**Healthy Child is a Happy Child: Designing an AI-based Coach for patient safety and mental wellbeing**

**Abstract 200 words**

# Introduction

Healthcare professionals work in a high-risk landscape, where a minor medical error or unsafe incidents may lead to substantial hazards to patient safety (Panagos & Pearlman, 2017; Stavroudis et al., 2008). . This challenge is often pronounced in intensive care units or operation theatres where this probability is quite elevated. Prevention of medical errors is widely acknowledged, given their potential repercussions on patient safety and mental wellbeing (Björkstén et al., 2016; Petrovic et al., 2022; Stavroudis et al., 2008). This requires cohesive teamwork among healthcare professionals, which emphasizes mutual understanding for diagnosis, the condition of patient and the plan of action (Panagos & Pearlman, 2017). Artificial Intelligence has emerged as a transformative force in the domain of healthcare, and can play an important role in improving healthcare.

AI has not only revolutionized the disease diagnostics (Ali et al., 2023), but also helped to improve mental and psychological health of an individual by different virtual assistants. These assistants are goal oriented and usually designed for self-monitoring and self-rehabilitation of an individual (Tsiouris et al., 2020). They allow real-time monitoring and provide (text or voice based) assistance to an individual for a healthier life style. However, none of the designed assistants were found to ensure safety and to avoid medical errors in neonatology, which may have life-lasting influence (Tsiouris et al., 2020).

In this study, we propose an artificially intelligent assistant - AI Coach, which aims to interact with medical teams to ensure patient safety in the domain of neonatology. This coach is preliminary designed to work together with the medical team, and may intervene during child birth if safety of any of the stakeholders is desired. The stakeholders may include doctors, residents, neonate, and family of neonate. This innovative coach will be designed using living labs methodology in which all stakeholders collaborate to design a prototype that can address complex phenomena in neonatology (Almirall et al., 2012; Leminen et al., 2012). Therefore, in this study, we address a) the framework of AI Coach, b) how a user will interact with the coach, and c) the test results of generated after using this coach.

The remainder of this paper is organized as follows. Section 2 explains the related work of AI Coach, while Section 3 explains the requirements and methodologies of the coach, while Section 4 presents the related interface design. Section 5 present the results for the coach evaluations, and the paper concludes in Section 7, after mentioning the limitations of the study in Section 6.

# Related work

This section lays the foundation for our study by exploring the interplay between neonatal care, medical errors, innovative methodologies, and virtual assistance in healthcare. First, we will explain the neonatal care and its inherent challenges, notably the medical errors associated to it. Secondly, we will explain the methodologies used for innovation. Lastly, we will explain the virtual assistance available in healthcare to improve the performance of healthcare sector.

## Medical Errors during Childbirth

Childbirth is a team work, in which different stakeholders work together as a team. These stakeholders may involve mother, family of a neonate, doctors and residents working on the floor, and the hospital management. The experience from all stakeholders, the ‘now’ moment of baby birth, and the conscious decisions made for the neonate are of prime importance, as they ensure the care and safety of an infant (Stavroudis et al., 2008), and the involved stakeholders (Profit et al., 2017).

Many departments are involved to ensure quality care during the child birth. For instance, maternity and gynae department who deal with pregnancy, labor, delivery and postnatal period (Petrovic et al., 2022), or neonatal care unity to ensure the safety of neonate (Profit et al., 2017). According to Profit et al. survey, few cases were reported with poor teamwork (Profit et al., 2017), resulting different medical errors (Stavroudis et al., 2008). This may have long-lasting impacts on a baby, and in severe cases, it might lead to fatality (Stavroudis et al., 2008). Therefore, on one side, it is essential to determine the state of the error to establish treatment plan is to be discussed if needed. While on the other side, preventive measures should be taken into account to ensure safety of all the stakeholders in healthcare (Björkstén et al., 2016; Petrovic et al., 2022; Stavroudis et al., 2008).

Stavroudis et al. also emphasize that a significant number of such errors are preventable by implementing a well-devised plan can mitigate the impact of these errors. In this context, consistent monitoring is regarded as playing a crucial role. Medical errors are not deliberately made, rather there are different factors mentioned by Björkstén et al. which may lead to a medical error. Forgetfulness and lack of attentiveness (399 cases among 772) was found to be the main cause for such inadvertent mistakes (See Table 3: (Björkstén et al., 2016)). Research shows that ‘patient safety’ and ‘just culture’ requires commitment and participation of all stakeholders including patient, healthcare team and the hospital management (Stavroudis et al., 2008).

