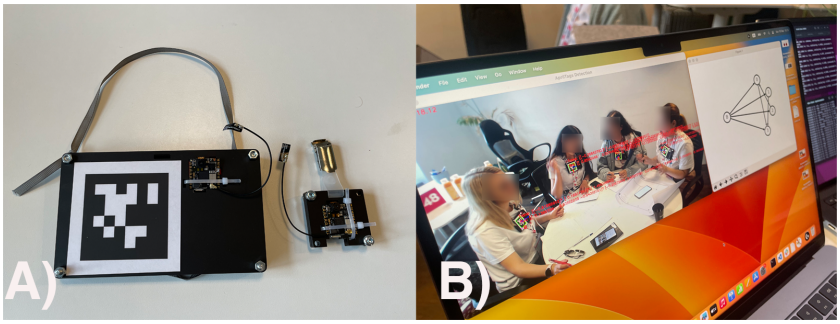

Figure 3. The Audio and Vision Badages in part A) and the badges in action during the hackathon in part B) [7]

Due to the limited number of badges and researchers familiar with the aforementioned system, only Team A was selected to be studied with this technology, with detailed reasons outlined in Section 3.1.2. As the hackathon lasted over two days and consisted of five workshops, it allowed testing different settings, as indicated in Table 3. A small room with a rectangular table and fixed chairs was used on the first day. A round table with movable office chairs was used on the second day. This change was based on the hypothesis that with movable chairs and round tables, people tend to face each other more when talking. Furthermore, different badges were used during the sessions to test their accuracy and general usability. A team from the University of Copenhagen continuously refined the badge parameters between sessions, as illustrated in Table 3. Changing the badges and environment settings helps to address RQ3 with a deeper understanding of different badges and their effectiveness in diverse situations.

Table 3. Data collection setting for the sensors

Set up	Devices	Environment	Parameters
Workshop 1	5 Vision Badges 5 Voice Badges	Square tables Fixed chairs	Looking Threshold: -0.8 Register Volume EQ: -20 dBm Recognize Volume EQ: -20 dBm Recognize Threshold: 0.3
Workshop 2	5 Regular Badges 5 Voice Badges	Square tables Fixed chairs	Looking Threshold: -0.95 Register Volume EQ: -20 dBm Recognize Volume EQ: -20 dBm Recognize Threshold: 0.3
Workshop 3	5 Regular Badges Jabra Microphone	Square tables Fixed chairs	Looking Threshold: -0.95 Register Volume EQ: -20 dBm Recognise Volume EQ: -20 dBm Recognise Threshold: 0.3
Workshop 4	5 Vision Badges 5 Voice Badges	Round tables Movable chairs	Looking Threshold: -0.85 Register Volume EQ: -20 dBm Recognize Volume EQ: -20 dBm Recognize Threshold: 0.3
Workshop 5	5 Regular Badges 5 Voice Badges	Round tables Movable chairs	Looking Threshold: -0.85 Register Volume EQ: -20 dBm Recognize Volume EQ: -15 dBm Recognize Threshold: 0.45

3.3 Methods for Analysis

Data collection resulted in observation notes from all four teams across all five workshops, 12 interviews, and 11 survey responses. Badges were used on Team A to gather quantitative information from all five workshops. At the start of the fifth session, the system failed to start; therefore, there were some missing data points from the beginning of that session. Before the data analysis, all the interviews that were conducted in Russian had to be translated into English. This was done, for the most part, by the interviewers themselves. After this, qualitative content analysis strategies proposed by Runeson [25] were supplemented with quantitative *(e.g., data collected by sensors)* analysis.

Qualitative content analysis was done by hypothesis generation, which Runeson [25] recommends that researchers be unbiased and open to whatever hypotheses are to be found in the data. First, a coding schema had to be developed. For this, an affinity diagram was created by following Plain's described affinity diagramming process [8]:

1. Tranfering all interviews and observation notes to the index cards. This was done

in an online tool called a draw.io[7] as the options for developing a large physical affinity diagram using sticky notes were limited.

2. Scattering the cards on a board.
3. Arranging the cards into groups of related ideas, issues, or topics. Although Plain states that this step should be done in a group, this was done by the author alone.
4. Developing titles for each group of cards. All the logical ones were assembled into a larger supergroup as shown in example Figure 4.

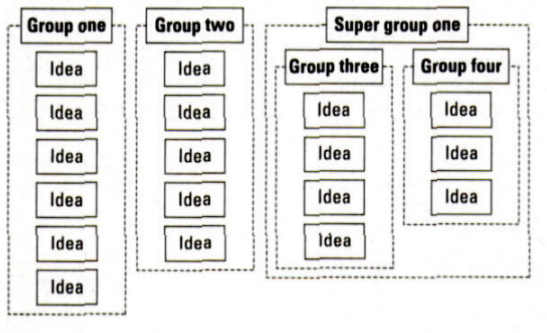

Figure 4. Affinity diagram example [8]

The affinity diagram was used to develop a coding schema for data. This schema could be applied to portions of text related to specific themes, areas, constructs, etc. Once established, all interviews and observations were systematically analyzed to formulate hypotheses. This involved coding responses and observer notes. The findings were then organized into a table as shown in Figure 5, with rows representing situations, notes, responses, etc., and columns representing research codes.

[7] https://www.drawio.com/

Index	Timestamp	C	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	O	G1	G2	G3	G4
1	09/09/2023 14:24	o									+	+										f				
2	09/09/2023 14:28	l												+								f			+	
3	09/09/2023 14:29	l											+	+								-				
4	09/09/2023 14:30	a												+							+	T				
5	09/09/2023 14:31	b									+											o				
6	09/09/2023 14:32	o																				p		+	+	
7	09/09/2023 14:33	r									+							+				i				
8	09/09/2023 14:34	a						+														c				
9	09/09/2023 14:35	t									+															
10	09/09/2023 14:37	i						+										+				C				
11	09/09/2023 14:39	o						+			+											o				
12	09/09/2023 14:40	n						+														n				
13	09/09/2023 14:41						+	+	+		+	+										v				
14	09/09/2023 14:44	&					+	+								+		+		+		e				
15	09/09/2023 14:45																	+				r			+	
16	09/09/2023 14:45	C	+											+								s				
17	09/09/2023 14:46	o						+														a			+	
18	09/09/2023 14:48	m									+											t				
19	09/09/2023 14:49	u										+										i				
20	09/09/2023 14:50																				+	o				

Figure 5. An example of research encoding

As all the observation notes were encoded in chronological order, they were compared to the quantitative data analysis provided by the mBox System (see Section 3.2.4). Out of all the encoded parts, the relevant and noteworthy findings allowing to answer RQ1 and RQ2 were further analyzed.

In addition, surveys were also analyzed to find answers to these questions. As some of the survey scales were added and modified, the reliability of these scales had to be validated. For that, Cronbach's Alpha was used, an index commonly used in studies to indicate instrument or scale internal consistency [31]. According to Taber [31], the characteristics of Cronbach's Alpha values have been described differently over the years. In general, a closer value to 1.00 means better internal consistency. Before survey analysis, all negative questions were negated, and then Cronbach's Alphas were calculated for each category proposed in Section 3.2.3. All the categories with internal consistency < 0.7 (shown in Table 4) were modified by dropping odd questions or responses to meet at least > 0.5 criterion. The leadership category was excluded from the analysis since it received only one response, as respondents could not identify clear team leaders.

Table 4. Survey reliability assessment using Cronbach's Alpha

Category	**Cronbach's α**
Team Processes	0.76
Teamwork	0.19
Goal Clarity	0.72
Learning	0.70
Mentorship	0.77
Other teams	0.37
Hackathon Setting	0.38

To address RQ3, the issues and challenges were marked down while finding the answers RQ1 and RQ2. Furthermore, different conversational patterns that emerged from the data were analyzed and explained via observation notes. This was also done vice-versa, meaning interview statements and observation records were backed by data collected by the sensors. All the positively and negatively matching cases were written down. The raw video and audio were also analyzed in some situations *(e.g., unclear which instrument explained the situation better)* to find the ground truth.

4 Findings

The following findings describe the team journeys as they unfolded. It helps to assess how collaboration evolves through the hackathon (RQ1). The findings are categorized based on the five workshops held during the hackathon. Each section commences with a concise assignment description, providing context for the interview and observation notes used to illustrate the findings. Each workshop section separately describes collaboration and mentorship aspects (RQ2), allowing for a time-wise comparison. To find aspects of collaboration that could be studied with the MMLA platform (RQ3), these findings are compared and complemented with data collected from the badges.

To give context to the following analysis, Team A consisted of four individuals with different ideas. A1 and A3 worked on one idea, while A2 and A4 had their own idea. All the ideas were connected to cafe or food catering. Team B initially had three members and grew up to six members, with some participants working on individual ideas. Team C had three members who worked on a single idea related to sustainable design. Team D had two members who worked on an idea related to food catering.

Conclusions are drawn, and a brief comparison with other teams is conducted, emphasizing the most noteworthy similarities and differences. This study primarily focused on Team A, on which technological means were utilized, and three other teams were used for comparison. Consequently, the ensuing findings offer insights predominantly about that team, supplemented with occasional observations from other teams.

4.1 Workshop 1 – What is the Problem?

Before the first workshop, the newly formed teams were given an assignment and instructions to make the most of it. After that, they had time to ask clarifying questions and moved to their collaboration spaces. The objective was to construct a matrix addressing the question *"What is the problem?"* by filling the following matrix cells:

- **Description of the Problem** should offer a clear understanding of the problem and relevant context, such as its relation to a product, service, or process.
- **Who is Affected** helps identify individuals or groups affected by the problem.
- **Causes of the Problem** analyses the root causes or contributing factors that lead to the problem in the first place, delving into underlying issues while maintaining conciseness.
- **Impact of the Problem** assesses both short-term and long-term consequences and implications of the problem, providing insight into the magnitude of the problem.

Each cell required a concise sentence that best answered the respective question, with no restrictions on sentence length.

4.1.1 Collaboration Dynamics Analysis

Team A started by sharing personal experiences related to their ideas. Communication was more question-driven towards A4 as her/his mother had participated in the program previously and had already opened a cafe. Many questions were about her/his mother's solution (*"A1 asked where A4's mother is working, A4 is telling about her/his mom's cafe"*, OBS). Soon after, two sub-conversations emerged as A2 and A4 discussed marketing, while A3 and A1 discussed the number of customers. Again, communication was led by more experienced participants as observed (*"A1 has a TikTok page for marketing, now s/he's sharing about the problems related to marketing on TikTok"*, OBS).

Each team member was taking notes for themselves during the discussion. No particular decision behind that was observed; this was instead an individual decision based on having different ideas. Furthermore, there were no signs of getting acquainted with each other, which may have happened previously during the team formation in the main hall.

The initial discussions did not align with the workshop goals. The general topics were mainly focused on solutions, for instance, getting the money, finding the space, and doing the marketing as observed (*"A4 still talking to A2 about unrelated things, mostly marketing of her/his mom's business."*, OBS). Inexperienced team members tried to benefit from these discussions (*"A2 says s/he's happy that A4 is in their team because s/he has so much experience. A1 smiled and agreed."*, OBS). This is also supported by interview statements, where A1 said (*"One participant [A4] had a mother who won the last contest, and this [A2] participant came with her/his own business, so there was something to learn and ask."*, A1). The same was noted by more experienced A2, who thought that others learned the initial steps from her/his (*"They were probably more interested in how to start, where to start, who to ask?"*, A2).

Halfway into the workshop, the discussion aligned more with the assignment, and participants took more notes. A1 started asking questions about the problems with the business idea. Because participants had different ideas, A2 and A4 tended to ignore such idea-related questions as observed (*"A1 asked everyone what the problem's reasons are. A3 responded, A2 and A4 ignored."*, OBS)

As A1 and A3 shared the same idea and were long-time friends (*"We met back in 8th grade. We are friends, yeah, we came here together."*, A1), they intended to communicate and work more as a pair (*"[...] A1 and A3 are mumbling to each other about the problems while in parallel writing this down"*, OBS), while A2 and A4 had more experience and therefore hoped between different discussions. Often, A2 worked alone on her/his idea, often with the help of a smartphone, and then asked A4 for feedback. When s/he was interviewed about her/his team's feedback, s/he said the following: (*"It was more useful for me to have interaction with the mentors."*, A2).

When the mentor joined the team, team-wise communication died, and participants either listened to the mentor or worked individually with their notes (*"Other keep taking*

notes on their own. The conversation is only mentor-A2", OBS).

After the mentor left, everyone took notes independently and did not communicate. For instance, A4's need for help was ignored (*"A4 throws the question out how to formulate her/his idea in one sentence."*, OBS). Communication restarted once participants finished taking notes, with A4 making different jokes. Furthermore, an off-topic discussion about fast food was noted. In general, there was a positive atmosphere. Although participants did not know each other beforehand, they felt confident to make sarcastic jokes (*"A1 says s/he was right from the beginning about her/his understanding of the problem and said in a funny way, 'I'll kill you all.'"*, OBS).

4.1.2 Mentorship Analysis

Mentor M1 joined the team in the middle of the workshop. S/he used question-driven guidance, as s/he kept asking questions related to problems. Furthermore, s/he challenged participants' answers by asking why questions. S/he used the 5 Whys framework, which s/he later shared with participants. After, s/he switched to just giving guidance as observed (*"Mentor was giving tips for a few minutes, all team members were just taking notes."*, OBS). This approach did not fit A1 as s/he expressed her/his opinions during the interview (*"[...] M1 was trying to ask more, to find out more, more like [...] to get to the bottom of it - and basically, there was no kind of mentoring, as I saw in M2, for example."*, A1). On the other hand, A2 preferred this approach more as s/he expressed (*"M1, s/he's the one who, with her/his leading questions, yes, s/he helped."*, A2)

Although the mentor M1 spent less than 10 minutes with a team, s/he managed to clarify the assignment as observed with A1 (*"Mentor left, A1 says that the team members initially made her/his understand the problem wrong, and s/he has to write everything anew."*, OBS). The same turned out from the interview with A4 (*"They [mentors] explained how to build these steps correctly [...] only I was more chaotic."*, A4).

4.1.3 mBox Analysis

In the collaboration analysis section 4.1.1, it was found that communication was mostly led by more experienced participants A2 and A4. It could be confirmed by looking at the time they spent speaking during that workshop. Figure 6 shows that A2 and A4 had bigger speech segments than A1 and A3. In fact, they both spoke roughly twice as much as A1 and A3, who mostly questioned their experience.

