

UNIVERSITY OF TWENTE.

# Making sense together

Multimodal learning analytics in a STEAM collaborative learning environment: An introduction to a proof-of-concept study

## Master's Thesis

Faculty of Electrical Engineering,  
Mathematics and Computer Science  
Masters in Interaction Technology  
Specialization in Intelligent Systems

Author: T. Laribi  
Academic supervisors: dr. IR. R.W. Van  
Delden, dr. M. Poel  
Industrial supervisor: prof.dr. M.T. Smith

## Abstract

The emergence of internet and e-learning platforms generated enthusiasm and curiosity among stakeholders in the education landscape and pushed traditional universities to rethink their strategies and methodologies. In the last decade, the rapid advances in data collection technologies and analytics techniques opened doors to a wide range of possibilities to drive this change. There is a growing excitement about the role of the field of learning analytics in guiding this transition to a better educational and learning experience, which lives up to the 21st century challenges. However, a broad range of learning analytics studies focus on online data generated from learning management systems, in which human-computer and human-human interaction is limited to and mediated by screen-based computers. In this thesis, we explore the development of a new *learning analytics* system, which focuses on the learning experience in a collaborative physical learning environment, using the learning analytics cycle as a methodology. The aim of the study is to contribute to inspiring the existing learning analytics approaches to expand their data collection to the physical environment, while addressing the ethical issues arising in the field. A background research led to the identification of the set of learning indicators that we can start measuring in the learning physical space, with an initial prototype system. The system, which collects the learners' data, was built and tested during two observational experiments using a system design methodology. The first experiment served to test the robustness, usability and performance of the designed prototype. Building upon the first experiment's feedback, the system was improved and used during the second experiment to collect learners' data, which was analysed using common LA data analysis techniques. Data analysis

resulted in a set of recommended interventions for the organisation running the physical educational space. These interventions aim at improving the implementation of the desired learning experience and optimising the learning environment for the learners. Throughout the design process of the LA system, we tried to take in consideration the common ethical issues that the field encounters.

# Contents

<b>Contents</b>	<b>IV</b>
<b>List of Figures</b>	<b>V</b>
<b>List of Tables</b>	<b>VI</b>
<b>Preface</b>	<b>1</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Research questions . . . . .	4
1.2 Overview of approach . . . . .	4
<b>2 Theoretical Background</b>	<b>6</b>
2.1 Learning in the 21st century . . . . .	6
2.2 Inferring Learning . . . . .	10
2.3 State-of-the-Art in <i>Learning Analytics</i> . . . . .	11
<b>3 Method and Material</b>	<b>19</b>
3.1 Learning Analytics methodology . . . . .	19
3.2 Context and Practice . . . . .	19
3.3 Map of Measurements: what do we measure? . . . . .	22
3.4 Instruments: How do we measure? . . . . .	26
<b>4 Implementation</b>	<b>37</b>
4.1 Phase 1 . . . . .	37
4.2 Phase 2 . . . . .	41
4.3 Ethics and Privacy Protocol . . . . .	51
<b>5 Evaluation and Findings</b>	<b>52</b>
5.1 Data collection experiments . . . . .	53
5.2 Data pipeline . . . . .	60
5.3 Data preparation and analysis . . . . .	64

5.4	Results . . . . .	71
5.5	Discussion . . . . .	82
<b>6</b>	<b>Conclusion</b>	<b>87</b>
6.1	Research questions . . . . .	87
6.2	Implications and future work . . . . .	91
	<b>References</b>	<b>92</b>
	<b>Appendices</b>	<b>102</b>
6.3	Appendix A1: Järvavecken experiment user manual . . . . .	102
6.4	Appendix A2: Konstfack experiment user manual . . . . .	103
6.5	Appendix A3: Participant consent form . . . . .	104
6.6	Appendix A4: Mentor consent form . . . . .	105

## List of Figures

2.1	The Learning Analytics Cycle . . . . .	14
3.1	Two BubbleNodes communicating through LEDs during Future Friday 2017 at KTH Kista Mentorspace. Available at: <a href="https://www.dagy.danderyd.se/nyheter/Future-friday-KTH">https://www.dagy.danderyd.se/nyheter/Future-friday-KTH</a> . . . . .	22
3.2	Map of Measurements . . . . .	25
3.3	Initial LA system diagram . . . . .	35
3.4	Final LA system diagram . . . . .	36
4.1	Final LSB prototype components . . . . .	43
4.2	Final LSB prototype with a BN board being soldered . . . . .	43
4.3	LF RFID system prototype testing with 4 antennas . . . . .	48
4.4	Integration of switches inside a chair cushion . . . . .	49
4.5	User's perspective of the chair cushion . . . . .	49
5.1	Järvavecken experiment setup . . . . .	55
5.2	Experiment setup . . . . .	58
5.3	Cross Industry Standard Process for Data Mining . . . . .	60
5.4	Example of WB raw data . . . . .	63

5.5	Example of WB preprocessed data . . . . .	63
5.6	IN Konstfack experiment raw data . . . . .	64
5.7	Social network graph visualisation, Konstfack experiment Batch1	73
5.8	Social network graph visualisation, Konstfack experiment Batch2	75
5.9	Social network graph visualisation, Konstfack experiment Batch3	76
5.10	Faults summary Step 1 . . . . .	78
5.11	Faults summary Step 2 . . . . .	78
5.12	Faults summary Step 3 . . . . .	78
5.13	Faults summary Step 4 . . . . .	78
5.14	Faults summary Step 5 . . . . .	79
5.15	Proficiency decision tree for 26 participants . . . . .	80
5.16	Proficiency decision tree for 32 participants . . . . .	81
5.17	Proficiency decision tree for 42 participants . . . . .	81

## List of Tables

3.1	Map of Measurements explanatory table . . . . .	26
3.2	Immediate needs and wants for the LSB system . . . . .	28
3.3	LSB Initial and final specifications . . . . .	29
3.4	Interaction Network System Immediate Needs and Wants . . . . .	32
3.5	IN initial and final specifications . . . . .	33
4.1	WB test operations on the DUT, Phase 1 and Phase 2 . . . . .	42
4.2	LSB displayed user feedback . . . . .	43
4.3	High Frequency RFID v.s Capacitive system comparison . . . . .	46
5.1	Study observations during the Järvavecken experiment . . . . .	53
5.2	Study observations during the Konstfack experiment . . . . .	56
5.3	BN technical complexity assessment . . . . .	66
5.4	Equation variable interpretation . . . . .	70
5.5	Node labels map . . . . .	71
5.6	Batch 1 graph centrality measures . . . . .	73
5.7	Batch 1 network measurements . . . . .	74
5.8	Batch 2 graph centrality measures . . . . .	75

5.9	Batch 2 network measurements . . . . .	75
5.10	Batch 3 graph centrality measures . . . . .	77
5.11	Batch 3 network measurements . . . . .	77



# Preface

This "dense" thesis report marks the end of my intense journey as a student. I would like to take some space (and some of your time) to put this thesis into context and use my takeaways from this research to reflect on my own school process.

My personal motivation for picking the topic of education stems from the fact that embedded in me is a deep attachment to school and learning, which is probably a wiring that was influenced by the environment I grew up in. However, I was NOT prepared for what was coming.

The dominant learning theory that I grew up with throughout my school years back in my country (for about 14 years) was heavily influenced by objectivism; using different types of reinforcement/punishment techniques to wire the student to behave and think in a strict manner about a given concept. The mentality that the teacher defines what is the right way to do and think about everything was pretty much present everywhere and still is at a large scale. Over the years, I unconsciously evolved to think that knowledge (and the only "credible" knowledge) is the bundle we are given at school that you have to acquire as a learner. And the manner in which it is acquired depends heavily on your ability as a learner to store and process information, and that it was not a shared responsibility in any way. I also grew up to view fields as their own independent entities of knowledge that seldom connect. The consequence of this conditioning on my view of the world at that time was that I started perceiving myself as a passive agent in society and the world at large. I started noticing that I was not interested in following the latest advances in the field I was in because I thought I would never have any influence over it anyway.

Later in life, I got the opportunity to go abroad seeking different educational challenges. At the first university where I studied for two years, I experienced a philosophy of learning that I feel was slightly different. Lectures and recitations were prescriptive and considerably guided in a slightly less objectivistic way than in my home country, where there was small room for questioning and debating concepts. On the other hand, we had a few

courses that were project-based but the instructional strategies were still majorly objectivistic in a sense where learning goals and paths were strictly defined in advance by the teacher. As small as it was, this slight shift in paradigm already made me feel uneasy as I could not identify why I was feeling that I had to learn differently and did not understand that it was intended. In addition, I was unsure about what is the proper way to manage that window of 'freedom' or responsibility over my learning that was given to me.

Following my bachelors', I studied for my masters at two different universities in two different countries which adopted a more progressive approach to education. Looking back, I think both universities were heavily constructivist in their instructional strategies, for the following reasons: there is more emphasis on identifying the context in which the target skills to be learnt are applied, there is a strong emphasis on the learner having an active role in manipulating information that they were exposed to, the presentation of information as 'recommendations' rather than facts, a focus on deriving learning from solving problems, and a focus on assessing transfer of knowledge and how it is applied to new situations. In theory, most learning philosophies share a lot of common definitions and bases, the differences remain in the interpretation which results in differences in instructional strategies. However, I felt that something was lost between the interpretation and the implementation process. Practically speaking, the way I experienced this type of constructivist environment as a learner was with a dominant feeling of confusion about what was expected from me (especially coming from years of experiencing objectivism). In addition, confusion around what constitutes basic "facts" in a field; which type of information needs to be kept as a "takeaway" or "principle" and which type of information is okay to let go of. As a learner, when I went for classes that were labeled 'introductory' to a certain field, my expectations were to 'receive' what constitutes a solid foundation (facts and principles) of that particular field/course, which I can build upon in a less guided manner in more advanced classes. My understanding now is that we have a very blurry picture of what constitutes 'foundations' to be taught in a more guided way, and we majorly assign this complex task and responsibility to the learner to construct. What was also less obvious to me is the understanding that knowledge content is not pre-defined and comes as recommendations, which practically can lead students to taking different directions in learning the same concept. Besides, there is little clarity around how aspects of the promoted learning process such as creativity, critical thinking, problem-solving are assessed.

At the university of Twente, I was informed by bachelor students about the

existence of Twente Education Model (TOM), which is a new model designed by the university that promotes an approach that goes through decreasing the guidance provided to students over years towards a student-driven learning approach (SDL). They also adopt a project-based approach that gathers different tracks to promote knowledge transfer and allow students to discover their interests. When I read about TOM and the SDL approach of Twente, only then I started understanding the expectations of the roles of both the learner and teacher, and the nature of interaction between them at the university of Twente. However, practically, it is still very hard to determine the exact limit of what constitutes enough/necessary guidance and what is the proper applied instructional strategy to adopt in order to account for the differences in background, culture, knowledge state, and interests of each student. During my bachelor years and during my thesis internship, I interacted with multiple schools and students, and tutored many students that were labeled as underachievers. My takeaway from this humble experience is that sometimes underachievement is mostly a consequence of how complex it is to understand and include all the factors that affect the learning process of a student in practical instructional strategies. There is an incredible number of nuanced variables that affect learning which makes it very hard for stakeholders to achieve an acceptable state where everyone can flourish and be satisfied. However, I think one thing we can start doing better is acknowledging that every school acts upon a learning philosophy(ies) and that the principles (chosen instructional strategies) that make that philosophy(ies) and their implications need to be understood and internalised by all stakeholders to at least build a common ground.

This reflection was intended to be focused on the relevant challenges encountered throughout my education journey. It does not mean that I did not experience the benefits of the mentioned theories of learning, as they would probably outnumber the less positive aspects. It is just a choice to convey a picture of one possible way (out of many) of how the misalignments in understanding the learning process might affect the learner at a practical level. In addition, I do not believe that I am an isolated case. Throughout my school years, I interacted with enough students to notice that there are shared patterns of the way learners are impacted.

I started this thesis with an extremely narrow vision of the educational system issues and grew up towards partly understanding the scary but exciting complexity of the world, as a consequence of a perseverant investigation of the different factors influencing the system. We have to accept that there is little chance that any developed 'vision' will remain to be complete. So we have to broaden our 'lens' and combine it with others' lenses to be able to act upon a better overview of this complex puzzle. I conclude this preface

with a citation from John Dewey that resonated with me during my whole thesis:

*“For in spite of itself any movement that thinks and acts in terms of an ‘ism becomes so involved in reaction against other ‘isms that it is unwittingly controlled by them. For it then forms its principles by reaction against them instead of by a comprehensive, constructive survey of actual needs, problems, and possibilities.”*

John Dewey

I hope this thesis will have some kind of meaning for someone to expand on.

# Chapter 1

## Introduction

*"The big problem we need to solve in higher education is changing the relation between research and practice... We do a lot of research to understand human learning but very little of that research about how people learn finds its way into practice ... In the other way, you can think of everything a teacher does as a small experiment."*

Candace Thille - Stanford University and Amazon.com

*"Value is being legitimately challenged, the quality of what people are learning, its relevance, the rigour with which it is being taught and the price which is ridiculous."*

Mitch Daniels - Purdue University

*"Sixty percent of the employers today say that the students lack self direction, collaboration, communication skills... given the incredible level of information that has to be assimilated today, you will always need to find yourselves being prepared to understand the process of learning and communicating."*

Bill Achtmeyer- EY-Parthenon.

The quotes above are some of the comments made during the 21st Milken Institute global conference 2018 <sup>1</sup> about higher education. The global conference brings together leading minds in various fields to discuss global concerns of our time.

---

<sup>1</sup>Global conference 2018, Milken Institute. Available at: <https://milkeninstitute.org/events/global-conference-2018>

The 21st century technological revolution that is transforming industries at a vertiginous pace, challenges the state of the current structure of higher education. Higher education has long been the source of the requisite credentials. However, industries and knowledge are irreversibly evolving more rapidly than the current system can accommodate. The emergence and increasing access to new models of education such as massive open online courses (MOOC), learning management systems (LMS) and other virtual environments opened doors for new models of education and new types of students. Learners in the digital age are more connected, have more of an active role in shaping the content of knowledge being published, which is more distributed and easily accessible by the majority. Consequently, informal learning is a significant component of our daily experiences (Siemens, 2004). Furthermore, questions are raised about the fundamental theory of learning; what does it mean to learn in the digital age? Is learning about acquiring knowledge, or is it the capacity to connect knowledge sources in a meaningful way, including human expertise?

Surveys of employers demonstrate that the digital age requires new standards, which the National Education Association recognized as the 21st century skills (National Education Association, 2010; Schooley, 2017; R. Lai and Viering, 2012). Examples of these skills are creativity, digital literacy, the ability to communicate ideas, effective collaboration, critical thinking, proactiveness, among others. Nevertheless, there is little consensus about what the new pedagogical models should encompass to support the student in acquiring these skills (Education technology solutions, 2019).

Ongoing discussions around the efficacy and quality of education are an opportunity to explore new instruments which allow a thorough understanding of the challenges faced by this industry. There is a need to gain a deeper understanding about what has been done in education so far, evaluate where we are at, and re-envision why, what, how, when and where we learn.

Although the effect of ICT on education is ever-growing, a World Bank review reports that it remains unclear how effective the current education technology (edtech) companies are and how technology can be used to best serve education in today's world ( World Bank World Development Report, 2018). What is less unclear is that the classroom of tomorrow will have little resemblance with the classroom of today.

Often times, the idea of the use of technology to improve education paints the picture of a classroom full of computers, ubiquitous computing devices, edtech kits, 3D printers, VR sets, and other material. However, technology is increasingly being used in the background to better understand the behaviour of individuals and their relationship to their environment, in order to provide insight and assist with decision making. In education, this field

is known as learning analytics (LA).

The use of data to assess the effectiveness and quality of educational organisations and drive decision making is not a new concept. The number of research publications, the dropout rate, the course completion rate, the student to teacher ratio, the retention rate are examples of such data. Nevertheless, these key performance indicators (KPIs) alone are not sufficient to address the complexity of the digital age, where processes need to evolve and adapt faster than ever.

The rise of Internet and internet of things (IOT) offers opportunities to look at the quantitative granular features that we have not had the opportunity to access before. On the other hand, artificial intelligence tools assist humans in handling the various dimensions that big data brings to the table, and explore patterns of behaviours that cannot be easily determined by traditional analysis given the complexity of the learning environment. Such tools hold the potential to help us understand the learning environment better, and possibly design a better list of metrics to effectively evaluate and understand the effectiveness and quality of educational organisations. Most research in the field of *learning analytics* focuses on online student trails and generic indicators to monitor and predict underachievements in order to provide the right support on time. However, according to recent theories of learning, individuals go through the learning process in a social environment that shapes and influences their behaviours and evolvement. The focus on generic online indicators **only** and digital traces such as, the number of clicks, the demographics, the frequency of logins and the activities on the platform ignores the offline social context of the learning activity. This method underlooks the influence of social interactions with instructors and peers and environmental conditions outside the online social nebula on the academic achievement. Other factors including, who is behind the screen, whether they are focused or distracted and whether they are taking a break made it difficult to establish a reliable replicable framework, particularly in collaborative learning (Kovanovic, Gašević, Dawson, Joksimovic, and Baker, 2016; Kovanovic et al., 2015; Winne et al., 2017). It is suggested that failing to consider contextual information can lead to major misinterpretations of results and limit the replication in broader settings of learning (Gašević et al., 2018).

An additional concern in the field of LA is the relevance of metrics used to measure learning. A well-grounded theory is necessary to advance this discipline in a consistent manner ( Dawson et al., 2015; Chen et al., 2015; Reimann 2016). As a matter of fact, Joksimović et al. (2016) bring to light inconsistent results in associating the learning performance with the measures in social networks. Moreover, there is a significant lack of research

on expanding the applications of LA to physical learning spaces. This kind of multimodal data is necessary to encompass the full picture of the learning process (Dowell et al., 2015; Hernández-García et al., 2015; Joksimović et al., 2016). Finally, little is known about the motive behind the focus on building predictive models to envision students' performance and what is meant by performance. What is less clear is the set of metrics that are used to determine this performance and the transparency of how the results are applied to improve the assessment of education.

The aim of this thesis is to contribute to existing learning analytics techniques by expanding the data collection to a physical learning environment, in a generalisable and ethical manner.

## 1.1 Research questions

We identified questions in the field of learning analytics that we aim to contribute to. Namely: the emphasis on online data, and the neglect of the social context and offline interactions. The research questions we aim to address are:

**RQ1** How do we derive learning indicators from a STEAM physical collaborative learning environment that can be used to infer learning?

**RQ2** Which instruments can generate quantitative data in a physical STEAM collaborative learning environment, in an ethical manner?

**RQ3** Which data analysis techniques are appropriate to analyse the data gathered from a physical STEAM collaborative learning environment?

## 1.2 Overview of approach

The thesis report starts with a background study in Chapter 2. The literature study encompasses relevant theories of learning explained in a simplified manner (section 2.1 and 2.2). These theories set a useful background to understand the learning environment and project the thesis deals with. The background also explores the state-of-the-art of learning analytics as a field and builds an overview of the learning analytics methodology, the relevant applications, the commonly used data analysis techniques in the field, and the limitations and ethical issues encountered (section 2.3). In chapter 3, we describe the methodologies used for the study, and the learning environment

(context) and activity characteristics in accordance with the learning theories described in the background (section 3.2). In addition, we define the learning metrics (section 3.3) which are used to determine the requirements for the design of the LA instruments used to collect the necessary data from the environment. The requirements are also included in this chapter (Table 3.2, Table 3.3, Table 3.4, and Table 3.5). Chapter 4 is a more practical description of the implementation of the LA instruments, which are utilised for data collection, as well as a description of two conducted observational experiments. The first experiment was used to perform a first iteration on the implemented instruments and test their robustness, usability, and performance. The instruments were improved based on the feedback from the first experiment, and used in the second experiment to collect the necessary data for the study. More practical details about the methodology used can be found in the next chapter. Chapter 5 focuses on analysing the data collected from the second experiment and applying relevant techniques to attempt to derive insight in accordance with the learning analytics cycle method. The chapter ends with outlining the results of the research and discussing them. Finally, the last chapter addresses the research questions and the implications of the study.

# Chapter 2

## Theoretical Background

### 2.1 Learning in the 21st century

The topic of learning has been a subject of fascination for many researchers and thinkers for decades. There have been many attempts to define learning. For instance, Shuell (as interpreted by Schunk, 1991) suggests the idea that: “*Learning is an enduring change in behavior, or in the capacity to behave in a given fashion, which results from practice or other forms of experience*”. On the other hand, Driscoll (2000) defines learning as “*a persisting change in human performance or performance potential. . . [which] must come about as a result of the learner’s experience and interaction with the world*”. Various theories of learning developed and evolved as an attempt to identify the origins of knowledge, and explain how human learning occurs. These theories provide important insights for us to understand the learning process better in order to identify the best ways to improve it. Perhaps the most common learning theories utilised in education are behaviourism, cognitivism, constructivism, and more recently connectivism. The following is an oversimplified explanation of these complex concepts in an attempt to set a necessary background to better understand the purpose of this study.

#### 2.1.1 Behaviourism

Behaviourism sees learning as a change in behaviour that occurs as a response to external environmental factors (stimuli). Hence, the aim of instruction in behaviourism is to shape the external behaviour of learners through reinforcement and repetition (rote learning). The teacher is given the responsibility to implement a process of reinforcement decisions to induce a desired change in the external behaviour of the learner resulting in the

targeted observable performance. In other words, the instructor is responsible for defining what is right or wrong for the learner. This conceptualisation of learning focuses on the external behaviour of the learner and the environment surrounding them. This theory is not interested in the internal behaviour and processes with which the learner acquires knowledge. The learner is characterized as reactive throughout the whole process. (Burton, J. K., Moore, D. M. (M.), Magliaro, S. G.,2004)

### 2.1.2 Cognitivism

Cognitivism came as a response to the dominant behaviourism paradigm in 1960s. Cognitivism theories used cognitive sciences to look into the learner's "*complex, cognitive processes such as thinking, problem solving, language, concept formation and information processing*" to draw useful insight about how learning occurs (Snelbecker, 1983). They focused on attempting to understand "*how the information is received, organized, stored and retrieved by the mind*"(Ertmer Newby, 1993).

Cognitivism states that knowledge acquisition is a result of a mental activity that comprises internal structuring and coding by the learner. In this theory, learning is viewed as a change between states of knowledge rather than change in response. From a cognitivism theory perspective, learning is the result of storing information in memory in an organised and meaningful structure, and that depends on both external and internal factors. Cognitive approaches argue that environmental and instructional conditioning cannot account for the whole learning process. Learning is perceived to be also dependent on the way learners structure, store, transform, retrieve information, and on the learner's internal value structure, beliefs, and attitudes (Winne, 1985). Consequently, learners are perceived to have a more active role during the learning process. On the other hand, instructors are viewed to be responsible in assisting learners to achieve the organisation and storage of information in the most optimal and effective manner. Cognitivist instructors use feedback (knowledge of results) to guide learners and support them in making accurate mental connections with previously learnt material (Thompson, Simonson, Hargrave, 1992).

### 2.1.3 Constructivism

The epistemology underlying behaviourism and cognitivism is primarily objectivistic: which states that the reality is external to the learner, that knowledge is acquired through experience, and that the aim of instruction is to project the reality structure onto the learner (Jonassen, 1991b). On the

other hand, constructivism is a paradigm that states that reality is internal and that knowledge is constructed by the learner from their experiences and their interpretation of the world (Perkins, 1991).

Constructivism perceives learning as the process of creating meaning from experience (Bednar et al., 1991). Therefore, every learner constructs their own meaningful reality. They build and refine their own unique lens from which they see the world by deriving meaning from their experiences (Jonassen, 1991a). Constructivism does not believe in the existence of some form of objective knowledge or reality that is mind-independent and that can be projected onto the learner. This theory acknowledges the existence of the real world but argues that what we know of the world also depends on our own interpretation of our experiences.

Learners are characterised as active agents in creating meaning as opposed to acquiring it, and there cannot be one objective meaning since there are as many possible meanings as there are possible personal experiences.

In order to achieve meaningful learning, constructivists argue these three crucial factors need to be taken in consideration: activity (practice), concept (knowledge), and culture (context) (Brown et al., 1989; Bednar et al., 1991). In such a learning environment, every action is perceived as *“an interpretation of the current situation based on an entire history of previous interactions”* (Clancey, 1986). Therefore, the meaning that the learner constructs about a concept constantly evolve with each new use into a more densely textured form. Consequently, it is crucial that the learning process occurs in a setting that is realistic and the tasks assigned be relevant to the student’s prior experiences. *“Understanding is developed through continued, situated use . . . and does not crystallise into a categorical definition”* (Brown et al., 1989, p. 33). In a constructivist learning environment, knowledge content is not pre-defined. Learners are encouraged to construct their own perspective and understanding about a concept before validating it through social negotiation. The instructor’s responsibility is to prepare the learner to use facts to construct their own perspective, and promote a setting that is as close as possible to a real-world situation with realistic problems to solve and collaboration with peers to get exposed to multiple perspectives (context). *“The role of instruction in the constructivist view is to show students how to construct knowledge, to promote collaboration with others to show the multiple perspectives that can be brought to bear on a particular problem, and to arrive at selfchosen positions to which they can commit themselves, while realizing the basis of other views with which they may disagree”* (Cunningham, 1991, p. 14).

### 2.1.4 Connectivism

Connectivism is a paradigm that was developed by George Siemens and Stephen Downes as a result of the proliferation of internet and the web which impacted the underlying conditions of the learning process that past theories are built upon (Siemens, 2004). This theory recognises that the digital age is connected, fast changing and mediated by technology, which puts the learners in learning environments with shifting core elements over which they have little control. Siemens states that *“Exponentially developing knowledge and complexification of society requires nonlinear models of learning (process) and knowing (state). We cannot sustain ourselves as learning/knowing beings in the current climate with our current approaches”* (Siemens, 2009, p. 3). At the heart of connectivism is the principle that knowledge is distributed across a network of connections, and that learning is a participatory self-organising process in which one learns from connecting specialised nodes or information sources. This paradigm states that knowledge can reside outside oneself (human and non-human appliances/ artefacts), and that the capacity to learn from these connections is more important than the current state of knowing. *“Experience has long been considered the best teacher of knowledge. Since we cannot experience everything, other people’s experiences, and hence other people, become the surrogate for knowledge. ‘I store my knowledge in my friends’ is an axiom for collecting knowledge through collecting people.”* (Karen Stephenson, undated)

Connectivism also tries to address the organisational knowledge management and transfer, which past theories did not account for: *“Knowledge that resides in a database needs to be connected with the right people in the right context in order to be classified as learning.”* (Siemens, 2004).

Research literature supports the idea that today’s learners have different thinking patterns and learn differently than previous generations as a result of the introduction of technological tools that students use (Barone, 2003; Brown, 2002). Siemens (2004) suggests that technology has rewired the learner’s brain. Prensky (2010) agrees that digital students prefer a higher level of interaction and activity, and that doing is more important than knowing for such learners.

## 2.2 Inferring Learning

Learning is a complex phenomenon that is not directly observable. Learning is inferential; we observe it through assessing its products and outcomes. However, inferring learning from learners' activities is challenged by the lack of consensus about a practical definition of learning.

Schunk (2011) describes a few relevant methods that are commonly used to assess learning products and outcomes. These methods are described as follows:

**Direct observations** They are instances of observable learner's behaviour that demonstrate learning. These behaviours are assessed by an observer, such as a teacher, to determine whether learning occurred (e.g.. An instructor observing a student that is solving a math equation to determine whether they are using the right procedure). A common challenge of this method is that there is a strong focus on behaviours that are observable. This results in two issues; the cognitive processes underlying the behaviour of the learner might be bypassed, and the absence of appropriate behaviour might be interpreted as though learning did not occur, which is not true for all situations.

**Written and oral responses** These methods are desirable and widely used indicators of learning due to their relative ease of use and their ability to generalise on the vast majority of educational material. In both methods, learning is assessed based on the quality of student responses to exam tests, homeworks, and reports. However, it is important to consider that there are factors other than learning which affect the performance of the learners in these tests. Therefore, they might not be the most valid reflection of learning.

**Self-reports** They are learner's assessments about themselves, through answering questionnaires, interviews, stimulated recalls, think-alouds, and dialogues. A common challenge with this method is the quality of responses and their level of correlation with the beliefs of the learner and their actual behaviour, which might significantly alter the quality of inference.

## 2.3 State-of-the-Art in *Learning Analytics*

In the last two decades, business analytics, big data and artificial intelligence increasingly evolved and led to the proliferation of applications in various fields. Examples of such fields are health analytics, social media analytics, and learning analytics. The term analytics refers to the scientific process that allows for the creation of actionable insights based on data examination (Picciano, 2012). Learning analytics is a new and promising research field that holds the potential to help disrupt the educational setting. We may argue that this field of research formally emerged at the first conference on *Learning Analytics and Knowledge* (LAK) in 2011. It is defined by the *Society of Learning Analytics and Research* (SoLAR) as the “*measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*” (Siemens, 2013).

Researchers have been exploring the possibilities to extract meaningful information from students’ online activities for a decade before the field of learning analytics became officially recognised. For instance, Newlin and Wang (2000) suggest that there exists a correlation between page hits and forum postings with online student performance. In a second study that was conducted in 2002, they explore the potential of using learning management systems in performance prediction through a study that involved 121 student in the field of psychology. In another paper, Macfadyen and Dawson (2010) used data mining on learning management systems to develop a proof of concept of an “early warning system” for educators. The published study reveals 15 variables that show correlation with the student grade. To further predict the student performance, they used regression modelling based on key variables such as the number of forum messages posted, number of email messages sent and the number of assessments completed. Their results indicate that the model helped justify 30% of the students’ grade variation. On the other hand, logistic modelling was able to correctly predict 81% of underachievements.