AI has revolutionized the way we approach healthcare, many analytical tools are being used by healthcare professionals to provide healthcare facilities to their patients. However, mostly tools can only help to detect the errors(or deviations) through data analysis. So, they might miss the ‘now moment’ that should be also considered for a conscious choice (Stavroudis et al., 2008). In the following sections, we will discuss state of the art methodologies to design such systems that be helpful to understand the ‘now moment’, and analyzing how such systems can be helpful to overcome the medical errors in neonatology.

## Current Methodologies for Innovation

There are several methodologies aimed to involve technical innovations particularly focusing on healthcare. Few well-known examples are user-centered design, in which users are engaged to design and evaluate a (user-intuitive) technology (Giacomin, 2014); or agile development, in which iterative and incremental development is focused among different team members (Manifesto, 2001). These methodologies are merged with clinical trials and studies to ensure the safety and efficacy of new technologies (Stepke, 2012).

Furthermore, several research-based methodologies are also addressed where research and technical innovations are integrated to provide solutions. These methodologies ensure that technical advancements are developed in tandem with the insights of researchers to produce effective solutions. For example, design science research (DSR) where technology-based solutions (or other artifacts) are created to solve some specific problem and thus evaluated in a controlled environment (Hevner et al., 2008). Another example is ethnographic research, where user behavior and needs are considered to design the relevant technological solution (Hammersley & Atkinson, 2007). Co-Creation allows different stakeholders to ideate and design an artifact (Sanders & Stappers, 2008). Living labs is a methodology closer to co-creation, where end-users and researchers cooperate to prototype and validate solutions in real-life contexts (Almirall et al., 2012; Leminen et al., 2012). This methodology is particularly focused to test innovations while analyzing users’ behaviors which can address the context of ‘now moment’ in a much better way.

## Virtual Assistance in Healthcare

From many years, pressure on healthcare have been increasing due to reduced availability or accessibility of healthcare professionals. To address this challenge, intelligent conversational agents or virtual assistants were developed. For example, Eliza was the first conversational agent ever designed (Weizenbaum, 1983), or Woebot, who helped as therapy chatbots (Fitzpatrick et al., 2017), and proven themselves for behavioral change. Ever since, there is a huge trend on using them with healthcare and improving their capabilities. Therefore, not only focus they on human understanding (through natural language processing), but also aim to understand the context of conversation through machine learning and artificial intelligence.

Unlike conventional agents virtual coaching system are known for their healthcare interventions. They not only offer an opportunity for mental wellbeing but it also allows the individuals to self-track and e-coach for their goals (Tsiouris et al., 2020). A system-human interaction can be possible by a conscious choice of users (by clicking certain option) or by monitoring sensors. For instance, in home coaches data maybe gathered by different sensory devices (e.g., Fitbit, Samsung, Polar, etc.), or in hospital where data maybe gathered by Virtual Reality (e.g., Microsoft Kinect, Nintendo, sensing gloves) or by medical equipment (e.g., ECG, vitals monitor, or accelerometers, etc.) (Tsiouris et al., 2020). In Neonatology, a profound emphasis is placed on the comprehensive care of the mother and neonate, and Artificial Intelligence can be play its role. For example, serious gaming has been used to strengthen the learning methodology for the childbirth (Bardelli et al., 2022; Hu et al., 2021). Similarly, root cause analysis can help to identify the errors and improve the action plans (Panagos & Pearlman, 2017). Treatment analysis and drug interactions have already been studied in clinical care (Panagos & Pearlman, 2017), however, the role of virtual assistants in neonatology is not studied so far (Tsiouris et al., 2020).

Moreover, conversational agents are designed to provide a textual interface to interact with its users. This which may need some time and attention to build the rapport and to achieve the goals, as it intends to understand the natural language. Similarly, Tsiouris et al. mentioned that different virtual coaches were designed to interact with a single individual (See Table 1 in (Tsiouris et al., 2020)), to overcome phobias and cognitive impairments (See Table 2 in (Tsiouris et al., 2020)). Such interactive agents are user-centered which are designed to fulfil certain user requirement and goals. However, in our knowledge, there is no virtual coach designed to avoid possible medical errors in neonatology, where team-based interaction is involved (See Section 2.1. and 2.2). Following section discusses the material and methods used in the study.

# Requirements and Methodologies of AI Coach

This section presents the whole process that lays the foundation for the development of AI Coach. Initially, we will discuss the design requirements for AI Coach, then its framework design, and then we present how does the coach can interact with its stakeholders.