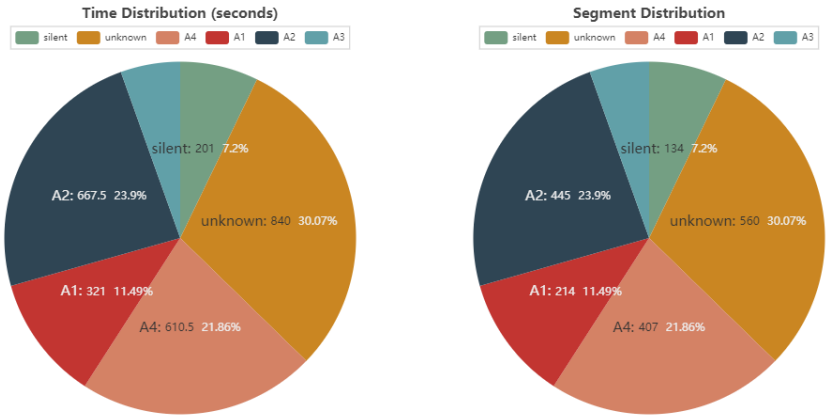

Figure 6. Workshop 1 speaker segments charts

Although observations indicated that participants had different ideas and tended to work more individually or in pairs (as observed with A1 and A3), analyzing how they communicated as a group is challenging. There were some indications that A1 and A3 could have communicated more as a pair (*"A3 and A1 are discussing the task they have just received"*, OBS). However, when examining the MMLA data, this hypothesis lacks support. For example, Figure 7 illustrates speakers' relations based on speech transitions during the workshop, allowing speakers and their follow-ups to be analyzed. The graph indicates that A1 and A3 did communicate, but interestingly, A1 had more extensive discussions with A2 and A4.

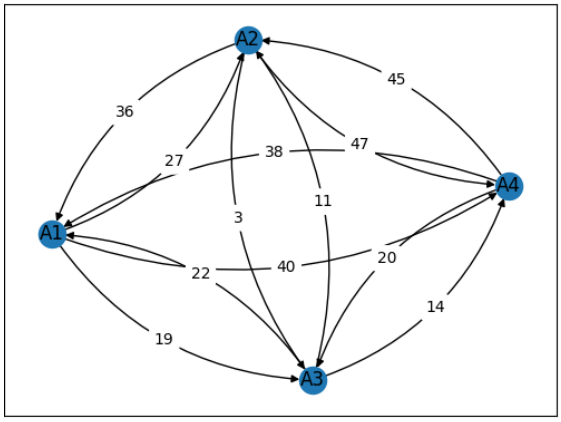

Figure 7. Workshop 1 speakers' relations graph

One of the outputs of MMLA sensors allows for examining speaker patterns. Many exciting patterns emerged, and considering the volume of work, they could be analyzed deeply in a separate study. Here, Figure 8 shows only the noteworthy patterns from

Workshop 1 audio analysis that help to explain or support collaboration and mentorship dynamics.

1) This pattern is taken from the start of Workshop 1 and shows a situation when A1 asks about potential clients, which is followed by A2 sharing her/his experience with marketing (*"A1 is thinking about clients who they could be. [...]. A3 says s/he does not know how to do marketing. A2 takes over and shares her/his experiences on how to do marketing: radio, billboards, etc."*, OBS). This pattern matches well with observations and raw video analysis, illustrating that A2 shared her/his experience in long speech segments.

2) Minutes after the previous pattern, A4 shared her/his mother's experience as seen on qualitative analysis (*"A2 says people are not allowed to give out business cards in Estonia. A4 is sharing how her/his mother solved the problems of business cards."*, OBS). It is clear that this pattern matches observations well and, similarly to the previous pattern, shows that A4 had long speech segments, indicating that s/he was sharing her/his experience.

3) Another pattern from the middle of the workshop that shows long A2 speech segments. Based on the previous two patterns, A2 probably shared some personal experiences. Indeed, this was the case according to the observations (*"A2 is sharing what is going wrong with her/his own project (also cafe, it means)."*, OBS). Although not noted in observations, it is evident that A4 somewhat initiated this as s/he had segments before and in the middle of A2's speech.

4) Although not immediately apparent, this pattern shows an external actor being present as for background noise, there are too many *"unknown"* speaker segments. Compared to the observation notes, this fact gets confirmed (*"Mentor came in. A4 is asking how to understand 'influence'. [...] A2 responded by talking about her/his existing cafe. [...] The mentor challenges A2's answers by asking further questions."*, OBS). This pattern shows that other participants listened when the mentor discussed a problem with A2.

5) In collaboration analysis, it was shown there was some ignorance as the A4 question remained unanswered at the end of the workshop (*"A4 throws the question out how to formulate her/his idea in one sentence. No one responded. S/he talks to herself. [...] No other discussion happens."*, OBS). The same situation can be seen from this pattern, illustrating that everyone was taking notes and did not respond to A4.

6) When the preceding pattern shows collective ignorance, this pattern is the opposite - a collective discussion. Collaboration analysis showed multiple cases when the team had collective discussions when someone asked relevant questions. This is one of the

cases when A2 asked for assignment clarification (*"A2 asked a clarifying question about the reason for the problem. Everyone jumped in to answer."*, OBS).

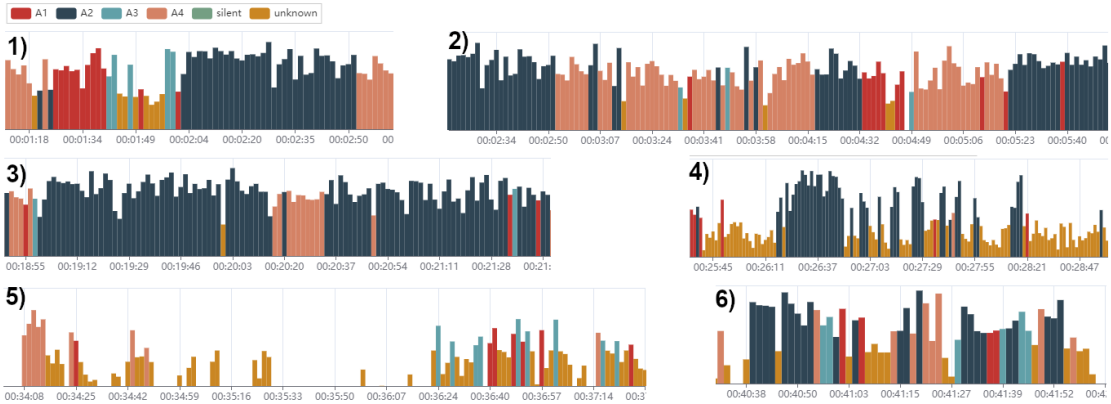

Figure 8. Workshop 1 speaker differentiation patterns

MMLA badges utilize visual detection to identify badges relative to each other. However, the format of this hackathon did not encourage active movement, leading participants to remain seated around the table predominantly. In Workshop 1, Team A was situated at the corner of a square table (environment setup was detailed in Section 3.2.4 in Table 3), and limited movement was observed throughout the hackathon. The positional details and trajectory of badges are depicted in Figure 9, showcasing the potential applications of the badges.

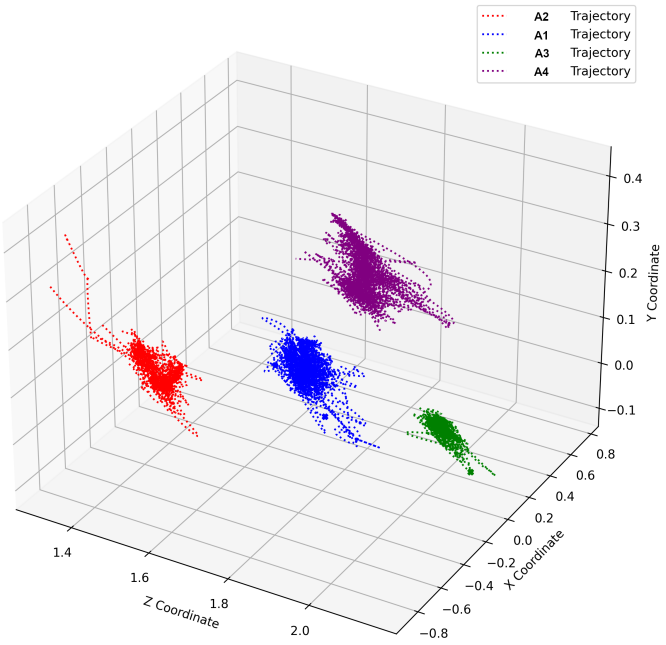

Figure 9. Workshop 1 3D Interpolated Trajectories of Participants

Compared to the video materials, it is apparent that participants leaned forward while taking notes. This may indicate the given back-and-forth movement in one spot. However, this does not provide insightful collaboration findings due to the hackathon type; therefore, spatial movements are not analyzed in the following workshops.

4.1.4 Comparison of Teams

As the previous sections focused primarily on Team A. The following paragraphs give insights into other teams' journeys throughout Workshop 1. In the last paragraph, their collaboration dynamics are compared in the form of a table.

Collaboration Dynamics in Team B Team B consisted of three individuals (B2, B4, B5) with different ideas. They started the workshop with a question-driven discussion around the assignment, led by B5. Due to coordination problems caused by having different ideas, B4 tried to direct the team to work together. During the interview, s/he also pointed out this as a primary problem (*"We had everyone pulling the blanket on their business. Teamwork did not work out because of that."*, B4). Much individual work was observed throughout the workshop, particularly with B2, who did not communicate much during the first half of the workshop. B4 kept focusing on the assignment and introducing new sections of it. Eventually, everyone decided collectively to discuss the focus of their ideas. During this collective discussion, B5 got distracted by the phone.

There was also a collective off-topic conversation, after which the team acknowledged that they had strayed from the topic. When the mentor walked into the room, participants glanced over and continued their discussion. The second mentor brought in additional participant B1, who joined the team halfway through the workshop. B5 was observed interrupting everyone, even mentors. This was also brought up during the interviews (*"S/he [B5] was taking time away from others; s/he did not even let mentors talk."*, B2).

Mentors had to break into the discussion as no-one was not asking anything. B4 led the communication with mentors by explaining the ideas. After getting the initial insight into ideas, mentors started clarifying the ideas using question-driven guidance. The whole team was listening to mentors and answered their questions. Mentors tried to focus on each idea equally, which was also brought out B4 during the interview (*"Mentor kept track of time, not so that one participant pulls the blanket on herself."*, B4). Participants did not have enough time to ask all the questions before the workshop ended.

Collaboration Dynamics in Team C Compared to other teams, Team C, a team of three participants, was the only one to start the workshop with a discussion about picking the idea. They even asked a mentor for help to pick the idea. Immediately, communication in three languages was observed: English, Russian, and Persian. C3 led the communication with her/his personal experience with design and colors. S/he confirmed that everyone understood the idea. This was followed by question-driven communication with C3, who explained the problem and potential solution. Due to the language barrier, Google Translate was used to communicate with C2. C1 decided to draw a grid on paper by herself; later, s/he expressed her/his decision to everyone that s/he would write. As two teams were in the same room, the mentor hopped between teams. Throughout the workshop, C3 developed the idea with her/his knowledge while C1 wrote things down. C2 was listening and often waited for the mentor's translation. C1 kept asking questions from C3 to focus more on what is essential, impact, etc. Although a language barrier caused a lot of confusion and incoordination, Team C kept a positive attitude and even laughed over the poor pronunciation when C3 tried to speak Russian. C1 and C3 often communicated as a sub-group and translated their ideas to C2 via Google Translate. They acknowledged this problem and apologized for it (*"C1 apologizes to C2 for not speaking English - reasons that it is hard to think in one language and then speak in another."*, OBS). As C2 did not understand what was written, s/he wrote a copy in Russian. All team members mentioned in the interviews that the language barrier was the main limiting factor (*"It was hard to be creative [...] also translate and say my meaning."*, C1) (*"If it were not for language barriers, maybe there would have been more time for us to talk."*, C2). Not only did C1 have problems understanding the Russian language, but s/he also had minor issues being interviewed in English, as many questions had to be reworded and simplified to her/his for understanding.

The mentor was present in the room all the time but had to switch between Team C

and Team D. It was imminent that the mentor's prominent role was to be an interpreter and solve the language barrier that both teams had, which was confirmed in the interviews as well (*"For us specifically the mentor was primarily an interpreter."*, C2). Despite that, s/he tried explaining the assignment and sharing domain knowledge. Whenever the mentor switched between the teams, s/he always asked for confirmation that everything was ok and s/he could switch. The mentor tried to scope the problem by asking questions about the root cause and addressing the assignment. C2 was more open with the mentor, using a louder tone, and often presented her/his ideas to the team through the mentor's interpretation.

Collaboration Dynamics in Team D Team D consisted of two participants. D1 started acquaintance with D2. Initially, they had off-topic conversations about their background. D1 led communication as D2 had issues understanding English. Therefore, they essentially used a mentor as a translator, similar to Team C. D1 dominated the communication throughout the workshop but often asked for D2's opinion. D1 shared a lot of personal problems behind the idea. Often, communication died off when the mentor switched to another team. During that time, participants just listened to the other team (*"We, for the most part, just sat there and did not realize what we were doing."*, D2).

The mentor's primary role was to be interpreted in the team's communication. As participants took no notes, the mentor guided them to write discussions down. The mentor used the questions-driven discussion to dig deeper into their idea. Once s/he fully understood it, s/he shared her/his view with D2 in Russian. Meanwhile, D1 disengaged and took notes independently.

Conclusions Table 5 compares various aspects of each team's workshop experience. Each team faced unique challenges, from communication issues and distractions to language barriers. The mentor's role varied, primarily as a translator for Teams C and D. Team A faced initial alignment issues but managed to get on track later. Team B struggled with coordination problems, and Team D faced challenges due to language differences. Team C faced a significant language barrier but maintained a positive attitude throughout the workshop.

Table 5. Collaboration comparison in Workshop 1

Aspect	Team A	Team B	Team C	Team D
Team Constellation	Team of 4, pair-work (idea-based), individual-work	Team of 3 (later 4), individual-work (especially B2)	Team of 3, pair-work (language-based)	2 individuals, individual-work (language-based)
Communication	Discussions led by experienced members, collective and pair discussion.	Question-driven, B5 being dominant, collective discussion.	Collectively picking an idea, language barrier, collective and pair discussion.	Mostly during mentor's presence, D1 being dominant, D2 behind language barrier.
Goal Alignment	3 different ideas, initial conversation deviated from workshop goals, later aligned.	Many different ideas, many deviations, B4 brought the focus back.	One idea, great alignment from the start.	One idea, minimal notetaking, mentor.
Interaction Dynamics	No acquaintance, positive atmosphere (sarcastic jokes), off-topic discussion, individual notes.	B2 tended to work alone, off-topic discussion, individual notes.	Aplogizies over language barrier, positive attitude.	Initial acquaintance, disengagement during mentor switches, off-topic discussion.
Mentor's Role	One mentor, improving team-wise communication, explaining the goal.	Two mentors, improving team-wise communication.	One mentor, interpreter, helping to pick an idea, explaining the goal.	One mentor, interpreter, explaining the goal, guiding to write things down.
Mentor's Approach	Question-driven guidance, 5 Why framework, questioning over guidance.	Breaking into the discussion, clarifying the idea, question-driven guidance, focusing on ideas equally.	Sharing domain knowledge, asking for confirmation, digging into the root cause.	Question-driven guidance.
Challenges and Issues	Initial unaligned communication	Coordination problems, distractions (phone).	Language barrier causing confusion and incoordination.	Communication challenges due to language. D1 disengaged during mentor switches.