After the introduction of the field of Learning analytics in 2011, researchers and institutions have been racing to harness the potential of this field in supporting students, teachers and stakeholders. Efforts are put into the expansion of the data scope through new digital sources and trackers, the development of new data analytics methods through advanced data mining tools, machine learning and AI, and the exploration of new objectives to help assess the learning activities (Blikstein and Worsley, 2016).

A literature review from Leitner, Khalil, and Ebner (2017) suggests that

the most utilised techniques in learning analytics are predictions, followed by outlier detection and distillation of data for human judgment. These techniques are believed to enable the identification of patterns of underachievement within learners, in order to provide the right support on time (Jayaprakash, Moody, Lauría, Regan, and Baron, 2014; Ali, Hatala, Gašević, and Jovanović, 2012; Clow, 2013). In recent years, LA systems are increasingly being integrated and used in online educational environments and higher education despite the lack of common methods and a unified framework of use (Arroway, Morgan, O’Keefe, and Yanosky, 2016).

One of the first implementation of an “early warning system” was in Purdue University as described by Kimberly (2012). The goal of the system is to identify risks of underachievement to allow an early intervention in order to improve the student retention at the university. Their predictive algorithm was based on student demographic data (age, gender, location), prior academic history (location, previous grades), student’s performance calculated by the sum of grades earned in the course, and student efforts determined by the student activity on the learning management system of the course. Based on these parameters, the system would classify the risk of underachievement into three categories that are communicated to the educator through a green, yellow or red light depending on the predicted rate. The university reported an improvement of its retention rate and student success rate after using the system (Arnold and Pistilli, 2012; Dietz-Uhler and Hurn, 2013).

The closest study to this thesis in terms of goal is most probably the PELARS project. This project used a set of ICT-based tools such as arduino-based embedded sensors, face tracking, hand movement tracking, position tracking, audio and video recordings, affective sensing and mobile-based self-documentation to collect multimodal data related to learning activities in STEM fields. The data is then processed using different data analysis methods such as machine learning classification models (Daniel, Emanuele, Giacomo, Mutlu, 2018). The results are used to assess the student learning progress in STEM activities in order to support teachers through recommendations.

Despite the positivity about the potential of LA to improve learning, there is still little evidence about demonstrated improvements in learning outcomes and optimisation of the learning environment (Lodge, Alhadad, Lewis, and Gašević, 2017). One challenge that the LA field might need to address more explicitly is the absence of well-grounded theory in the development of LA systems (B. Chen, 2015; S. Dawson et al., 2015; Reimann, 2016). In the past few years, more researchers have attempted to contextualise the learning indicators used by LA systems. This operation resulted in including learners’

social interactions and information exchange as an important indicator of learning (Siemens, 2004; Illeris, 2018; Bell, 2010). There is empirical evidence about the benefit of strategies that promote collaboration in a learning environment on the learning process (Bernard et al., 2009; Garrison and Arbaugh, 2007; Lambert and Fisher, 2013).

### 2.3.1 The Learning Analytics cycle

The learning analytics cycle is an iterative process. In this study, we adopt a four component model shown in in Figure 2.1 (Chen et al., 2012; Clow, 2012b; Siemens, 2013).

**Learners** element one is about the data collection from the learning environment. The main questions to be addressed are; Which learning environment is used? What is the pool of learners we observe? What do we identify as learning?

**Data** The second element includes data preprocessing and preparation for analysis. In this part, we primarily discuss which data from the use case can be produced and how the data is cleaned to serve as a proper analysis base.

**Metrics** The third element deals with the analysis and interpretation using various techniques such as machine learning models. In this section, we are concerned with the type of analytical instruments used to identify patterns, how the data can be visualised, interpreted and turned into meaningful information signifying learning.

**Interventions** In the last element, we try to formulate Insightful action to drive effective decision-making based on the analysis findings. In addition, we provide feedback about the system performance in order to refine and improve it.

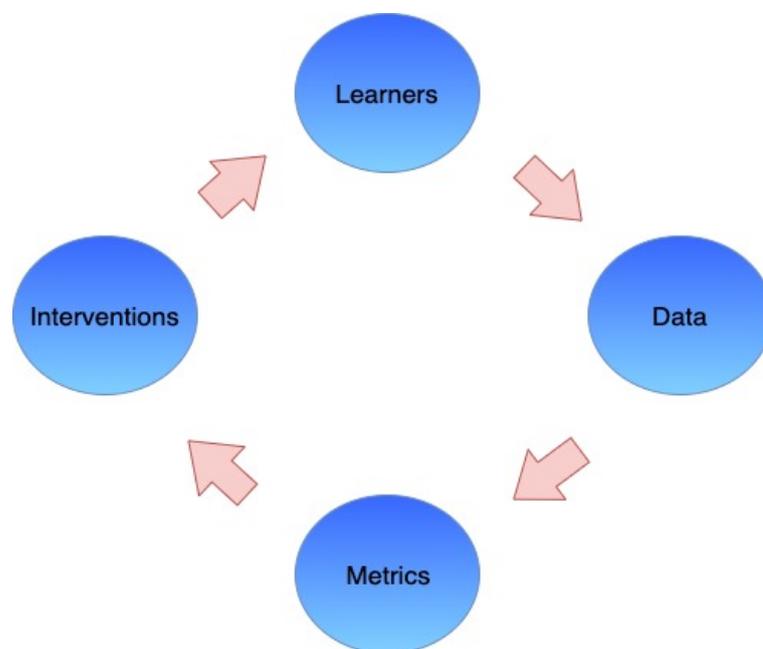


Figure 2.1: The Learning Analytics Cycle

### 2.3.1.1 Insightful Action: closing the feedback loop

Learning analytics aims at helping students, teachers and stakeholders in decision-making towards improving the learning environment. This step is concerned with displaying the relevant gained insight from data analysis in a way that makes sense to the concerned parties (Duval, 2011; Atif, 2013; Verbert et al., 2013; Bodily and Verbert, 2017). These insights are presented in a form of dashboards and other tools that illustrate the students' activities and behaviors, students' level of engagement, underachievement predictions, personalised feedback and recommendations, and pedagogic effectiveness assessment (Arnold and Pistilli, 2012; Fritz, 2011; Lockyer, Heathcote, and Dawson, 2013; Mor, Ferguson, and Wasson, 2015; Ruipérez-Valiente et al., 2017; Kuzilek, Hlosta, Herrmannova, Zdrahal, and Wolff, 2015; Persico and Pozzi, 2015; Bakharia et al., 2016; Misiejuk and Wasson, 2017; Wong, 2017).

### 2.3.2 Data analysis techniques

The learning analytics field applies sophisticated data analysis techniques to identify useful patterns of knowledge in the educational datasets. The common analysis methods that are used include statistics, social network analysis (SNA) (Misiejuk and Wasson, 2017), data mining techniques such as

supervised and unsupervised machine learning, and association rule mining (Liu, 2006; Han and Kamber, 2006).

### 2.3.2.1 Statistical methods

It consists of applying general statistical operations (average, mean, standard deviation, etc...) on generic data such as number of clicks, frequency of logins, number of posts and messages to provide a basic quantitative overview of the student's interaction with the machine. These numeric indicators are then presented to educators in the form of reports, tables, charts or graphs to facilitate their interpretation.

### 2.3.2.2 Supervised learning

Different classification and regression algorithms have been widely used in the field of learning analytics in the past years (Macfayden and Dowson, 2010; Romero and Ventura, 2013; Gašević et al., 2015; Tempelaar et al., 2015). Classification modeling is the process of creating a function that categorises data into predefined discrete classes and using the same function to predict the class of data with unknown labels. Common classification methods include neural networks (NN), decision trees, k-nearest neighbor (KNN), naive Bayes classification, and support vector machines (SVM). Regression models are mainly used for continuous numeric predictions. As mentioned earlier, predictive modeling of academic risks of underachievement might be the most common application among the LA field (Leitner et al., 2017; Gardner and Brooks, 2018). The objective of such predictions is to forecast a future event such as risks of failure in a specific class in order to allow educators to intervene with the proper support to reduce the probability of occurrence.

### 2.3.2.3 Unsupervised learning (Clustering)

Clustering is the process of organising the data into groups (clusters) that share similar characteristics based on how close the objects are in space. Unlike classification, the class label of the training set in clustering is not predefined. We can identify three main clustering methods: density-based methods, partitioning methods, and hierarchical methods. In learning analytics, clustering is used to create personalised recommendations of resources for students based on choices of other students with similar goals and performance (Romero and Ventura, 2010; Avella, Kebritchi, Nunn, and Kanai, 2016). It can also be used to help organise the course material into different groups depending on the level of difficulty, among other applications.

#### 2.3.2.4 Association rule mining

Association rule mining is the process of exploring interesting correlations and associations within large sets of data. For instance, it can be used to discover how the choices made by a student in selecting a set of courses for the semester are related. Moreover, It can also be used to predict the factors that lead to academic success or failures of students in order to enhance their performance (Kumar, Varun and Chadha, Anupama, 2011).

#### 2.3.2.5 Social Network Analysis (SNA)

SNA is a quantitative study and visualisation of the interactions and relationship evolution between knowledge-possessing entities within a social structure (Wasserman and Faust, 1994; Borgatti et al., 2009). These entities form a network and are usually referred to as nodes. They can be machines, websites, humans, groups, organisations or any entity that holds the potential of being grouped in terms of a relationship. The links that connect the nodes to form the network are referred to as ties or edges.

In order to identify key actors in the network, the following parameters might be useful to calculate:

**Degree centrality.** Determines how well a node is connected to the rest of the network.

**Closeness centrality.** Indicates how close a node is to every other node in the Network. It is a useful indicator of how information spreads across the network.

**Betweenness centrality.** A measure of a node's influence in the network. Nodes with the highest betweenness centrality tend to be the gatekeepers given their important role in connecting different subgroups within the network.

**Eigenvector centrality.** A measure of a node's importance. Eigenvector centrality is an indicator of how well a node is connected to the influential nodes within the network.

**Clustering coefficient.** In a network that is not fully connected, certain nodes that possess certain common characteristics or relate to each other in a particular way tend to form communities or clusters. The clustering coefficient is an indicator of how integrated or fractured a network system is.

**Network Connectivity/Cohesion.** A measure of the minimum number of nodes to remove before the network breaks into components. It is an indicator of how well the network is connected.

**Network density.** A measure of the number of existing connections within the network out of all possible connections. It is considered an indicator of the network robustness.

In our research project, we aim to use the SNA method to study the human-human and human-machine interactions within the learning environment.

### 2.3.3 Ethics and privacy

Learning analytics systems deal with large datasets that are beyond the human capabilities to analyse. Concerns about ethics and privacy are increasingly raised in the LA research community.

There is a need for policies and guidelines about the most ethical way to conduct analytics research (Macfayden and Dawson, 2012; Drachsler and Greller, 2016; Ferguson, Hoel, Scheffel, and Drachsler, 2016). Such a framework is necessary to prevent data misuse, protect the user's identities and determine the boundaries of possible decisions to be made on the basis of the acquired knowledge. Needless to say, these ethical concerns are not unique to higher education, comparable debates are also present in any research implying humans as participatory agents (Cooper, 2004). For instance, questions remain about transparency around the clear purpose behind the data collection, the methods used, which biases are inherited from the used techniques, whether and which third parties or companies are involved in data processing, whether the data is transferred to institutions or legal entities without explicit consent, and how the results and the acquired knowledge are used (Boyd and Crawford, 2012; Slade and Prinsloo, 2013; Rubel and Jones, 2016; Tsai and Gašević, 2017; Tsai et al., 2018). Most importantly, who owns the data and which party is the most appropriate to have control over it remains unsolved at this moment in time (Pardo and Siemens, 2014; Petersen, 2012; Prinsloo, 2013; Drachsler and Greller, 2016). Moreover, it is important to ask to which degree the traditional consents fail to protect and inform students about the complexity of the new tools and methods utilised for data manipulation and any future scenario that might result from LA development (Slade and Prinsloo, 2013; Mantelero 2014; Roberts, Chang and Gibson, 2017).

The LA field struggles to overcome such challenges for now. Consequently, cases of data misuse such as student labeling, categorising, and bias that results in discriminatory practices have been reported ( Beattie, Woodley and Souter, 2014; Tsai and Gašević, 2017).

# Chapter 3

## Method and Material

### 3.1 Learning Analytics methodology

The learning analytics cycle is an iterative process described in section 2.2.1 ( Figure 2.1). Learning analytics systems must be regularly improved and refined based on the continuous knowledge acquired through data (Clow, 2012a). This chapter addresses the first element of the adopted LA methodology. It focuses on describing the reasoning behind the process of building the necessary instruments to generate learners' data.

From the constructivism and connectivism perspectives described in section 2.1.3 and section 2.1.4, it is emphasised that learning occurs within a social setting, and that context influences the learning process. Therefore, it is essential to account for the characteristics of the learning environment setting and the tasks portrayed to the learner during the learning activity. The environment and the project activity that will serve as a basis for this study are analysed and briefly described in this chapter (section 3.2).

### 3.2 Context and Practice

#### 3.2.1 Learning Environment: *context*

In this project, learners' data is collected from a physical learning environment and turned into meaningful information through the LA cycle. The learning environment within which the study is conducted is called the **Kista Mentorspace**.

The Kista Mentorspace (Mentorspace) is an organisation that consists of a physical learning space, currently located at the Royal Institute of Technology (KTH) in Stockholm, Sweden.

The Mentorspace can be perceived as a space that tries to complement the educational curriculum with hands-on learning activities that help learners act out the theoretical concepts known by the learners. The type of formal activities promoted in the space are applied educational seminars to teach student how to use the tools available to them (performed in a guided manner), and hands-on realistic projects that vary in complexity. The learning theories underlying the approach with which the projects are conducted vary between objectivism and a more constructivist approach depending on the level of complexity of the project at hand and the learner's knowledge state. The Mentorspace social network consists of the triple helix model of innovation; which relies on the interactions between three sectors; academia, industry and government to promote economic and social development. The Mentorspace approach involves working with schools, governmental organisations, communities and industry. Most projects that the Mentorspace offers come from one of the triple helix model nodes. The application of this model ensures that the project activities are well grounded in reality and are tied to a real-world situation. The diversity in project topics promotes multidisciplinary learning and knowledge transfer as the learner gets exposed to multiple angles and applications of the concepts they learn at university. Moreover, it maximises the chances of meeting different learners' personal interests (relevance to the learners' personal experience). In addition, the Mentorspace encourages learners to use their constructed knowledge to benefit the organisations and communities around them. Therefore, the Mentorspace promotes the idea that organisations (including the Mentorspace organisation) and communities might also learn and grow as a result of the student's learning process and how they use their knowledge within society.

Constructivism acknowledges the role of guidance and the social aspect in the learning process. For instance, Jonassen (1991a) described three levels of knowledge acquisition; introductory, advanced, expert. Jonassen argues that constructivism is best suited for advanced knowledge in ill-structured domains, and that introductory stages require more support in a more objectivistic approach. This view is also implied in the notion of environments as the zone of proximal development by Vygotsky (1978). Bruner (1985) states that *"learning for him [Vygotsky] involved entry into a culture via induction by more skilled members"* and that *"a mentor is the highest level of educator"* (Vygotsky, 1981). Mentorship is an essential mindset and a core concept in the Mentorspace. Any individual in the Mentorspace who

appears to be more knowledgeable in a particular area is encouraged to provide guidance and share their experience with other individuals. Mentors may come from various organisations or industries and they may be from any age group. Therefore, the Mentorspace does not make any intervention without its mentors.

During Mentorspace workshops and events, the mentors are trained on the level of guidance and instruction they are ought to provide while assisting the participants in implementing their projects. They act as a bridge between the theory and practice through their knowledge of applied science. On the other hand, they help the participants to understand the relevance of the concepts they learn and derive meaning from them, with respect to society and their own lives. They achieve this by engaging with them in topics of their [participants] personal interests.

It is observed in the Mentorspace that mentors provide a source of guidance that often allows students to surmount barriers placed by the school and/or society, and to perceive within themselves talents that have been concealed. The results observed throughout the years are students who are digitally more literate, mindful of the relevance of their study and career, and are able to demonstrate understanding and mastery of academic concepts through practical accomplishments.

This thesis project is an opportunity for the Mentorspace to turn such observations into concrete quantitative data and improve their processes to achieve greater value.

### 3.2.2 Use case: *practice*

The project chosen for the experiments is an early version of what is called a *BubbleNode* (Figure 3.1).

The BubbleNode (BN) is a simple embedded computing system designed to be used in schools or other educational environments to support class content related to computational thinking. The design of the BN lends itself to be used in a wide range of school classes and subjects, not just ones focused on computing. The other interesting point about the BN is that it is not a typical 'black box' technology, since every BN part is a basic electronic component that has a function that can easily be explained in a simple manner. It is not an end in itself since it is used as a tool to accomplish a bigger task. The intent of building a BN is for participants to ask 'could I use it to do ...'. For example, add interactivity, motion, communication, or 'intelligence' to some other activity or object. The BN can easily be connected to external devices and is programmed to communicate with other

BNs through an IR LED.

The BubbleNode is provided as a kit during the experiments and is intended to be assembled by students or teachers that want to be introduced to computational thinking and information technology. For this research, we chose the simplest version of the BubbleNode and divided the building process into five functional sections for participants, that we will elaborate upon in a later section.

The advantage of using the simplest version of the BubbleNode for this research project is its low complexity and simple circuitry that is suitable for quick prototyping. One other important benefit of using the BN is that it needs no formal engineering experience requirements or know-how.



Figure 3.1: Two BubbleNodes communicating through LEDs during Future Friday 2017 at KTH Kista Mentorspace. Available at: <https://www.dagy.danderyd.se/nyheter/Future-friday-KTH>

### 3.3 Map of Measurements: what do we measure?

In contrast to the conventional learning analytics research method, in which learners' data is collected from existing online digital systems such as educational learning management systems, our aim is to expand the possibilities

to the physical learning environment. This process requires the design of suitable instruments that enable the translation of the learning activities taking place in the Mentorspace into quantitative metrics. The PELARS project attempted to collect such multimodal data using, among other devices, cameras, audio recordings, computer vision systems, facial and hand tracking systems (Ruffaldi, Dabisias, Landolfi, and Spikol, 2016). We plan to create a model which is as free from personal data as possible. For current and future Mentorspace projects, the process must be replicable and generalisable.

In order to reach such a model, it is necessary to identify the measurable indicators that support the development of the LA system architecture.

In this section, we attempt to answer the first research question:

**RQ1** How do we derive learning indicators from a STEAM physical collaborative learning environment that can be used to infer learning?

Modern theories of learning state that learning is a process that is affected by both internal and external factors. The internal factors are related to the learner's cognitive architecture, whether it is processing and storing information, creating meaning, or the ability to make connections between knowledge sources. On the other hand, the external factors encompass the learning environment conditions (including social factors, instructional strategies, pedagogical intent, and access to information sources), which was expanded to online learning environment with the emergence of the web. For this study, we focus on the physical learning environment only.

The learning indicators that are chosen to be measured for this initial study are decided upon the questions that are of most interest to the Mentorspace. The Mentorspace organisation would like to use the answers to these questions to optimise the process of learning for their users. The chosen metrics might appear to be high level and are certainly subject to further study and iterations and to change in the future.

We quickly came to the conclusion that the system for this study is bi-partite; one aspect measures the human-human and human-machine interactions that take place during our educational workshops between the participants and the knowledge resources provided by the Mentorspace; human expertise (mentors), and artefacts (access to tools and material). We want to know how people and machines interact in the space and how the space resources are used by the individuals to achieve learning. We call the system component responsible for capturing these interactions the *Interaction Network System (IN)*.

The second part of the system acts as a direct observer, looking for indicators of learners' behaviour and cognitive process during the learning activity, through analysing the interaction between the learner and the educational project on which they work. While recording the necessary data, the system component engages in the interaction as well through giving feedback to the participant about their progress. Thus, it is used as a resource to both the mentor and the participant throughout the educational activity. This part we call the *Learning Space Box (LSB)*.

The following map figure illustrates the requirements to which the system architecture is bound. It is crucial to keep the distinction between the IN and the LSB in mind when reading the map. We will elaborate on the specifications of both subsystems in the next section.

The map is organised in a way such that the circles represent the indicators to be directly captured from the learning environment by the LA system. The colour code indicates which metrics are measured by which subsystem; IN or LSB. The resulting data from the measurement of these metrics will be combined at a later stage. This combination of metrics is illustrated by the green circles and is explained in Table 3.1. In that table, questions that we attempt to answer using a combination of the collected data are listed according to the group number on the map. We also added a column to explain how the results are used for transparency purposes.

It is important to note that these measurements are not intended to establish new ways of classifying individuals or in any way to compromise their privacy. The goal is to provide insight into the academic impact of the Mentorspace and to support the organisation in improving the quality of their value proposition. It is therefore necessary to ensure that our system architecture is consistent with this vision and reduces the likelihood of being misused (section 2.2.3).

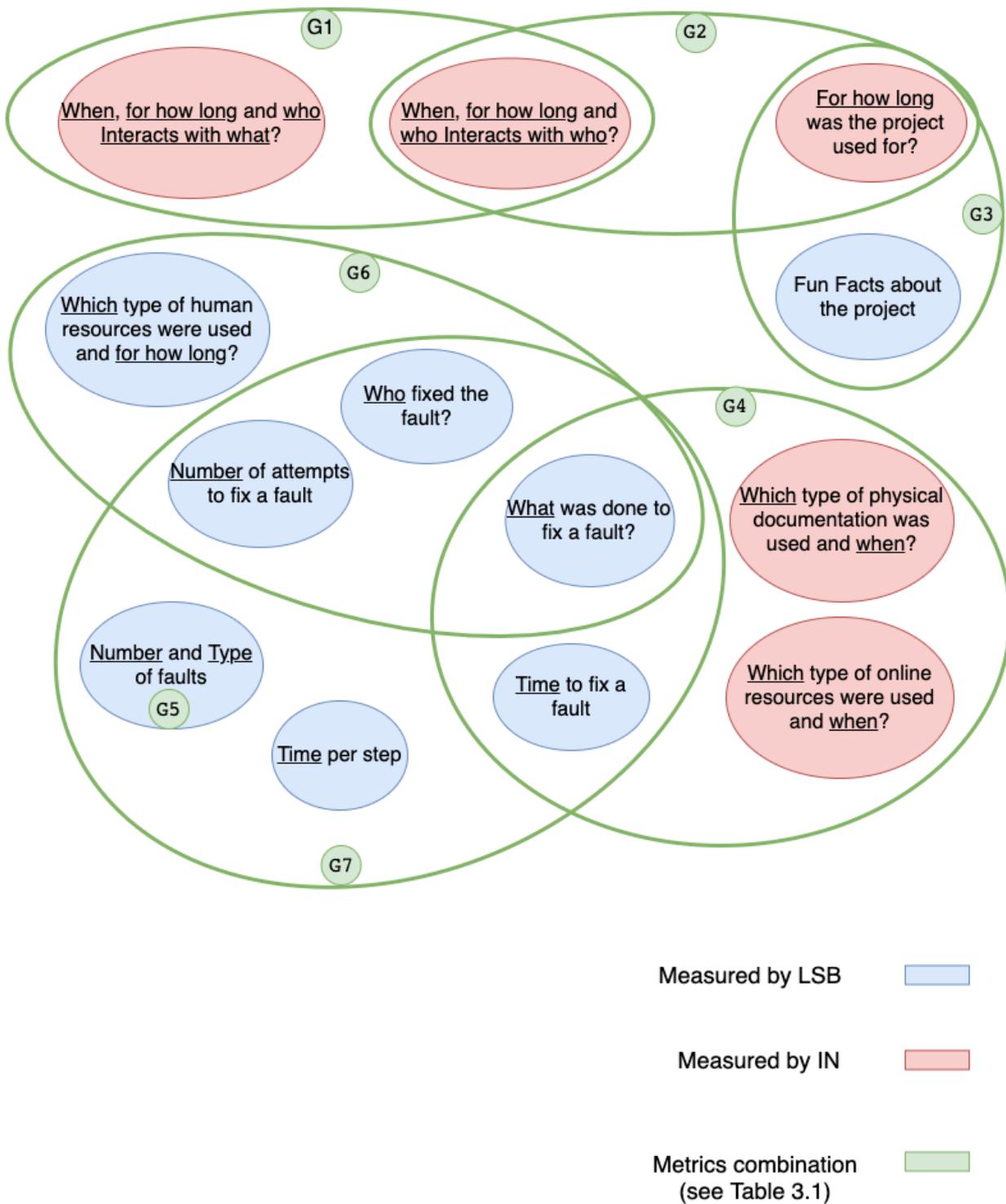


Figure 3.2: Map of Measurements

Group	Organisational (Business) Questions	Purpose
G1	How are the Mentorspace resources used?	<ul style="list-style-type: none"> <li>- Identify which provided artefacts are mostly used and why.</li> <li>- Understand which mentors are the most active and why</li> <li>- Identify to which degree the human and non- human resources are involved in the learning process. This helps improve the resources provision and the mentor training.</li> <li>- Understand where in the space are the mentors mostly needed. This helps with effective resource distribution and mentor training. (i.e. a concentration of users around the 3D printer over a significant amount of time might mean a strong interest in this artefact which means that some mentors need to acquire 3D printing skills.)</li> </ul>
G2	How is the resulting project used?	<ul style="list-style-type: none"> <li>- Understand whether the participants used the project with their peers. This helps identify to which extent the projects meet the interests of the user and triggers their curiosity.</li> </ul>
G3	Did the user demand to know more about the project after it was built?	<ul style="list-style-type: none"> <li>- Understand to which extent the environment conditions are engaging and promote reflection and critical thinking</li> <li>- Learn about the user's interests to match them with better projects.</li> </ul>
G4	How is the documentation used to solve the assignment challenges?	<ul style="list-style-type: none"> <li>- Understand which combination of resources is most effective for different users.</li> <li>- Understand the skill gap (i.e. if a classroom of users struggle with a particular challenge related to a concept that is supposed to be learnt in their school class, then there is a possibility that the goals of the class are poorly achieved. This will help design specialised projects and improve the mentor training. to target the skill gap.</li> <li>- Assess the level of engagement of the user (i.e. if heavy traffic is perceived towards social media platforms, this might indicate that the project is not engaging enough)</li> </ul>
G5	What is the level of complexity of a particular project for a particular group of users?	<ul style="list-style-type: none"> <li>- Identify which technical aspects represent significant challenges for the users. This helps understand the skill gap and improve the mentor technical training to meet the challenges. It also enables a better control over the level of difficulty of the project assignments by incorporating selected levels of challenges.</li> </ul>
G6	Are the reflection and collaboration process between the mentor and user effectively working?	<ul style="list-style-type: none"> <li>- Understand to which extent the environment conditions promote analytical thinking, perseverance, collaboration with the mentor and/or peers?)</li> </ul>
G7	Are the Mentorspace processes resulting in user proficiency	<ul style="list-style-type: none"> <li>- Understand whether the activities help participants build a skillset</li> <li>- Identify which skills users acquire over time with particular project assignments and how the process can be optimised</li> </ul>

Table 3.1: Map of Measurements explanatory table

### 3.4 Instruments: How do we measure?

The system architecture is entirely based on the identified metrics we want to measure (Figure 3.2). It acts as a tool for data collection.

In this section, we attempt to answer the second research question of this thesis:

**RQ2** Which instruments can generate quantitative data in a physical STEAM collaborative learning environment, in an ethical manner?

In other words, we want to explore how would the two subsystems, LSB and IN, look like?

### 3.4.1 Methodology

The methodology followed to design the system is described by Karl T. Ulrich and Steven D. Eppinger in industrial product design and development (2016). The method consists of identifying the user needs and wants from which intermediate statements are derived. The intermediate statements present a high level view of the system requirements. These requirements are then used to map the initial system specifications that picture the ideal system in the absence of constraints such as time and resources. The idea is to then try to achieve the initial system specifications with a prototype. The resulting prototype then builds the picture of the final system specifications that take in consideration the challenges and limitations encountered during the prototyping process.

An iterative design process is used to refine the system throughout the implementation phase in order to achieve the highest possible system quality within the time limit of the thesis project.

### 3.4.2 The Learning Space Box

In the previous section, we described the LSB as the part responsible for recording the interaction between the participant and the educational project being built. The LSB also takes part in the interaction by providing feedback to the participants and mentor about the project progress. The LSB acts both as an observer and a resource.

#### 3.4.2.1 LSB intermediate statements

Intermediate requirements of the LSB system are established as follows

Needs and Wants	Intermediate Statements
Time per Step/Time to fix a fault	LSB will measure time
Number and type of faults	LSB will specify tests to perform on a DUT LSB will allow the user to select tests for the DUT LSB will record and report number and type of faults
Number of attempts to fix a fault?	LSB will record the number of attempts to fix a fault
What was done to fix a fault?	LSB will detect which actions were taken to repair the DUT
Which type of online resources were used and when?	LSB will direct participants to useful online resources resources
Which type of human resources were used and for how long?	LSB will record the proximity and orientation of Mentorspace human resources with respect to participants in the space
Who fixed the fault	LSB will detect who repaired the DUT
Fun facts about the project	LSB will show extra information about the project
	LSB will communicate with the DUT
	LSB should be portable
	LSB UI should be intuitive and easy to use
	LSB must have a good battery life
	LSB data must be secure
	LSB should be engaging
	LSB must keep the user's identity unknown

Table 3.2: Immediate needs and wants for the LSB system

From table 3.2, we now imagine the LSB system to be a device that is constantly connected to the **device under test** (DUT) being built, in order to collect the necessary data stated in Figure 3.2. The LSB captures the data through its internal circuit board that we call the Wobble board (WB). Therefore, the scenario is pictured as follows:

1. During the building phase, WB is connected to the DUT through the LSB interface. In this particular use case, the DUT is the BubbleNode board.
2. The DUT and the WB work together to continuously assess the internal state of the DUT while it is being built.
3. Based on the DUT internal state, the WB then provides appropriate feedback or recommends resources to the user through the LSB interface when requested.