## Design Requirements

User-centered coaches or agents involve their end users throughout development cycle. So, the ultimate success of such systems depends whether the end users will find such applications useful. Involving users can be helpful to reduce the design and the development time. This can result in more accurate assessments regarding user requirements. System specific requirements can be gathered from interview sessions, system analysis, or use-cases (Rosson & Carroll, 2002). Following are the fundamental concepts which were considered to design and implement AI Coach. The aim of AI Coach is to provide assistance during childbirth while acting as a team-player. For Design following approaches were considered.

### Formulation of Design Outline

Living Labs thrives on engagement and collaboration to align the system design with the real-world requirements (Almirall et al., 2012; Leminen et al., 2012). Interviewing and iterative discussions can be a useful tool used for system design, as they allow a direct interaction with the involved stakeholders. Also, It helps to gain a deeper understanding of the needs and expectations of the users (Almirall et al., 2012).

The interviews were conducted in a (semi-)structured way, where narration of critical events (like child birth) was considered to extract scenarios (See (Doornkamp et al., 2022; Weigl et al., 2022; Xu et al., 2022)), and to formulate the design of AI Coach. Thus, the involved stakeholders helped us to extract and identify the user requirements as per the neonatology environment. It is to be noted, that the focus of these interview sessions was a) what are the possible situations that can lead to such medical errors (termed as scenarios), and b) how can we design a system that can ensure proper execution of protocol. The aim was to avoid the errors that can be caused by forgetfulness (Björkstén et al., 2016).

### Attachment Theory and Human Emotions

According to attachment theory, the human coach interaction should be based on different (human-based) factors. For instance, virtual system should already be aware of the desires and goals of its user(s). Similarly, to make certain decisions, the system should be aware of the human emotions and the related environment. Over the period of time, such a system should learn to be ‘like-minded’, so that a relationship can be maintained with the user (Roelofsma, 2013).

While discussing about Neonatology, this theory of secure attachment can be extended to the emotional bonds among the team members, where team members are comfortable with interdependence and trust each other's abilities, leading to effective communication, cooperation, and problem-solving. To have this harmony and just culture, all of the team members should have adequate ‘shared mental models’, to observe and communicate each other’s actions.

### Shared Mental Models

Neonatology department has necessary and sufficient content complexity, where all team members interact together. Due to multiple factors, this collaboration could result in medical errors (See Section 2.1). To ensure patient safety, all the team members must have a shared understanding of diagnosis to develop any mutual plan of action. In other words, they should have ‘shared mental models’ (SMM) to communicate and deicide for a certain plan of action. They explain how individuals in a group develop their specialized knowledge (through mental models) and rely on each other’s expertise to interact and learn to share and enhance their mental and cognitive abilities, thus enhancing the team coordination and performance (Mathieu et al., 2000).

Computational models allows inferring or deducing cognitive states of an individual, which can predict user(s) or team(s) behaviors (Treur, 2020; van Ments et al., 2021). Van Ments et al, have explained how doctor and nurse can interact together as a team to perform the related medical task (i.e., tracheal intubation of a patient) through shared mental models. Similar behaviors are also addressed in other studies (Doornkamp et al., 2022; Weigl et al., 2022; Xu et al., 2022). These behaviors can be facilitated by trivial secondary tasks like pressing a button, or answering in a specific way, e.g., through questionnaires (Clark & Maguire, 2020) which can help in better understanding of human behaviors and tendencies.

Shared mental models allows effective team collaboration, which can improve decision-making regarding a task. Also, when combined with Living Labs, it emphasizes the importance of collaborative understanding among the participants’ behaviors and preferences. This can also help to design a supportive system. So, our goal was to design an interactive virtual assistant who could interact as a ‘virtual team member’, i.e., AI Coach. This may represent a visualization tool that can act as a teamwork facilitator (Redlich et al., 2017). That is, it can address different reasoning mechanisms (e.g., human factors) to interact with the users in relevance with the medical protocol, and its execution.

## Framework Design of AI Coach

In IT, a framework refers to a structure that consist of many phases or components working together to act as a foundation for the software design and implementation. This article introduces a novel Human-AI based framework to support medical teams in different healthcare settings. This is a **M**odel**-V**iew**-C**ontroller based architecture, in which the AI Coach is designed to be extended as per user needs. It has ability to offer assistance to variety of end users, which can be expanded depending upon the processes, user behaviors and culture(s) of an organization like medical healthcare. The framework is mainly composed of two components, i.e., AI Coach and the stakeholders of a system (See Figure XXX).

A diagram of a server

Description automatically generated

**Figure** **XXX**: Block Diagram of AI Coach system

### Stakeholders

AI Coach system is a software application that aims to interact and serve two type of stakeholders in particular, i.e., end users and experts.