4.2 Workshop 2 – What is the Solution?

Before the second workshop, participants were assigned the next task. Similarly to the first one, they were asked to draw a matrix to answer the question, "What is the solution?"

by providing succinct responses in the following four cells:

- **Solution Description** provides a clear and concise description of the proposed solution, which will address the problem identified in the previous workshop.
- **What Do You Need** outlines the resources, skills, tools, technologies, etc., required to implement the solution.
- **What is the Impact/Outcome** brings out both short and long-term positive changes and benefits of the solution for the users.
- **How Does it Differ From Existing Solutions** highlights the unique competitive advantage over the existing solutions, demonstrating innovation.

Participants were instructed to fill each cell with one concise sentence that encapsulates the solution to the best of their ability without a specified word limit.

4.2.1 Collaboration Dynamics Analysis

From the start of the workshop, the discussion was immediately related to the assignment as observed (*"A3 and A1 are discussing the task they have just gotten."*, OBS). Similarly to the previous workshop, the communication was led by more experienced participants, A2 and A4. For instance, A2 shared where her/his restaurant gets the raw materials (*"A2 said they are ordering through Barbora."*, OBS). A1 used question-driven communication to gather knowledge and guidance from A2 (*"A1 is thinking about where to get a workspace for her/his cafe. S/he asked A2 where s/he got that."*, OBS).

Constellation stayed the same, where A1 and A3 worked more as a pair, while A2 and A4 worked more individually on their ideas from the start of the workshop. A2 and A4 were often observed conversing, which may reflect A2's preference to collaborate with more experienced participants (*"I would turn to those who have experience."*, A2). On the other hand, A3 tended to speak only with A1 (*"A3 is so far being quiet overall, [...] s/he speaks quietly to A1."*, OBS).

For the first time, having different ideas and sub-conversations confused collaboration, as A1 noted (*"A1 shared that s/he noticed that when the team members are talking to each other, they are confusing each other. Everyone agreed and kept writing down stuff."*, OBS).

Occasionally, there were some disagreements. For instance, when A1 was concerned with how lawyers work, A3 disagreed with A4's explanations (*"A3 said 'No, it is not true' and was explaining her/his experience."*, OBS). These disagreements and confusion around the assignment caused A4 to express her/his tiredness (*"A4 says s/he's tired and needs to rest. [...] A4 shared s/he has a headache."*, .) These signs were not observed among the other Team A members.

When the mentor joined the team, communication died off. A1 and A3 occasionally discussed some things quietly around their idea, but they mostly listened to the mentor. When the mentor was not paying attention to a particular participant, e.g., guiding someone else, the participants tended to get easily distracted (*"A4 is [...] distracted drawing. A1 is talking to the mentor. A2 started doing something on her/his phone."*, OBS)

4.2.2 Mentorship Analysis

There were many signs of confusion before the mentor joined the team (*"They do not say about the mentor specifically, just like 'I am confused, I need explanation in other words.'"*, OBS). Usually, A1 asked for assignment clarification from her/his teammates, which, in some cases, was followed by collective discussion. Once mentor M1 joined the team, A1 was the first to ask a question (*"Mentor came in, and A1 happily said they were waiting for her. A1 asked what [...] means."*, OBS).

Mentor M1 joined in the second half of the workshop. Since participants already had some questions, the first mentorship minutes were participants' question-driven. There was no collective communication but a sub-conversation between the mentor and idea authors. Later on, mentor M1 used challenging questions towards A2 to help investigate the problem more deeply. This confused A2 even more, after which the mentor encouraged her/him as observed (*"The mentor tells her/his 'Do not worry, it is better to fail now than when you invest the money'"*, OBS). When the mentor finished speaking with A2, s/he asked whether someone had something to ask.

Similarly to the first workshop, Team A misunderstood the given assignment. When A1 asked for clarification about influence, the mentor said that the answers were wrong as observed (*"The mentor says the answers are wrong, and A1 is again like, 'Exactly, girls, [...] I coped from you!'"*, OBS). After the mentor's clarifications, there was not much confusion (*"A1 is sharing that it is a bit clearer now."*, OBS).

4.2.3 mBox Analysis

Similarly to the preceding workshop, collaboration analysis showed that A1 and A3 worked more as a pair. The speakers' relationships graph in Figure 10 does not support this finding. From A3's perspective, it is apparent that s/he communicated most with A1, who had, in fact, extensive conversations with other participants as well. There is no apparent sub-group of communicators.

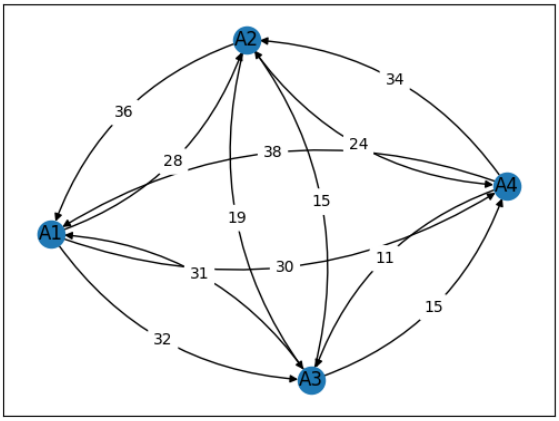

Figure 10. Workshop 2 speakers' relationships graph

Figure 11 shows the second workshop's most noteworthy communication and interaction patterns.

1) The first pattern shows feedback and guidance between participants. By observer, it was noted that A1 was sharing a problem, to which others answered (*"A1 shared with A3 that s/he's afraid they will not have enough staff. A4 listened to it and chipped in some comments. A2 is writing something on her/his own."*, .) A2 worked independently, which matches this pattern, as no A2 comments are recorded.

2) Another interesting thing to look at is disagreements. In the second workshop, A3 disagreed with A4, after which s/he explained her/his viewpoints (*"A1 asked A3 about lawyers because A3 told her/his smth about it before. A3 is responding to her. [...] A4 is now sharing her/his experience with it. [...] A3 said, 'No, it is not true' and was explaining her/his experience"*, OBS). This observation matches well with the conversation pattern. Like the previous workshop, when someone shared their experience, the speech was more prolonged than simple comments and questions.

3) In the second workshop, a collective discussion was also present, as shown in this pattern. According to the observations, Team A was discussing the challenges related to the cafe (*"A1 and A3 discussing some other challenges related to the cafe. [...] A4 is joining the discussion."*, .)

4) Not every communication and interaction pattern is apparent and comparable with observation. Mentors did not wear badges, so participants' badges detected their voices, therefore labeled as *"unknown"*. On the other hand, random background noise or changes in temper or pitch were also occasionally detected as *"unknown"* speaker. This pattern makes it unclear when the mentor joined the team. Therefore, observations must complement this data (*"Mentor came in, A1 happily said they*

were waiting for her.", OBS). The mentor's presence is only apparent when (s)he has longer speech segments, as shown in the second half of this pattern.

5) In some cases, patterns allow for investigating collaboration more deeply than qualitative analysis. In this pattern, an A1 discussion with a mentor can be seen. Occasionally, A3 chips in some comments. Based on the previous findings that A1 and A3 had the same idea and tended to work more as a pair, this fits the picture well. On the other hand, in the observation notes, only participant A1 was noted to be discussing with the mentor (*"A1 is asking her/his question now. The mentor asks additional questions [...] A1 is answering.*", OBS). When using video materials as a ground truth, it is apparent that A3 also said some short sentences to the mentor.

6) This pattern shows a two-sided conversation, where A4 and A2 do not participate. Compared to the observations, in immanent, A4 was listening to that conversation, while A2 was busy on her/his phone (*"A2 shortly shared an idea. Now, A3 is explaining the problem further. A4 is now listening, and A2 is still on her/his phone.*", OBS).

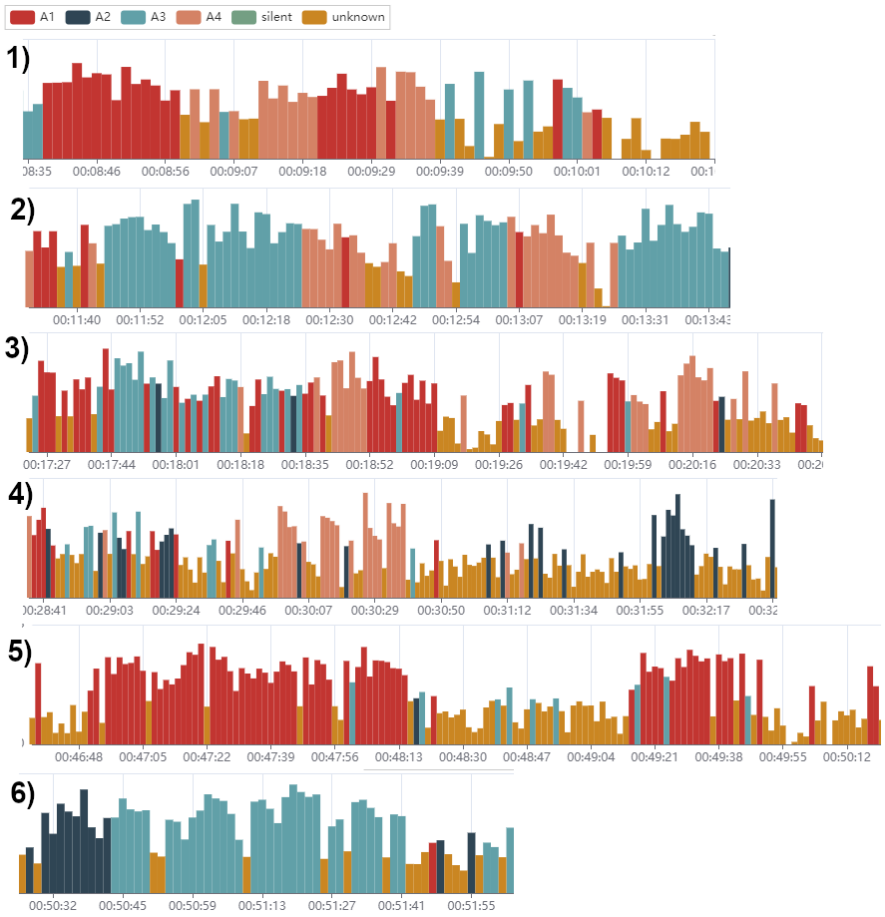

Figure 11. Workshop 2 speaker differentiation patterns

4.2.4 Comparison of Teams

As the previous sections focused primarily on Team A. The following paragraphs give insights into other teams' journeys throughout Workshop 2. In the last paragraph, their collaboration dynamics are compared in the form of a table.

Collaboration Dynamics in Team B At first, there was a collective decision on whether the team should divide the tasks. Nothing was decided when B2 initiated the acquaintance with B1, who joined the team at the end of the previous workshop. It resulted in a collective discussion around B1's idea. When the mentor walked into the room, the discussion continued as no-one did not switch to mentor. Eventually, the mentor engaged in the discussion herself and offered help. During the communication with the mentor, everyone deviated from the assignment as no one filled their matrices on paper. Again, dominant speakers were B4 and B5. At the end of the workshop, only

B5 managed to complete the assignment.

When the mentor broke into the discussion, s/he initiated the introductory round first, opening her/his background. After that, there was a feedback carousel, where the mentor switched between different idea authors, listed them, and gave feedback. Occasionally, participants asked questions and shared their experiences with mentors. Team B was in a separate room like Team A, but the mentor was with the team for almost the whole workshop. All the interviewed participants were happy with the mentor's accessibility. After the workshop, mentors had lunch with the team (*"Some mentors were sitting with us, even eating together."*, B3).

Collaboration Dynamics in Team C C1 and C3 arrived 5 minutes late. C3 continued to share the knowledge behind the colors. The mentor was used as an interpreter during the collective discussions, which C1 mostly led. S/he tried to scope the project and often alluded to what was written on the paper (*"C1 was like a leader, [...] s/he reviewed our activities."*, C3). In interviews, it turned out that C1 was a school teacher and was used to leading people. Again, similarly to the first workshop, whenever the mentor was not present translating, Team C used Google Translate, both text and voice options, to translate and communicate with C2. Two language-based sub-conversations were occasionally present: C1 and C3 discussed in Persian, while the mentor and C2 spoke in Russian. The collective discussion was done in English.

In this workshop, the mentor was like a part of a team. S/he listened to ideas and developed them further with her/his knowledge. Notably, the mentor had years of experience in web development, to which Team C's idea was related. When s/he switched between the teams, s/he constantly checked for progress and asked for potential problems – s/he was generally very supportive and friendly.

Collaboration Dynamics in Team D The workshop started with experiments to communicate despite having a language barrier. D2 shared her/his daughter's picture, followed by an off-topic discussion about family. This was not a successful attempt, and later, the team communicated only during the mentor's presence. Both participants took notes individually, but D2 appeared more distracted. They explained their work via the mentor's translations.

The mentor did not have enough time to guide the team as s/he constantly switched between two teams to translate. Occasionally, s/he asked if the team was ok. Interestingly, halfway through the workshop, D1 and D2 started looking at each other when the mentor translated their speech. Previously, they looked at the mentor. As the mentor saw major collaboration problems due to the language barrier, s/he clarified that the teams did not need to work together (*"Mentor clarifies that teams do not need to work together if they do not want to."*, OBS).

Conclusions Table 6 showcases teams' collaboration comparison in workshop 2. For instance, Team A showed strong teamwork with consistent engagement, guided by a mentor providing supportive clarification. Team B relied heavily on individual efforts, facing challenges aligning goals and limited mentor interaction. Team C encountered language barriers, relying on an interpreter for communication and project scoping. Team D faced significant language challenges, resulting in distractions in the mentor's absence.

Table 6. Collaboration comparison in Workshop 2

Aspect	Team A	Team B	Team C	Team D
Team Constellation	Consistent, teamwork, minor pairwork.	Consistent, much individual work.	Minor language-based pair-work.	Some individual work
Communication	Communication led by experienced participants. Sharing personal experiences. Question-driven. Discussions diminished when the mentor joined. Collective and pair-wise discussions.	Collective discussion.	Led by C1. Collective and pairwise discussion (language-based).	Discussion only during mentor presence (language barrier)
Goal Alignment	Immediate engagement with the assignment.	Focus on discussion, not on the actual assignment.	C1 scoping the project, often bringing focus to the assignment.	-
Interaction Dynamics	Some signs of fatigue.	Collectively planning to divide tasks. Getting acquainted with a new member.	Using Google Translate.	Off-topic discussion about family.
Mentor's Role	Assignment clarification. Guidance. Encouragement to fail.	At first, the mentor was ignored. Being present the whole workshop.	Interpreter. Discussing and developing the idea.	Interpreter. Leading discussion.
Mentor's Approach	Participant's question-driven guidance, challenging with questions.	Breaking into the discussion. Feedback Carousel.	Constantly checking for progress. Asking for potential problems.	Asking for potential problems.
Challenges and Issues	Incoordination, some disagreements, and distractions during mentorship.	Only B4 managed to finish the assignment.	Team being late. Language barrier.	Language barrier. Distractions when mentor not present (D2).