Besides an intuitive user interface, the WB should have a long battery life. Moreover, It is necessary that the whole process keeps the user's identity

unknown.

### 3.4.2.2 LSB initial specifications

From the intermediate statements, we derive the following initial system specifications

Intermediate Statements	Initial System Specifications	Final System Specifications	Unit
LSB will measure time	Real time clock, resolution of 0.1 Accuracy of 100	Real time clock, resolution of 0.1 Accuracy of 100	s parts per million
LSB will specify tests to perform on a DUT	LSB provides test vectors to the DUT	LSB provides fixed test parameters to the DUT	list
LSB will allow the user to select tests for the DUT	LSB has a capacitive touch screen 2 rows x 20 columns of letters	LSB has an LCD display  LSB has a menu select device  LSB has a menu select button	list  quadrature encoding binary
LSB will record and report number and type of faults	LSB will record test vector results from the DUT	LSB will record fixed test parameter results from the DUT	list
LSB will record the number of attempts to fix a fault	LSB communication channel to DUT	LSB communication channel to DUT	asynchronous serial
LSB will know which actions were taken to repair the DUT	LSB records all step data from DUT	LSB records all step data from DUT	list
	LSB has non volatile memory of 65 535 bytes	LSB has non volatile memory of 32 767 bytes	list
LSB will direct the participant to useful online resources	LSB communicates with a database of information using A2211 wireless		list
LSB is portable	LSB width=44, length=63, height=15 LSB weight= 200	LSB width=44, length=63, height=15 LSB weight= 200	mm g
LSB UI is intuitive and easy to use	LSB user interface is a menu base	LSB user interface is a menu base	list
LSB has good battery life	LSB battery life is 2000 LSB batteries are 3AA Alkaline batteries	LSB battery life is 2000 LSB batteries are 3AA Alkaline batteries	h list
	16 bit low power MSP430 microcontroller	16 bit low power MSP430 microcontroller	list
	LSB has an adjustable frame for DUT	LSB has a fixed frame for the DUT	list
		LSB is in various bright colours	list
	LSB communicates with IN using universal serial bus (USB)		list

Table 3.3: LSB Initial and final specifications

When connected to the DUT, the WB provides test vectors. Based on the test vectors, the DUT conducts a series of diagnostic tests that measure signals at specific points on its circuit board. At each assembly step, the DUT calculates an identical fixed number of such points and reports the resulting data back to the WB for analysis. The DUT does not know which

assembly step the participant is building or the significance of the data itself. Based on the raw data reported by the DUT, the WB extracts information to answer the questions set on the map of measurements. The WB also keeps track of the serial number of the DUT. This means that it is easy to identify which collected data belongs to which DUT (in case the participants share the LSBs) while keeping the participants' identity unknown. The LSB has a non volatile memory, which means that all the collected data is locally stored and safeguarded against power cycling or disconnection with the DUT. Using the current circuit design of the WB, the LSB battery life is calculated to hold 2000 hours, which is approximately equivalent to 83 days. This guarantees the LSB functionality for several events.

The user interface comprises a menu-based touch screen. The user will be given the option of either requesting feedback about the DUT at a given time, requesting more resources to help them solve a particular problem, or showing interest in learning more about the project with "Fun Facts".

With respect to making the LSB engaging, we chose to focus on a useful aspect that was witnessed during the Mentorspace educational workshops, for the time being. From the team's experience, we found that it is often challenging for the user to keep the DUT board steady when soldering parts together. Thus, the LSB will allow the participant to use it as an adjustable supporting platform for the DUT. The LSB will have a frame that can be modified to fit the DUT shape for different activities. In the context of this thesis, the DUT frame is ring-shaped to fit the BN circular circuit board.

### 3.4.2.3 LSB final specifications

Due to the limited time allocated to achieve the thesis objective, we were restricted to reduce the prototype to a more limited minimum viable product (MVP). Therefore, in line with the available resources, allocated time and budget, we carefully selected the features to be implemented in the short term to deliver a fair final proof of concept. The final system specifications are shown in Table 3.3.

In view of the fact that the experiments are being carried out on the same basic version of a BN, we found it unnecessary at the moment to invest time in developing test vectors. For the purpose of this prototype, we speculate that it is unnecessary to conduct exhaustive tests on the board. Test vectors are therefore regarded as a future enhancement and will not be implemented at this stage. The WB will supply the DUT with fixed, specific test parameters.

With regard to the LSB recommendation system, it is considered to have

the least priority at this level. Other changes such as the display, memory capacity and adjustable frame were based on the budget and the accessibility of the resources within the timeframe.

The WB circuit is designed to support DUTs other than the BubbleNode through adaptation of the firmware. This implies that it is possible, at this early stage, to use the same circuit board for various scenarios involving embedded system projects with a certain range of complexity.

### 3.4.3 The Interaction Network System

The interaction network, as described above, is the part responsible for investigating how participants interact with other individuals and with the Mentorspace resources in order to achieve intellectual growth.

Interaction is a complex concept that is understudied. A few papers provide insight into the philosophy of interaction and attempt to distinguish between types of interaction. However, It is still challenging to solidly identify what constitutes interaction itself from observable actions. In this research, the scenarios that are construed as a type of interaction that is desired to be recorded during the Mentorspace activities include not so obvious bodily actions. An example of such scenarios is a mentor standing next to their mentee and watching their building process while mentally or verbally supporting them or assisting them.

In order to encompass all possible scenarios of what we view as interaction, we choose to build upon Webster's definition and define interaction in the context of this study as "*a kind of action and/or influence that occurs as two or more entities have an effect upon one another*". Human-human interaction is thus the state of one or more individuals having any effect (verbal, physical, or psychological) on one or more individuals. Similarly, human-machine interaction is the state of one or more individuals having any effect on one or more machines and vice versa.

#### 3.4.3.1 IN intermediate statements

Based on the map of measurements (Figure 3.2), intermediate system requirements for the interaction network are identified as follows:

<b>Needs and Wants</b>	<b>Intermediate Statements</b>
When, for how long and who interacts with who/what?	IN will detect the distance and orientation of a person with respect to another person or object IN will detect the real-time position of Mentorspace resources in the space
How and for how long was the project used	IN will time stamp actions
Which type of physical documentation is used and when?	IN will record which Mentorspace non-human resources the user comes in contact with and when
Which type of online resources are used and when?	IN will record traffic towards online resources
	IN has to be invisible to the user
	IN will keep the user's identity unknown

Table 3.4: Interaction Network System Immediate Needs and Wants

We speculate that it is sufficient to acquire the position and orientation of individuals at a given time with respect to other individuals and non-human Mentorspace resources. The intermediate statements are open to expand in the future to encompass deeper levels of interaction details, if deemed necessary, while protecting the individual's privacy.

In addition, IN leverages from the Mentorspace server to collect data about the online traffic of the participants without revealing individuals' identity. The goal is to extract information about the potential knowledge gap in the provided resources and the level of engagement of participants.

From a user perspective, the IN is almost invisible and individuals in the space will not need to wear or carry anything that might alter the natural flow of interaction.

### 3.4.3.2 IN initial specifications

From the intermediate statements stem the following initial IN system specifications:

Intermediate Statements	Initial System Specifications	Final System Specifications	Unit
IN will detect the distance and orientation of a person with respect to another person or object	IrDA compliant Infrared transceiver	not implemented	list
	Badges with a 16 bit low power MSP430	not implemented	list
	The badges communicates with LSB using USB microcontroller	not implemented	list
IN will detect the real-time position of the Mentorspace resources in the space	2AA batteries per badge	not implemented	list
		Chair cushion format switch buttons	binary
		Width = 20, length = 20, height = 6	cm
		minimum weight = 20	Kg
		Chair switch controller uses USB to communicate with the rest of the IN	list
		Switches use 16 bit MSP430 microcontroller	list
		RFID compliant EM4001 Strandard	list
		RFID operates at 125	KHz
		EM4001, 40 bit tags	list
		Antennas spaced at approximately 15 on x and y axes	cm
		Antenna's diameter= 12	cm
		8 multiplexed antennas per reader	list
		Antenna coil inductance = 1.337	mH
		Antenna wire thickness 30AWG	list
		Forward link budget=10	cm
		Positional resolution =15	cm
		IN system controller: Lenovo yoga Thinkpad, Type 20CD-000MMS,	list
	OS Linux Debian version 8		
	All data is kept locally on laptop	All data is kept locally on laptop	list

Table 3.5: IN initial and final specifications

The technologies we chose to use for tracking the position and orientation are infrared (IR) and radio-frequency identification (RFID).

RFID is used to locally track the real-time position of individuals while they are moving in the space. The RFID reader uses antenna multiplexing to continuously detect the RFID tags carried by individuals in their badges. The position of the tracked individual/object is deduced by associating the position of the antenna in the space and the tag identifier it detects at a given time.

Individuals and machines are also equipped with IR transceivers that are the size and shape of a usual badge worn during Mentorspace activities. IR transceivers communicate when they are within a certain distance range and angle. IR communication will add an essential dimension regarding the possible scenarios of interaction taking place when a person is standing in a certain area of the space. For example, when detecting that a mentor is standing next to a participant, it is difficult to know whether they are interacting with the mentee and the technology they are building, or speaking to another individual across the room on the opposite side, if we do not

have the orientation data. Such information is important for a correct data interpretation.

This IN system design only differentiates between individuals that are mentors, Mentorspace staff, or participants. The identity of individuals is therefore kept unknown.

### 3.4.3.3 IN final specifications

Similar to the LSB, we were significantly constrained to narrowing down the prototype due to the time required to implement all the specifications. For this reason, we chose the IN prototype features that can best be combined with the already selected LSB prototype features, in order to demonstrate a fair proof of concept.

In the context of the thesis, the interaction network only tracks the real time position of individuals in the space. We have realised that there are various potential technologies that could fulfil the requirements. After experimenting with different methods discussed in the implementation chapter, we settled on the final system specifications stated in Table 3.5.

The following figures demonstrate the flow of communication between different components of the LA system. Figure 3.3 represents the system in accordance to the initial system specifications to be implemented in the near future. Figure 3.4 represents a picture of the system implemented in the context of this study.

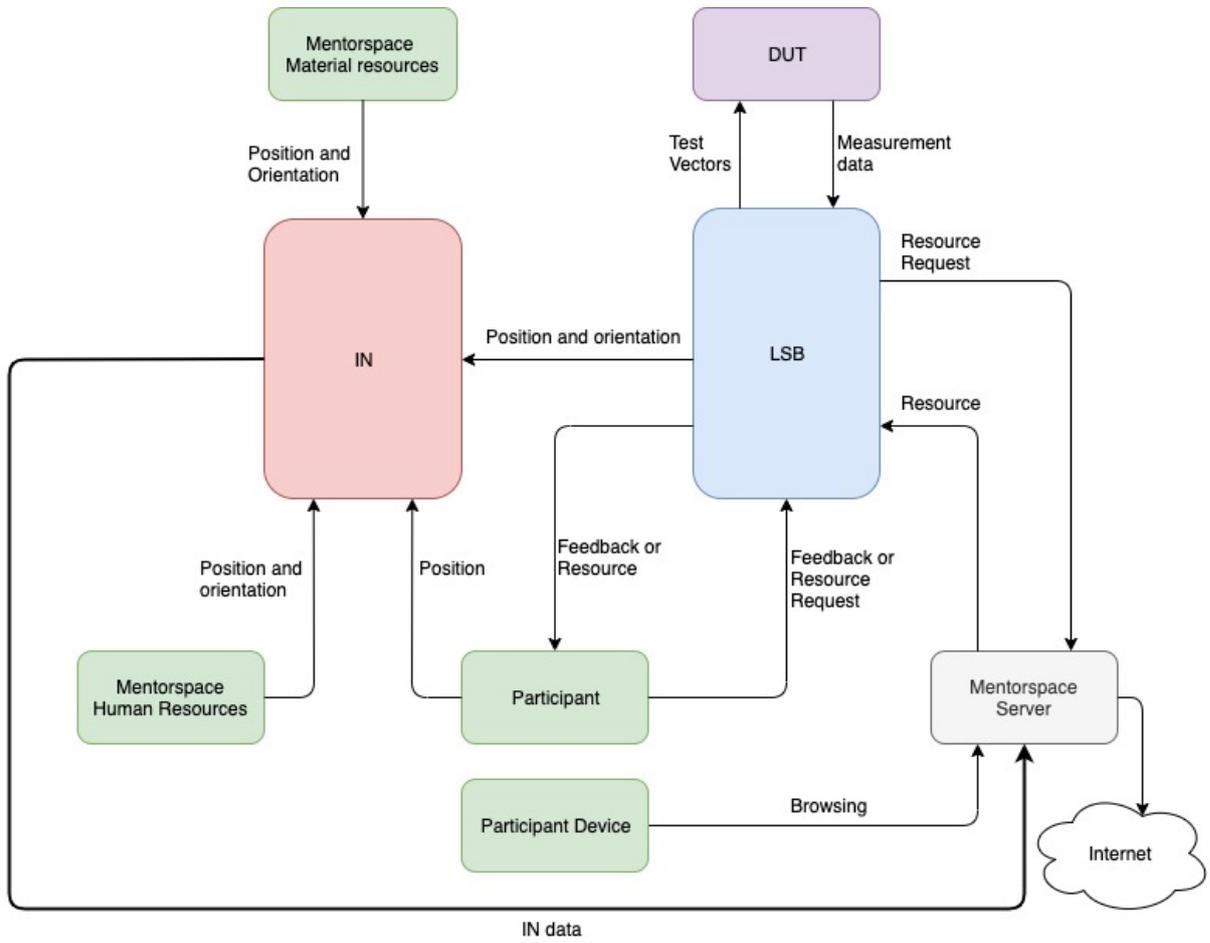


Figure 3.3: Initial LA system diagram

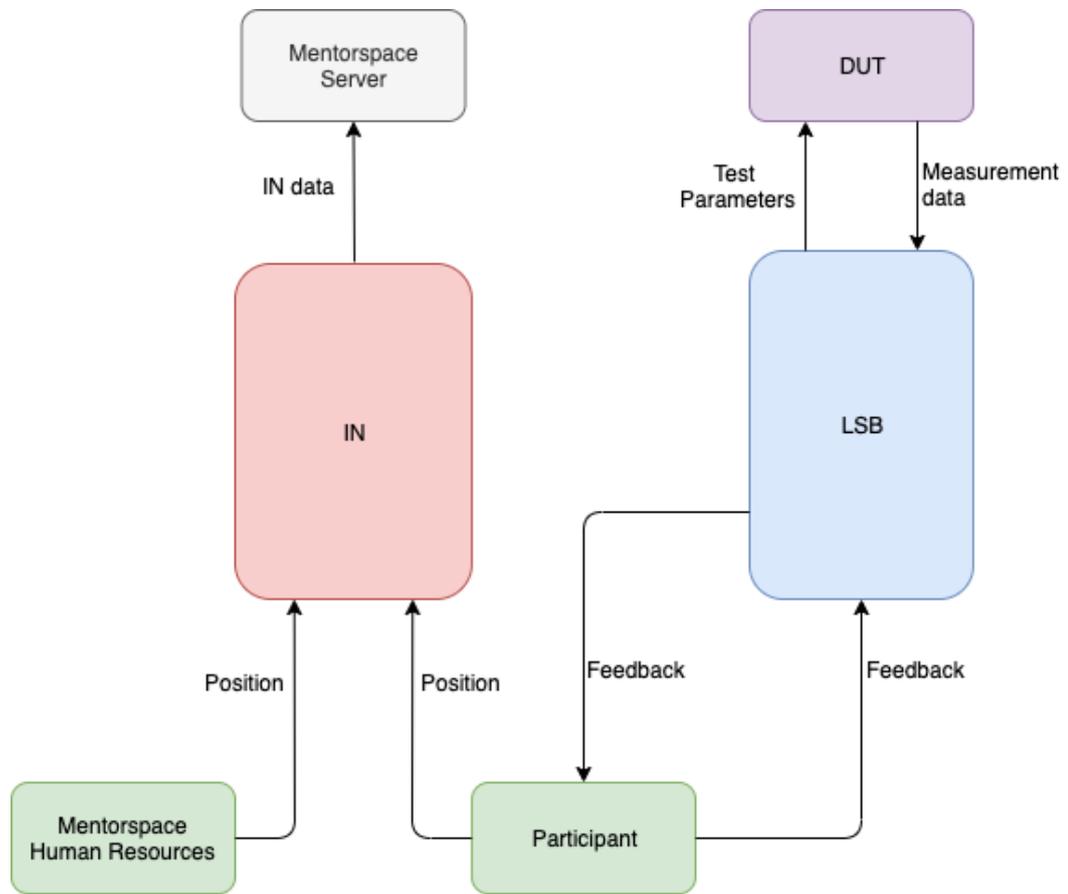


Figure 3.4: Final LA system diagram

# Chapter 4

## Implementation

The objective of this chapter is to provide an overview of the implementation process of the LA instruments used to collect the adequate data from the learning environment.

The process to build the system was divided into two phases as follows:

**Phase 1** This phase concerns the implementation of the first LSB prototype, and testing both the technology and usability during the first experiment.

**Phase 2** This phase comprises LSB enhancements based on the feedback from the first experiment, the implementation of the IN first prototype, and the testing of the full system during the final experiment.

Due to some unforeseen challenges, team members were reduced to two individuals, myself included, during the establishment of the system specifications. Consequently, it was necessary to adapt the thesis scope to encompass logistical tasks as well.

### 4.1 Phase 1

#### 4.1.1 LSB Implementation

##### 4.1.1.1 Building the WB

After a few weeks of circuit design and testing, the complete WB circuit schematic was carefully designed using KiCad free software based on the system specifications stated in Table 3.3. The schematic was then used to

mill out the first WB prototype on a copper clad PCB using the Mentorspace milling machine. Subsequently, the prototype board was manually built and tested along with the WB software.

The WB software allows it to communicate hand-coded test parameters to the DUT via asynchronous serial communication, receive signal state parameters from the DUT and execute an algorithm to determine the DUT status at a given time. At this stage, the data reported by the DUT depends on two things; the steps completed by the user on the DUT and the faults existing on the board as it is built. The role of the WB is to understand the data reported by the DUT with respect to the faults, and provide feedback to the participant and the mentor through the LSB interface indicating the region of the DUT in which a fault occurred. Another piece of software is embedded in the DUT's microcontroller. This part ensures the reception of the test parameters from the WB, the execution of an algorithm that performs the tests required to determine the DUT signal states, and sending them to the WB where they are converted into meaningful information.

For the moment, the WB is able to perform a diagnosis on the DUT at the end of each step of the project only. We define a step as *"an arbitrary delimitation of the project building process"*. The BN building process is divided into five functional blocks (Appendix A1 and Appendix A2) that need to be finished in order. The goal of this division is to ensure that the participants solder the parts together in a meaningful way. This helps them understand how a certain combination of electronic components results in a function such as data transmission, data reception, power, etc. The steps are designed as follows:

**Step1.** The WB checks the power section of the BN. This step involves the battery, the power switch, and the bypass capacitors. At the end of this step, the WB is able to check whether the right power is available to the DUT.

**Step2.** The WB checks the processor section of the BN that involves the microcontroller component. At this stage, the WB can only check whether the component works and communicates. It does not yet diagnose pin shorts.

**Step 3.** The WB diagnoses the three color LED section of the BN. The WB checks for the direction of the LED but it does not yet diagnose pin shorts.

**Step 4.** This step concerns the infrared transmission section. At this stage, it is only possible to diagnose transistor pin shorts related to the base resistor.

**Step 5.** The last step consists of the infrared receiving section. For this phase, the WB is programmed to diagnose short circuits related to the gain setting resistors, and whether the OpAmp is placed in the right direction.

A total of six LSBs are manually built for the first experiment; four for the participants and two additional LSBs as a backup.

#### 4.1.1.2 LSB Interface design

The LSB interface comprises a display through which the participant interacts with the menu and receives feedback, a rotary encoder which serves as a selector, and a button to validate the selected function on the menu. In addition, the LSB is equipped with a ring-shaped DUT frame that acts as a support for the participant when soldering their project together.

The user interaction flow for this phase is as follows:

1. The LSB is booted and requests the user to select the appropriate step to test.
2. The participant uses the rotary encoder to scroll through the existing step numbers.
3. The user selects the desired step number by pushing the button.
4. The LSB requests the user to connect the LSB to their BN board, turn on their BN and push the LSB button again.
5. Following the user action, the LSB displays the appropriate feedback for the chosen step.

In this phase, the algorithm for user feedback is implemented only for Step 1 due to time constraints. Step 1 concerns the power functional block of the BN consisting of two capacitors, a battery holder, batteries, and a switch. For the remaining steps, the LSB is programmed to provide a positive default feedback.

The ring-shaped DUT support was designed, 3D printed and tested in the Mentorspace. However, due to the tight time schedule, the LSB boxes in

this phase were bought at an electronics store and manually modified to be equipped with a display, button and rotary encoder.

## 4.1.2 Experiment logistics

### 4.1.2.1 Experiment Environment

In line with the Mentorspace spirit, the first experiment is an intervention at the Järvaveckan event in Stockholm. Järvaveckan is an annual gathering held at Spånga Idrottsplats. The goal of the event is to bring together communities, politicians and social actors from different industries to raise awareness about social issues and make the various parties work together to prioritise these issues on the political agenda.

The Mentorspace was called to represent KTH in the event. We responded to the request by suggesting a workshop about building a BubbleNode. Our intervention was scheduled to take place on June 16th from 3pm to 8pm. The workshop was a free drop-in event welcoming youth from all backgrounds.

### 4.1.2.2 Participant manual

The target audience for this workshop are youth who are beginners with no required engineering experience or scientific prerequisites. Therefore, the manual (Appendix A1) was designed in a way that divides the construction of the project into relevant blocks of tasks that are described in a concrete manner, given the level of guidance needed by the target participants (Jonassen, 1991a).

During the activity, the participants are provided with a BubbleNode kit, soldering material, and the manual. The manual contains a drawing of the BN board and the electronic components needed for each step. It puts a name on each component and helps locate where the components should be soldered on the board. The manual in this phase is mostly visual and contains very little text given the audience age diversity.

### 4.1.2.3 Mentor Training

The mentors who volunteer to participate are trained on the project to build and the event schedule prior to each activity.

Three days before the event, the mentors were invited to a training session of a couple of hours. During the session, we went through the event content

and schedule and mentoring tips. In addition, mentors were given the opportunity to build a BN using the LSB prototype in order to go through the same experience as their mentee and build expertise themselves. Finally, the training ended with a discussion about what mentorship means, the level of guidance to be provided during the workshop, and a feedback session about the utility of the LSB.

## 4.2 Phase 2

### 4.2.1 LSB enhancements

#### 4.2.1.1 WB enhancements

Building upon the feedback received from the first experiment, further diagnostic operations are hand-coded during this phase. Table 4.1 shows the WB algorithm enhancements implemented compared to the first phase.

It is important to note that the diagnostic tests performed in the context of this thesis are limited, in view of the fact that they are hand-coded and focused on the most probable and common faults observed during the workshops. In future enhancements, test vectors will provide a higher level of precision and include a larger pool of faults that may not be obvious.

<b>Step</b>	<b>Phase 1</b>	<b>Phase 2</b>
1	Battery on	Battery on/ Battery cable direction
		Capacitor and switch on
		Pin shorts
2	Microcontroller on Microcontroller communication	Microcontroller on
		Microcontroller communication
		Microcontroller direction
		Microcontroller pin shorts
3	LED on/ LED direction	LED on/ LED direction
		LED pin shorts
4	Transistor one pin short	Transistor all pin shorts
		Transistor direction
		IR LED direction
		IR LED pin shorts
5	OpAmp direction	OpAmp direction
		OpAmp pin shorts
		Phototransistor direction
		Phototransistor pin shorts

Table 4.1: WB test operations on the DUT, Phase 1 and Phase 2

#### 4.2.1.2 LSB Interface enhancements

In order to improve the LSB interface and increase the prototype robustness, adapted boxes are designed and 3D printed for the LSB components to be tightly sealed together.

With respect to the menu based UI, user feedback is implemented for the rest of the steps in this phase (Table 4.2). For successful steps, the LSB displays a positive feedback : 'Looks Great! Good Job!'. The goal of the user feedback is to push the participant to investigate the fault by indicating the area in which it occurs without giving away the answer. We aim at helping the user focus their attention and resources on the right problem while saving them time on unnecessary troubleshooting. Further research on the best level of feedback to be given in order to optimise the learning

opportunity is necessary.

More functions are implemented in the LSB menu such as the possibility to reset the date and time and the ability to check the battery percentage of the LSB at any given time.

Step	User Feedback
1	No power or BN
2	Check processor pins or direction
3	Check 3 color LED pins or direction
4	Check transistor pins or direction
	Check transistor and LED pins
5	Check phototransistor pins or direction
	Check phototransistor and opamp pins

Table 4.2: LSB displayed user feedback

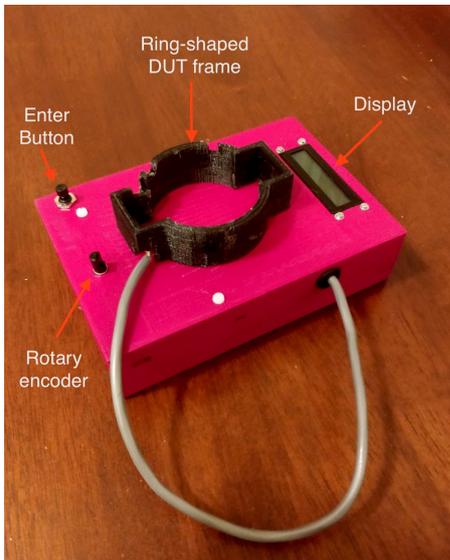


Figure 4.1: Final LSB prototype components

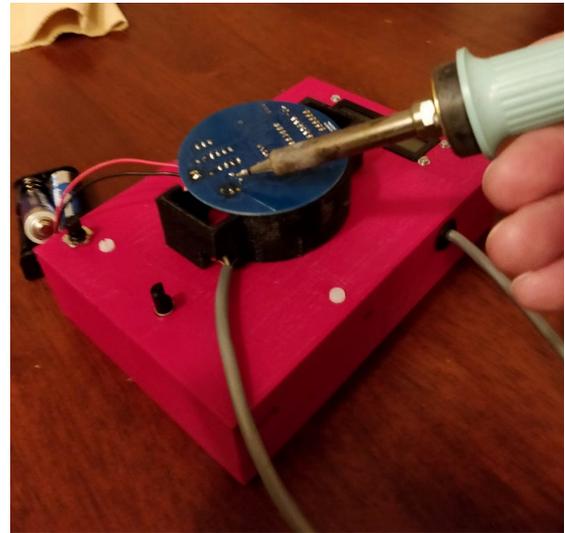


Figure 4.2: Final LSB prototype with a BN board being soldered

## 4.2.2 IN implementation

The interaction network in the context of this thesis consists of monitoring individuals' real-time position during the experiment. This section describes the prototyping process that led to the final IN system based on the system specifications stated in Table 3.5.

### 4.2.2.1 Cricket

We first looked at existing indoor positioning systems which may potentially be adapted to our application and vision. Crickets, a system developed at the Massachusetts Institute of Technology (MIT) appeared to be an enticing prototyping choice.

Cricket is an indoor positioning system which offers location information in the form of space identifiers, position and orientation. Cricket is a group of beacons that are mounted on the ceiling of the space to be monitored. Such beacons periodically transmit location information using an RF signal. They also emit an ultrasonic pulse which enables listeners, that are attached to the target devices to be monitored, to estimate the distances to nearby beacons. Distance measurements are then used to estimate the listener's location. The system promises to also provide indoor orientation via the Cricket compass. However, it is not clear whether the latter is commercially available.

Cricket appeared to be a suitable technology to prototype the IN. However, Sameshima and Katz (2009) from Carnegie Mellon University have thoroughly documented their experiments with Cricket and ultrasound technology. The paper reports a distance short of 5 to 9 percent with the error increasing proportionally to the distance increase between a cricket and a listener. The study states that ultrasonic sensors like crickets are sensitive to noisy environments. Another reason that makes crickets not suitable for our application is the fact that people and objects act as an obstruction to the crickets: *'If two people come together for a conversation, the mere presence of another person might block a clear view of the listener network. Similarly if an object equipped with a cricket goes below a table, it essentially disappears from sight.'*

From the study, it appears that the IN application requires a level of accuracy that crickets cannot offer. Therefore, we chose to not invest time

and resources in moving forward with this option.

#### 4.2.2.2 High Frequency RFID v.s Capacitive system

We proceeded with exploring two positioning system ideas that can quickly be prototyped. The first approach is to use a 13.56 MHz RFID system. The antenna coils of such a system are deployed on the floor on which the mentors would walk while carrying an RFID tag in the internal bottom of their shoe. The antenna that detects the tag would report the tag indicator while the system records the reporting time.