**End users** are the stake holders who interact directly with the system to achieve their goals, learn and have an pleasant experience via usage of technology. AI Coach is designed to serve neonatology in Erasmus MC, for example during baby delivery/care, or intensive care unit. Therefore, it not only has the capability to adhere proper neonatal protocol execution but it can also foster emotional and social wellbeing of the people who are actually on the floor, e.g., patient, their family, or the medical team working together in neonatal care. Understanding the time constraint the coach is designed to fulfil its objective through minimum user engagement. These users are further discussed in Section 3.2.

While **experts** are the stakeholders which work in different domains and interact to provide support in a meaningful activity. They can from different domains like psychology, artificial intelligence, health and safety sciences, who interacted with each other to ensure patient safety and just culture. The main focus of experts is to brainstorm and develop the algorithm of the coach. As a result, the behavior of AI Coach behavior is adaptive by nature. This is accomplished by developing (shared) mental models (van Ments et al., 2021). Further details can be found in Section 3.2.

Current role of AI Coach is to serve neonatology department in Erasmus MC, Rotterdam. So, it can provide support for neonatal care and wellbeing of patient(s) and medical team, where the possible end users are; baby, patient, their family and the staff on the floor.

### AI Coach System

AI Coach is an AI based virtual assistant whose design is inspired by the principles of sociotechnical systems (i.e., with the focus on social systems and technical systems in healthcare). Therefore, primarily it is human centered and have a social impact on its users. Still, it is quite flexible to address any changes in the social landscape over the time (within an organization). Therefore, all the stakeholders are considered for human interaction with the AI Coach.

AI Coach is designed with a deep understanding of needs of the medical team who is working on the floor in the neonatology department (See Section 3.1.1). This can address different social situations/aspects, capabilities and behaviors in different situations like child birth. For instance, if a baby is not breathing, AI Coach will assist to achieve the particular goal of helping the baby to breath (Xu et al., 2022). Similarly, if a resident wants to speak up during any medical process the coach will help him/her to speak up in a certain situation (Doornkamp et al., 2022). Therefore, end user interaction is adaptive which follows through the series of choices which may have been (un)consciously made by the end users.

Similarly, the design of AI Coach is quite flexible, which means that an organization (like Erasmus MC) can design the usage of AI Coach as per needs (via their expert users). This may address changes in environment, technology or organization, which may evolve over time. In other words, the stakeholders can become an architect of AI Coach. As a result, this may not only effect the assistance process of the coach, but it may also affect the roles of the involved stakeholders and the coach itself. In the section below, we describe the components and functionality of AI Coach.

#### Components of AI Coach System

In this section, we will describe technical aspect of AI Coach system, i.e., how it was designed to offer functionality to its end users. AI Coach is a server based virtual assistant, that is designed based on requirements were formulated through the outlines from the staff of neonatology department in Erasmus MC. Few important considerations were taken into account during the design time, i.e.,:

1. Job is physically demanding, which may lead to forgetfulness, stress or influence on the personal skills of a resident or a doctor
2. Job has a time constraint, they have to respond in a quick and timely manner
3. Baby birth is a team work, and safe execution of protocol ensures no errors in the procedure
4. Different situations may lead to protocol deviations (medical errors) which may lead to poor mental wellbeing, and sufficient support is desired in this.

Key stakeholders were identified and invited to take part in the developmental process of AI Coach for neonatal care. Different stakeholder groups were identified through scoping interviews, therefore, human-computer interaction mainly involved two types of actors, i.e., the residents or the doctors. AI Coach had to considered the rest of the stakeholders for assistance, i.e., the neonate, mother, father (or a family member). To offer a deeper analysis, we undertook formal stakeholder analysis, where individual behaviors and related support was modeled along with relevance to the organization and available resources.

Interviews reflected that, shared contribution of all stakeholders ensure better performance and improved safety in neonatology department. We also considered different kind of data (e.g., protocols, user information) to dive into the possible situations that may cause poor adherence of protocol. As a first step, the protocols were investigated along with the interviews, to see the possible deviations in the healthcare protocol of neonatal care (Madar, Roehr, Ainsworth, Ersda, et al., 2021), which helped us to extract use-cases. *Use-case* is a term that describes the functional requirements and explain how stakeholders can interact with the system. Table XXX mention few examples in which each use-case is addressed through possible scenarios. i.e., a success scenario and when AI Coach predicts that support is desired or intervention is necessary.