4.3 Workshop 3 – Problem-Solution Fit

The third workshop focused on identifying the Unique Value Proposition (UVP). Participants were guided in the main hall to address the questions, "What is your offer?" and

”Who is your customer?” They were tasked with crafting a concise sentence to describe both what they do and what they offer.

4.3.1 Collaboration Dynamics Analysis

This workshop was relatively short and started with A1’s confusion around wording her/his idea’s uniqueness. Everyone tried to help her/his as observed (*”A3 is trying to help her/his [A1] figure it out. A4 is listening, offered a few words as examples.”*, OBS). As A1 expressed her/his confusion for minutes, it turned into a collective discussion as 5 minutes later it was observed (*”They are brainstorming an idea together. A1 is writing down some ideas.”*, OBS).

After some discussion, A4 suggested searching for information about competitors, after which A1 and A3 started using their smartphones. As discussion died, A2 and A4 started using their phones as well, but more as a distraction (*”Everyone is on their phones, A4 is texting, A2 is scrolling, A1 and A3 are googling ideas similar to theirs.”*, OBS). Before the workshop ended, A2 and A4 shared their sentences about ideas. This communication was initiated by A2 when s/he asked A4 to share what s/he had written.

In the end, an off-topic conversation was observed (*”A2 and A3 are talking about something related to finding a space for the cafe.”*, OBS).

4.3.2 Mentorship Analysis

Due to the fact Team A was in a separate room, there was no mentorship available on a question basis, as it would have been in the main hall. In this workshop, no mentors visited the team. This was also noted in the interview with A1 (*”Well, the first day there were problems because at the end they almost did not come to us anymore, I think there was more communication with mentors on the first floor [main hall] [...] .”*, A1). Compared to the previous workshop, there were no signs from participants that could have shown the need for mentorship. This may be related to the fact that this workshop only lasted for 15 minutes.

4.3.3 mBox Analysis

In the third workshop, a Jabra microphone was tested (environment setup was detailed in Section 3.2.4 in Table 3). This resulted in unexpectedly poor results, with many speaker detection problems. The observer noted that the mentor was absent during that workshop, yet sensors labeled 37.21% of speech as *”unknown”* as shown in Figure 12. The preceding analysis shows that this label is often connected to a mentor’s speech. The possibility that the observer did not note down a mentor’s presence is excluded after viewing the video materials. Because of that, most of the audio logs and patterns derived from these are inaccurate.

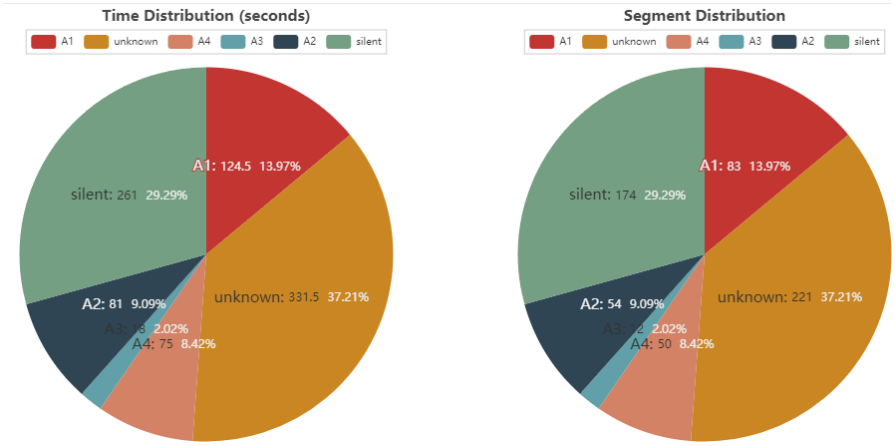

Figure 12. Workshop 3 speaker segment charts

Speaker differentiation did not work well for this workshop, so the deeper communication pattern analysis could not be done, yet there is one interesting situation. Figure 13 shows almost 6 minutes of communication patterns from that workshop. At first glance, it is evident that there are huge gaps of silence. Furthermore, only some segments of A1 and A4 speech are recorded, meaning that others were totally passive or ignored this communication. Timestamped comparison with observations shows clearly that everyone was on their phone, mostly distracted. Only A1 and A3 were somewhat focused on the assignment, as they were searching the web for similar companies (*"A2 still on her/his phone. A1 asked A3, 'Any other ideas?'. [...] A4 suggested googling it. [...] Everyone is on their phones, A4 is texting, A2 is scrolling, A1 and A3 are googling ideas similar to theirs."*, OBS).

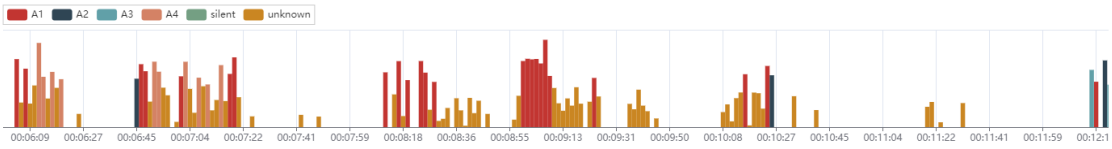

Figure 13. Workshop 3 speaker differentiation patterns

4.3.4 Comparison of Teams

As the previous sections focused primarily on Team A. The following paragraphs give insights into other teams' journeys throughout Workshop 3. In the last paragraph, their collaboration dynamics are compared in the form of a table.

Collaboration Dynamics in Team B In this workshop, B4, B5, and B1 worked as a group, while B2 was passive during the conversations and occasionally used her/his phone as a distraction. Eventually, B5 also got distracted by the phone. Only B4 took notes and worked on her/his idea. After 15 minutes, the organizer asked them to go downstairs and present the work due to the lack of presenters. B4 expressed her/his desire to go to the main hall with the whole team because other teams were seated there.

No mentors visited the team during this workshop. As other Team B participants were happy with the access to mentors, B4 brought out the lack of mentors during the day 1 (*"We did not have enough of them on day 1, and when they came [...] [we] spent less time on empty conversations.*", B4).

Collaboration Dynamics in Team C The team was again late to the workshop. When they finally gathered, C3 read out the work s/he had done in the main hall from her/his smartphone. S/he had used a note-taking app. As the description was too long, C1 handed out the paper to C3 for her/him to write with a big smile. C2 also proposed her/his idea using Google Translate, but C1 refused to do it as s/he thought it was too long. The mentor immediately defended C2 and stated that this assignment had no word count. Later, C3 again shared her/his domain knowledge with colors.

Collaboration Dynamics in Team D During this workshop, both D1 and D2 worked individually and communicated only during the mentor's presence. When the mentors were not with them, they often got disengaged from the task and just listened to another team.

Mentor constantly switched between two teams to translate. D1 asked the mentor some clarifying questions. No other guidance was observed.

Conclusions Table 7 compares the collaboration aspects of different teams during the third workshop. It shows that Team A exhibited collective discussion but struggled with focus, often affected by smartphone distractions. Team B worked as a sub-group and as individuals but faced distractions. Team C stood out for its excellent focus on assignments and often leveraged tools like Google Translate. Team D only engaged during the mentor's presence. Mentor accessibility varied, with challenges noted for Team A.

Table 7. Collaboration comparison in Workshop 3

Aspect	Team A	Team B	Team C	Team D
Team Constellation	-	Sub-group and one individual (distracted).	-	Individual work
Communication	Collective discussion. Significant gaps of silence.	-	-	Only during the mentor's presence.
Goal Alignment	Lack of focus on assignment.	One participant focused on the task.	Good focus on assinment.	-
Interaction Dynamics	Collectively brainstorming for A1. Use of smartphones. Off-topic conversation.	-	Sharing domain knowledge. Using Google Translate. Passing work (C1 to C3).	-
Mentor's Role	Not present. Poor accessibility noted.	Not present.	Interpreter. Defending opinions.	Interpreter. No guidance. Answering questions.
Mentor's Approach	-	-	-	-
Challenges and Issues	Smartphones as distractions.	Smartphones as distractions.	Being late.	Disangagned during mentor's absence.

4.4 Workshop 4 – Business Model Canvas

The fourth workshop focused on introducing the Business Model Canvas (BMC). Participants were all given large A3-format BMC templates to fill. Although the standard BMC consists of 9 main topics, due to limited time, participants were asked to focus on the following sections:

- **Customer Segments** provides a clear understanding of whom the solution creates value for and who the most important customers are.
- **Value Propositions** defines the value the solutions provide to customers, outlines which problems the solutions address, and specifies the needs the solutions satisfy.
- **Channels** assesses how the solutions reach the customer.

If participants had additional time, they were allowed to proceed with other BMC sections in a particular order.

4.4.1 Collaboration Dynamics Analysis

The fourth workshop started with A2 being absent, as s/he consulted with a different team in the main hall. Once all members were present, they started a collective discussion about BMC. They discussed the possibilities of working as a team around one idea. Still, they could not decide and eventually used rock-paper-scissors to pick an idea as observed (*"A1 suggested A2's idea, A2 suggested they do someone else's idea. A1 and A4 played rock-paper-scissors [...] ."*, OBS).

Once the idea was picked, the team used question-driven discussion towards A4 to define her/his customer segment. As they discussed different segments, they had different off-topics communication, not related to the idea, but rather personal perspective to things (*"A1 and A4 discussed who of them likes winter sea [...]"*, OBS).

When different customer segments were discussed, everyone started working on their ideas again, and A2 got distracted by a phone. Similarly to the previous workshops, there were two sub-groups: A1-A3 and A2-A4 (*"Discussion between A2 and A4 about pizza prices. [...] A1 and A3 continue talking about their customer segment."*, OBS). The discussion kept being question-driven and sharing personal experience (*"A2 asked A1 and A3 about their experience with packing lunches."*, OBS). Later, the observer also noted that the discussion was mostly about the A1-A3 idea (*"Although A4's idea was chosen [...], most of the discussion is about the A1-A3 idea."*, OBS).

Throughout the workshop, there was a positive and supportive atmosphere, as different members kept making jokes (*"A joke about someone's salary. [...] more jokes about how people eat while at work. [...] A lot of joking around today, related to the teams' ideas"*, OBS). This may have positively affected A3, who noticeably spoke more than the day before (*"A3 commented that s/he's talking a lot today, and yesterday s/he was nervous."*, OBS). A1 supported A3's comment (*"'Yes, yesterday I was the star, you are the star today'"*, OBS).

Occasionally, there was a general discussion around A2's restaurant and its pricing. Again, this discussion was question-driven and more like a sharing professional experience as noted (*"Discussion about pricing in A2's restaurant. Everyone asking follow-up questions and giving suggestions to A2."*, OBS).

Participants used smartphones to search for competitors. This was done more as individual work: searching and filling out BMC based on findings. During this search, participants did not participate in discussions and were more zoned out, as observed (*"A2 is writing down stuff while looking at her/his phone. [...] has not joined any discussion."*, OBS).

4.4.2 Mentorship Analysis

No mentors were helping Team A during this workshop. The first notes about the need for a mentor appeared roughly in the middle of the workshop once A4 expressed the

need (*"A4 said they now need to wait until someone comes and tells them what to do later."*, OBS). Almost at the end of the workshop, A2 asked the observer about missing mentors (*"A2 just asked if any mentor would come today."*, OBS).

4.4.3 mBox Analysis

In qualitative analysis, a situation arose where A3 believed s/he spoke more on the second day, a claim that A1 graciously acknowledged (*"A1 responded to that, 'Yes, yesterday I was the star, you are the star today'"*, OBS). Summing up the speech segments reveals that on day 1, A3 had 366 segments (549 seconds), while A1 had 772 segments (998 seconds). On day 2, A3 had 410 speech segments (615 seconds), while A1 had 398 segments (597 seconds). This results in A3 being 11.34% more talkative, while A1 spoke 40.18% less on the second day. There is clear evidence that A1 spoke less on the second day, but confidently confirming that A3 spoke more on the second day is challenging due to the marginal difference.

The fourth workshop's speakers' relations graph, shown in Figure 14, shows roughly similar findings as preceding workshops. No clear sub-group, dominant speaker, and any signs of ignorance. A3 spoke slightly more with A1, but this relationship is not mutual.

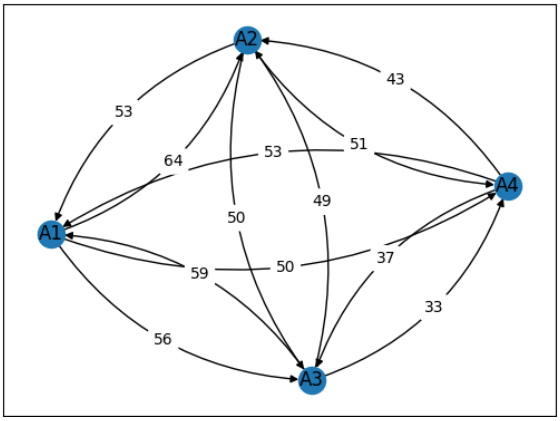

Figure 14. Workshop 4 speakers' relations graph

Time-wise, the fourth workshop was the longest and provided valuable insights into collaboration dynamics. Figure 15 shows the most significant patterns that occurred during that workshop.

1) This pattern shows a collective discussion around the A2 idea once the team decided to focus on only one idea. It matches well with the observations notes (*"A2 is showing on her/his paper what s/he wrote for her/his idea as an example for others. Everyone is listening [...] A1 suggested to A4 about searching how food is organized at races."*, .) In observation notes, A3 was never mentioned, but the pattern shows

some fragments of her/his speech. Further video material analysis showed that A3 actually chipped in some comments.

2) Not always do MMLA sensors show the reality, as there are some limitations. This pattern shows a situation where there were two simultaneous conversations. At first glance, it seems like the team was collectively discussing, but observations did not support this (*"A1 and A3 are discussing their customer segment. [...] Dicussing between A2 and A4 about pizza prices."*, OBS).

3) As previously analyzed, A3 spoke more than the previous day. It is interesting to note when A3 thought s/he had spoken more. This happened after a given conversation pattern and was first marked by an observer. A3 did manage to make everyone listen to her; when s/he shared her/his experience (*"A2 asked A1 and A3 a question about the price of their product, everyone was talking about that. A3 is talking more today than yesterday, mostly related to her/his experience with eating at work. [...] S/he [A3] talks a lot about her/his experiances"*, OBS). While not clearly obvious from the observations, there was a collective discussion around A3's experience, as other participants also had some speech segments.