The second approach is the use of capacitance. The system consists of deploying capacitive plates to cover the surface area of the experiment floor. On the other hand, the mentors wear a metal foil or a flexible plate on the internal bottom of the shoe. The detected voltage variation that is due to capacitance variation resulting from the superposition of the shoe plate on a floor plate will indicate the presence of a mentor.

Table 4.3 shows a rough evaluation of the pros and cons of each system before prototyping.

In order to help decide on the technology to use, quick prototypes of both systems are implemented. With respect to the high frequency (HF) RFID system, we encountered challenges related to finding the right capacitance value for impedance matching. It was not possible to achieve a high enough inductance, which significantly limited the forward link budget range.

Concerning the capacitive system, the prototype appeared to be overly sensitive to the shoe material and the blood chemical composition of the individual which is hard to accurately predict. The prototype gave significantly different values depending on the mentioned factors. At this stage, it appears that the capacitive prototype is too unstable. Therefore, we consider that it is the least reliable option.

	HF RFID	Capacitive plates	Favoured technology
Underlying physics	Inductive transfer of magnetic field using an inductive (coil) antenna	Voltage representing charge between plates (capacitor)	Uncertain
Limitation of coupling	Strength of the magnetic field which is proportional to the number of turns of the antenna wire. Limited by area covered by the multiple turns.	Size of effective capacitance formed between a person's shoe and a plate on the floor. Limited by the low frequency of data transfer (about 1.2Khz)	Capacitance
Switching	The system requires a matrix of antennas. The antennas are switched using fast recovery diodes.	The system requires a matrix of floor plates. The signal induced in the plates needs to be buffered with isolation amplifiers. The outputs of the amplifiers are switched using low resistance analog switches.	RFID
Ease of implementation	Coils require lots of magnet wire (labor intensive to build). The large inductance of the coils requires a complex impedance matching network to be designed. The matching network will need to be implemented such that it can be adjusted to account for environmental factors (i.e. floor composition, shoe composition)	Capacitor plates are easy to construct (shoe insert and a floor plate). The signal buffering and switching solution takes several hours to design but is easily replicated.	Capacitance
Reliability	Impedance matching may need to be adjusted for all antennas in the matrix for every event location depending on the floor composition. Risk of antennas being damaged or detuned during an event.	Small capacitance makes signal detection difficult. This may be solvable by using signal reconstruction methods based on looking for signal transitions rather than signal states. It is unknown whether the environment introduces other significant noise sources. The effect of plates near the human body on the signal performance is not yet verified.	Uncertain
Cost	Estimate for individual antenna cost is about 4 euros (wire, transmission line, components for the matching network). The cost for the RFID reader is about 50 euros.	Estimate for individual capacitance plate is about 2 euros (capacitive plate material, buffering circuits, analog, switches, and wiring). A digital processing board is needed.	Uncertain
User factors	Easy to use. The user only needs to put an RFID tag on the inner bottom of one shoe. The system uses 13.56 Mhz non-ionizing RF radiation, mostly in the form of a time varying magnetic field.	User needs to put metal foil or flexible plate on the inner bottom of one shoe. The user needs to wear active driver electronics including batteries either on top of the shoe, or around the user's ankle, which might affect the ease of walking. The system does not use RF energy.	RFID
Time to implement	A conservative time estimate to build each antenna with a matching network is about 1 hour. Extra time is needed to build and test the antenna switch and the matching network. A conservative time estimate to build a switch that can accommodate 4 antennas is about 3 hours.	A conservative time estimate to build analog switches and signal buffers for a switch that can accommodate 4 floor plates is about 3 hours. A conservative time estimate to build a digital processing board to run the system is about 4 hours per board. The number of digital processing boards is to be determined.	Uncertain
Ease of deployment	Antenna assemblies need to be arranged on the floor while connected to the switching circuits. The antennas may be affected by the floor material or material under the floor such as steel reinforcement or beams. The coils are made out of very thin wire and can be damaged in transit. The antennas may be perceptible to people walking on them.	Capacitive plates need to be arranged on the floor while attached to the switching circuits. Plates can be designed in a way where floor material does not affect the operation to a large degree. Hard to damage in transit. The plates are very flat, and hard to feel when walking on them.	Capacitance

Table 4.3: High Frequency RFID v.s Capacitive system comparison

### 4.2.2.3 Low Frequency RFID

Finally, we decided to look at low frequency (LF) RFID systems that are commonly used in the food industry. Due to its low frequency (125 KHz), the LF RFID system prototype allowed the achievement of the needed inductance (1.337 mH). To attain this result, an antenna coil of 71 turns with a diameter of 12 cm is manually built using magnet wire. The rest of the RFID prototype circuit is built using an ID-3LA RFID reader. The forward link budget range achieved with this system prototype is 10 cm. This means that the mentor foot can be detected at 10 cm distance from the antenna, which is more than enough for our application.

Due to resources and time constraints, only four ID-3LA RFID readers were available for the experiment. Each RFID reader can handle eight antennas. However, during the testing process, an accidental short circuit damaged one of the RFID readers. Therefore, we were constrained to use only three RFID readers for the experiment. 24 RFID coils were manually wound and tuned using a Fluke automatic LCR meter type PM6303A.

The system was tested all together at the office building. Prototype tests show that the system could successfully detect the tag in my shoe across all the antennas. In order to ensure no negative surprises on the experiment day, the antennas were additionally tested at the experiment venue a week prior to the event. The floor composition of the venue did not affect the forward link budget range as it successfully detected the tag in my shoe at a range of 10 cm.

From the participant's point of view, the system on the floor is covered by two rugs with a dimension of 120x180 cm each. The system will be near invisible to the individuals in the venue. Each coil was thoroughly sealed with duct tape to increase the prototype robustness. Given the thickness of the magnet wire used (0.2 mm) and the number of turns of each coil, the antennas are hardly perceptible to people walking on the rug.

The deployed system covers an area of 240x180 cm.

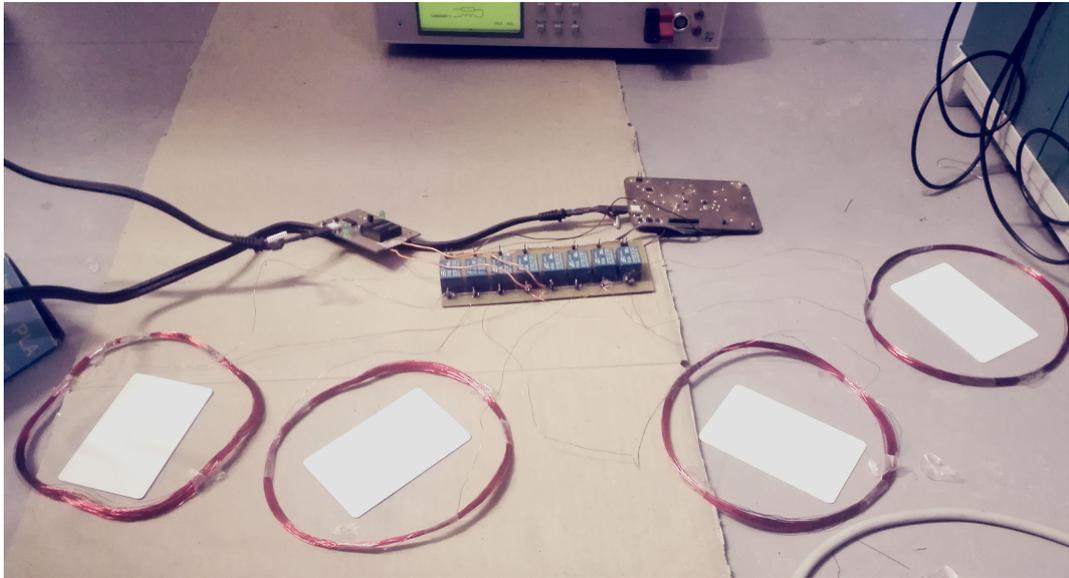


Figure 4.3: LF RFID system prototype testing with 4 antennas

#### 4.2.2.4 Cushion switches

In view of the absence of IR transceivers in the context of this thesis, a simple prototype system was implemented to automatically detect the presence of a participant at a given time. During the experiment, participants spend most of their time on their chairs to build this project. Therefore, a simple switch circuit that connects to the rest of the IN system is integrated in the chair cushions.

The switch detects any object of a minimum weight of 20 Kg putting pressure on the chair cushion. This is a limited short term solution as it does not detect the presence of a participant in areas outside the chair cushion. However, we consider that it is enough to achieve a proof of concept for the IN system.

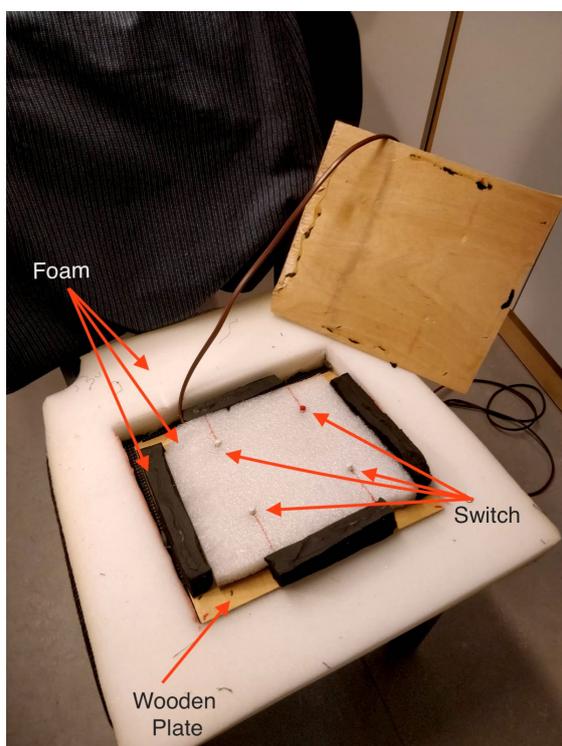


Figure 4.4: Integration of switches inside a chair cushion



Figure 4.5: User's perspective of the chair cushion

### 4.2.3 Experiment logistics

#### 4.2.3.1 Experiment Environment

For this study, we contacted Sweden's largest university of art, craft and design, Konstfack university. Konstfack offers a wide range of courses and research activities in areas of art, design, craft, education, and visual communication. They are interested in re-envisioning the relationship between the individual, the community and the environment they live in, which significantly intersects with the Mentorspace vision.

In recent years, they have been seeking to enrich their offer with interactive art courses. While discussing the possibilities of collaboration with the school, a professor explained that the majority of students feel that their lack of background in technology acts as a barrier to making their art more interactive. Therefore, the school was enthusiastic about organising a first workshop in partnership with the Mentorspace to bring together tech students and art students. We both agreed that the BN workshop was a good first step towards breaking the ice with technology and connect the

participants with KTH students in order to encourage a further collaboration and facilitate access to tech expertise.

The workshop was planned to be a free drop-in event for three hours at the art school.

#### **4.2.3.2 Participant manual**

In this phase, the participant manuals are modified to accommodate the acquired feedback from phase 1 and fit the new target group (Appendix A2).

We worked closely with a Konstfack student and professor to agree on a good way to design the manual for an artistic audience. From our discussions, it appeared necessary to add a theory of operation to the manuals. The objective behind the theory of operation is to teach the students about the possibilities that technology offers and inspire them to use it as a tool in their creations. In addition, it allows them to build basic knowledge and vocabulary in technology with which they can begin sharing ideas about combining art and technology with the mentors.

The manual keeps the same functional block division and level of task description. However, effort was put into framing the text in a more engaging way that fits the background of the art students.

#### **4.2.3.3 Mentor Training**

Our pool of mentors for this study consisted of KTH students with a background in embedded systems and human-computer interaction. In accordance with the training in phase 1, mentors were walked through the event schedule, and the objectives of the activity with respect to the art university on one hand, and the LA research project on the other hand. Mentors have engaged in building their BN using the final LSB prototype and the participant manual. They were encouraged to provide feedback about the material presented throughout the training.

### 4.3 Ethics and Privacy Protocol

This section explains the protocol about obtaining permission and consent from the involved parties to exercise the research activity and collect data in compliance with the General Data Protection Regulation of the European Union (GDPR) and local Swedish regulations.

Since the beginning of the project, we made sure to architect a system that minimises any risks of violation of the participant's privacy without constraining the research objectives. The data collected is free of any personal data and does not enable individual identification.

In order to ensure the approval of the research project, representatives of the organisations we collaborated with were contacted in advance to be granted permission to approach participants with our research project. In addition, a written informed consent is prepared for the mentors and the participants before the start of the mentor training and the activity respectively (Appendix A3; Appendix A4). For underage groups, parents or guardians are approached with the consent form before the activity. Through the consent form, the participants and mentors are informed of the research objectives, the risks of participating, the nature of the data collected, and their right to not participate.

The data collected through this research cannot be used for purposes other than the goals stated in the form. If the researchers wish to use it otherwise, they must seek renewed consent.

# Chapter 5

## Evaluation and Findings

In the previous chapters, we described the learning environment and the use case for this study. We proceeded with defining the metrics for the learners' data which we measure through the LA system. Next, we described the prototyping process of the LA system that is used for data collection. In this chapter, we will address the rest of the learning analytics cycle (Figure 2.1), while attempting to answer the second and third research questions. First, we present the data collection experiments that are carried out at the Järvavecken event and Konstfack school of art and design. The objective of the Järvavecken experiment is to perform a first iteration on the LSB prototype achieved during phase 1. This iteration helps us acquire necessary feedback on the quality and usability of the LSB to be implemented during phase 2 of the implementation process. In addition, it helps test the accuracy of the data collected and the robustness of the prototype. The experiment at Konstfack represents a second iteration on the LSB prototype after the implementation of the feedback from the first experiment, and a first iteration with respect to the IN prototype achieved during phase 2. The data collected during this experiment is used for the data analysis and interpretation from which we hope to derive useful insights for intervention (Figure 2.1).

## 5.1 Data collection experiments

### 5.1.1 Study 1: Järvavecken event experiment

#### 5.1.1.1 Methodology

The study design follows an observational design approach. This experiment is used as a first iteration to test the accuracy and robustness of the WB prototype and the usability of the LSB interface design. The feedback derived from this experiment will be used to enhance the LSB system during phase 2 (section 4.2).

Tested LSB features	Researcher notes
Time per step	Time recording everytime a participant operates an LSB.
Type of fault	Real time observed fault and course of action to solve it.
User feedback	Real time feedback provided by the LSB display.
Display, Enter button, rotary encoder functions	Usability challenges encountered by the user when operating LSB.
Ring-shaped LSB frame	Usability challenges while using the LSB ring-shaped support frame.

Table 5.1: Study observations during the Järvavecken experiment

In view of the fact that our current system does not record data during steps but at the end of the steps, the **time to fix a fault** is the sum of the time per step of all steps that contribute to fixing a particular fault. Therefore, the accuracy of the time per step implies the accuracy of the time to fix a fault in the context of the current system design. Similarly, the **number of attempts to fix a fault** is derived from the number of times the user operates the LSB until a positive feedback which indicates that the fault is fixed. The reason is because the user operates the box after each attempt to fix their board problem as a tool to validate their solution through the LSB feedback. Finally, the **number of unique faults** during the building process is easily derived from the information about the kind of fault at each step if accurate.

Therefore, if we verify the accuracy of the prototype at measuring the metrics presented in Table 5.1, the rest of the metrics indicated in blue on the map (Figure 3.2) are easily derived.

Although the LSB display feedback is only implemented for Step 1 in this study, the raw data recorded by the WB for each step is more detailed and encompasses all steps (Table 4.1). This enables us to compare the raw data to the researcher's observations.

### 5.1.1.2 Participants

A total of 30 participants attended the Järvaveckan experiment. The age of the participants' ranged from 12 to 19 years. Underage participants were accompanied by their guardians. All participants had no prior experience of building circuit boards. Of the total number of participants, 13 individuals in groups of four participated on the table equipped with LSBs.

### 5.1.1.3 Experiment setup

The organisers reserved two tents for the Mentorspace event. One table with four chairs was put in each tent. The tables were similarly equipped with soldering material to be used and shared by the participants. Only one table was supplied with four LSB devices, which is the only table considered relevant for this experiment.

30 identical BN kits and manuals were prepared for the experiment. Four mentors took part in the event. Each table was assigned two mentors who were trained beforehand (section 4.1.2.3). In front of the tents, two KTH ambassadors were assigned the role of promoting the workshop by presenting a complete BubbleNode and inviting people to build their own.

A rework station, also known as 'the hospital', was mounted in the corner of one of the tents. The hospital is a small table with advanced tools that enable advanced rework operations to be carried out in order to repair faults that could not be fixed by the mentor and their mentees.

One researcher was responsible for taking notes about the observations that will serve as ground truth for the evaluation of the LSB prototype (Table 5.1).

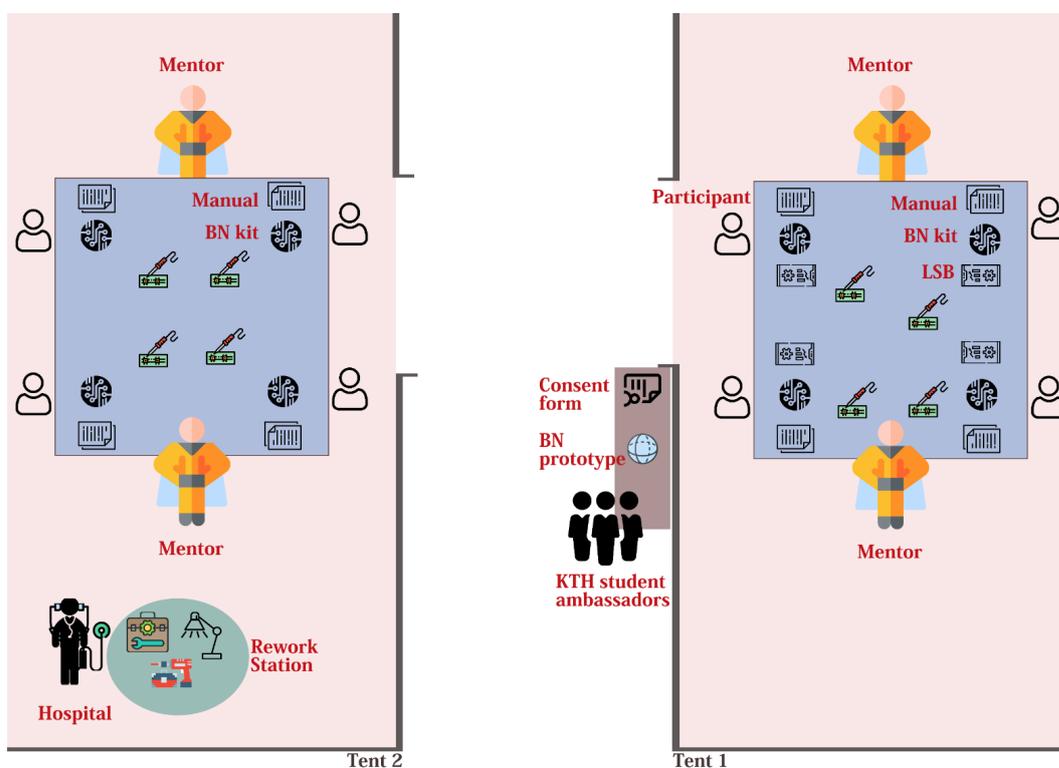


Figure 5.1: Järvavecken experiment setup

#### 5.1.1.4 Experiment procedure

The experiment was designed in a way that does not interfere with the regular flow of interaction of the Mentorspace activities. The objective is to provide insight into the Mentorspace processes. Therefore, it is important to not alter the learning environment established by the Mentorspace.

When an individual is keen to build a BN with the mentors, consent is required from them or the guardian, depending on the participant's age. The participant is then directed to take a seat at one of the tables. Priority is given to the table equipped with the LSBs as we want to maximise the data collection. Once sitting, the mentor provides the participant with a kit and a manual, and begins introducing the activity. At the end of each step, the participant is asked to use the LSB device. While the feedback is only reliable for step 1, it is important that the LSB collects data for the remainder of the steps so that we can later compare it against the

observations.

The mentor continues to accompany the mentee on every step explaining the principle of the functional blocks and component names and function. Mentors are encouraged to engage in discussions with the participant about their passions and interests. They are also encouraged to inspire the participant to question and reflect upon how the BN connects to the concepts they have acquired in school and how it can be used to achieve goals of their interest.

The mentors are asked to only intervene in the building process when the participant encounters a problem that is sufficiently complex to not be able to solve it on their own. In situations where the problem needs advanced intervention, the mentor is encouraged to communicate with 'the hospital' for support. After the participants are done building the BN, they are asked to test it among themselves with other BNs by bringing them close together to send and receive data.

## 5.1.2 Study 2: Konstfack school experiment

### 5.1.2.1 Methodology

Similar to the previous experiment, this study follows an observational design approach. This experiment aims at testing the LSB prototype enhancements (Table 4.1 and Table 4.2) and the IN system implemented during phase 2, while collecting the necessary data to be used for data analysis (section 5.2).

Tested prototype features	Researcher notes
Time per step	Time recording everytime a participant operates an LSB.
Type of fault	Real time observed fault and course of action to solve it
User feedback	Real time feedback provided by the LSB display.
Display, Enter button, rotary encoder functions	Usability challenges encountered by the user when operating LSB.
Ring-shaped LSB frame	Usability challenges while using the LSB ring-shaped support frame.
When, for how long and who interacts with who	flow of interactions between individuals with timestamps
When, for how long and who interacts with what	flow of interactions between individuals and the LSB with timestamps

Table 5.2: Study observations during the Konstfack experiment

With respect to the LSB, the same testing methodology as in the first experiment is used. In order to test the current IN system, it is enough to record which individuals are interacting with each other at a particular time, which individuals are interacting with which LSB at a given time, and where in the space the interaction occurs. These observations are then evaluated against the raw data collected by the LA system in section 5.2.

### 5.1.2.2 Participants

This experiment was attended by 16 individuals from the art school. Participants were a mix of students and lecturers from different tracks. A total of 8 participants were involved in the table with the IN and LSB devices.

### 5.1.2.3 Experiment setup

A week prior to the event, the IN system is tested with the venue floor in order to make sure to reduce the risks of detuning the antennas. On the event day, a hallway in front of the art school entrance and a room were booked for the entire day. The hallway was used for the experiment and the room was reserved for a focus group with the participants.

Two tables equipped with four chairs, soldering material and 4 LSB devices per table were placed on each side of the hall. Only one table was equipped with the IN system.

Given the fact that the current IN system does not measure orientation, the tables were placed in a way that maximises the distance between the participants so that if a mentor works with a mentee, they will be constrained to physically be at a distance range that allows the system to accurately associate the mentor with the right participant. For example, if mentor A is helping mentee 1 at a given time and they are called to help mentee 2, they will have to physically move closer to mentee 2 as the distance between the mentees is significant enough to prevent the mentor from helping mentee 2 while being closer to mentee 1. This constraint is engineered in order to minimise the IN system error with respect to associating the position of a mentor with a particular participant at a given time. Two 180x120 cm tapestry are placed on each side of the table covering the IN system.

Further in the hall, the hospital station with advanced tools is set up to help fix advanced faults.

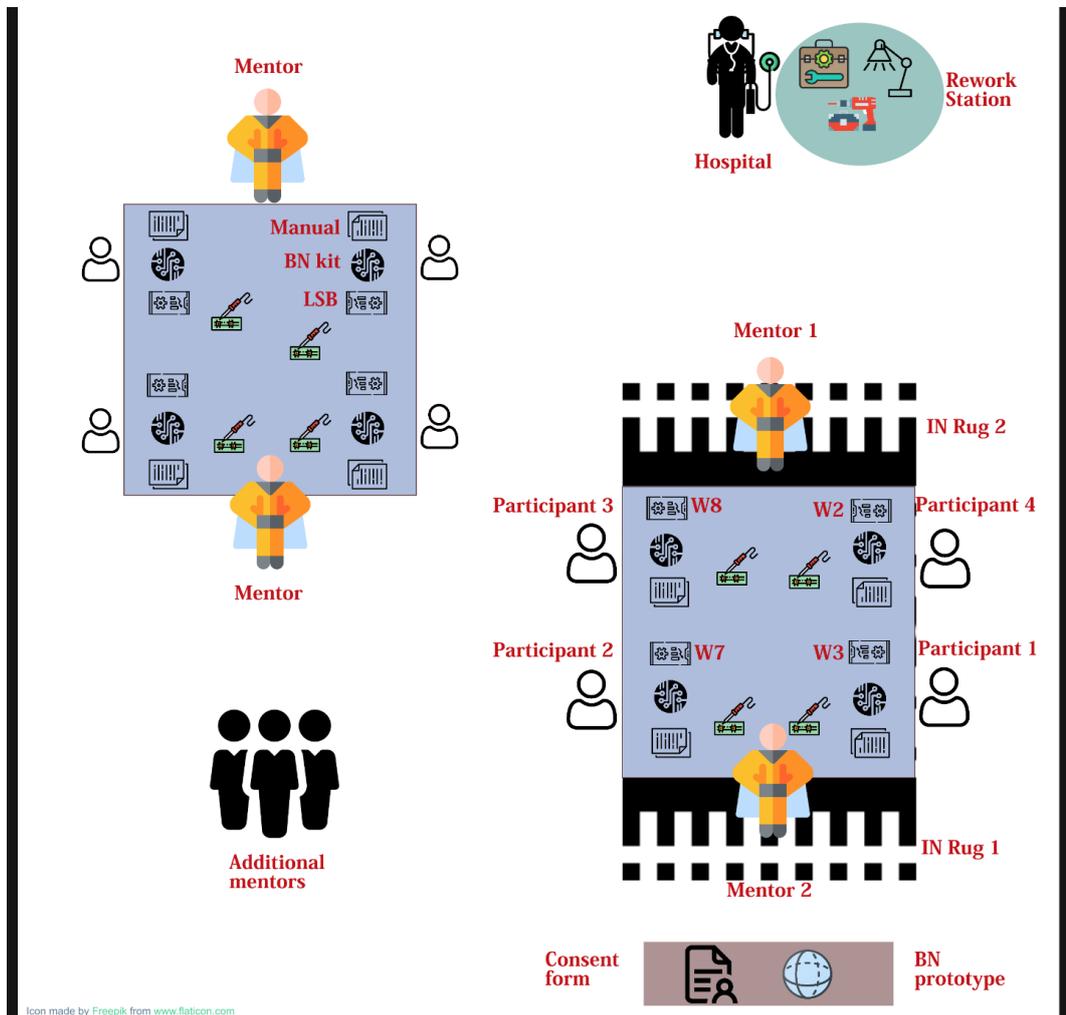


Figure 5.2: Experiment setup

#### 5.1.2.4 Experiment procedure

Before a participant is allowed in, they are asked to read and sign a consent form (Appendix A3). The participant is then directed to one of the tables, prioritising the one with the IN system to maximise data collection. Once a participant is seated, they are provided with a kit and a manual. The mentor starts engaging with the participant to begin the activity. Building upon the past experiment, mentors walk the participants through the material provided including the LSB devices and their use. Participants were encouraged to use the LSB at the end of each step and reflect upon its feedback before moving to the following step.

The only constraint in this experiment was for the mentors to remember to "always stay on the rug" when dealing with the table equipped with the IN system. The mentors were given this instruction to maximise the chances of recording a full picture of the interactions occurring around the table. In line with the Mentorspace processes, mentors were not restricted to helping specific participants and were encouraged to engage in the flow of interaction spontaneously.

During the entire experiment, two individuals were assigned to take notes about real-time human-human and human-LSB interactions related to the table equipped with the IN system. Another individual was assigned to take notes about the LSB devices behaviour and verify whether the provided feedback matches the real state of the participant's project.

After a participant is done building and testing their BN, they are invited to the focus group room.

## 5.2 Data pipeline

Real world data is likely to be incomplete and noisy. Therefore, a preprocessing step to correct inconsistent data and filter noise for an accurate analysis and interpretation is necessary (Daniele Ravi, Charence Wong, Benny Lo, Guang-Zhong Yang, 2016). In this section, the process of moving from raw data to deriving meaningful information to answer our objectives will be explained. The Cross Industry Standard Process for Data Mining (CRISP-DM) illustrated in Figure 5.3 will be used as a framework for data handling (Wirth Hipp, 2000).

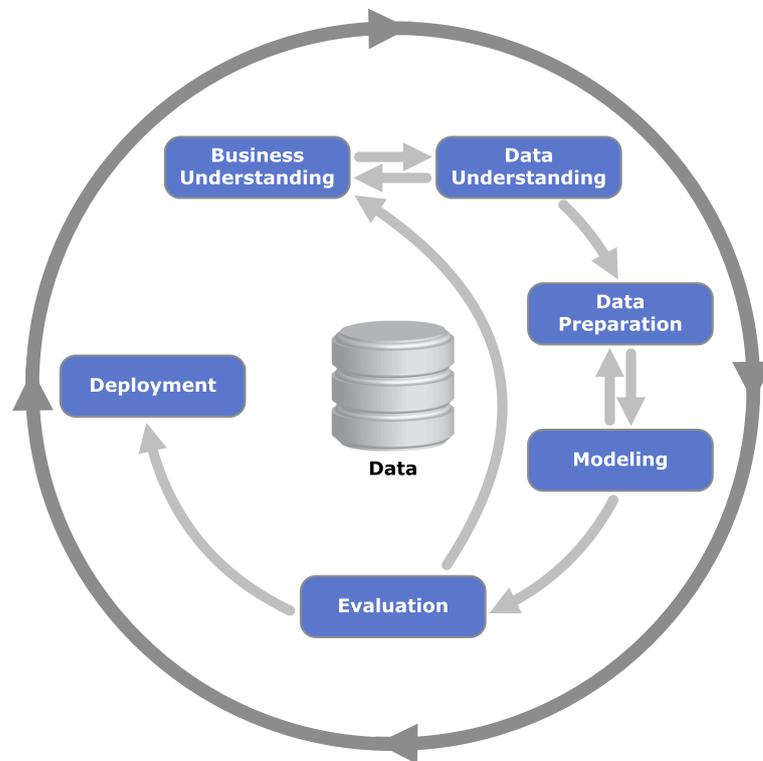


Figure 5.3: Cross Industry Standard Process for Data Mining

### 5.2.1 Business understanding

In this section we will address the business questions from the map of measurements that the Mentorspace wishes to answer with the collected data (Figure 3.2 and Table 3.1). We will then proceed with looking at the acquired data and try to understand how to properly meet the objectives.