4) Similar pattern to the previous one, where A2 has long speech segments. This time, it is noted that there was a collective discussion as everyone was asking follow-up questions (*"Everyone asking follow-up questions and giving suggestions to A2"*, OBS). From the conversation pattern, it is imminent that A1 did not participate in this discussion and only later is confirmed by observer (*"A1 is looking up something on her/his phone, but s/he is nodding to what A2 is telling."*, OBS).

5) This is another collective discussion around A1 and A3 pricing. Interestingly, A2 is not observed to take part in this conversation (*"A4-A1-A3 continue talking about A1/A3's pricing"*, OBS), but clearly A2 said something during this discussion. Furthermore, according to observations, A2 was on her/his phone (*"A2 listened and went on her/his phone."*, OBS). Video analysis shows that A2 was on her/his phone, but s/he spoke a little and was aware of the topic being discussed.

6) The same situation happens in this pattern, as seen in the previous one, but more towards the end of the workshop. A3 was sharing her/his experience, while A2 was noted to be on her/his phone (*"A3 is sharing her/his experiences about a pizzeria in Estonia. A1 jumped in and also commented on something. [...] A2 writes down stuff while looking at her/his phone."*, OBS). Although A2 was working individually on her/his phone and taking notes, s/he did participate in the collective discussion, as the video material analysis confirmed.

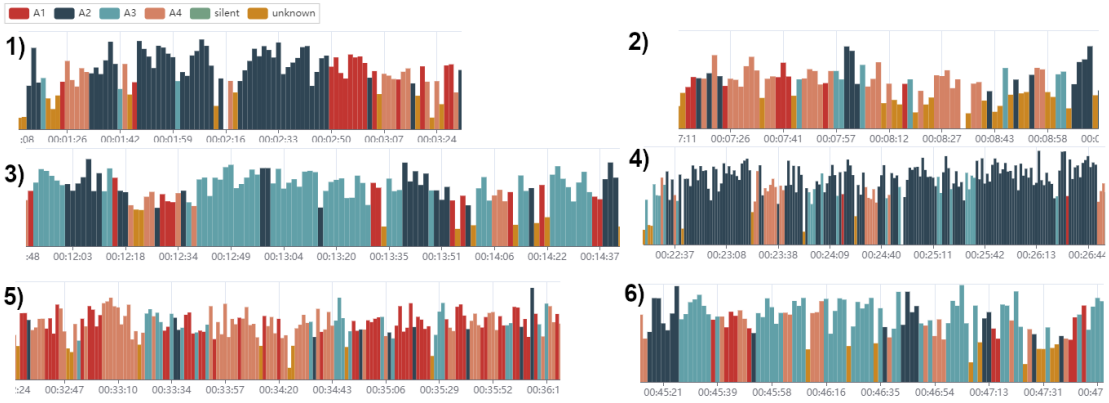

Figure 15. Workshop 4 speaker differentiation patterns

4.4.4 Comparison of Teams

As the previous sections focused primarily on Team A. The following paragraphs give insights into other teams' journeys throughout Workshop 4. In the last paragraph, their collaboration dynamics are compared in the form of a table.

Collaboration Dynamics in Team B The fourth workshop started with a new team member, B6, resulting in a team of 5. Everyone was involved in the collective discussion around B6's idea. Only B4 was occasionally distracted by her/his phone. After five minutes of acquaintance and a new idea introduction, B4 took the lead in communication and brought the discussion back to the assignment. B4 went to the main hall to take pictures of the BMC examples. Meanwhile, the discussion went back to B6's idea. When B4 returned, s/he guided others on how to fill the BMC. Eventually, the discussion died as B4 was working alone on the assignment, B1 got distracted by her/his email, and B6 was just passively sitting and listening to others. Two sub-conversations emerged as B4 asked questions from B2, while B5 and B1 discussed their idea. These conversations naturally merged into collective discussions where participants brought up potential competitors. Only B6 did not participate in this discussion. During the interviews, both B2 and B4 brought this up as a problem (*"We were already trying to avoid pitfalls that we were stepping on day one by talking in circles about the same things. New people came in and did not understand anything."*, B4). To make matters worse, the lead mentor accidentally added a new teammate, B3, to Team B, making it a team of six *(The decision behind that was not recorded)*. B2 immediately started the acquaintance round with new member B3. As with all new members, B3 had her/his idea, which s/he presented to the team. B4 was not interested in new ideas and kept working alone, using the smartphone to search for information. The collective discussion got off-topic after B6 tried to bring it back to the assignment. B4 was happy with that decision and retook the lead communicator

role as s/he explained the assignment. Two mentors entered the room, but the team kept discussing the assignment.

Mentors had to break into the discussion as no one asked anything from them. While Team A had clear signs of the need for mentorship, Team B, on the other hand, was doing well without mentors. At first, one of the mentors clarified the task. Later, mentors kept directing the conversation with their questions and narrating examples.

Collaboration Dynamics in Team C Again, participants were late to the workshop. As C3 was missing, C2 surfed the web on her/his phone. Once C3 arrived 8 minutes late, teamwork started. Initially, C1 and C3 worked as a sub-group; C2 was excluded and worked independently. C1 explained the assignment to D1, who was late but also situated in the same room. As C2 kept working alone, C1 decided to see what s/he had written. Again, Google Translate was used to communicate. It was observed that Team C wrote semantically wrong things in their BMC, showing that they did not understand the assignment.

The mentor joined in the second half of the workshop and immediately asked about the status, then s/he switched to Team D. While s/he was guiding the other team, Team C participants kept listening to the mentor's guidance. When the mentor switched back to Team C, questions about the assignment were thrown at her/him. S/he decided to show illustrations from the phone to explain BMC.

Collaboration Dynamics in Team D Interestingly, only D1 showed up to the workshop while also being late. D1 interacted with Team C to understand the assignment. S/he grabbed the BMC paper from the main hall and moved closer to Team C. Once the mentor joined her, s/he confirmed that s/he was okay with working alone. Interestingly, D1 was working on some previous assignments as s/he took notes on a different paper, not the BMC one.

Conclusions Table 8 offers insights into the collaboration dynamics comparison of the fourth workshop based on various aspects. It shows that Team A operated more as a team with minor deviations, fostering a supportive atmosphere. Team B, with new members, faced challenges of passivity and distractions but showed signs of leadership. Team C initially struggled with language barriers and later benefited from mentor intervention. In the case of Team D, only D1 was present. S/he faced challenges such as solo work and late attendance. The mentor's role varied, with Team B benefitting from directing discussion, while Team C saw the mentor more as an interpreter. For Team D, the mentor helped to onboard with the task as s/he was late to the workshop.

Table 8. Collaboration comparison in Workshop 4

Aspect	Team A	Team B	Team C	Team D
Team Constellation	Working as a team. Two sub-groups.	Two new members. Some individual work (B4)	At the start, a pair and one individualist.	D1 being solo. D2 did not show up.
Communication	Collective discussion. Some sub-group discussion. Sharing personal experiences.	B4 leading the discussions. A lot of collective discussion. Some sub-group conversations.	A lot of pair discussion (language-based). Collective discussion on mentor's presence.	-
Goal Alignment	Good alignment, minor deviations. Clarity issues at the end.	B4 bringing the focus on the task.	Wrongly written BMC	Working on previous assignments.
Interaction Dynamics	Minor off-topic discussion. Supportive and positive atmosphere. Some joking.	Getting acquainted with new members. Off-topic discussion.	Used Google Translate. Guiding Team D.	-
Mentor's Role	Not present. Explicitly asking for a mentor.	Directing discussions. Clarifying assignment.	Joined in the 2nd half. Interpreter. Clarifying assignment.	Clarifing task. Giving guidance. Encouraging to work alone.
Mentor's Approach	-	Breaking into the discussions, as no one asked questions. Directing conversations. Giving narrating examples.	Showing illustrations on phone. Checking for the need for help.	-
Challenges and Issues	Need for mentorship.	New members being passive. Distractions (smartphones).	Being late. Language barrier.	Being late. Being solo.

4.5 Workshop 5 – Roadmap of Actions

The last workshop was more generic. In that workshop, participants were allowed to continue working on their BMC. Additionally, they were required to create a roadmap of actions and formulate a one-minute pitch for their solutions. These pitches were presented after the workshop.

4.5.1 Collaboration Dynamics Analysis

Similarly to the previous workshop, A2 was late for 5 minutes to talk to a few mentors in the main hall. S/he expressed her/his disappointment that their team was in a separate room and could not talk to other people and mentors. This time, the mentor was with Team A from the beginning of the workshop and never left until the end.

While different participants were talking to mentors, others were listening or taking notes individually (*"The mentor-A2 discussion continued, A3 was listening to it, A1 and A4 are taking notes."*, OBS). This pattern continued throughout the workshop, with minor exceptions when A1 and A3 quietly discussed some things around their idea (*"A1 and A3 are meanwhile talking to each other about their idea."*, OBS).

Occasionally, participants used their smartphones with unknown purpose (*"A2 is looking up smth on her/his phone"*, OBS). At the beginning of the workshop, A2 and A4 mostly used phones, but at the end, A1 and A3 were noted to be using phones. Even the mentor was on her/his phone before that workshop ended.

A friendly atmosphere was generally noted, as the mentor made a couple of jokes while providing feedback to A2, which A2 seemed to like as s/he laughed (*"The mentor [...] made a joke, A2 laughed"*, OBS).

After the workshop, participants had a chance to present their ideas in the main hall, which subgroups of A1 and A3 did. The idea was presented only by A1, but s/he talked as a group (*"S/he talked about 'we, us'"*, OBS). As this decision was not noted during the workshop, it was later asked in the interview. A3 answered: (*"When there are a lot of strangers, I start stuttering too much, so it is not my thing."*, A3), while A1 described the decision more as experience-based (*"I am more communicative originally, and I have experience in some oratory, I have experience in, for example, TikTok."*, A1)

4.5.2 Mentorship Analysis

Mentor M2 started the session by confirming that s/he would be there to help when needed and participants should carry on what they have done so far. First, s/he gave feedback to the team's BMCs, which was the previous workshop's assignment. Then, s/he reworded and clarified the previous assignment in the main hall. After, A1 had a chance to ask for feedback from mentor (*"A1 asked the mentor about her/his idea"*, OBS).

Mentor M2 preferred to listen to ideas and give feedback; during that, s/he tried to make sure that all could talk by switching to different ideas (*"Mentor said, 'Let's switch to A1 and A3'. A1 read out what they wrote. The mentor gave feedback."*, OBS). While listening to ideas, the mentor also took notes on a piece of paper. This pattern continued throughout the workshop. Different members bounced off their ideas to mentors and listened to feedback; based on that, they made improvements and reiterated the process (*"A1 is reading out another option. The mentor is giving feedback. The mentor asked A1*

to read out the sentence again.", OBS).

4.5.3 mBox Analysis

In the fifth workshop, the MMLA sensors failed to start on time, so there was a lack of data for deep analysis. The issue was fixed 5 minutes before the workshop ended. One conversation pattern could still be extracted. Figure 16 shows a situation when the mentor was helping A2 and then switched to A1 (*"A1 and A3 are half on their phones, half making notes. The mentor is interacting with A2 to figure out what A2 should write. The mentor asked A1 and A3 to share what they wrote. A1 is reading it out. [...] The mentor is giving feedback. The mentor asked A1 to read out the sentence again.*", OBS). This pattern matches fully with the observation notes and shows that when the mentor communicated with the participants, others were listening, taking notes, or distracted by phone.

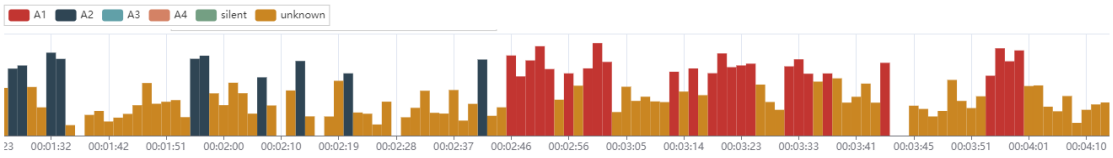

Figure 16. Workshop 5 speaker differentiation pattern

4.5.4 Comparison of Teams

As the previous sections focused primarily on Team A. The following paragraphs give insights into other teams' journeys throughout Workshop 5. In the last paragraph, their collaboration dynamics are compared in the form of a table.

Collaboration Dynamics in Team B B5 and B1 were absent for 5 minutes while others worked on the assignment. At first, everyone was working individually on the task. B6 was distracted by the phone. B1 initiated the discussion to get feedback, which B4 and B5 also gave. B2 and B3 were listening. B1 led the whole discussion, while others just asked clarifying questions. When the mentor joined a couple of minutes later, Team B offered her/his to sit down with them. When the workshop ended by the organizer, participants stated that they did not calculate the time. B2 kept discussing with the mentor while others backed things up.

Similarly to the previous workshop, no one asked questions from a mentor who listened to the discussion. Eventually, B2 initiated communication with the mentor. Then, a question-driven mentorship was observed, where participants were asked different questions from the mentor. When the questions got off-topic (*"B3 is checking with the mentor about next week's assignment"*, OBS), the mentor directed the conversation back

to the current assignment. Later, the mentor helped B4 shape her/his idea, after which B4 left the room for a few minutes.

Collaboration Dynamics in Team C Similarly to the previous workshops, participants were late to the workshop. It was imminent that they had spent time in the main hall as C2 arrived with the coffee. C2 wrote independently, while C1 and C3 worked more as a pair. The signs of decision-making and reviewing the work seen on the previous day with C1 were not noted. In this workshop, communication was mainly led by C3.

This time, the instructions in the main hall were given only in Russian, and the mentor initially clarified them to Team C. S/he showed the slides from her/his phone. As the mentor constantly switched between two teams, C1 often asked for confirmation about what s/he had written. Almost at the end, the second mentor entered the room. C3 quickly pitched the idea and explained the science behind the colors. The second mentor asked clarifying questions, was generally receptive, and agreed to the problem. S/he did not guide or give any particular feedback.

Collaboration Dynamics in Team D D1 was present from the previous workshop, and D2 also showed up. It turned out that D2 had a language exam previously. Immediately after D2 had put down her/his stuff, s/he left the room. After five minutes, s/he entered the room together with C2. Interestingly, s/he sat beside her, across from her/his team member. This situation is shown in Figure 17). There was no communication between D1 and D2 throughout the workshop, as they worked ultimately separately across the table. This sitting constellation did not promote teamwork. Only some interaction between D2 and C2 was noted when D2 asked for task clarification. Similarly to the previous workshop, D1 kept working on something unrelated to BMC, while D2 grabbed a canvas from the main hall and wrote things down there.