Given the limitations of the implemented prototype, we will only address a subset of the questions outlined in Table 3.1.

**G1** *How are the Mentorspace resources (human and non-human) used?*

One way to improve the organisation processes is to identify the needs and demand of its users. Answering this question will provide the organisation with insight about how to better allocate their resources in the future, how to train their human resources better and which interest trends are observed within different groups that use the space. In the context of this thesis, the users are the event participants only. However, Mentorspace users can be students, parents, industry workers and other types of individuals that are interested in developing their ideas within the organisation. Mentorspace resources are considered to be either human resources (mentors and Mentorspace staff), or non-human resources (material resources such as 3D printers, milling machines, laser cutters among others.). In the context of this project, the only Mentorspace non-human resource provided is the LSB. We consider that if we can answer this question using the participants, mentors and LSBs as representatives of the mentioned Mentorspace elements, it is enough of a proof that it is possible to expand to other resources by tagging more elements in the future.

**G5** *What is the level of complexity of the project to the user groups?*

It is important for the organisation to have control over the level of difficulty and challenges provided by the projects they design for the users. The reason is because the Mentorspace aims at teaching individuals skills that are relevant to their background and aspirations. Having an idea about the level of perceived complexity to different user groups (i.e. school class levels) will help the Mentorspace identify the skill gaps to target with adapted challenges implemented in the project. Moreover, it will help the organisation improve their mentor training to properly address the gaps.

**G7** *Are the Mentorspace processes resulting in a higher user proficiency over time?*

This question is about understanding how much user groups are cultivating their skillset throughout the Mentorspace projects. It is about identifying whether a certain combination of organisational resources results in developing proficiency at certain tasks/concepts over time.

**G6** *Are the reflection and collaboration process between mentors and users effectively working?*

In order to optimise the learning process of individuals in the space, it is also crucial to understand the level of involvement of the mentors in the process: Are they contributing at all to the individual's development? Are they helping with technical challenges too much that they limit the individual's opportunity to learn? Are they providing value that goes beyond the project boundaries?

Being able to answer such questions will help the organisation understand how they can train their mentors better to optimise the collaboration process with individuals in the space, and identify patterns of behaviours about how individuals collaborate differently together towards a common goal.

## 5.2.2 Understanding the data

Now that the objectives are defined, we will take a look at the raw data collected throughout the experiments.

### 5.2.2.1 Raw LSB data

Each time the user requests feedback from the LSB device, it stores data related to the step which the participant is at in memory blocks as series of hexadecimal values. These values represent signal values reported by the DUT (Example in Figure 5.4). They are carefully analysed by the WB and transformed into meaningful information (Example in Figure 5.5)

The information comprises the memory block number at which it is stored, the project step number, the unique DUT serial number, the exact time and date of use, the DUT battery status at that given time, and the output feedback provided to the user.

For instance, Figure 5.5 shows that the user working on the DUT number 100109 successfully completed step 3 at 17:43:27, encountered a challenge in step 4 that is related to the transistor component, then successfully corrected the fault and completed step 4 at 18:03:50.

```

Block 7:
0x5
0x0
0x3
0x5
0x4
0x2
0x0
0x56
0xa
0xa
0x13
0x6
0x10
0x10
0x0
0x1
0x80
0x27
0x81
0x27
0x90
0x5
0x91
0xf
0xa2
0x1
0xb9
0xa3
0x0
0x61
0xa4
0x2
0x68
0xa5
0x1
0xb2
0xa6
0x1
0xb2
0xa7

```

Figure 5.4: Example of WB raw data

```

Block 11:
This is data for step 3
DUT serial #: 100109
Time: 17:43:27
Date: 2019-09-26
DUT battery: 3.13 volts
Participant received no battery or power warnings.
Participant received the following advisory:
Looks Great!
Good Job!

Block 12:
This is data for step 4
DUT serial #: 100109
Time: 17:54:44
Date: 2019-09-26
DUT battery: 3.12 volts
Participant received no battery or power warnings.
Participant received the following advisory:
Check transistor
pins or direction

Block 13:
This is data for step 4
DUT serial #: 100109
Time: 18:03:50
Date: 2019-09-26
DUT battery: 3.15 volts
Participant received no battery or power warnings.
Participant received the following advisory:
Looks Great!
Good Job!

```

Figure 5.5: Example of WB preprocessed data

### 5.2.2.2 Raw IN Data

Figure 5.6 presents an 8 seconds snapshot of the IN raw data obtained from the experiment at Konstfack. The reported data comprises the time at which a tag is detected in the first column, the antenna number at which the tag was detected in the second column, the detected tag indicator in the third, fourth and fifth column, and the chair status illustrated in the last column (0 = presence of a participant on a chair, 1 = no participant). Note that the last column is composed of four digits which correspond to the status of each chair (4 chairs). During this experiment, three readers were used. Column 3,4 and 5 are associated with reader 0, reader 1 and reader 2 respectively. For instance, a mentor with a tag indicator 6300D7F05A was detected in the location of antenna R3 of reader 2 at 16:52:12, while mentor 61005B8C26 was at the location of antenna R4 of reader 1 at the exact same time. Knowing the location of the antennas at the venue allows us to deduce the location of the detected individuals in the room. In order to maximise the chances of detecting individuals accurately, each

mentor was provided with two tags to put in both of their shoes. For instance, 6300D7F05A and 6300D7609C indicators represent the same individual. This was a safety measure given the limited number of antennas deployed in the experiment. This is consistent with the observation in the data given the fact that both tags are detected in the same location within 1s time interval.

```

16:52:10,R0,0000000000,0000000000,0000000000,0000
16:52:11,R1,0000000000,6300D7609C,0000000000,0000
16:52:11,R2,0000000000,0000000000,0000000000,0000
16:52:12,R3,0000000000,6300D7F05A,0000000000,0000
16:52:12,R4,61005B8C26,0000000000,0000000000,0000
16:52:13,R5,0000000000,0000000000,6400DBE62E,0000
16:52:13,R6,0000000000,0000000000,0000000000,0000
16:52:14,R7,0000000000,0000000000,0000000000,0000
16:52:14,R0,0000000000,0000000000,0000000000,0000
16:52:15,R1,0000000000,6300D7609C,0000000000,0000
16:52:15,R2,0000000000,0000000000,0000000000,0000
16:52:16,R3,0000000000,6300D7F05A,0000000000,0000
16:52:16,R4,61005B8C26,0000000000,0000000000,0000
16:52:17,R5,0000000000,0000000000,6400DBE62E,0000
16:52:17,R6,0000000000,0000000000,0000000000,0000
16:52:18,R7,0000000000,0000000000,0000000000,0000

```

Figure 5.6: IN Konstfack experiment raw data

### 5.3 Data preparation and analysis

This section is an attempt to answer the last research question:

**RQ3** Which data analysis techniques are appropriate to analyse the data gathered from a physical STEAM collaborative learning environment?

We will go through each question from the objectives mentioned in section 5.2.1 and argue about the best way to achieve a good enough answer with the data collected. Note that this section is not about how to achieve high model accuracy as the collected data is not enough to achieve high performance and reliability. The goal is to show that data generated from the real world through this LA system holds valuable information that can be analysed and derived using traditional LA data analysis techniques.

### 5.3.1 First objective

**G1** *How are the Mentorspace resources (human and non-human) used?*

An appropriate method to provide insight about this question based on the available data is a technique that illustrates the connections between different elements of the space, explores the evolution patterns of the relationship between them, determines the social influences within the space, identifies communities and how they form, and predicts future links for a better resource allocation. Such a method is known as Social Network Analysis (SNA).

#### 5.3.1.1 Social Network Analysis

In the social network graph, the nodes represent the elements interacting during the experiment (mentors (M), participants (P), LSBs (W), hospital (H)). The edges indicate that there is an interaction recorded between two connected nodes. In this graph, we assume that the interaction relationship is reciprocal each time. Therefore, the edges are undirected (symmetric relationship). The edges are equipped with weights representing the percentage of time of interaction between the connected nodes.

To establish the necessary database for the graph, a python script is created to extract all the associations between individuals from the IN data and organise them into a csv file. The interactions between individuals and the LSBs are added manually from the observation notes. In the future, IR transceivers will automatically generate this piece of information.

Since there were two batches of 4 participants, P(m)a represents participants from the first batch and P(m)b from the second batch. (m) is the chair number.

The network graph is created using the python library NetworkX. In view of the fact that we do not have many nodes to handle in the context of this project, we did not have to worry about the proper layout algorithm for a clear visualisation. In the future, more sophisticated tools are necessary to handle the complexity of larger datasets and add interactivity with the graph.

The following parameters are calculated and presented in section 5.4: degree centrality, closeness centrality, betweenness centrality, clustering coefficient, network connectivity/cohesion, and network density.

### 5.3.2 Second objective

**G5** *What is the level of technical complexity of this project with respect to the user groups?*

The answer to this question is derived from the number and type of faults recorded in each step for each participant over time. A simple histogram chart may be a good first step to visualise such data. The type of fault at each step is deduced from the feedback provided by the LSB ( Figure 5.5). The number of faults is easily deduced from the same dataset. To prepare the dataset for the histogram, a python script is written to extract all types of faults and their corresponding frequency into a csv file. Matplotlib python library and Seaborn are used to plot the chart.

An early profiling of the project technical complexity is established based on simple calculations described in Table 5.3. As the project targets beginners with no prior experience, no project requirements are set. Therefore, the calculations are based on all the possible technical mistakes one can make per step. The assigned weights in Table 5.3 are proportional to the level of difficulty added by each element. The weights are arbitrary for the purpose of this thesis and need to be refined with further research. This first assessment results in an initial ranking of the project steps with respect to their level of difficulty.

Step	Soldering connctions		Parts		Pin shorts		Part backwards		Swapped resistors		Relative Complexity	Normalised complexity
	Number	Weight	Number	Weight	Number	Weight	Number	Weight	Number	Weight		
1	12	1	4	2	3	2	1	5	0	3	26	0.55
2	14	1	1	2	12	2	1	5	0	3	45	0.80
3	4	1	1	2	3	2	1	5	0	3	17	0.30
4	9	1	4	2	3	2	2	5	2	3	34	0.70
5	16	1	5	2	7	2	2	5	2	3	56	1.00

Table 5.3: BN technical complexity assessment

### 5.3.3 Third objective

**G7** *Are the Mentorspace processes resulting in a higher user proficiency over time?*

From the map of measurements and its table (Figure 3.2 and Table 3.1), we speculate that proficiency might be correlated with the evolution of the following factors over time:

- Time spent per step

- Time spent on fixing a fault
- Number of attempts to fix a fault
- Number and type of faults per step
- Who contributed to fixing faults
- What was done to fix a fault

In the context of this thesis, the current LSB is limited to recording data at the end of each step only. Therefore, the last factor, What was done to fix a fault, could not be quantitatively captured by the current system. Future systems will allow for data to be recorded during the whole building process. One way to answer G7 is to use a classification model. Our main goal at this stage is to gain insight into the collected multimodal data and explore possibilities to exploit the data and the information it might offer. For simplicity reasons, we choose to use a binary classifier. For this task, we think that a decision tree is a suitable option for the following reasons:

- It is simple to understand and can be visualised, which helps build insight about the data.
- It is interpretable and explainable in contrast with other classifiers such as artificial neural networks.
- It accepts both numerical and categorical data, which is especially helpful given the type of our variables.

On the other hand, decision trees present some disadvantages; they are hard to generalise well as they are prone to overfitting particularly when some classes dominate, and instability can occur when there are small variations in the data.

#### 5.3.3.1 Decision tree data preparation

The dataset is composed of an instance (row) for each participant and 25 features based on the measured metrics illustrated in G7 of the map of measurements (Figure 3.2) for each step (5 measured metrics x 5 steps). Given that it is a binary classification, the target variable can take two possible values: "Proficient" or "Not proficient".

A participant is considered to be getting proficient if:

- They need less help from the mentor over time.

- They persevere when a fault occurs and end up solving it either on their own or through collaboration with the mentor.
- They cause less faults of the same nature over time (i.e. less short circuits).
- They spend less time on faults of the same nature over time.

The feature values are extracted from the LSB data for each step. The fifth feature, Who contributed to fixing the fault, is manually added and deduced from looking at the IN data and observation notes.

A first assessment of our dataset shows three main challenges:

- Too small of a sample with respect to the number of features.
- A significantly imbalanced dataset (80% of instances are labeled as "proficient", 20% are labeled as "Not proficient")
- Categorical values

### **Small sample and Imbalanced dataset**

Our sample of participants is limited to the number of users that took part in the experiment. undeniably, such a small number will result in overfitting and may not be enough to extract meaningful information.

The imbalanced dataset where the classes are not represented equally is a direct consequence of the small sample we have. This may result in significant performance problems with the predictions. In order to fix the issue, more data collection is required. However, due to time constraints within this project schedule, we will use a resampling technique.

Oversampling is a method that allows to add instances from the under-represented class to balance the dataset. The minority class in our current dataset is the class of 'Not proficient'. We may need to use oversampling to create more instances from both classes, in order to increase the size of the dataset and achieve better results.

For this task, we use the synthetic minority oversampling technique (SMOTE). This technique uses the K-nearest neighbours of each minority class, selects one of them randomly and computes linear interpolations to create a new instance with the minority class. We used this technique to avoid creating copies of existing instances which may increase the chances of overfitting.

### **Categorical values**

Categorical values are converted into numerical values using one hot encoding. The reason we use one hot encoding instead of label encoding is that the latter considers the higher the categorical value the better the category, which may result in wrong outcomes. One hot encoding creates a column for each possible category and assigns it a value of 0 or 1 indicating the presence or absence of a particular category, which results in a more accurate outcome. In our case, the resulting dataset grows to 36 columns.

### 5.3.4 Fourth objective

**G6** *Are the reflection and collaboration process between mentors and users effectively working?*

From the map of measurements and its table, the collaboration process is considered to be correlated with the following factors:

- Who contributed to fixing a fault.
- What was done to fix a fault.
- Number of attempts to fix a fault.
- Which resources were used and for how long.

One way to look at the answer to this questions is to use curve fitting to construct a mathematical formula which represents how well the collaboration process went during an event. The formula would depend on the factors considered to be correlated with the collaboration process (Figure 3.2 and Table 3.1). One approximate example such a formula is the following:

$$Z_0\left[\left(\frac{b_{00}}{c_{00}} + \frac{b_{01}}{c_{01}} + \frac{b_{02}}{c_{02}} + \dots\right)W_0 + \left(\frac{b_{10}}{c_{10}} + \frac{b_{11}}{c_{11}} + \frac{b_{12}}{c_{12}} + \dots\right)W_1 + \left(\frac{b_{20}}{c_{20}} + \frac{b_{21}}{c_{21}} + \frac{b_{22}}{c_{22}} + \dots\right)W_2 + \dots\right] + Z_1[\dots] + \dots = \sum_m \left(\sum_p \left(\sum_n \frac{b_{pn}}{c_{pn}}\right)W_p\right)Z_m \quad (5.1)$$

Variable	Interpretation
b	Weight of the action undertaken by the participant to fixing a fault
c	Weight of the action undertaken by the Mentorspace staff to fixing a fault (mentors, hospital)
W	Weight of the fault being fixed
Z	Weight of the step m (depends on the complexity of the step)
m	Number of steps of the project
n	Number of attempts to fixing a unique fault
p	Number of unique faults within a step m

Table 5.4: Equation variable interpretation

Each step of the project has a different weight depending on its complexity with respect to other steps. Within a step, each unique fault has a different weight depending on the level of difficulty of solving it. For instance, having to fix a short circuit between two components on the BN is less difficult than having to replace a processor. Now, each unique fault can be solved after 'n' number of attempts. These attempts are actions undertaken by either the participant on their own, a collaboration between the mentor/hospital and the participant, or by the mentor/hospital alone. The fraction  $\frac{b_n}{c_n}$  represents a factor of collaboration between the participant and Mentorspace staff. The factor is inversely proportional to the amount of help the participant receives from the Mentorspace staff to solve the problem.

Note that with the current system, it is not yet possible to collect data about the different actions undertaken to solve each fault within a step. However, we expect this level of detail to be acquired with the future system as the LSB will be collecting data during the whole building process. The establishment of the exact formula and the right values for the weights is outside the scope of this thesis.

## 5.4 Results

### 5.4.1 Social Network Analysis

In view of the dynamic nature of the experiment, the results are shown in the form of snapshots of three distinct points in time, which we call 'Batches', representing the evolution of the events at Konstfack. Batch 1 shown on Figure 5.7 demonstrates the interactions between the participants and the Mentorspace resources during the first group of participants. Batch 2 in Figure 5.8 reflects the interactions during the transition period between the two groups of participants when certain participants from group 1 are finishing their boards while some participants of group 2 join the table. Batch 3 illustrated in Figure 5.9 shows the interactions during the second group of participants.

Table 5.5 is a support table to help interpret the node labels.

In all the graphs, the visualisation is such that the node size varies with the betweenness centrality and the node color with the node degree. Besides, the edge thickness is directly proportional to the edge weight, which represents the percentage of interaction time between two connected nodes during a specific batch.

<b>Label</b>	<b>Interpretation</b>
M(n)	Mentor number n
H	Hospital
W(p)	LSB number p
P(m)a	Participant of chair m of group 1
P(m)b	Participant of chair m of group 2

Table 5.5: Node labels map

#### 5.4.1.1 SNA first batch

The high local clustering coefficients ranging from 0.4 to 1.0 from table 5.6 suggest that Mentor 1, Mentor 2 and the LSBs play a key role in creating clusters of nodes that are tightly grouped together. From the graph visualisation, it is easy to identify the two distinct clusters formed by each mentor, their mentees and the LSBs associated with them. It is important to note that having a high local clustering coefficient does not necessarily make the node more important. This importance of a node is determined by the

centrality measures.

The degree centrality values suggest that Mentor 3 is the most popular Mentorspace resource in this Batch (0.55), followed by Mentor 1 with a value of 0.45 and Mentor 2 (0.36). This results from the fact that Mentor 3 is directly connected to six nodes within the graph. Mentor 1 is directly connected to five nodes and Mentor 2 is directly connected to four other nodes.

On the other hand, the closeness centrality values show that all the provided Mentorspace resources are connected to all the network nodes at a similar level. This suggests that most of the Mentorspace resources in this experiment are in a relatively fair position to initiate the flow of information, with Mentor 3 being the quickest at spreading the information.

Despite the popularity of the three mentors indicated previously by the degree centrality measure, the betweenness centrality values show that Mentor 3 plays the most critical role by far. From the graph, Mentor 3 is positioned as a bridge between two subgraphs within this network configuration. If information were to be passed from one cluster to another, Mentor 3 would be vital to ensure the flow of information between the two groups. The removal of Mentor 3 or any change in its position would result in a disconnection between the two clusters which explains the value of the network cohesion. In contrast, LSB 8 has a betweenness centrality of 0.0. Consequently, the removal of W8 would not alter the connection of any pair of nodes, which is reflected in the graph visualisation.

The fairly high average network clustering coefficient and low density value indicates that the robustness of the network could be improved.

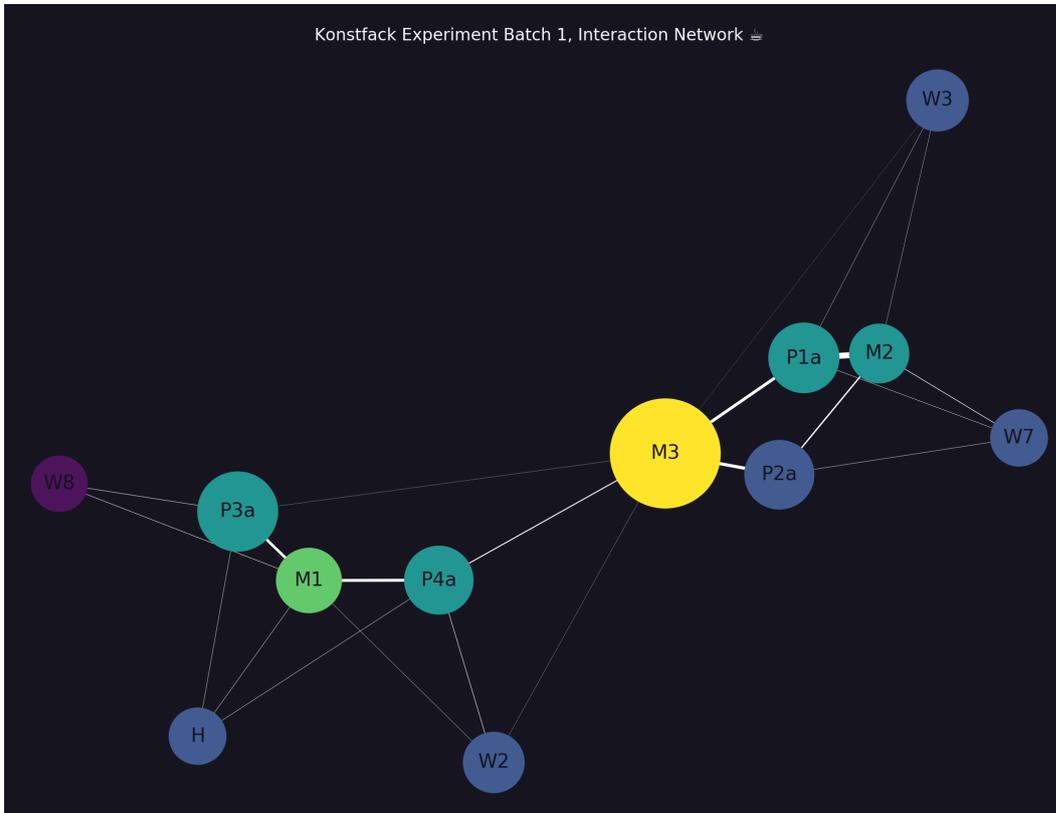


Figure 5.7: Social network graph visualisation, Konstfack experiment Batch1

Local clustering coefficient	Degree centrality	Closeness centrality	Betweenness centrality
W8 (1.0)	M3 (0.55)	M3 (0.69)	M3 (0.58)
H,W2,W3,W7 (0.67)	M1 (0.45)	W2 (0.52)	M1(0.07)
M2 (0.5)	M2 (0.36)	W3 (0.5)	W3(0.042)
M1 (0.4)	H,W2,W3,W7 (0.27)	M1 (0.46)	W2 (0.036)
M3 (0.13)	W8 (0.18)	H (0.42)	M2 (0.024)
		M2,W8 (0.41)	H, W7 (0.006)
		W7 (0.39)	W8(0.0)

Table 5.6: Batch 1 graph centrality measures

Average clustering coefficient	Cohesion	Density
0.53	1	0.33

Table 5.7: Batch 1 network measurements

#### 5.4.1.2 SNA second batch

The cohesion result from batch 1 indicates that the elimination of one particular node, which was identified as Mentor 3 due to its high betweenness centrality, is sufficient to divide the network into subgraphs. It is reflected in this snapshot as Mentor 3, who previously played a crucial role, appears to no longer hold a bridge position. This results in two disconnected clusters of nodes.

The high local clustering coefficients of the LSBs with a value of 1.0 followed by Mentor 1 and Mentor 2 with a value of 0.33 demonstrate that these nodes have a fairly strong tendency to form associations or clusters. This is easily seen in the graph as each mentor forms a community with their assigned mentees and the LSBs associated with them.

In this batch, the most popular Mentorspace resources are Mentor 2 with a degree centrality of 0.44 followed by Mentor 1 with a degree centrality of 0.33. It is shown on the graph that these nodes are the most connected to their relative clusters compared to other nodes.

On the other hand, the betweenness centrality of all nodes is significantly low, which suggests that none of the nodes is vital to the rest of the network. If information needed to be disseminated across the network, initiation would be necessary on both ends of the network since there exists no link between the two clusters. This disconnection is well represented by the network cohesion score.

The network density score of this batch is lower than the first batch which indicates a poorer network robustness.

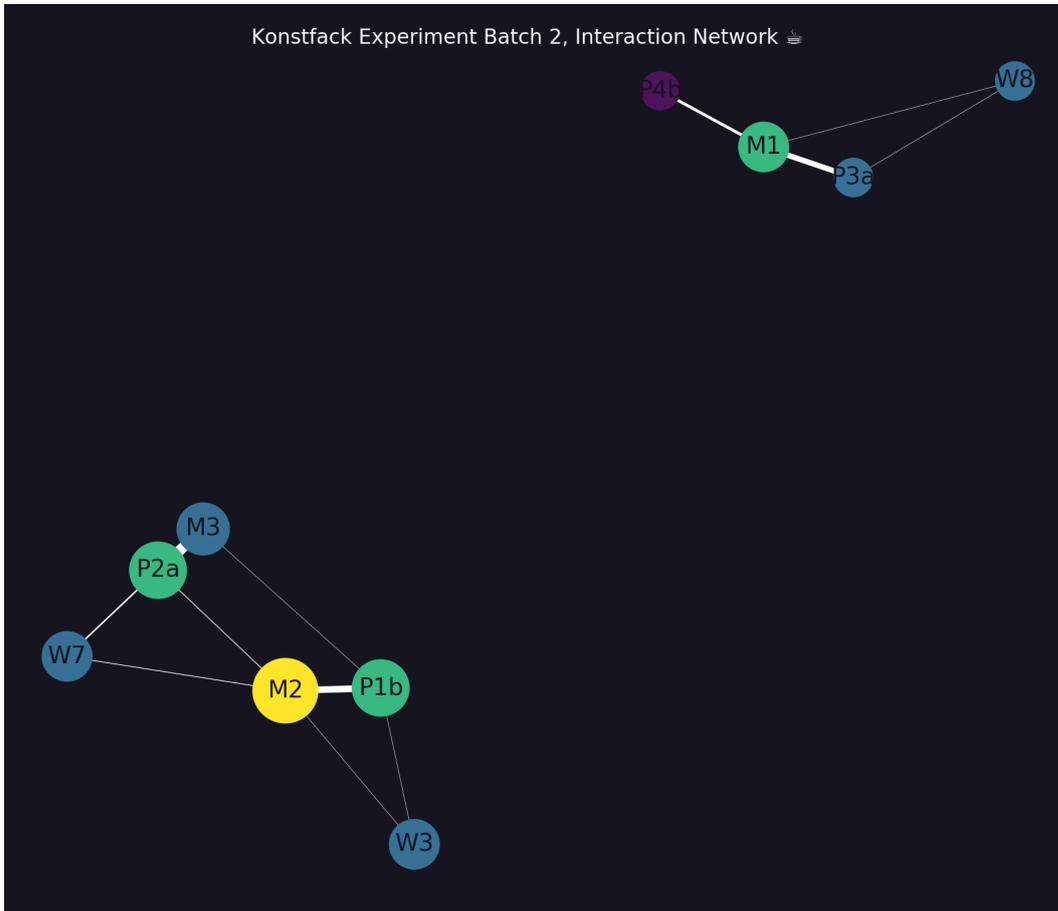


Figure 5.8: Social network graph visualisation, Konstfack experiment Batch2

Local clustering coefficient	Degree centrality	Closeness centrality	Betweenness centrality
W3, W7, W8 (1.0)	M2 (0.44)	M2 (0.46)	M2 (0.1)
M1, M2 (0.33)	M1 (0.33)	M3, W3, W7 (0.35)	M1(0.06)
M3 (0.0)	M3,W3, W7, W8(0.22)	M1(0.33)	M3(0.01)
		W8 (0.25)	W3, W7, W8 (0.0)

Table 5.8: Batch 2 graph centrality measures

Average clustering coefficient	Cohesion	Density
0.53	0	0.27

Table 5.9: Batch 2 network measurements

### 5.4.1.3 SNA third batch

The network configuration and the results of the final batch are very similar to the first one. One difference is that the hospital 'H' appears to replace Mentor 3 with its position as a bridge, which reflects on its high betweenness centrality compared to other nodes in the network. This indicates that H plays a vital role in connecting the network and ensuring the flow of information across it.

The formation of two distinct clusters is driven by the same set of nodes as batch 1, which is demonstrated by their high local clustering coefficient. The fairly high average clustering coefficient in this batch and the low density score suggest that the network is poorly connected.