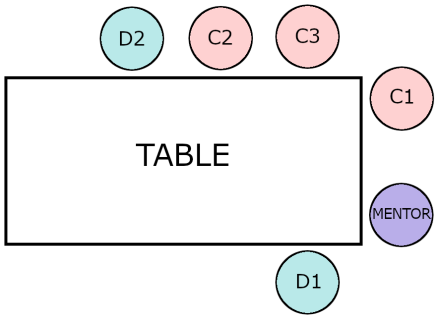

Figure 17. Workshop 5 Team D and C sitting plan

The mentor did not have to translate for Team D, as there was no communication between participants. S/he helped to clarify the assignment for both Russian-speaking participants (D2, C2).

Conclusions Tabel 9 summarizes key aspects of four teams in the fifth workshop. Team A focused on collective efforts with mentor-guided communication but faced challenges like distractions. Team B initially emphasized individual work, had mentor-driven communication, and faced time-related challenges. Team C showed a pair-work and minimal communication. Team D predominantly worked as individuals, as the seating constellation did not promote teamwork. Mentor roles varied across teams, involving feedback, guidance, and discussion management.

Table 9. Collaboration comparison in Workshop 5

Aspect	Team A	Team B	Team C	Team D
Team Constellation	Some pair-work, but mostly collective.	First, a lot of individual work.	C1 and C3 worked as a pair. C2 took individual notes.	Only individual work.
Communication	Communication was mostly led by a mentor. Some discussion in a pair.	Question-driven towards B1.	-	No team-wise communication. Some communication with Team C.
Goal Alignment	Mentor helped to keep focus.	-	-	D2 worked on previous assignments. D1 on the given one.
Interaction Dynamics	Friendly atmosphere.	-	-	No interaction. Members were sitting far away.
Mentor's Role	Being present the whole workshop. 2nd mentor joined later. Providing feedback on previous assignments.	Directing back off-topic discussion. Answering questions. Giving guidance.	Confirming written artifact.	-
Mentor's Approach	Distracted by a phone. Feedback-oriented. Incorporating humor into feedback. Directing participants to work. Iteratively developing ideas.	Question-driven mentorship, sitting with the team.	Asking clarifying questions.	-
Challenges and Issues	A2 being late. Distractions (smartphones). Disappointment about the separate room.	B5 and B1 being late. Lack of time.	Being late.	Sitting constellation.

5 Discussion

The following sections first answer proposed research questions (Sections 5.1, 5.2 and 5.3) and then show the implications for practice (Section 5.4) and organizing hackathons (Section 5.5). The last two sections discuss limitations encountered during the research (Section 5.6) and using the MMLA system (Section 5.7).

5.1 Evolution of Collaboration Dynamics – RQ1

When addressing RQ1, the research findings revealed interesting dynamics across diverse teams. Initial team formations and interactions showed various forms of collaboration styles, with some teams exhibiting a preference for individual work while others engaged more in a collective discussion. The impact of language barriers emerged as a prominent factor influencing collaboration for some teams, leading to the use of tools such as Google Translate and distinct sub-group conversations.

Over the two days, teams demonstrated adaptability in response to new challenges as new team members. The influence of experienced team members on communication dynamics became apparent, with leadership roles evolving and specific individuals taking the lead in discussions, where they guided and shared personal experiences. Despite occasional distractions and deviations from the workshop goals, positive team atmospheres persisted for all teams, characterized by humor and mutual support. The emergence of sub-groups within teams indicated complicated collaborations, with individuals leaning towards those with shared ideas or language capabilities. Distinct phases of individual work, collective discussions, and moments of disengagement highlighted the dynamic nature of collaboration throughout the hackathon.

As the hackathon progressed, teams struggled to focus on the assigned tasks. Distinctive communication patterns, such as off-topic conversations, interruptions, and the use of smartphones, reflected on the collaboration dynamics. External factors, such as room arrangements and team sizes, also influenced the teams' ability to work collectively. In the case of Team B, it became apparent that new team members who joined on the 2nd day did not onboard well. They did not contribute to the teamwork and had a minor influence over the collaboration dynamics as, most of the time, they just sat and listened.

5.2 Mentors' Role in Collaboration Dynamics – RQ2

Findings related to RQ2 showed that external actors, especially mentors, did play a crucial role in shaping collaboration dynamics. There were instances where mentors actively engaged with teams, employing question-driven approaches and providing guidance. In these examples, collaboration dynamics tended to improve as mentors managed to clarify assignments and resolve any misunderstandings teams may had. In some cases, mentors even iteratively refined the ideas and solutions of teams.

Good accessibility and continuous presence of mentors did positively influence collaboration. Many participants noted that mentors fostered a supportive and inspiring atmosphere. In some cases, it came with the cost of time, and teams did not manage to finish their tasks. The absence or limited availability of mentors hindered collaboration, leading to disengagements and unaddressed questions.

For some teams, the mentor's primary role was to be an interpreter, which also influenced collaboration in a good way in overcoming language barriers.

Collaboration dynamics were not affected by other teams as there was a lack of interactions with other teams. However, there was one instance when the participant sought help from the other team.

5.3 The Usage of MMLA Platform – RQ3

As no previous studies have utilized the mBox, an open MMLA Platform developed by the University of Copenhagen, the usability and potential aspects for future research offer valuable insights for both platform development and potential future studies. Therefore, it is essential to address RQ3 by showing different aspects of using on-person wearable MMLA sensors.

Wearable sensors enabled the quantitative analysis of speech segments, offering insights into who spoke when and how much. Comparing speech segments helped identify dominant speakers within the team. Furthermore, visualizing these segments revealed patterns indicating collective discussions and potential subgroups within the team. Longer segments were often associated with team members who shared personal experiences or explained workshop tasks. Patterns effectively visualized distractions, mainly attributed to smartphones when no one spoke for an extended period. However, according to the observations, in some cases, everyone was writing and, as a result, not actively communicating.

Audio logs facilitated the creation of speakers' relationship graphs illustrating the collaboration relations between speakers. These graphs were employed to detect sub-groups and central speakers. However, the findings from these graphs contradicted observations, as observations often led to a distorted understanding of individuals and sub-groups. On the contrary, derived graphs exhibited no signs of pair-work or speakers' ignorance.

The sensors frequently detected mentors' presence but labeled them as "unknown" speakers. The MMLA sensors encountered difficulty detecting speakers in specific workshops or picked up excessive background noise, particularly evident in the third workshop, where different microphones were used. Mentor speech segments prominently influenced team discussions and dynamics, as participants tended to listen to the mentor and, therefore, had shorter speech segments related to question-driven communication towards the mentor.

As the audio and video logs collected by the sensors are timestamped, it allows for studying day-to-day variations in participants' speech. This aspect was particularly intriguing in the case of A3, who believed s/he spoke more on the second day, while the analysis revealed only a marginal difference. Different aspects like spatial awareness could be explored further with all the timestamped insights these sensors provide.

While not extensively explored in this study due to the nature of the chosen hackathon, the mBox System demonstrated good spatial awareness. This was briefly examined in the analysis of the first workshop, where badge trajectories were represented as a 3D graph. The graph depicted participants' movements around the table, offering insights into their seating arrangements and exit times. However, this analysis did not significantly contribute to the collaboration findings due to the hackathon format. This aspect could offer more valuable insights in a hackathon setting where participants are required to move or change their positions actively.

5.4 Implications for Practise

This study explored the MMLA Platform's possibilities to study hackathons on a larger scale. Falk et al. have identified scaling up as one of the six possible research fields in their paper [5]. Additionally, Falk et al. proposed using sensor-based technologies for future work [5]. With these considerations, this research contributes to future investigations, revealing various aspects and limitations within this research gap.

As it became apparent, multiple aspects of collaboration could be researched with this technology. This thesis used collected audio logs well and derived useful communication patterns and speaker transition graphs from these logs. As this methodology proved to work well, it could be used in future research with a large-scale hackathon scaling up participants-wise. Furthermore, spatial awareness of the MMLA Platform was explored only briefly but showed promising findings. Therefore, this aspect of the MMLA Platform could also be used on a bigger scale, but the possible reliability issues must be considered.

Several workshop analyses indicated that data collected via the MMLA Platform lacks interpretation and context, a gap addressed by this thesis through qualitative data collection, mainly through observations. Future studies must explore how the MMLA Platform can be utilized and interpreted without relying on qualitative findings. Once this challenge is addressed, the technology can be employed to study various aspects on a larger scale. While this research focused on collaboration aspects, future studies may extend to other fields using this technology.

This study provided a unique perspective on collaboration dynamics, revealing that participants predominantly focused on personal goals due to the unconventional hackathon setting. This collaboration study could be replicated in a more traditional setting, such as a makethon[8], where all participants collaborate towards a team goal.

[8]A type of hackathon where participants create physical artifacts

In this scenario, spatial awareness within the MMLA Platform could be fully utilized. Additionally, due to the lack of interaction among teams in separate rooms, future investigations may explore questions such as when participants seek communication with other teams, the number of teams each participant communicates with, the average time participants spend collaborating with other teams, etc. Similar questions could be addressed about mentorships, focusing on identifying mentorship needs, determining when a team is stuck and requires assistance, optimal time mentors should spend with a team, indicators of sufficient mentorship, methods to prevent disengagement during mentorship, etc.

5.5 Implications for Hackathon Organizers

As multiple aspects of collaboration were researched, it provided numerous implications for organizing hackathons. Firstly, it is apparent that multilingual teams cannot collaborate effectively. Therefore, a dedicated mentor who serves as an interpreter must be attached to such teams. In general, these situations must be avoided. It may be even better to combine participants with different ideas who can speak the same language. The participants themselves even proposed this during the interviews.

As teams were located in separate rooms due to research-related privacy concerns, participants tended to feel isolated and often lost track of time. In some cases, signs of goal misalignment were seen. Mentors also visited these teams less frequently than those seated in the main hall. If not necessary and privacy concerns could be overcome, then large enough rooms should be found to accommodate all the teams.

From communication analysis, it turned out that the most dominant speakers usually had the most experience. While they may not lead the team, they often contribute to the success of the collaboration with their knowledge; in a sense, they are comparable to mentors. Although participants showed interest in continuing their work only with experienced members, it is advisable to organize team formations so that experienced participants are shared equally across the teams. This avoids inequality across the teams and may provide better outcomes from a hackathon.

As participants were asked to form teams around the same ideas, they still intended to work on their individual goals, with one team as an exception. This often resulted in uncoordinated teamwork and collaboration issues, as some signs of ignorance were spotted. Furthermore, mentors had to grasp multiple ideas and use the Carusell approach to treat and devote time to all participants equally. Future hackathons should promote working around particular team goals rather than individual goals. This enhances teamwork and may produce better outcomes as mentors could devote more time.

5.6 Research Approach Limitations

This study explored a medium-sized hackathon with 15 participants, noting limitations tied to its specific theme and untraditional format. It recognizes potential outcome variations when applied to hackathons of different scales, themes, and formats. Limitations include constrained research environments, with teams experiencing limited interaction and restricted access to mentorship. The use of supportive researchers introduces biases, seen in varied interview approaches and note-taking styles, despite attempts to standardize procedures. Language barriers, mainly participants speaking Russian, complicate the study, requiring translations and potentially introducing interpretational differences. The findings are presented cautiously, acknowledging constraints in the specific research setting and urging a critical perspective from readers. These limitations are detailed in the following paragraphs.

Hackathon selection may be considered one of the most influential limitations. This study only focused on one specific medium-sized hackathon, fostering 15 participants. The selected hackathon had a particular theme and had more of an untraditional style – an idea garage. The outcomes and conclusions drawn from this study may vary when applied to hackathons of diverse scales, themes, and formats. Expanding this research to encompass a broader range of events with varying settings, designs, and participant dynamics could yield additional insights and a more comprehensive understanding. Multiple teams were studied to mitigate this, and the primary focus was on the MMLA analysis to explore RQ3.

The research environment was a limiting factor for many teams, as participants noted during the interviews the lack of interaction with other teams and poor accessibility to mentorship. Furthermore, being in a separate room also cut off some information shared in the main hall during the workshops. Participants had to take photos of the slides to use during the workshops. To address this limitation, some teams got dedicated mentors, plus organizers moved along the rooms to share important information.

As this study employed supportive researchers who helped to conduct the interviews and observe teams during the workshops, it may have also introduced personal biases. Every interviewer used somewhat different wording and reordered the questions based on the interviewees' responses. To address this challenge, a general briefing was done beforehand. It turned out to be insufficient, as researchers tended to take notes with different levels of comprehensiveness. Some observations focused more on the speech, while others concentrated on higher-level collaboration. Not only did supportive researchers introduce personal bias, but also analyzing the results *(e.g., developing an Affinity Diagram without group)* and providing the findings may have instituted interpreter bias. Therefore, all the findings were discussed and illustrated by participants' perceptions in the case of interviews and by notes from observers in the case of observations.

Additionally, it turned out that language was an influential limiting factor. Most participants spoke Russian, so the whole event was organized in that language. This

meant that all the materials *(e.g., consent forms, interview questions, etc.)* had to be translated from English. A professional paid language office did this to mitigate the risks of linguistic inconsistency. Furthermore, as most of the teams spoke Russian/Ukranian, this study had a limited pool of researchers who could understand and observe such teams. More than half of the interviews were also conducted in Russian, meaning the notes and transcriptions had to be translated back to English. This, on the other hand, may have introduced different interpretations. To mitigate this, interviews were translated by the same researchers who conducted them. Even the teams who spoke English had trouble answering questions during the interviews. Therefore, the interviewer reworded and simplified the questions according to specific needs. This could be considered an organizational limitation, as this study did not influence the team's creations.

The applicability of findings is constrained to this specific research setting, and variations in hackathon scale, theme, or design may yield unexplored and potentially conflicting results, limiting generalizability. Despite efforts to mitigate interpreting bias in qualitative data analysis, some level of bias may persist. Because of these limitations, readers should look at this study and its findings cautiously and critically.

5.7 mBox Limitations

This foundational study serves as a cornerstone for MMLA's continued evolution [6, 7] to unlock the full potential of this technology. When working with these sensors and their base stations, some apparent limiting factors appear, which may have affected the outcomes of this research. In the following paragraphs, the apparent ones are brought up.

The most significant limiting factor was that the developed mBox system [7] could not detect speakers simultaneously. This means that sub-conversations were not detectable in quantitative data as the loudest speaker (with the highest confidence value) was recorded at any given time. Whenever a participant tended to speak quietly, and any other participant spoke simultaneously, her/his speech was not recorded. After the hackathon, this factor was soon detected and paid attention to. When writing this, a research group from the University of Copenhagen has already addressed this issue and revisited the MMLA system, which can now detect multiple speakers simultaneously. To mitigate this, speaker patterns were always compared to the observations during analysis, and raw video and audio recordings were watched whenever there was doubt about multiple simultaneous conversations.

In workshop 3, this research tried out a Jabra conference microphone instead of decentralized badge microphones. This resulted in poor speaker detection as 37.21% of speech was unknown. Not only did the system not detect speakers, but it also tended to mix up some speakers, giving out false information. Plus, any background noise disturbed the results even further. Overcoming this limitation was difficult. Thus, the mBox analysis for the given workshop was not detailed so as not to provide unreliable results.

Mentor detection via badges must be addressed in future studies. This research implied observations and badges to detect whenever a mentor joined the team, as following the conversation patterns does not give this information. Although the mentor's speech was nicely recorded as *"unknown"* speaker, it was not always clear whether this was a real mentor, some background noise, or, in the case of Workshop 3, an undetectable participant's speech. Therefore, in future studies, the detection parameters must be tuned even further, or mentors should get a separate audio badge, as the current solution is not independent and cannot be deployed without observations. This study mitigated this factor by comparing observations to map the mentor's presence precisely. In case of doubts, raw video and audio recordings were analyzed.

As this research was a cornerstone for future studies with this technology, it became apparent that the current mBox solution [7] is challenging to scale. Badges are too large and bulky to be carried and deployed on masses. Furthermore, with each team, the current approach requires at least one researcher who would be responsible for setting up the base stations, giving out and collecting the badges, and fixing any issues on the fly. As this technology is still in the prototype phase [7], scalability and robustness are being addressed by a team from the University of Copenhagen in future revisions. In this study, badges were only given to one particular team, who only had to wear these badges during the workshops. A research assistant was always present near the team, monitoring the data collection and addressing issues whenever needed.

As seen, there were many limiting factors when using this technology. Some of these may have significantly impacted the collected data *(e.g., issues with simultaneous speaker detection)*. Although these sensors helped to reduce the bias and human factors in this research, it is apparent that new issues arose. Therefore, readers should exercise discernment and consider the findings critically, as they may not represent an absolute or definitive truth.

6 Conclusion

This study explored collaboration dynamics in a hackathon, where participants were assigned to work on business models in teams. It employed conventional data collection instruments, such as observations and interviews, and combined the results with findings from the Multimodal Learning Analytics Platform's sensors. These sensors were developed by a team from the University of Copenhagen. Four teams were studied with qualitative instruments; one wore additional MMLA badges and sensors. Qualitative data was encoded using affinity diagrams and then juxtaposed with quantitative findings, such as audio and video logs. The same logs were also seated into illustrative conversation patterns and graphs, further deepening the findings. The main results are derived from one team on which all instruments were applied. Although other teams were also researched thoroughly using conventional methods, the findings from these teams are provided more conclusively.

Exploring collaboration among diverse teams revealed different teamwork styles influenced by language barriers and individual goal alignment. Leadership roles evolved around experienced members, who tended to lead discussions. Positive team atmospheres prevailed despite occasional distractions, marked by humor and mutual support. As the hackathon progressed, teams faced challenges maintaining focus on assigned tasks, which was evident in communication patterns such as off-topic conversations and smartphone use.

External actors, particularly mentors, played a crucial role in shaping collaboration dynamics, providing guidance, and clarifying assignments. The continuous presence of mentors positively impacted collaboration, fostering a supportive atmosphere, though at times hindering task completion. Mentors were often described as interpreters because they helped overcome the language barrier for some teams. Limited interaction with other teams had minimal impact on collaboration, except for one instance of seeking help. However, in interviews, isolation from other teams was primarily brought out.

The usage of the MMLA Platform showcased numerous potential applications for this type of technology. Speaker segments were employed to analyze dominant speakers, distractions, sub-groups, and collective discussions. Additionally, longer segments were linked to specific situations, mainly when individuals shared their stories or experiences. The audio logs, from which speaker segments were derived, proved to be easily usable and, through a simple script, were transformed into speakers' relationship graphs. These graphs offered further insights into sub-groups and central speakers. Interestingly, these findings often contradicted the observation notes, as no clear sub-groups were identified. Labeling external actors as *"unknown"* facilitated the detection of mentor presence with high time accuracy, providing intriguing insights into collaboration dynamics, where participants tended to speak less and listen more. Given that the audio and video logs are timestamped, it enabled exploring collaboration and communication changes across workshop-to-workshop variations. Furthermore, certain aspects, such as spatial

awareness, were only briefly touched upon in one of the workshops. However, even this limited exploration yielded believable results and could be used in future studies.

As this thesis serves as a foundational piece for future studies utilizing this technology, several areas and fields could be further explored. Firstly, the MMLA Platform provides an opportunity to delve into various aspects of collaboration and communication. For example, future studies could investigate participants' movements during the hackathon, a possibility briefly mentioned in Section 4.1.3. Communication patterns and speech transcription also offer a rich avenue for researching communication dynamics. This study only scratched the surface of pattern analysis. For instance, the same patterns could be fed to machine learning algorithms to extract valuable real-time data for organizers, such as when the team is sitting too quietly and needs mentorship. Notably, since the time of writing this paper, a team from the University of Copenhagen has successfully upgraded sensors to detect different speakers simultaneously. This advancement opens the door to more in-depth analyses of sub-groups, individual work, interaction with mentors, question-driven communication, and more.

The primary objective of this thesis, namely to explore the possibilities of MMLA technology, underscores the need for replication on a larger scale. Evidently, the exploration should be extended to a broader context to fully leverage this technology's potential.

Acknowledgments

I had many supporters without whom this thesis would not have been finished. In this section, I acknowledge the most important people related to this study who helped me to collect data or provided valuable research insights.

Firstly, I would like to thank my supervisor, **Alexander Nolte**, for proposing this idea and selecting me to work on it. It has been a fantastic journey where I met interesting people and learned many new skills. Plus, big thanks for guiding me through the process and always being present whenever I needed help.

Special acknowledgment to **Daniel Spikol** and his team from the University of Copenhagen, creators of the MMLA platform, who deployed this system during the hackathon. Their generosity in sharing the collected data has significantly enriched the depth of my research. I would also like to thank **Zaibei Li** for providing valuable insights in interpreting some of the quantitative findings.

A pivotal part of the hackathon phase was the remarkable team of observers. I sincerely thank **Viktoriia Abakumova** and **Kateryna Kubrak** for observing foreign language-speaking teams and conducting insightful interviews with them.

I express my deepest gratitude to **Garage48's team** for organizing such a fantastic event and allowing me to study collaboration and mentorship aspects during it. I admired how they operated like a close-knit family, fostering a positive and supportive work environment for themself and the participants.

Lastly, I would like to thank all eager participants for participating in this research. And those who were not chosen but still gave consent to be studied. I was amazed by the number of consents I got.

Karl Rapur

References

[1] Jeanette Falk Olesen and Kim Halskov. "10 Years of Research With and On Hackathons". In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. DIS '20. Eindhoven, Netherlands: Association for Computing Machinery, 2020, pp. 1073–1088. ISBN: 9781450369749. DOI: 10.1145/3357236.3395543.

[2] Alexander Nolte, Ei Pa Pa Pe-Than, Abasi-amefon Obot Affia, Chalalai Chaihirunkarn, Anna Filippova, Arun Kalyanasundaram, Maria Angelica Medina Angarita, Erik Trainer, and James D. Herbsleb. "How to organize a hackathon – A planning kit". In: *CoRR* abs/2008.08025 (2020).

[3] Joann Keyton. "Collaboration". In: *The International Encyclopedia of Organizational Communication*. Mar. 2017. ISBN: 9781118955604. DOI: 10.1002/9781118955567.wbieoc026.

[4] Claire Scoular, Daniel Duckworth, Jonathan Heard, and Dara Ramalingam. *Collaboration: Definition and Structure*. Australian Council for Educational Research, 2020. URL: https://research.acer.edu.au/ar_misc/39/.

[5] Jeanette Falk, Alexander Nolte, Daniela Huppenkothen, Marion Weinzierl, Kiev Gama, Daniel Spikol, Erik Tollerud, Neil Chue Hong, Ines Knäpper, and Linda Bailey Hayden. "The Future of Hackathon Research and Practice". In: *ArXiv* abs/2211.08963 (2022).

[6] Paulo Blikstein and Marcelo Worsley. "Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks". In: *Journal of Learning Analytics* 3.2 (Sept. 2016), pp. 220–238. DOI: 10.18608/jla.2016.32.11. URL: https://learning-analytics.info/index.php/JLA/article/view/4383.

[7] Zaibei Li, Martin Thoft Jensen, Alexander Nolte, and Daniel Spikol. "Field report for mBox: Designing an Open MMLA Platform". In: *LAK24: 14th International Learning Analytics and Knowledge Conference*. 2024.

[8] Craig Plain. "Build an affinity for KJ method". In: *Quality Progress* 40.3 (2007), p. 88.

[9] Alexander Nolte, Linda Bailey Hayden, and James D. Herbsleb. "How to Support Newcomers in Scientific Hackathons - An Action Research Study on Expert Mentoring". In: *Proceedings of the ACM on Human-Computer Interactio* 4.CSCW1 (May 2020). DOI: 10.1145/3392830.

[10] Alexander Nolte, Ei Pa Pa Pe-Than, Anna Filippova, Christian Bird, Steve Scallen, and James D. Herbsleb. "You Hacked and Now What? - Exploring Outcomes of a Corporate Hackathon". In: *Proceedings of the ACM on Human-Computer Interactio* 2.CSCW (Nov. 2018). DOI: 10.1145/3274398.

[11] Joanna F. DeFranco, Christopher S. Eagle, James Bret Michael, John Viega, and Jeffrey Voas. "Hackathons 101". In: *Computer* 54.5 (2021), pp. 65–69. DOI: 10.1109/MC.2021.3064435.

[12] Gerard Briscoe and Catherine Mulligan. *Digital Innovation: The Hackathon Phenomenon*. London, UK: Queen Mary University, 2014.

[13] Maciej Rys. "Invention Development. The Hackathon Method". In: *Knowledge Management Research & Practice* 21.3 (2023), pp. 499–511. DOI: 10.1080/14778238.2021.1911607.

[14] *Guides — DevPost*. Accessed on May 15, 2023. URL: https://info.devpost.com/guides.

[15] Lydia I. Marek, Donna-Jean P. Brock, and Jyoti Savla. "Evaluating Collaboration for Effectiveness: Conceptualization and Measurement". In: *American Journal of Evaluation* 36.1 (2015), pp. 67–85. DOI: 10.1177/1098214014531068.

[16] Wendy Mendes, Albert Richard, Tähe-Kai Tillo, Gustavo Pinto, Kiev Gama, and Alexander Nolte. "Socio-technical constraints and affordances of virtual collaboration – A study of four online hackathons". In: *Proceedings of the ACM on Human-Computer Interaction* 6 (2022), pp. 1–32.

[17] Tähe-Kai Tillo. "Does It Make Sense to Hack Online? – A Multiple Case Study on Team Collaboration in Remote Hackathons". MA thesis. University of Tartu, 2021.

[18] Erik H. Trainer, Arun Kalyanasundaram, Chalalai Chaihirunkarn, and James D. Herbsleb. "How to Hackathon: Socio-Technical Tradeoffs in Brief, Intensive Collocation". In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 1118–1130. ISBN: 9781450335928. DOI: 10.1145/2818048.2819946.

[19] Ei Pa Pa Pe-Than, Alexander Nolte, Anna Filippova, Christian Bird, Steve Scallen, and James D. Herbsleb. "Corporate hackathons, how and why? A multiple case study of motivation, projects proposal and selection, goal setting, coordination, and outcomes". In: *Human–Computer Interaction* 37 (2020), pp. 281–313. URL: https://api.semanticscholar.org/CorpusID:221060392.

[20] Alar Leemet. "Hackathons for Corporate Innovation – a Longitudinal Study". MA thesis. University of Tartu, 2020.

[21] Cleo Schulten, Alexander Nolte, Daniel Spikol, and Irene-Angelica Chounta. "How do participants collaborate during an online hackathon? An empirical, quantitative study of communication traces". In: *Frontiers in Computer Science* 4 (Sept. 2022). DOI: 10.3389/fcomp.2022.983164.

[22] *Major League Hacking*. Accessed on June 2, 2023. URL: https://mlh.io/.

[23] *About us — Garage 48*. Accessed on November 7, 2023. URL: https://garage48.org/about.

[24] *Empowering Women Idea Garage — Garage 48*. Accessed on November 8, 2023. URL: https://garage48.org/events/the-empowering-women-idea-garage.

[25] Per Runeson and Martin Höst. "Guidelines for conducting and reporting case study research in software engineering". In: *Empirical Software Engineering* 14 (2009), pp. 131–164. URL: https://api.semanticscholar.org/CorpusID:207144526.

[26] Bruce A. Reinig. "Toward an Understanding of Satisfaction with the Process and Outcomes of Teamwork". In: *Journal of Management Information Systems* 19.4 (2003), pp. 65–83. DOI: 10.1080/07421222.2003.11045750.

[27] John Sawyer. "Goal and Process Clarity: Specification of Multiple Constructs of Role Ambiguity and a Structural Equation Model of Their Antecedents and Consequences". In: *Journal of Applied Psychology* 77 (Apr. 1992), pp. 130–142. DOI: 10.1037/0021-9010.77.2.130.

[28] Anna Filippova, Erik Trainer, and James D. Herbsleb. "From Diversity by Numbers to Diversity as Process: Supporting Inclusiveness in Software Development Teams with Brainstorming". In: *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*. 2017, pp. 152–163. DOI: 10.1109/ICSE.2017.22.

[29] Tomoko Nemoto and Temple. "Developing Likert-Scale Questionnaires". In: *JALT2013 Conference Proceedings*. 2014.

[30] Edwin Olson. "AprilTag: A robust and flexible visual fiducial system". In: *2011 IEEE International Conference on Robotics and Automation*. 2011, pp. 3400–3407. DOI: 10.1109/ICRA.2011.5979561.

[31] Keith Taber. "The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education". In: *Research in Science Education* 48 (Dec. 2018), pp. 1–24. DOI: 10.1007/s11165-016-9602-2.

Appendices

I Consent Form

Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods

Consent to Act as a Participant in a Research Study

Dear Hackathon participant

We appreciate your interest in participating in our upcoming hackathon event. Before you proceed, we kindly request your consent to participate in this research study. This form outlines the purpose, procedures, and any potential risks involved. Your participation is entirely voluntary, and your input will contribute to valuable insights for our research.

Study title: Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods

Principal Investigator: Alexander Nolte

Co-Investigators: Karl Rapur, Daniel Spikol

Introduction: Our understanding of how small groups work together during hackathon events has been based on observations, interviews, and surveys. These research tools provide only a partial picture of what happens at these events and are difficult to scale up to cover multiple groups. Our research aims to gain a new understanding of how people cooperate and collaborate in these hackathon events by combining qualitative methods with new quantitative tools. These new tools use digital sensors to capture human interactions; for instance, people's movement around the spaces, who they work with, and patterns of conversations. These devices take the form of intelligent badges that people wear that use technology to recognise other badges and capture with whom and where people spend time working. Additionally, the conversation patterns, such as the amount of time people speak. These new devices use onboard technologies with real-time data analysis to capture the patterns, not the people. By combining the qualitative methods of observations, interviews, and surveys with these new technologies, we aim to capture and gain an understanding of how people work together to create innovation in these events.