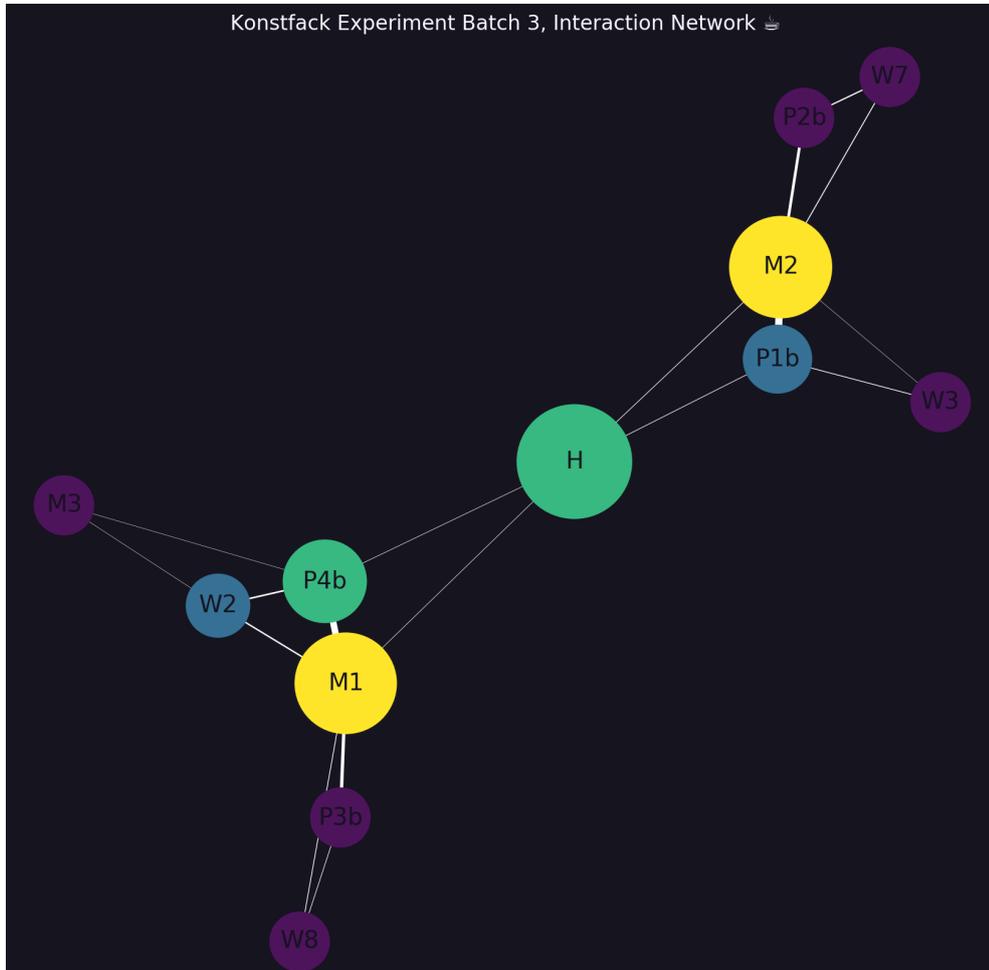


Figure 5.9: Social network graph visualisation, Konstfack experiment Batch3

Local clustering coefficient	Degree centrality	Closeness centrality	Betweenness centrality
M3, W3, W7, W8 (1.0)	M1, M2 (0.45)	H (0.61)	H (0.55)
W2 (0.67)	H (0.36)	M1 (0.55)	M2(0.39)
H, M1, M2 (0.33)	W2 (0.27)	M2(0.52)	M1(0.38)
	M3, W3, W7, W8 (0.18)	W2(0.41)	W2(0.027)
		W8 (0.38)	M3,W3, W7, W8(0.0)
		M3, W3, W7 (0.37)	

Table 5.10: Batch 3 graph centrality measures

Average clustering coefficient	Cohesion	Density
0.73	1	0.28

Table 5.11: Batch 3 network measurements

### 5.4.2 Histogram chart

The results of the experiment are presented in the form of one histogram per step. Each histogram presents the distribution of the frequency of faults per kind of fault for each participant.

The results indicate that the first most complex step for the participants is the final step. The second most complex step is step 2 followed by step 4, step 1 and step 3.

The graphical representation of the faults shows that the first most popular technical challenge in this project is to be aware of which components have a particular direction. On the other hand, the second most popular mistake among participants is creating short circuits. Some participants appear to make the same fault across most steps.

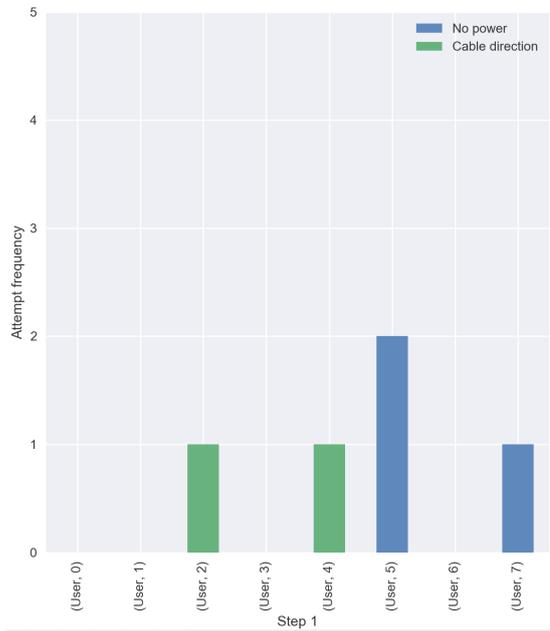


Figure 5.10: Faults summary Step 1

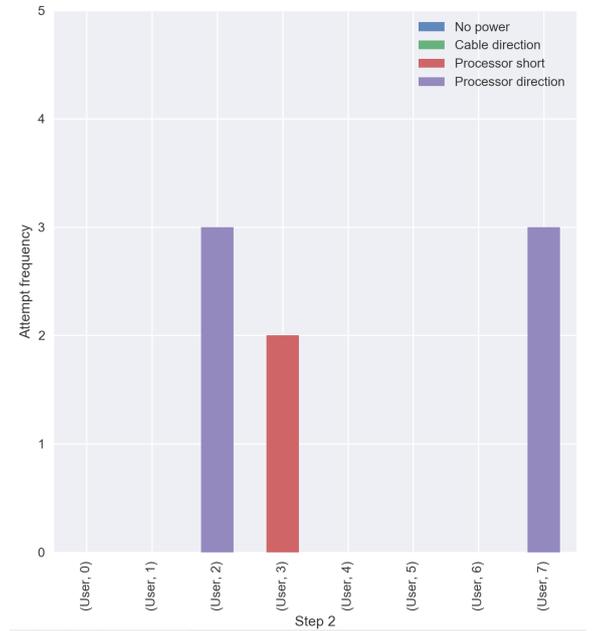


Figure 5.11: Faults summary Step 2

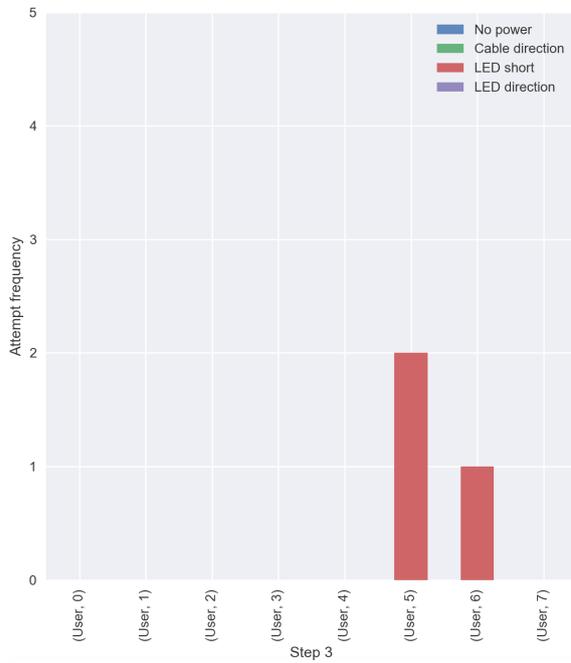


Figure 5.12: Faults summary Step 3

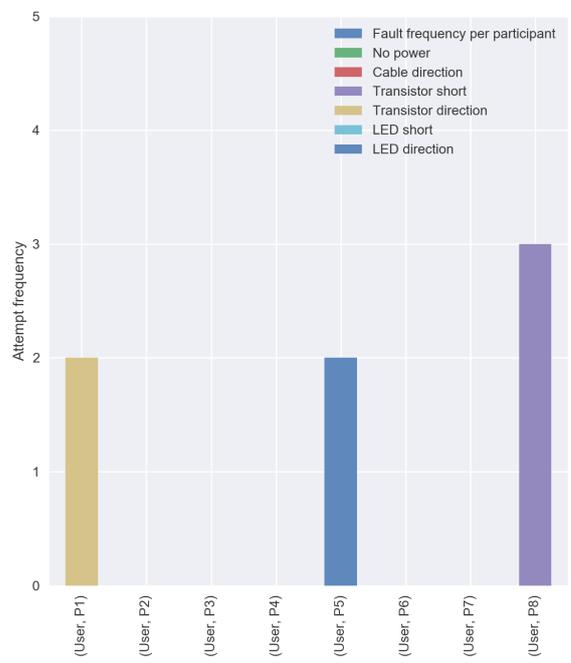


Figure 5.13: Faults summary Step 4

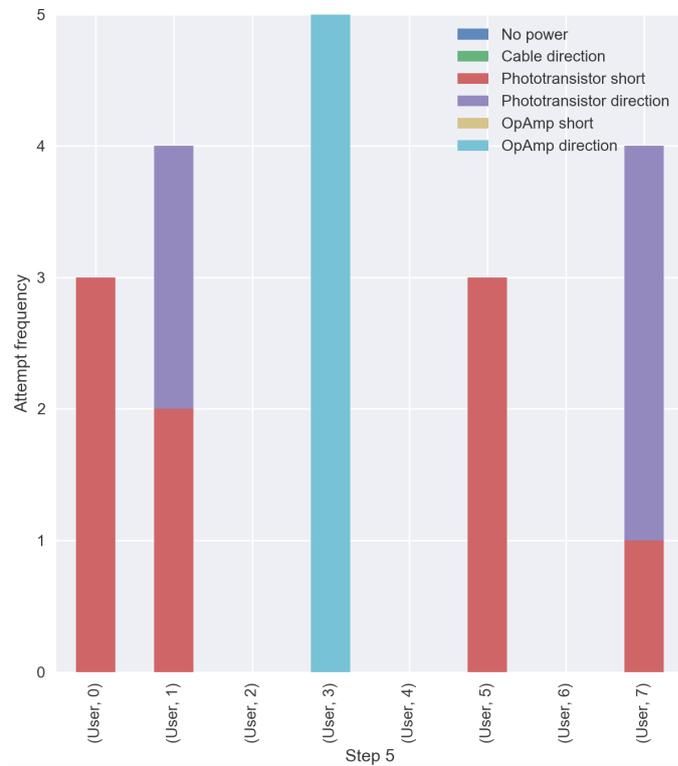


Figure 5.14: Faults summary Step 5

### 5.4.3 Binary classification tree

#### 5.4.3.1 Small sample and imbalanced dataset

The accuracy achieved with the imbalanced dataset for 8 participants is either 1.0 or 0.0 depending on the training and test set. It suggests that the model is overfitting when the test set is composed of the dominant class (Proficient), and underfitting when the test set contains the least dominant class (Not proficient). The decision tree graph demonstrates that the most significant feature to split on is "whether a participant caused a fault in the last step" (Step 5).

#### 5.4.3.2 Oversampling results

The dataset size was increased to 26 participants (Figure 5.15), 34 participants (Figure 5.16), and 42 participants (Figure 5.17), using an oversampling technique in order to achieve a balanced dataset. The decision tree in Figure

5.15 achieves a model accuracy of 87.5% . The decision making process can be summarised as follows:

- In the presence of a fault during the final step of the project, if the fault is solved by the mentor alone without a collaboration with the participant, then the participant is classified as not getting proficient.
- In the presence of a fault during the final step of the project, if the fault is not solved by the mentor alone and the participant spent a small amount of time at solving it, then the participant is classified as getting proficient.
- In the presence of a fault during the final step of the project, if the fault is not solved by the mentor alone and the participant spent too much time at solving it, then the participant is classified as not getting proficient.

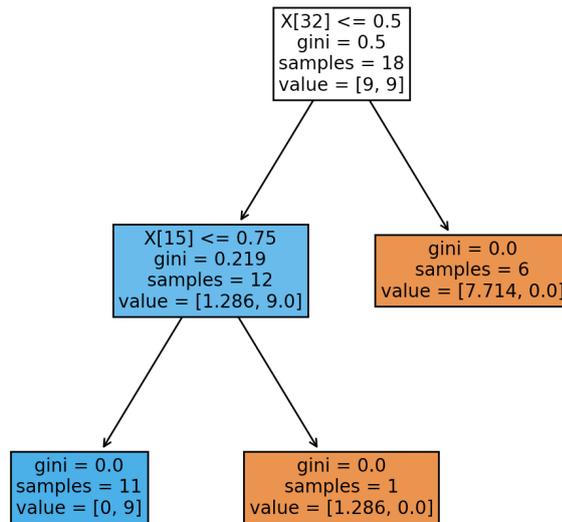


Figure 5.15: Proficiency decision tree for 26 participants

The decision trees in Figure 5.16 and Figure 5.17 achieve a model accuracy of 90% and 92.3% respectively. The algorithm applies the following rules:

- In the presence of a fault during the final step of the project, if the fault is solved by the mentor alone without a collaboration with the participant, then the participant is classified as not getting proficient.

- In the presence of a fault during the final step of the project, if the fault is not solved by the mentor alone and the participant solved the issue after a few attempts, then the participant is classified as getting proficient.
- In the presence of a fault during the final step of the project, if the fault is not solved by the mentor alone and the participant solved the issue only after too many attempts, then the participant is classified as not getting proficient.

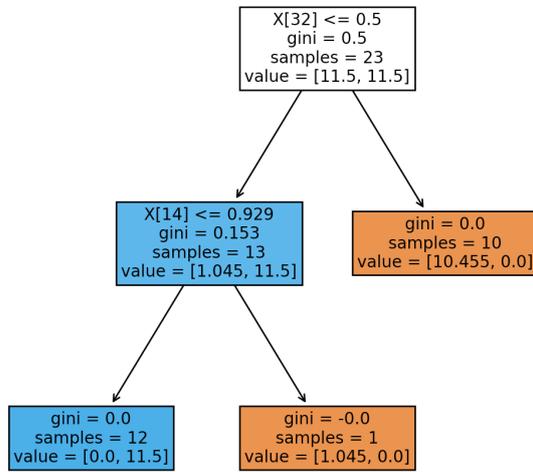


Figure 5.16: Proficiency decision tree for 32 participants

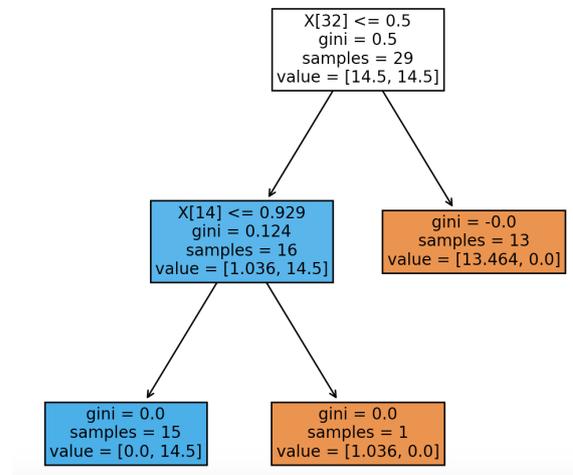


Figure 5.17: Proficiency decision tree for 42 participants

All three decision trees indicate that the classification relies mostly on the events happening during the last step of the project. Particularly, on how much help the participant receives from the mentor, how much time they spend on solving the faults, and after how many attempts.

## 5.5 Discussion

Throughout the report, we started the learning analytics cycle ( Figure 2.1) by explaining the learning environment and learners' data to be collected in section 3.2 and 3.3. We followed by giving a detailed overview of the LA system we designed to collect such data in section 3.4. Next, we performed data analysis using specific techniques described in section 5.3. In this section, insights are derived from the results presented in section 5.4 to propose interventions aimed at improving the shortcomings of the current Mentorspace processes. However, in order to close the loop of the LA cycle, the effectiveness of the suggested interventions needs to be tested with further experimentation after implementation.

### 5.5.1 Results discussion

#### 5.5.1.1 Social network analysis

The results of the SNA show a maintained high clustering coefficient of the LSBs and M1 and M2 throughout the experiment. We believe it stems from the fact that both mentors agreed to divide the table among themselves at the beginning of experiment: M1 decided to mentor participants from chair 3 and chair 4 while M2 took chair 1 and 2. The LSBs were intentionally assigned to specific chairs: W2 to chair 4, W3 to chair 1, W7 to chair 2, W8 to chair 3. The research observations indicate that this division is maintained throughout the entire experiment which explains the formation of two constant clusters across all batches on the graph.

In comparison, mentor 3 interacts more dynamically with both sides of the table during batch 1, moves to M2's group during batch 2 and finally returns to M1's group during the last batch. Observations show that during the natural division of tasks that occurred in the beginning of the experiment, M3 did not assign themselves to any mentee and acted as an additional mentor that helped mentees in need when their mentor was busy. At the beginning of the final batch, M3 had to leave the experiment for a personal emergency, which explains their lack of involvement in the final batch on the graph.

The SNA results show that H intervened a number of times throughout the experiment; mainly for mentor 1 and their mentees during batch 1, as well as for both mentors and their mentees during batch 3. From the thickness of the edges on the graph, it is clear that the interactions occurred in every case for a short time. We believe the reason is because H plays a slightly different role than M3 in this experiment. H only intervenes when the faults are too

advanced to be resolved with simple instruments (i.e. Soldering a processor backwards). H interacts with those in need only to gather information on what went wrong. Then, they take the board away to the rework station for troubleshooting. Finally, they return to the individuals in question to explain the problem.

The thickness of the edges on the graph show that some participants interacted with mentors more than others. In one particular case, the observations revealed an interesting insight. During batch 2, M3 seems to have interacted with P2a for a considerable amount of time. However, observations indicate that half of the time, M3 is standing near P2a while mostly concentrating on their phone without paying attention to the progress of P2a. This remark reinforces the idea that more features need to be added to the IN system in order to detect each individual's orientation and inclination in the space (Table 3.5). Further research and testing with the additional system features is required to avoid such misleading results in the future.

Throughout the experiment, the two distinct clusters appear entirely disconnected from each other. However, the research observations reveal that M1 and M2 interacted frequently through speech and involved their mentees in the conversation. Therefore, the clusters are more cohesive than seen on the graph. We believe it is because the current IN system relies completely on the individuals' location in the space to decide on interactions. We think that a large portion of the problem will be solved with the additional IN features we plan to implement in the near future (Table 3.5). Furthermore, we have ideas for how to capture data about when a person uses speech without recording the speech content. Further research and development must therefore be undertaken to investigate ways to encounter this issue.

The static nature of the SNA performed might be misleading when the activity being analysed is dynamic. One example of an erroneous insight is that H appears to have the highest betweenness centrality during the last batch. This suggests that H is a vital node which links the network and, when necessary, is a key element for spreading information across the nodes. We believe this is true provided that the node holds the same bridge role throughout the duration of the batch. However, H intervenes only twice during the batch for a short amount of time then returns to the rework station. For more accurate insights in the future, a dynamic SNA is therefore required.

Finally, Individuals appear to spend more time interacting with each other than with the LSBs. We believe it stems from recording the interactions with the LSBs only at the end of each step. In reality, the participants also use the LSBs as a platform to keep their board steady during the building process. We think that once the planned features in Table 3.3 are

implemented, this problem will be solved as interactions will be constantly recorded even between the steps.

### 5.5.1.2 Histogram

The results show that the ranking of the step complexity of this project is in line with the technical complexity calculated in section 5.3.2. We believe that it shows that the Mentorspace has the right methodology to assess the technical level of difficulty of this project and can therefore adapt it to different user groups depending on the future learning goals. Future experiments with a bigger sample are required to test this hypothesis and determine the appropriate weights for different faults in order to reach a more accurate assessment of technical complexity.

The histograms convey that the most popular fault is related to the components' direction. Besides, multiple participants appear to make this same mistake across most steps. These results indicate that additional attention must be put on explaining the necessity of placing certain electronic components in the right direction and the reasoning behind it.

Moreover, it is apparent that the second most significant technical challenge that participants struggle with is short circuits. From the experiment observations, we noticed that most mentors do not extensively explain the right techniques to adopt in order to minimise the chances of causing short circuits.

Finally, when we take a look at the histogram and SNA results, we see a correlation between the intervention of H and the participants' faults. For instance, during the first batch (Figure 5.7), we notice an interaction between H and P4a. Comparing the time at which the interaction is recorded by the IN system with the LSB data, we deduce that P4a is dealing with the final step at that time. Hence, we deduce that it is highly likely that H stepped in to fix the OpAmp that is put backwards by the participant, which is considered a fault that requires advanced tools to solve. The research observations confirm this course of action. This correlation is shown to be consistent through all the intervention cases. We believe this consistency shows that the LSB and the IN system are in phase with respect to data collection.

### 5.5.1.3 Decision Tree

The results of all the decision trees are based on the final step (Step 5). We believe it is due to the fact that most of the faults made by the participants in this experiment are concentrated in the last step, which shows on the

histogram as it is the most complex one. The histogram results support this statement (section 5.4.2).

The feature selection made by the decision trees provide interesting insights. The feature chosen to have the highest weight states that each participant who has their faults in the final step fixed by their mentor only are not considered to have developed proficiency. We think it is a fair decision because all the possible types of faults tackled in step 5 are introduced in the previous steps. Therefore, there is a high likelihood that the participant encountered them previously. If not, we believe they should have acquired sufficient skills throughout the steps to solve the issue themselves or through a collaboration with mentors or participants. However, more experiments are needed to assess this hypothesis. If further research reveals the opposite, we must look at other causes related to Mentorspace processes and products. The results show that the second most significant feature is related to either the amount of time spent on fixing a fault or the number of attempts to fix the faults in the final step. When labelling the dataset, we assumed that proficiency is also related to how fast a person can troubleshoot a problem or how fast they can solve it (section 5.3.3.1). However, further research is required to investigate this assumption among others ( i.e. What is the accurate limit of time to consider when it is 'enough time' or ' enough number of attempts' to make a judgment about proficiency? Is encountering the same type of fault a significant factor in improving proficiency? etc...)

#### 5.5.1.4 Interventions

Based on the results from the experiments, the proposed interventions call for improving **information content and sharing**, and **encouraging more collaboration** between individuals to improve the cohesion of the network. The suggested interventions are summarised as follows:

**Shared material** One way to improve the current network connectivity is to share the resources provided by the Mentorspace such as soldering irons, manuals, computers, 3D printers and more. We expect the material sharing to also initiate interactions among participants which may lead to further collaboration especially when tackling similar challenges.

**Mentor training** The SNA results suggest that a good way to increase the network connectivity is to train mentors to be more dynamic like M3 and help different mentees during the same session. However, we learned across Mentorspace interventions that it is necessary for

mentors to accompany the same mentees throughout the events. The reason is that it gives them time to build trust and social bonds which gives confidence to mentees to open up and share more about themselves. This is particularly important when working with sensitive categories of individuals. The Mentorspace cares about teaching their mentors to be a source of inspiration and good advice for their mentees as mentoring is also about sharing our struggles in life to learn more about how others deal with them and get useful guidance. While preserving this proclivity to work in clusters of individuals, we suggest to have additional 'free' mentors like M3 that would endorse the role of a medium between individuals and provide help when needed.

Mentors have the power to encourage good practice. Therefore, we suggest to train the mentors to encourage their mentees to interact and collaborate with other participants in addition to sharing resources with them.

With respect to tackling the project complexity better, the histogram results indicate a need for better technical practices. For future mentor trainings, we recommend to integrate a section around the best soldering practices to be taught to the mentees in the beginning of future events in order to help them avoid short circuits. Moreover, we suggest that the mentors be made aware of the necessity to teach the mentees about the meaning of the direction of certain electronic components.

**Strategic information distribution** On the basis of previous recommendations, the exchange of material and resources is expected to increase their connection to the various nodes in the space. The shared materials thus become essential elements for the dissemination of information. This implies that it would be beneficial to concentrate knowledge and information in such shared resources such as LSBs, 3D printers, computers among other resources.

Besides the shared material, the additional mentors that are regarded as 'free' mentors are expected to also be key elements for the effective distribution of information as demonstrated by the role of M3 in our experiment. We suggest that those mentors be either trained or chosen to carry the target information and possibly have advanced expertise to increase their value.

# Chapter 6

## Conclusion

In this thesis, we present an introduction to a proof-of-concept study for what we consider to be a multimodal learning analytics approach in a STEAM collaborative learning environment. While this approach attempts to inspire a new way of applying the learning analytics approach to a physical learning setting, it also tries to achieve this aim in a way that might address the ethical issues encountered in the field, as no personal identity data about learners is stored or used throughout the study.

### 6.1 Research questions

**RQ1** *How do we derive learning indicators from a STEAM physical collaborative learning environment that can be used to infer learning?*

One of the main challenges of this project was to identify an initial set of granular metrics that can provide answers about ways to optimise the learning environment. We developed an initial high level map of data trails to capture (Figure 3.2). We believe these indicators might have the potential to complement metrics derived from learning management systems and used extensively in the LA research community (i.e. number of clicks, message content on forums, screen time etc...).

In view of the limited time required to implement the necessary system to capture all the stated indicators, some data trails are not measured in this thesis project. Nevertheless, those recorded suggest insights that can motivate effective strategies to enhance the learning experience (section 5.5.1.4). In the near future, we would like to explore and refine the map further with the following recommendations:

- Closing the feedback loop; Test the suggested interventions and use the feedback to review the tested metrics and refine them.
- Implement the full system (Figure 3.4) to record the rest of the learning indicators on the map.
- Situate and define proficiency, and interaction in a way that is more grounded in relevant learning theories and philosophies of education.
- Definition and iteration over the environmental and cognitive metrics that are considered to have influence on the learning process in a collaborative environment.
- Identify which types of interactions are correlated with the learning process and experience, and need to be monitored.
- Deeper investigation about the ethical implications of the designed instruments; when does too much impersonal data become personal data. The current system needs no personal data to derive useful insight about the learning experience. However, too much data collection over time might reveal patterns of behaviours that can uncover one's identity. Research is needed to identify ways of avoiding data misuse scenarios without, if possible, compromising the system performance.

Besides enabling effective decision making, we believe that learning analytics holds the potential to advance learning sciences. Therefore, we hope to use the the present and future system as the main research instrument to learn more about the points raised above, through more experiments and iterations.

**RQ2** *Which instruments can generate quantitative data in a physical STEAM collaborative learning environment, in an ethical manner?*

The data collection was done through a hardware infrastructure that we designed from scratch to meet the project needs with minimal costs. The infrastructure comprises an IN system that measures the real-time human-human and human-machine interactions in the learning environment (section 4.2.2), and an LSB system that monitors the interaction between the learner and the artefact being created (section 4.1.1 and section 4.2.1). Due to time and resource constraints, we implemented and tested a limited version of the system (Figure 3.4), which is an incomplete version of the full system we aim to create (Figure 3.3).

When we started thinking about the LSB design, we imagined it to be more than a passive system that collects data about learners to be used for organisational purposes. We designed the LSB with the expectation that it becomes a useful active **personalised** support system for the learners to directly benefit them. During the experiments, we were positively surprised to see that learners developed a sense of excitement plugging the LSB into their board at the end of each step and looking forward to get guidance from the feedback about how they performed. Some learners verbally expressed to us how useful they find that added level of guidance. They stated that it was precise enough to provide useful guidance but intriguing enough to give the opportunity to the learner to troubleshoot their board and reflect on the problem. Moreover, the ring-shaped frames were extensively used during the experiment and proved to be an appreciated support to prevent the electronic board from wobbling when being soldered. In the future, further research about the LSB interface design is required to determine the right constructive level of feedback to best support the learners in their journey for intellectual growth.

On the other hand, the LSB measurement results showed consistency with the experiment observations and with the IN system measurements. We could identify similar suggested patterns of learners' behaviour from the data analysis results of both the IN and LSB and retrace them back to recorded observations (section 5.5.1). For instance, some perceived interactions on the SNA results could be easily linked to results shown on the histogram. In the current system, the LSB records data at the end of each project step only. This results in some limitations in the measurements. For instance, when a participant causes faults during the building process, all we perceive is the final consequences at the end of the step. In other words, it is not possible to determine the course of action that led to the faults. Although exhaustive testing was performed to cover most possible scenarios of faults, this manual approach is limited in terms of data accuracy and sustainability, especially for complex projects. In the future system (Table 3.3), we expect the LSB to be recording data about the DUT status during the whole building process. Moreover, test vectors will be implemented to perform more exhaustive testing on the DUT and exactly identify what went wrong at which point in the process.

The IN system developed in this project relies entirely on the real-time position of individuals in the space. When we designed the system, we anticipated the shortcomings related to lack of information about mentor's orientation and physical inclination with respect to objects and other individuals in the space, and the real-time position of participants. Moreover,

we faced a few technical problems that led to restricting the floor surface covered by the IN system. Therefore, we designed the experiment in a way to minimise risks of erroneous data, without compromising the Mentorspace usual event dynamic. The data collected was consistent and accurate with respect to time and position of the mentors in the space (Figure 5.6). However, some data about the position of the participants on one of the chairs presented a few inconsistencies that were corrected using experiment observations. The future IN system will use transceivers instead in order to monitor the position of the participants.

Future development of the IN will include information about the real-time orientation and physical inclination of mentors in the space, the real-time position and orientation of the participants in the space, and the real time position and orientation of material resources in the space.

**RQ3** *Which data analysis techniques are appropriate to analyse the data gathered from a STEAM collaborative learning environment?*

Social network analysis was used to visualise the human-human and the human-machine interactions during the experiment. This approach offered a quantitative method to assess the properties of individuals and understand the distribution of resources in the space. The analysis revealed that some mentors have the proclivity to work with specific participants in a way that disconnects them from the rest of the network. On the other hand, mentors that are not assigned individuals tend to play a key role in information exchange and network cohesion. However, from the Mentorspace experience, we believe that it is of paramount importance to preserve both types of practices while redistributing the resources in a manner that improves connectivity and encourages information exchange. .

A histogram chart was used to provide a visual overview of the perceived technical complexity of different project steps. The histogram provided insights into the most common technical challenges faced during the experiment. The chart reveals that many participants appear to repeat the same mistake over time. Consequently, we identified key practices to be added to future mentor trainings to overcome these specific deficiencies and accelerate the participants' proficiency.

Finally, we used a decision tree to help determine whether participants get proficient over time with a particular Mentorspace resource configuration. We soon realised that the sample of participants is too small to derive any useful recommendations for interventions. In the future, we plan to

investigate the relationship between the selected features using methods such as Pearson correlation coefficient or Principal Component Analysis. Such techniques will help avoid redundancy of information and refine the map of measurements. In addition, we would like to experiment with unsupervised learning models to learn more about the factors that influence proficiency in a collaborative learning environment.

## 6.2 Implications and future work

We designed our LA data collection system in a way that can be used on different computer-based projects. For future work, we aim at implementing the rest of the features of the desired LA system and experiment with different projects in a way that is more grounded in theories of learning. We also have further ideas about adapting the system to screen-based environments while maintaining our strict ethical approach.

The implications of this thesis vary and collaboration with other fields is necessary to achieve better results. For the LA community, this study might have the potential to help reach a generalisable ethical LA system based on multimodal data from a combination of online and physical learning environments. On the other hand, learning sciences and learning psychology might benefit from using these types of instruments as an extra set of lenses and an extension of their methods to understand the physical learning environment better. This process might result in new strategies and perspectives for the instructional design field which in turn might inspire new perspectives of teaching and provide teachers with more tools, such as being able to exactly determine the appropriate level of guidance needed by students. For teachers, The LSB can be used as a formative assessment tool as a way to introduce personalised learning to classrooms, provided that it can provide personalised feedback and useful insight to the teacher about the learner's progress in a physical setting. From a connectivism point of view, it is suggested that the interaction between humans and computers affect the process of learning and even the thinking patterns of learners. Therefore, design fields such as interaction design can join to explore the design of new ways of interacting with technology to improve the learning process.

## References

Accenture.com. IT'S LEARNING. JUST NOT AS WE KNOW IT. How to accelerate skills acquisition in the age of intelligent technologies. [online]

Arroway, P., Morgan, G., O'Keefe, M., and Yanosky, R. (2016). Learning Analytics in Higher Education. Research report. Louisville, CO: ECAR.

Alan Rubel and Kyle M. L. Jones (2017). Data Analytics in Higher Education: Key Concerns and Open Questions, *University of St. Thomas Journal of Law and Public Policy* 11, 1: 25–44.

Atif, Amara and Richards, Deborah and Bilgin, Ayse and Marrone, Mauricio (2013). Learning Analytics in Higher Education: A Summary of Tools and Approaches. 30th Annual conference on Australian Society for Computers in Learning in Tertiary Education, ASCILITE 2013.

Ali, L., Hatala, M., Gašević, D. and Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers and Education*, 58(1), 470-489. Elsevier Ltd.

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., and Dawson, S. (2016). Recipe for success: Lessons learnt from using xAPI within the connected learning analytics toolkit. In D. Gasevic, G. Lynch, S. Dawson, H. Drachler, and C. P. Rosé (Eds.), *Proceedings of the sixth international conference on learning analytics and knowledge* (pp. 378–382). New York, NY: ACM. doi:10.1145/2883851.2883882

Bauer-Wolf, J. (2018). Overconfident Students, Dubious Employers. *INSIDE HIGHER ED*. [online]

Bednar, A. K., Cunningham, D., Duffy, T. M., Perry, J. D. (1991). Theory into practice: How do we link? In G. J. Anglin (Ed.), *Instructional technology: Past, present, and future*. Englewood, CO: Libraries Unlimited.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892-895. doi:10.1126/science.1165821

Billy Tak Ming Wong (2017) . Learning analytics in higher education:

an analysis of case studies: *Asian Association of Open Universities Journal*, Vol. 12 No. 1, pp. 21-40

Beattie, Scott and Woodley, Carolyn and Souter, Kay (2014) Creepy Analytics and Learner Data Rights. In: *Australasian Society for Computers in Learning and Tertiary Education*, 23 - 26 November 2015, Dunedin, New Zealand.

Blikstein, P., and Worsley, M. A. B. (2016). Multimodal Learning Analytics and Education Data Mining: Using computational technologies to measure complex learning tasks. *The Journal of Learning Analytics*, 3(2), 220-238.

Boyd, Danah and Crawford, Kate. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society*. 15. 662-679.

Bodily, Robert and Verbert, Katrien. (2017). Trends and issues in student-facing learning analytics reporting systems research. 309-318.

Brown, J. S., Collins, A., Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.

Burton, J. K., Moore, D. M. (M.), Magliaro, S. G. (2004). Behaviorism and Instructional Technology. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (p. 3-36). Lawrence Erlbaum Associates Publishers.

Chen, Hsiu-chin and Chiang, Roger and Storey, Veda. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*. 36. 1165-1188. 10.2307/41703503.

Cooper, A. (2012). A Framework of Characteristics for Analytics. *CETIS Analytics Series*, 1(7).

Clow, D. (2012a). The Learning Analytics Cycle: Closing the Loop Effectively. Paper presented at the Proceedings of the 2nd 115 international conference on learning analytics and knowledge.

Clow, D. (2012b). The learning analytics cycle. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12,

134-134. doi:10.1145/2330601.2330636

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), pp.683-695.

C. Wright, Mary and McKay, Timothy and Hershock, Chad and Miller, Kate and Tritz, Jared. (2014). *Better Than Expected: Using Learning Analytics to Promote Student Success in Gateway Science*. *Change: The Magazine of Higher Learning*. 46. 28-34. 0.1080/00091383.2014.867209.

Conde-González, Miguel and Hernández-García, Ángel. (2015). Learning analytics for educational decision making. *Computers in Human Behavior*. 47. 1-3. 10.1016/j.chb.2014.12.034.

Chen, Guoqing and Liu, Hongyan and Yu, Lan and Wei, Qiang and Xing, Zhang. (2006). A new approach to classification based on association rule mining. *Decision Support Systems*. 42. 674-689. 10.1016/j.dss.2005.03.005.

Clancey, W. J. (1986). Review of Winograd and Flores' understanding computers and cognition: A favorable interpretation. (STAN-CS-87-1173) Palo Alto, CA: Department of Computer Science, Stanford University.

Cooper RW, Frank GL, Hansen MM, et al. Key ethical issues encountered in healthcare organizations: the perceptions of staff nurses and nurse leaders. *J Nurs Adm*. 2004;34(3):149–156

Department for Business Innovation and Skills (BIS), corp creator. (2015) Graduate labour market statistics : January to March Q1 2015.

Dietz-Uhler, B.; Hurn, J. E. Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective. *Journal of Interactive Online Learning*, Volume 12, Number 1. USA. 2013.

Dowell, Nia M.; Skrypnyk, Oleksandra; Joksimovic, Srecko; Graesser, Arthur C.; Dawson, Shane; Gašević, Dragan; Hennis, Thieme A.; de Vries, Pieter; Kovanovic, Vitomir (2015). Modeling Learners' Social Centrality and Performance through Language and Discourse. *International Educational Data Mining Society*, Paper presented at the International Conference on Educational Data Mining (EDM)

Dragan Gašević et al., Lim, C. P., and Tinio, V. L. (Eds.). (2018). *Learning*

analytics for the global south. Quezon City, Philippines: Foundation for Information Technology Education and Development.

Drachsler, Hendrik; Cooper, Adam; Hoel, Tore; Ferguson, Rebecca; Berg, Alan; Scheffel, Maren; Kismihók, Gábor; Manderveld, Jocelyn and Chen, Weiqin (2015). Ethical and privacy issues in the application of learning analytics. In: 5th International Learning Analytics and Knowledge Conference (LAK15): Scaling Up: Big Data to Big Impact, 16-20 Mar 2015, Poughkeepsie, NY, USA.

Drachsler, Hendrik and Greller, Wolfgang. (2016). Privacy and Analytics – it’s a DELICATE Issue. A Checklist for Trusted Learning Analytics.. 10.1145/2883851.2883893.

Duval, Erik. (2011). Attention please! Learning analytics for visualization and recommendation. ACM International Conference Proceeding Series. 9-17. 10.1145/2090116.2090118.

E. Arnold, Kimberly and Pistilli, Matthew. (2012). Course signals at Purdue: Using learning analytics to increase student success. ACM International Conference Proceeding Series. 10.1145/2330601.2330666.

E. Ruffaldi, G. Dabisias, L. Landolfi and D. Spikol, "Data collection and processing for a multimodal learning analytic system," 2016 SAI Computing Conference (SAI), London, 2016, pp. 858-863. doi: 10.1109/SAI.2016.7556079

Ferguson, Rebecca (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6) pp. 304–317

Ferguson, Rebecca and Hoel, Tore and Scheffel, Maren and Drachsler, Hendrik. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of Learning Analytics*. 3. 5-15. 10.18608/jla.2016.31.2.

Fritz J (2011) Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2), pp 89-97

George Siemens , Ryan S. J. d. Baker, Learning analytics and educational data mining: towards communication and collaboration, *International Conference on Learning Analytics and Knowledge*, April 29-May 02, 2012,

Vancouver, British Columbia, Canada

Gibbs, G. (2010) *Dimensions of Quality*. The Higher Education Academy.

Gardner, Josh and Brooks, Christopher. (2018). Coenrollment networks and their relationship to grades in undergraduate education. 295-304. 10.1145/3170358.3170373.

Gansemer-Topf, Ann and Saunders, Kevin and Schuh, John and Shelley, Mack. (2004). A Study of Resource Expenditures and Allocation at DEEP Colleges and Universities: Is Spending Related to Student Engagement?

Gasevic, Dragan and Dawson, Shane and Siemens, George. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*. 59. 10.1007/s11528-014-0822-x.

Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques*. San Francisco, CA: Elsevier.

Jayaprakash, Sandeep and Moody, E.W. and Lauria, Eitel and J, Regan and Baron, J.D.. (2014). Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics*. 1. 6-47.

Joksimovic, Srecko and Dawson, Shane and Manataki, Areti and Kovanovic, Vitomir and Gasevic, Dragan and Friss de Kereki, Inés. (2016). Translating network position into performance: Importance of centrality in different network configurations. 10.1145/2883851.2883928.

Jonassen, D. (1991b). Objectivism vs. constructivism: Do we need a new philosophical paradigm. *Educational Technology Research and Development*, 10.

Jonassen, D. H. (1991a). Evaluating constructivistic learning. *Educational Technology* 31(9), 28-33.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*. <https://doi.org/10.1177/2053951714528481>

Kuzilek, Jakub; Hlosta, Martin; Herrmannova, Drahomira; Zdrahal, Zdenek and Wolff, Annika (2015). OU Analyse: analysing at-risk students at The Open University. *Learning Analytics Review*, LAK15-1 pp. 1–16.

Liu, B. (2006). *Web Data Mining*. Berlin Heidelberg: Springer.

Long, P. and Siemens, G. (2014). Penetrating the fog: analytics in learning and education. *Italian Journal of Educational Technology*, 22(3), 132-137. Ortona, Italy: Edizioni Menabò - Menabò srl.

Liaqat Ali, Marek Hatala, Dragan Gašević, Jelena Jovanović, A qualitative evaluation of evolution of a learning analytics tool, *Computers and Education*, Volume 58, Issue 1, 2012, Pages 470-489, ISSN 0360-1315, <https://doi.org/10.1016/j.compedu.2011.08.030>.

Lester, J., Klein, C., Rangwala, H. and Johri, A. (2017). *Learning Analytics in Higher Education*. ASHE Higher Education Report, 43(5), pp.9-135.

Lockyer, L and Heathcote, Elizabeth and Dawson, Shane. (2013). Informing Pedagogical Action Aligning Learning Analytics With Learning Design. *American Behavioral Scientist*. 57. 1439-1459. 10.1177/0002764213479367.

Mantelero, Alessandro (2014) The Future of Consumer Data Protection in the E.U. Rethinking the 'Notice and Consent' Paradigm in the New Era of Predictive Analytics. *Computer Law and Security Review*, 2014, 30, 643-660.

Macfadyen, L. p. and Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: a proof of concept. *Computers and Education*, 54 (2), 588-599.

McKay, Tim and Miller, Kate and Tritz, Jared. (2012). What to do with actionable intelligence: E2Coach as an intervention engine. *ACM International Conference Proceeding Series*. 10.1145/2330601.2330627.

Macfadyen, L. and Dawson, S. P. (2012). Numbers are not enough. why e-learning analytics failed to inform an institutional strategic plan. *Educational Technology and Society*, 15 (3), 149-163.

Mor, Yishay and Ferguson, Rebecca and Wasson, Barbara. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action: Learning design, TISL and learning analytics. *British Journal of Educational Technology*. 46. 221-229. 10.1111/bjet.12273.

Nguyen, Quan and Tempelaar, Dirk and Rienties, Bart and Giesbers, Bas. (2016). What learning analytics based prediction models tell us about feed-

back preferences of students. *Quarterly Review of Distance Education*. 17. 13-33.

Nunn, S., Avella, J., Kanai, T. and Kebritchi, M. (2016). Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review. *Online Learning*, 20(2).

Palotai, P. (2017). E-Skills and Jobs in the Digital Age. [online] Brussels. Available at: <https://ec.europa.eu/epale/fr/content/e-skills-and-jobs-digital-age>.

Picciano, A. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Online Learning*, 16(3).

Pardo, Abelardo and Siemens, George. (2014). Ethical and Privacy Principles for Learning Analytics. *British Journal of Educational Technology*. 45. 10.1111/bjet.12152.

P. A. Ertmer, T. J. Newby (1993). Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective. *Journal of Teacher Education*.

Petersen, R. J. (2012). Policy dimensions of analytics in higher education. *EDUCAUSE Review*, 47(4), 44-49.

Perkins, D. N. (1991). Technology Meets Constructivism: Do They Make a Marriage? *Journal of Educational Technology*

Persico, Donatella and Pozzi, Francesca. (2014). Informing learning design with learning analytics to improve teacher inquiry: Informing LD with LA to improve teacher inquiry. *British Journal of Educational Technology*. 46. 10.1111/bjet.12207.

Rhiannon Williams on iNews (2018). Microsoft: Governments must invest in education to close digital skills gap. [online]

Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice*, 2, 1–13

R. Lai, E. and Viering, M. (2012). Assessing 21st Century Skills: Integrating Research Findings. National Council on Measurement in Education Vancouver, B.C.

Roberts, Lynne and Chang, Vanessa and Gibson, David. (2017). Ethical Considerations in Adopting a University- and System-Wide Approach to Data and Learning Analytics. 10.1007/978-3-319-06520-5\_7.

Romero, C., Ventura, S. (2010) Educational Data Mining: A Review of the State-of-the-Art. *IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 40 (6), 601-618

Romero, C., Ventura, S., Pechenizky, M., Baker, R. (2010) *Handbook of Educational Data Mining*. 2010. Editorial Chapman and Hall/CRC Press, Taylor and Francis Group. Data Mining and Knowledge Discovery Serie

Schooley, Robert (2017). Why are Soft Skills Missing in Today's Applicants. *Murray State Theses and Dissertations*. 42.

Schunk, D.H. (1985). Participation in goal setting: Effects on self-efficacy and skills of learning disabled children. *Journal of Special Education* , 19 , 307-317.

S. Sameshima, H. and P. Katz, (2009). Experiences with Cricket/Ultrasound Technology for 3-Dimensional Locationing within an Indoor Smart Environment

Siemens, G. (2004). Connectivism: A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning*, 2.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380- 1400. doi:10.1177/0002764213498851

Slade, Sharon and Prinsloo, Paul (2013). Learning analytics: ethical issues and dilemmas. *American Behavioral Scientist*, 57(10) pp. 1509–1528.

Snelbecker, G. (1983). *Learning theory, instructional theory, and psychoeducational design*. New York: McGraw-Hill.

Solutions, E. (2019). Edtech policy is failing globally and needs a 'critical update', finds new report by Varkey Foundation's Atlantis Group. [online] Education Technology Solutions.

Spikol, D., Ruffaldi, E., Dabisias, G. and Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*, 34(4), pp.366-377.

The National Education Association (2010). *Preparing 21st Century Students for a Global Society: An Educator's Guide to the "Four Cs"*.

Thompson, A., Simonson, M. R., Hargrave, C. P. (1992). *Educational technology: A review of the research*. Washington DC: Association for Educational Communications and Technology.

T. Elias, "Learning Analytics: Definitions, Processes, Potential," (2011)

Tsai, Y., and Gašević, D. (2017). Learning analytics in higher education - challenges and policies: a review of eight learning analytics policies. LAK.

Ulrich, K. T., & Eppinger, S. D. (2016). *Product design and development*.

Varun Kumar, Dr and Anupama, Chadha. (2011). An Empirical Study of the Applications of Data Mining Techniques in Higher Education. *International Journal of Advanced Computer Sciences and Applications*. 2. 10.14569/IJACSA.2011.020314.

Verbert, Katrien and duval, erik and Klerkx, Joris and Govaerts, Sten and Luis Santos, José. (2013). *Learning Analytics Dashboard Applications*. American Behavioral Scientist.

Wong, Jacqueline and Baars, Martine and de Koning, Björn and van der Zee, Tim and Davis, Dan and Khalil, Mohammad and Houben, Geert-Jan and Paas, Fred. (2019). *Educational Theories and Learning Analytics: From Data to Knowledge: The Whole Is Greater Than the Sum of Its Parts*.

Alvin Y. Wang, Michael H. Newlin (2000). Characteristics of students who enrol and succeed in psychology web-based classes. *Journal of Educational Psychology*.

Alvin Y. Wang, Michael H. Newlin (2002). Predictors of web-student performance: The role of self efficacy and reasons for taking an on-line class. *Computers in Human Behavior*.

Winne P.H. et al. (2017) Designs for Learning Analytics to Support Information Problem Solving. Informational Environments.

World Bank World Development Report (2018). Learning to Realize Education's Promise. Population and Development Review.

Wasserman, S., and Faust, K. (1994). Social Network Analysis: Methods and Applications (Vol. 8). Cambridge University Press.

## Appendices

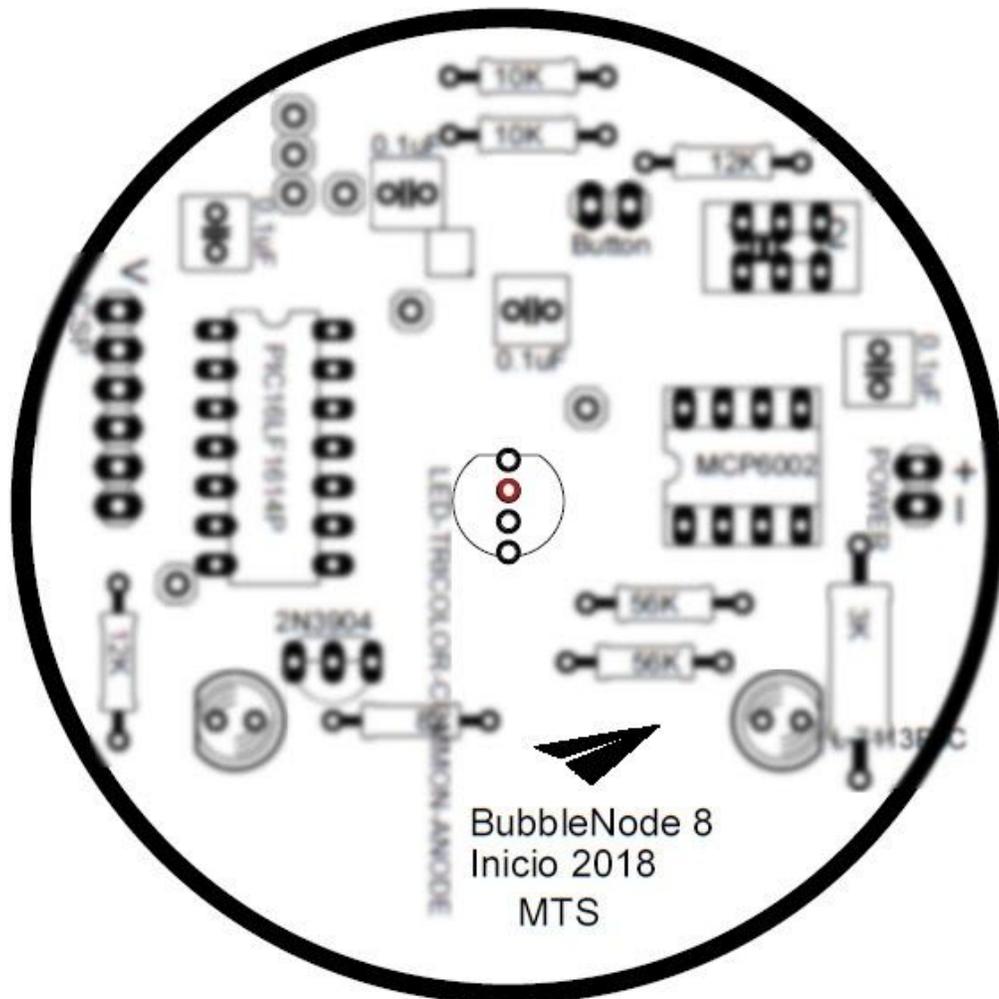
### 6.1 Appendix A1: Järvavecken experiment user manual

# *Bubble Node Instructions*





### Step 3. Three Color LED

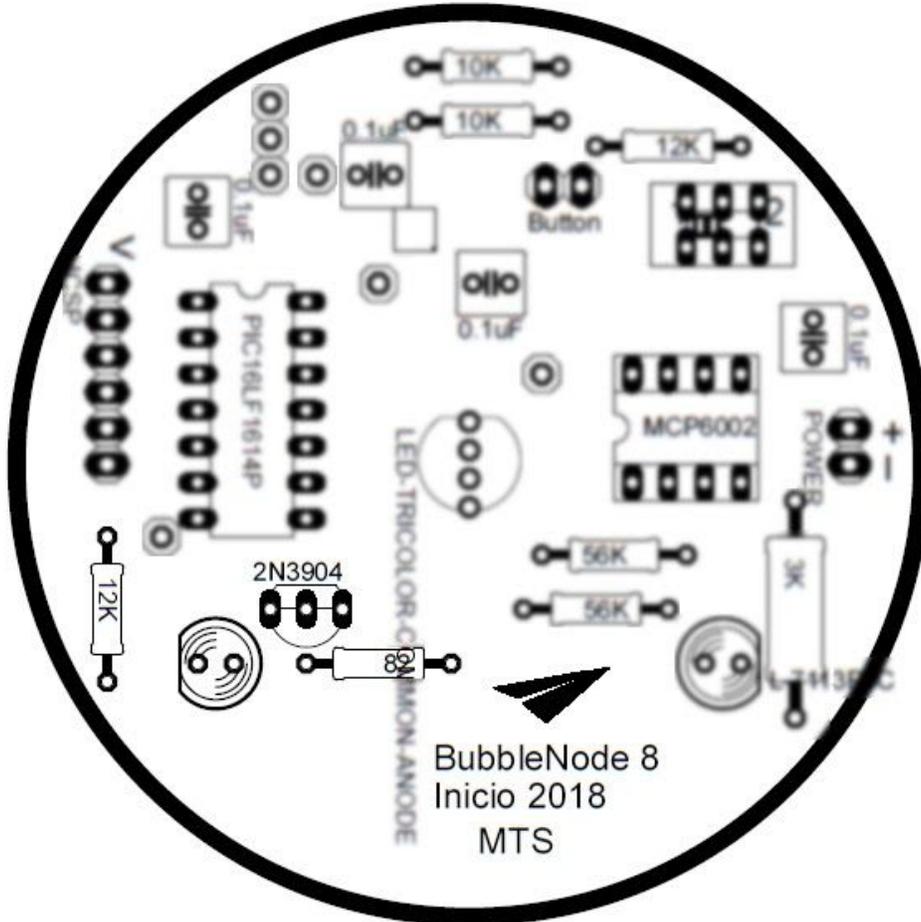


RGB LED



Long leg goes into Purple hole

## Step 4. Data transmitter



Transistor  
2N3904



Make sure the  
transistor is in  
the right  
direction



IR LED

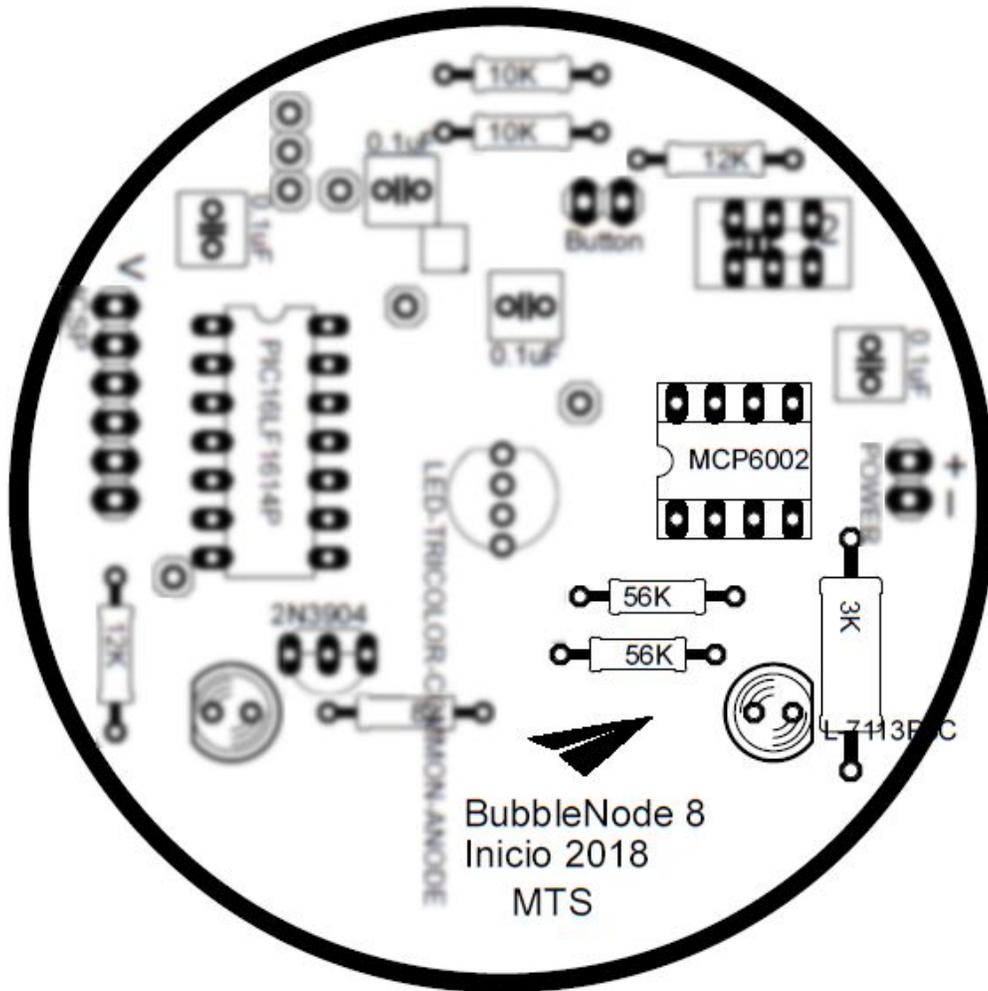


Resistor  
12K $\Omega$



Resistor  
82 $\Omega$

## Step 5. Data Receiver



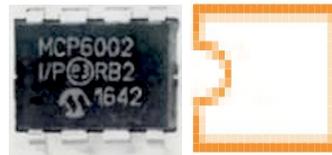
Phototransistor



Resistor  
3K $\Omega$



Resistor  
56K $\Omega$



OP Amp  
MCP6002

Make sure the  
OP Amp is in  
the right  
direction

## **6.2 Appendix A2: Konstfack experiment user manual**

# BUILDING YOUR BUBBLENODE



# CONTENTS

INTRODUCTION .....	3
STEP 1: POWER .....	4
STEP 2: PROCESSING .....	6
STEP 3: 3 COLOUR LED .....	8
STEP 4: DATA TRANSMITTING .....	10
STEP 5: DATA RECEIVER.....	12
RESISTOR COLOUR CODE.....	14
CONCLUSION .....	15

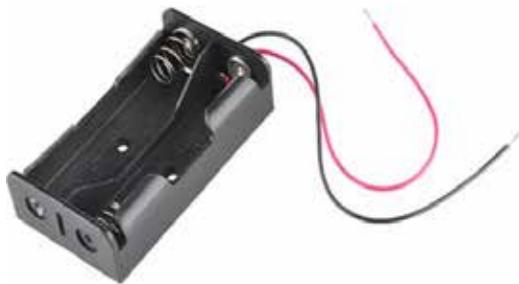
**BUILDING A BUBBLENODE IS EASY,  
AND WHAT'S VERY COOL IS THAT YOU  
BUILD IT IN ALMOST EXACTLY THE  
SAME WAY THAT A LOT OF  
ELECTRONIC PRODUCTS ARE MADE.  
BUBBLENODES ARE MADE BY  
CONNECTING PARTS TOGETHER ON A  
PRINTED CIRCUIT BOARD. THE PARTS  
ARE CONNECTED TO THE BOARD BY  
USING A METHOD CALLED  
"SOLDERING". DON'T WORRY IF YOU  
HAVE NEVER SOLDERED BEFORE. IT'S  
EASY AND WE WILL SHOW YOU HOW  
TO DO IT.**

## STEP 1: POWER

EVERYBODY NEEDS TO HAVE POWER, IT DOESN'T MATTER WHO YOU ARE YOU NEED ENERGY TO GET THROUGH YOUR DAY. THE BUBBLENODE NEEDS ENERGY TOO. IN THE BUBBLENODE'S CASE, ENERGY COMES FROM BATTERIES. BATTERIES CONVERT CHEMICALS INTO ENERGY.