Content of the study: This study is conducted by professional researchers from the University of Tartu and the University of Copenhagen. We will ask you to fill out several surveys, we may recruit you for interviews, and we will observe the event with trained researchers. You may be recruited to wear smart badges, which capture data only in selected work areas. Plus, we will capture video and audio for the group work activities to build these systems and ensure the devices and data analysis are accurate. We will use these recordings for ground truth for the systems and the data analysis.

Financing: The research is financed by Garage48 via the grant “Study of hackathons and other open-innovation grassroot-events” (LLTAT23087)

Participation requirements: Anyone 18 or older participating in the “Empowering Women Estonia 2023” hackathon organised by Garage48.

The expected duration of the study: You will be observed during the hackathon, and after that, you will be interviewed for about 45-60 minutes with a follow-up online survey. Interviews are expected to take place at most a week.

Risks and Benefits: The risks associated with this research are no greater than those ordinarily encountered in daily life. There are no direct benefits to participants, but the researchers anticipate potential societal benefits being derived from their research.

Privacy and Confidentiality: The research project follows European laws (GDPR) and the national rules in Estonia and Denmark for data handling of information and participants' rights. The investigators will follow the following procedure to protect participants' identities during this study: The original audio and video files will be only accessible to the Principal and Co-Investigators. The audio and video files will be processed, potential identifiers will be removed or aggregated, and the original files will be deleted after analysis. All the files will be stored in the University of Tartu's servers and won't be shared with third parties nor sent to foreign countries.

Your data and consent form will be kept separate. Your consent form will be stored securely and will not be disclosed to third parties.

By participating, you understand and agree that the data and information gathered during this study may be used by the University of Tartu for publication purposes. However, any identifiable information will not be mentioned in any such publication or dissemination of the research data and/or results. All research records will be preserved until the end of the year 2023. Aggregated data will be archived by the Principal Investigator.

Questions about the Study: If you have any questions, comments, or concerns about the study either before, during, or after participation, please contact Principal Investigator Alexander Nolte via email alexander.nolte@ut.ee.

If you have any questions about your rights and ethics of research, feel free to contact the Research Ethics Committee of the University of Tartu via email eetikakomitee@ut.ee or telephone at +372 737 6215.

For data protection-related inquiries, please contact the Data Protection Inspectorate at phone number +372 5620 2341.

Voluntary Participation: Your participation in this research is voluntary. You may discontinue participation at any time during the research activity. This does not include the data analysis phase. Your decision regarding whether to participate in this study will not result in any loss of benefits to which you are otherwise entitled.

Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods

Consent to Act as a Participant in a Research Study

Informed Consent Form

I, _____, have been informed about the above-mentioned study, and I am aware of the purpose and methodology of the research and its risks.

☐ I confirm my consent to participate in the study.

☐ I confirm my consent for the processing of my personal data.

I know that additional information regarding any questions arising during the study will be provided to me by _____.

Signature of the participant: _____

Date: _____

Name of the person providing information to the participant: _____

Signature of the person providing information to the participant: _____

Date: _____

II Observation Guidelines

Observation Guide

The table below presents a concise one-page observation guide for the "Empowering Women Ukraine 2023" event. The primary objective of this observation is to gain insights into the various forms and strategies of team collaboration employed by participants. Additionally, the observation aims to evaluate the impact and influence of external actors, including other teams and mentors, on project development and team dynamics throughout the event. This observation guide will assist in comprehending the collaboration dynamics and the role of external support in achieving successful outcomes during the hackathon.

Feel free to use any note-taking tool, but don't use online tools *(e.g. Google Docs)* to avoid possible data leaks.

NB! Please timestamp your observation notes.

On Linux, it's recommended to bind a BASH script to a hotkey. This works in most text editors. For that, please follow the tutorial:
https://softhints.com/linux-mint-hotkey-paste-insert-timestamp/

On Mac, it's recommended to use the Automator script. It works with any text editor. Please follow the tutorial for Automator: https://explog.in/notes/inserttimestamp.html

On Windows, the recommended setup is Notpad, which has F5 as a timestamp shortcut.

Timestamping allows us to analyse how team collaboration changes during the hackathon. For example, each time you note down task allocation by timestamp, we can see how task allocation changes over the course of the event. For us, it's a valuable by-product.

You don't need to categorise your notes. They can simply follow a timeline as you take them.

Category	Description	Observer Notes
Task Division & Prioritization	Assess the processes, strategies, and considerations involved in decision-making and the distribution of work responsibilities.	• **How are skillsets acknowledged?** *(e.g. observables introduce themselves and their skills)* • **How are tasks divided?** *(e.g. self-pointed, delegated by sb, collectively divided, skill-based)* • **How are tasks prioritised?** *(e.g. time-based, mentor recommendations, simplicity)* • **What specific roles are assigned?** *(e.g. team leader, workshop reflector, info searcher)*
Team Constellation	Understand how and when subteams emerge and when they sync.	• **How do the teams split up?** *(e.g. task-based, natural, decision-based)* • **How big are the emerged sub-groups?** *(e.g. 2 + 3 members, individual workers)* • **How does the team sync up their work?** *(e.g. discuss the individual findings)*
Leadership	Understand how the leader is appointed and his/her role in team success.	• **How does the *de facto* leader arise?** *(e.g. emerge, self-selection, election, being talkative)* • **How does the leader manage team collaboration?** *(e.g. divides tasks, asks mentors, guides, motivates, encourages, does follow-ups)*
Collaboration & Communication	Understand how people work together towards a common goal and how information is shared within the observed group.	• **How are decisions made?** *(e.g. collectively, delegated by leader)* • **What off-task communication happens?** *(e.g. family, work, lunch, sport)* • **How does the team encourage and support each other?** *(e.g. sharing food, verbal encouragement, hugging)* • **What communication forms are used?** *(e.g. verbal, visual, body language, digital)* • **What side activities happen?** *(e.g. eating, visiting others, singing, browsing web)*
Challenges	Assess disagreements, confusion, trust issues, diversity and inclusion problems, lack of skills, materials, etc.	• **What conflicts and challenges appear?** *(e.g. lack of skill, internal conflict, verbal dispute)* • **What was the situation during which conflicts arose?** *(e.g. team was working in subgroups, mentor was guiding, leader was out of the room)* • **How are conflicts handled?** *(e.g. via mentorship, via leadership, collectively solved)*
Mentorship Influence	Acknowledge mentorship's influence over the team dynamics and potential outcome.	• **Why does the team seek guidance?** *(e.g. collective decision, feeling stuck, having questions)* • **How does a mentor approach?** *(e.g.voluntarily, by asking)* • **How does a mentor communicate?** (e.g. speaking, drawing, body language), • **What help do mentors give?** *(e.g. team coordination, project scope, business)* • **How does the team react to mentor feedback?** *(e.g. redefine scope, divide work, ignore, ask extra)*
External Influence	Understand other teams' and individuals' influence over the team dynamics.	• **How does the team communicate with other teams?** *(e.g. seeking for help, off-topic conversation, status update)*
Hackathon Setting	Assess the hackathon setting and facility feasibility.	• **How does the team navigate in space?** *(e.g. sb gets lost)* • **How does the hackathon agenda affect teamwork?** *(e.g. focusing on reflections, not listening others, having too little time)*

III Interview Guidelines

Interview Guide

The following interview questions explore teams' collaboration patterns, leadership and influence of external actors. The questions are aligned with the research objectives. The interview is meant to take **45-60 minutes**. Feel free to question and take notes in any suitable language.

NB! Please record the audio with the most suitable software for you. Bear in mind that both the interviewer's and the respondent's voices should be recorded. Screen recording is not necessary.

Some of the recommendations for software are Audacity (https://www.audacityteam.org) and OBS Studio (https://obsproject.com/). They work on Mac, Windows and Linux.

Guidelines for a successful interview

1. Welcome the interviewee
2. Thank him/her for taking part in the interview
3. Introduce yourself and the purpose of the interview.
4. Asks for permission to record. *(If not given, explain it may take longer since you have to take more comprehensive notes.)*
5. Conduct the interview by following the given questions.
6. Ask to fill out the survey **(5-10 minutes)**
    - LINK WILL BE HERE
7. Thank him/her for his/her input.

Interview Questions

Team collaboration

1. Did you know any of your team members before the hackathon?
    - How did you know them?
2. What was done with the project before this hackathon?
    - Did you participate in the prior work? How?
3. How did you start with the project?
4. How did you decide what to do?
5. How were tasks assigned among team members?
6. What are you particularly proud of in your team collaboration?
7. Were there any challenges where teamwork could have been improved?
    - How were these challenges resolved?
    - *If not mentioned, ask about conflicts and challenges you observed.*
8. How would you approach teamwork differently in a similar project?

Leadership

1. How would you characterise the leadership within your team?
2. Did you ever feel a lack of leadership during the hackathon?
    - Could you bring out one specific instance?
3. Could you name a specific leader within your team? You are allowed to mention yourself as a leader.
   Mentions someone else
    - What were the key strengths and weaknesses of him/her?

   Mentions himself/herself
    - Why did you decide to become a leader?
    - Do you have any previous background in leading people? Where?
    - Who else could have been a good leader besides you?
4. How would you improve your team's leadership in future hackathons?

Individual

1. What do you feel you contributed? What was your role?
    - Could you name some specific things that you did individually?
    - Could you name some specific skills you put into practice?
2. How did your personal goals or objectives align with your teams?
3. Did you ever feel like you were not being listened to or your opinions weren't considered?
    - Could you bring out any specific instance of this?
4. What did you learn from the hackathon?
5. Looking back, what would you do differently in future hackathons to maximise your learning?

Mentorship

1. How would you describe the role of mentors during the hackathon?
2. Could you share a specific instance where a mentor played a crucial role in helping you overcome a challenge or providing valuable insights?
3. Did you find mentors' guidance helpful? Please elaborate.
4. Were there any challenges or limitations in interacting with mentors during the hackathon?
    - How could these be addressed in future events?

Other teams

1. Did you interact with other teams during the hackathon?
    - Did you receive any assistance from them? How did they help your team?
2. Did you assist other teams? How did you help them?

Hackathon setting

1. Was there a workshop where you felt you had too much or too little time?
    - How did it influence the teamwork?
2. Did you ever feel your team had too few or too many members?
    - Could you bring out some specific instances?
    - How did it influence the teamwork?
3. Did you keep working outside the workshops (e.g. during the reflections or lunch)?
    - What was the reasoning behind it?
4. How did you decide who represents the team during the workshop reflections?
5. How did the facility help your team's collaboration? Was everything easily accessible, or did you feel lost?

IV Survey Guidelines

Post-Survey Questions

As an attendee of the Empowering Women Idea Garage hackathon from September 9-10, 2023 (https://garage48.org/events/the-empowering-women-idea-garage), you are invited to complete a survey regarding your experience and opinions about this event! The survey aims to gain insights into the overall participant experience during the hackathon, focusing on teamwork, leadership, learning, mentorship, interactions with other teams, and the hackathon setting. The questions are mostly multiple choice / agree-disagree and will take approximately **10 minutes** to answer. The data collected will be used to improve future hackathon events and tailor them to participants' needs and expectations.

Teamwork		
What is your full name?	Free text	
Would you describe your team processes as more:	• (1) Inefficient to (5) Efficient • (1) Uncoordinated to (5) Coordinated • (1) Unfair to (5) Fair • (1) Confusing to (5) Easy to understand	*(Semantic differential scale - Satisfaction with Process [5])*
Please indicate your level of agreement with the following statements related to TEAMWORK	• I always felt free to voice my comments during the hackathon • Everyone had a chance to express their opinions • I had a language barrier while expressing my opinions • The team members participated very actively during the collaboration • The team climate was warm and positive • The team members trusted and supported one another	*(Likert scale - Perceived participation [5])* • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree
Please indicate your level of agreement with the following statements related to GOAL CLARITY	• I was uncertain of my duties and responsibilities in this team. • I was unclear about the goals and objectives for my work in this team. • I was still determining how my work related to the overall objectives of my team	*(Likert scale - Goal Clarity [5])* • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree

Leadership		
Did your team have a leader?	• Yes, I was the team leader • Yes, there was a team leader. • No, there wasn't a clear leader.	*(If yes, then open the following question.)*
Please indicate your level of agreement with the following statements related to LEADERSHIP	• The team had a clear and effective leader • The team leader effectively communicated the project's goals and objectives. • The team leader effectively motivated and inspired team members. • The team leader played a crucial role in dividing tasks. • The team leader effectively managed conflicts or disagreements within the team.	*(Likert scale - Perceived participation [5])* • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree
Individual		
Please indicate your level of agreement with the following statements related to LEARNING	• I felt that my skills were adequately utilised during the hackathon. • My problem-solving abilities improved as a result of participating in the hackathon. • I was able to learn from the experiences and expertise of other participants. • I received constructive feedback on our hackathon project, which I found helpful. • I found the overall hackathon experience to be valuable for my personal and professional development.	*(Likert scale - Perceived participation [5])* • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree
External actors		
Please indicate your level of agreement with the following statements related to MENTORSHIP	• We had access to mentors during the hackathon. • I felt comfortable asking questions and seeking help from mentors during the hackathon. • Mentors helped us to overcome challenges and roadblocks. • Mentors supported us in scoping our project. • Mentors showed interest in us beyond the project we were working on. • Mentors encouraged us to think critically and approach problems from different angles • Mentors were professionals in their field of industry related to the hackathon.	*(Likert scale - Perceived participation [5])* • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree
Please indicate your level of agreement with	• Other teams helped us in our efforts. • I helped other teams in their work.	*(Likert scale - Perceived participation [5])* • Strongly disagree

the following statements related to OTHER TEAMS	• I had off-topic conversations with other teams.	• Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree
Hackathon setting		
Please indicate your level of agreement with the following statements related to HACKATHON SETTING	• The hackathon venue was easily accessible and navigable. • The hackathon venue provided sufficient space for all participants. • The hackathon organisers effectively communicated important updates and information throughout the event. • The hackathon agenda was well-planned and thoughtfully arranged. • There were opportunities for networking and building professional connections.	(*Likert scale* - Perceived participation [5]) • Strongly disagree • Somewhat disagree • Neither agree nor disagree • Somewhat agree • Strongly agree

V Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Karl Rapur**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce for preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Analysis of Collaboration Dynamics in Hackathons: Combining Quantitative and Qualitative Methods,

   supervised by Alexander Nolte.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe on other persons' intellectual property rights or rights arising from the personal data protection legislation.

Karl Rapur
January 3, 2024