THE FIRST STEP IS TO BUILD POWER. FOR THIS REASON, YOU WILL BE CONNECTING THE FOLLOWING COMPONENTS:

**THE BATTERIES  
TO PROVIDE  
POWER**

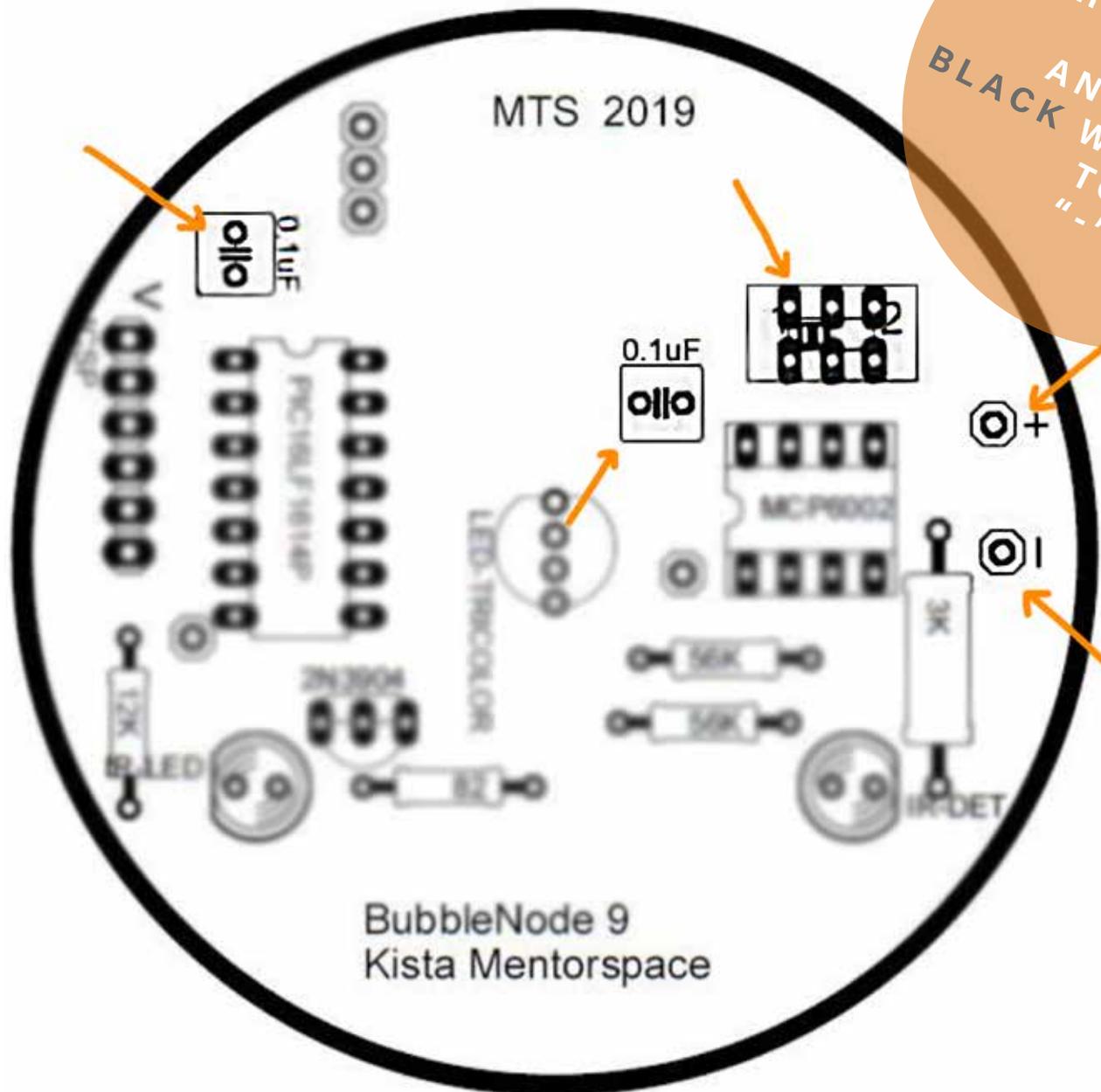


**THE CAPACITOR  
TO STORE ENERGY  
FOR LATER USE.**



**THE SWITCH  
TO TURN THE  
POWER ON AND OFF**





MAKE SURE THE RED WIRE GOES TO "+" AND THE BLACK WIRE GOES TO "I"

MTS 2019

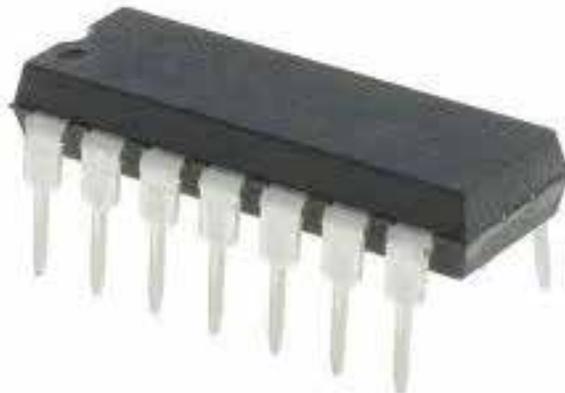
BubbleNode 9  
Kista Mentorspace

## STEP 2: PROCESSING

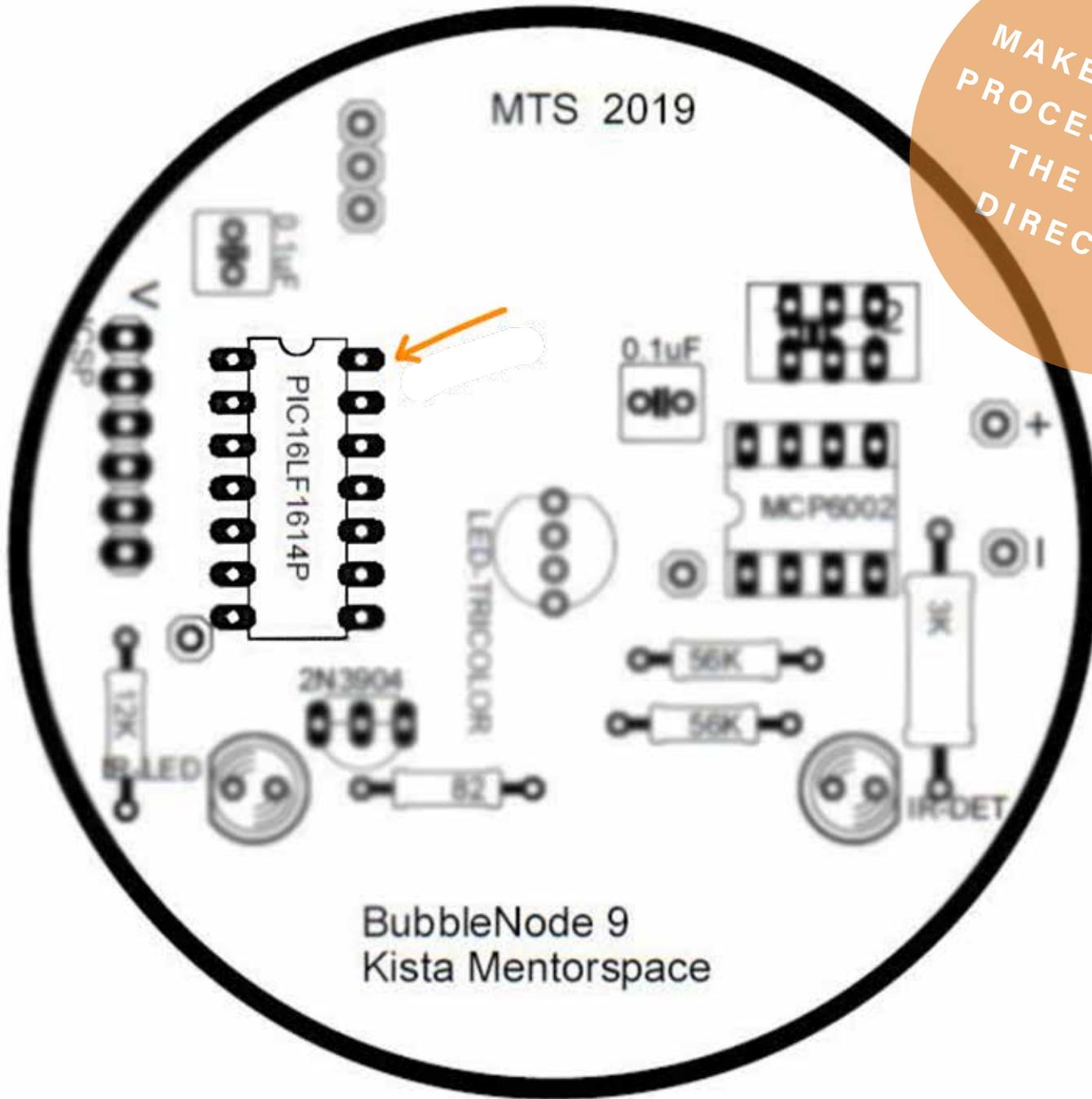
**PROCESSING MEANS TO COMPUTE SOMETHING. YOU CAN VIEW THE PROCESSOR LIKE A LOW RESOLUTION BRAIN. EVERY ELECTRONICS GADGET NEEDS A BRAIN, BUT THE QUESTION IS; WHAT DOES IT KNOW? IT KNOWS HOW TO SHARE DATA WITH OTHER OBJECTS, HOW TO DO ARITHMETIC, AND IT KNOWS HOW TO USE THE DATA TO MAKE THINGS HAPPEN.**

**WITHOUT THIS, NOTHING ELSE WILL FUNCTION.**

### **THE PROCESSOR PIC16LF1614P**



MTS 2019



MAKE SURE THE  
PROCESSOR IS IN  
THE RIGHT  
DIRECTION!



BubbleNode 9  
Kista Mentorspace

**STEP 3:  
3 COLOUR LED**

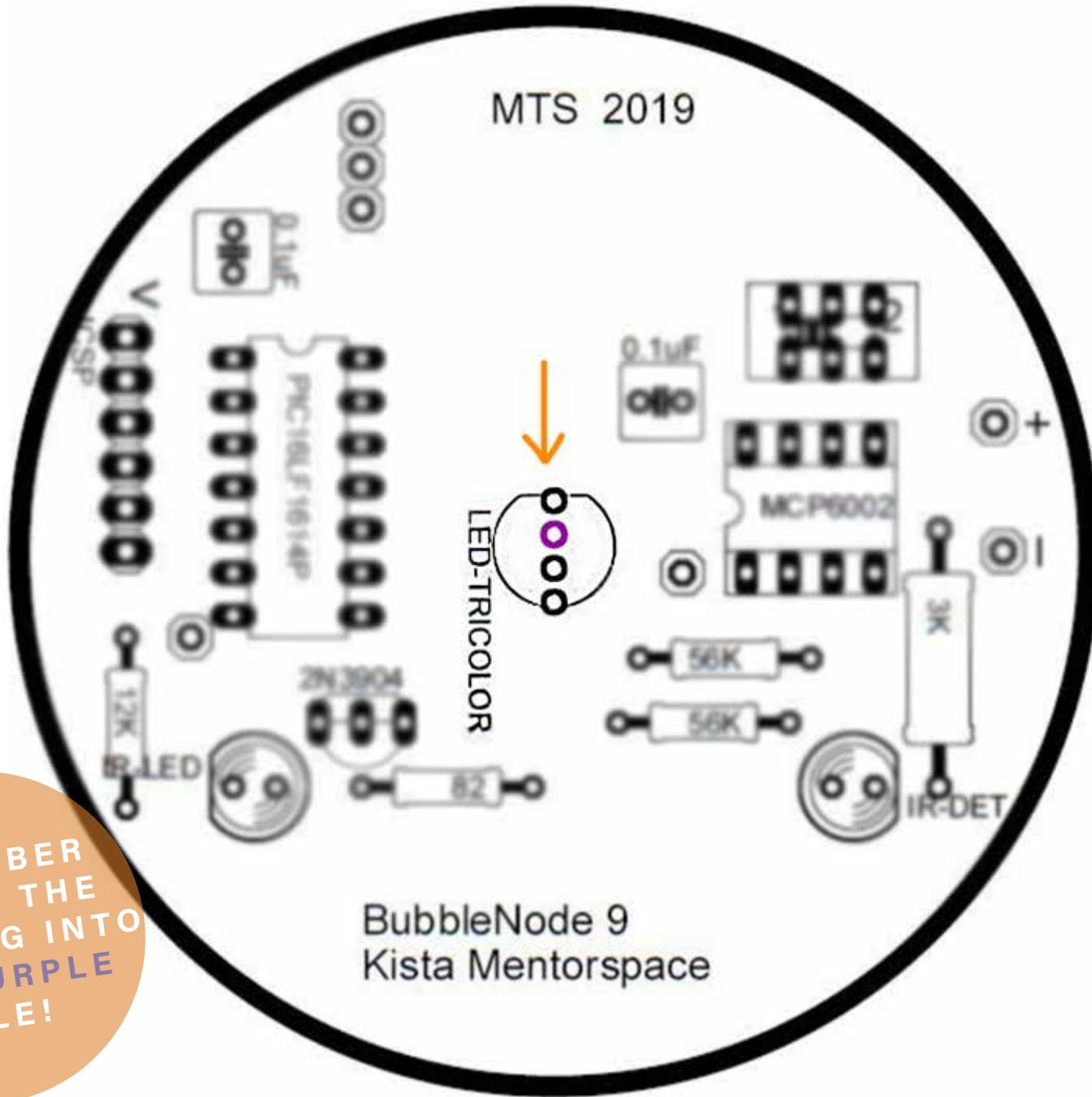
**ONE OF THE THINGS THAT YOU CAN LEARN IS HOW OBJECTS CAN INTERACT WITH PEOPLE. AS AN ARTIST, YOU KNOW YOU CAN DO THAT IN MANY VISUAL WAYS. THE BUBBLENODE INTERACTS WITH YOU BY USING LIGHT WITH DIFFERENT COLOURS, THAT IS THE JOB OF THE THREE COLOUR LED.**

**IT LIGHTS UP IN RED, GREEN AND BLUE AND BY MIXING THESE COLOURS YOU CAN MAKE ANY OTHER COLOUR YOU WANT, AS YOU PROBABLY KNOW.**

**RGB LED  
LED-TRICOLOUR**



MTS 2019



REMEMBER  
TO PUT THE  
LONG LEG INTO  
THE **PURPLE**  
HOLE!

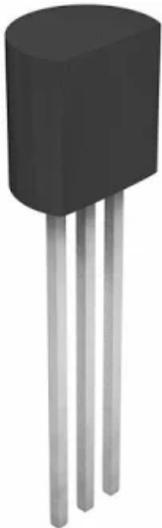
BubbleNode 9  
Kista Mentorspace

**STEP 4:  
DATA TRANSMITTING**

**IN ADDITION TO INTERACTING WITH PEOPLE,  
THE BUBBLENODE CAN INTERACT WITH OTHER OBJECTS  
LIKE ANOTHER BUBBLENODE. IT DOES IT THROUGH  
SHARING DATA. IT USES INFRARED LIGHT THAT YOU  
CANNOT SEE, BUT THE OTHER BUBBLENODE WILL.**

**TO ACHIEVE THAT, YOU WILL BUILD A DATA  
TRANSMITTER BY PUTTING TOGETHER THE FOLLOWING  
ELEMENTS:**

**A TRANSISTOR  
2N3904**



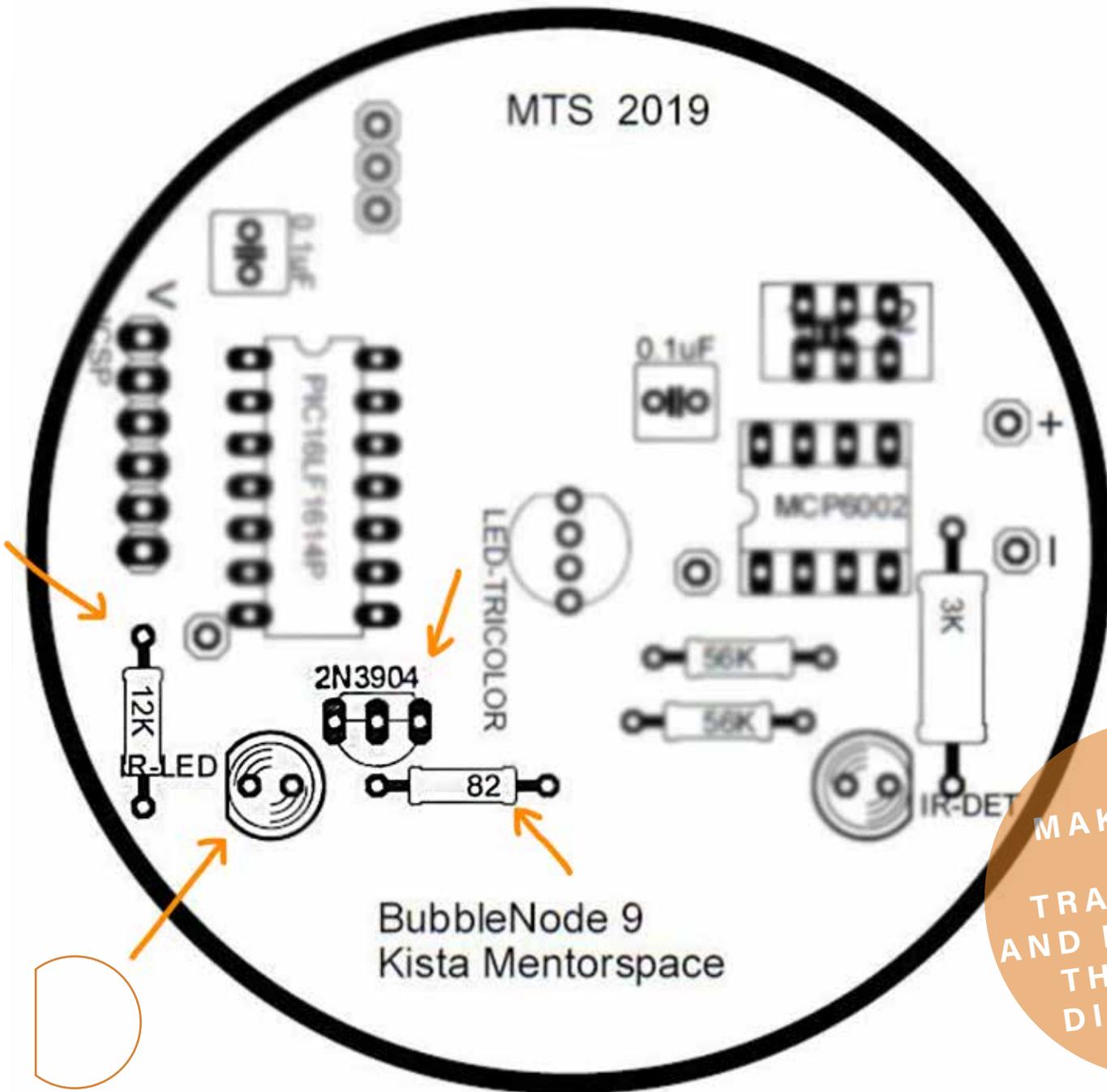
**AN IR LED**



**TWO RESISTORS  
12K  $\Omega$  AND 82K**



MTS 2019



BubbleNode 9  
Kista Mentorspace

MAKE SURE  
THE  
TRANSISTOR  
AND LED ARE IN  
THE RIGHT  
DIRECTION!

**STEP 5:  
DATA RECEIVING**

**NOW, YOU WILL BUILD THE PART OF THE BUBBLENODE THAT HEARS THE DATA SENT FROM ANOTHER BUBBLENODE. HERE, THE CIRCUIT WILL RECEIVE THE DATA AND TRANSLATE IT TO THE PROCESSOR THAT YOU PUT IN STEP TWO. IN OTHER WORDS IT IS A WINDOW TO THE OUTSIDE WORLD.**

**TO BUILD YOUR DATA RECEIVER, YOU HAVE TO PUT TOGETHER THE FOLLOWING ELEMENTS: AN OPAMP, A PHOTOTRANSISTOR AND TWO RESISTORS.**

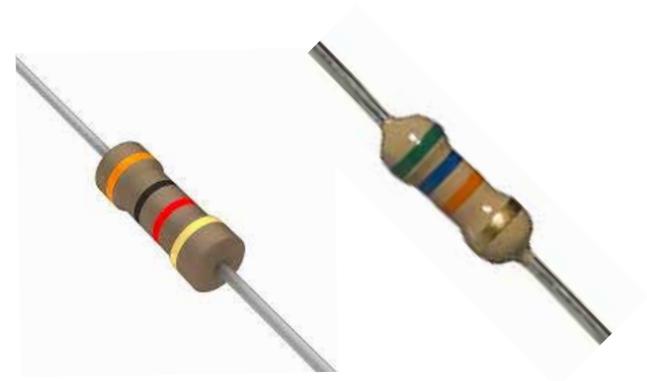
**AN OP AMP  
MCP6002**



**PHOTOTRANSISTOR**

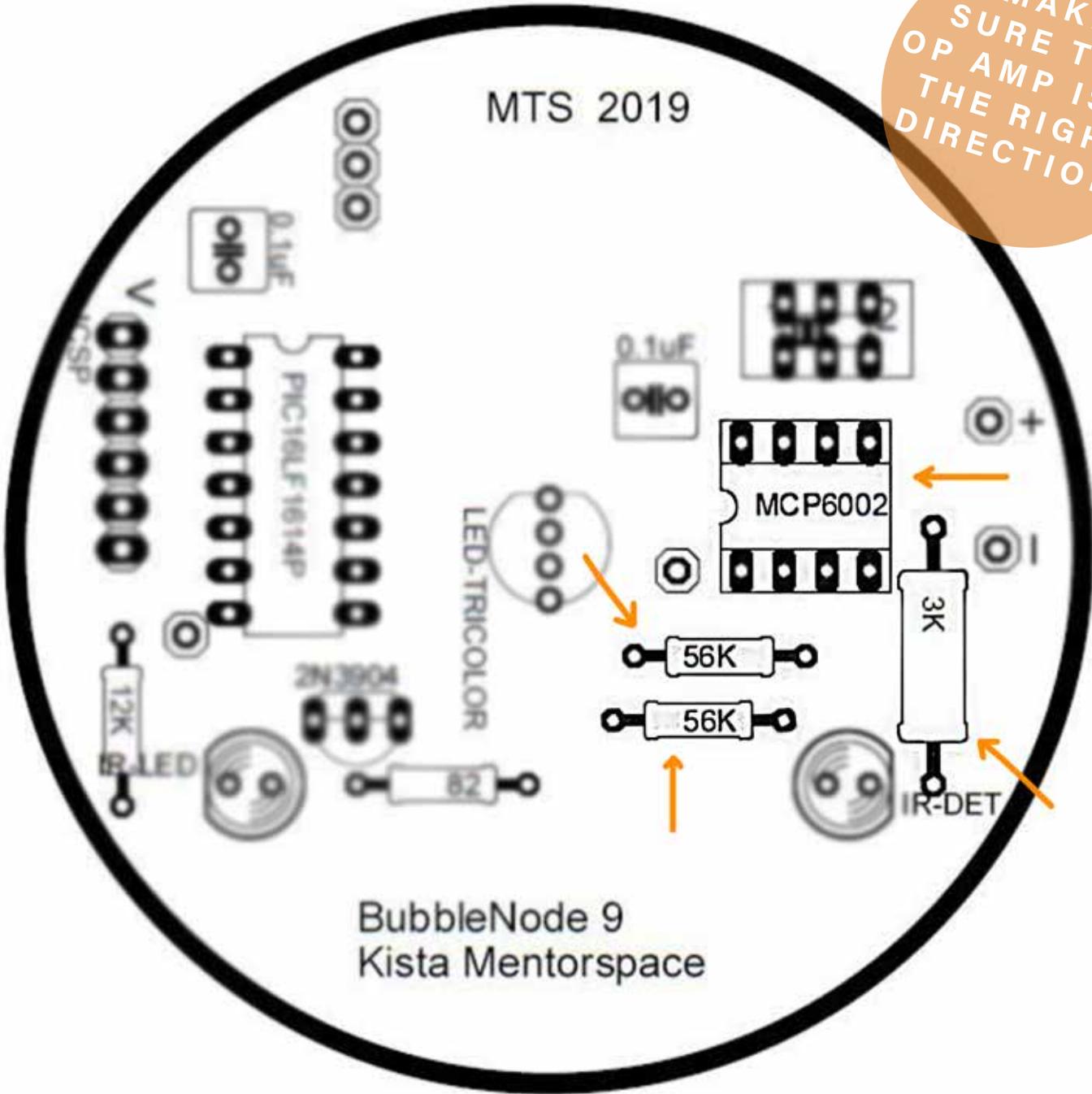


**TWO RESISTORS  
3K  $\Omega$  AND 56K  $\Omega$**



MTS 2019

MAKE SURE THE OP AMP IS IN THE RIGHT DIRECTION!



BubbleNode 9  
Kista Mentorspace

RESISTOR COLOUR  
CODE



YELLOW VIOLET RED

4 7 2

4 7 X 10<sup>2</sup>

4K7 Ω

Colour	Digit	Multiplier	Tolerance
Black	0	1	
Brown	1	10	± 1%
Red	2	100	± 2%
Orange	3	1,000	
Yellow	4	10,000	
Green	5	100,000	± 0.5%
Blue	6	1,000,000	± 0.25%
Violet	7	10,000,000	± 0.1%
Grey	8		± 0.05%
White	9		
Gold		0.1	± 5%
Silver		0.01	± 10%
None			± 20%

**THE END! GO DISCOVER YOUR  
IDENTITY COLOR! WHEN YOU PUT ALL  
THE STEPS TOGETHER, THE BUBBLE-  
NODE WILL INTERACT WITH YOU AND  
WITH OTHER BUBBLENODES.  
EVERY BUBBLENODE HAS ITS  
IDENTITY COLOR, GO DISCOVER IT BY  
INTERACTING WITH OTHERS!**

**NOW, ENJOY IMAGINING HOW YOU  
CAN INTEGRATE SUCH A  
TECHNOLOGY IN YOUR ARTWORK!**

### **6.3 Appendix A3: Participant consent form**

# Consent Form

## Purpose

This study investigates the quality of the Kista Mentorspace processes and impact it has during the educational workshops it offers. As part of this study, you will be asked to participate in a workshop and a focus group and answer structured and open-ended questions.

## Participants' Rights

I understand that my data will be kept in the strictest of confidence and will be available only to the researcher. No one will be able to identify me when the results are reported and my name will not appear anywhere in the written report. Please do not share other people's identities or responses from the focus group with others to maintain the anonymity of the participants outside of the focus group. I also understand that I may skip any questions or tasks that I do not wish to answer or complete. I understand that the consent form will be kept separate from the data records to ensure confidentiality. I may choose not to participate or withdraw at any time during the study without penalty. I understand that my data will not be linked to me personally in any way. I understand that in any report on the results of this research my identity will remain completely anonymous and will not be involved in any data manipulation of any kind.

If I am uncomfortable with any part of this study, I may contact Nina Laribi, master student at the University of Twente [t.laribi@student.utwente.nl]

I understand that I am participating in a study on my own free will.

## Consent to Participate

I acknowledge that I am at least eighteen years old, and that I understand my rights as a research participant as outlined above. I acknowledge that my participation is fully voluntary.

**Print Name:** \_\_\_\_\_

**Signature:** \_\_\_\_\_ **Date:** \_\_\_\_\_

## 6.4 Appendix A4: Mentor consent form

# Consent Form

## Purpose

You have been invited to participate in a study under the direction of Nina Laribi. The purpose of this experiment is to investigate the quality of the **Kista Mentorspace** processes and impact it has during the educational workshops it offers.

You will be asked to assist and mentor students from another university to build an embedded system. Data will be collected around the interactions between the participants involved. This data is independent from anything that might reveal your identity.

## Benefits and Risks

Your participation may benefit you and other mentors by helping the **Mentorspace** improve their processes and mentor training. However, no risks are anticipated beyond those experienced during a regular event.

- I..... voluntarily agree to participate in this research study. I understand that even if I agree to participate now, I can withdraw at any time or refuse to answer any question without any consequences of any kind.
- I have had the purpose and nature of the study explained to me in writing and I have had the opportunity to ask questions about the study. I understand that all information I cause for this study will be treated confidentially.
- I understand that in any report on the results of this research my identity will remain completely anonymous and will not be involved in any data manipulation of any kind.
- I understand that the data generated in this study is completely independent of any details which may reveal my identity or the identity of people I speak to .
- I understand that I am free to contact any of the people involved in the research to seek further clarification and information.

## Contact

If you have any questions or concerns regarding this study, please contact:

Nina Laribi, University of Twente [t.laribi@student.utwente.nl]

I understand this information and agree to participate fully under the conditions stated above

**Signature of participant**

**Date**

.....

.....

**Signature of researcher**

**Date**

.....

.....