Table usecases: Scenarios (use cases) in neonatal care

|  |  |  |  |
| --- | --- | --- | --- |
| **Use-case (step)** | **Success Scenario** | **Possible Intervention desired** | |
| **When** | **Possible Causes of Failure** |
| Access neonate breathing and tone | Neonate is breathing adequately and tone is good | Baby is not breathing or needs ventilation support | Mask is not appropriate  Sufficient FiO2 levels are not provided |
| Communication Behaviors | Members of medical team can communicate properly | Poor Communication due to different possible factors | Residents/team members can’t communicate |
| Family of Neonate may get affected | Delivery makes everyone happy | Development of postpartum depression in parents | Adequate support isn’t provided |

Once the requirements and expectations of the stakeholders collected, we designed the basic layout of AI Coach System that can support these use cases. The system layout is divided in two parts, server and a controller. Expert users are allowed to access and design the knowledgebase for AI Coach, that play role in the computational logic of the AI Coach, whereas, end users (i.e., medical team) interact through terminals to monitor and achieve these goals.

##### Knowledgebase for AI Coach

For every conversational agent or virtual assistants a related knowledgebase is vital, as it provides a basis for human computer interaction. Section 2.3 indicate how assistants or conversational agents can use domain knowledge to interact with its users. Here, the knowledgebase of system consist of ‘Support Knowledgebase’ and ‘Protocol Knowledgebase’ (See Figure XXX).

To generate knowledgebase, we modeled the collected requirements which can help the team to follow the medical protocols, through shared mental models approach (van Ments et al., 2021). The modeled scenarios reflected how different stakeholders can interact with each other to achieve a success scenario. For instance, Xu et al. explains how team should interact to ensure baby can breathe normally (Xu et al., 2022), or how mother can face postpartum depression (Weigl et al., 2022), or how can speaking up behavior of a resident can ensure risk free environment and help in proper execution of protocol (Doornkamp et al., 2022). For this, we also explored when AI Coach needs to intervene and ensure the patient safety and mental wellbeing by different support related messages. The interactions between stakeholders and the support related messages are considered as knowledgebase resource for the human interaction with AI Coach.

The above mentioned studies explore how social elements can be integrated into the design of AI Coach. For example, we studied how a healthcare practitioner may respond in case of development of postpartum depression among parents for a behavioral change. A similar design methodology was considered for the interaction of AI Coach, which aims to reflect the similar behavior towards a family member(s), i.e., to provide support to avoid postpartum depression (Weigl et al., 2022). Thus, the AI Coach interacts to avoid postpartum depression or help if it is developed in father (or a family member). If certain factor is found AI Coach notify the support messages from the support knowledgebase using visual and audio support. The aim of support knowledgebase is to use support messages to make users aware of what is happening or about to happen. Few example some support messages can be seen in Table supmsgs. Therefore, these models not only ensures that system meets its intended goals (adherence of protocol and mental wellbeing by monitoring), but they also manifest the possible behaviors of AI Coach, which may vary as per scenarios. The dynamic behaviors of coach are discussed below, while the related interfaces will be discussed in Section 3.2.

Table supmsgs: Example Support messages that can be used for human coach interaction

|  |  |  |
| --- | --- | --- |
| **State** | **Meaning and Context** | **Example Support Messages** |
| epidural\_used | If epidural is administered: It might increase the chances of depression | * If epidural is used, then there are more chances of development of postpartum depression |
| birth\_dev | Deviation from actual birth plan: This may elevate the fears regarding the deviation, which may lead to stress | * As there is a deviation in birth protocol, an elevated fear or stress levels is expected * you know if there is a deviation in birth, this can increase the fear levels in the patient and can lead to stress too * Support is desired as there is a deviation in normal birth protocol |
| stress | Stress: This may lead to postpartum depression and father needs support | * In stressful environment, there are more chances to develop postpartum depression. * Father needs support too, as there are high chances of developing postpartum depression. |

##### Framework Logic and Human Coach Interaction

A basic human-coach interaction enables medical professionals to monitor their progress and avoid any possible medical errors. So, AI Coach helps to avoid forgetfulness or ensure mental well-being of the people on the floor. We aimed for a meaningful and limited interaction, so minimal attention and understanding is required from the medical team with a complete functionality.

For communication model, we considered not only the input from a user, but also considered medical protocols and the data related to desired support. Thus the coach analyzes a scenario based on input from the user and then communicates with the user with respect to the different use cases addressed through the shared mental models (e.g., see Table usecases). This communication can benefit the team in different ways, for instance, for protocol monitoring, action impact assessment, future proposals and reminder notification (See Figure clogic).

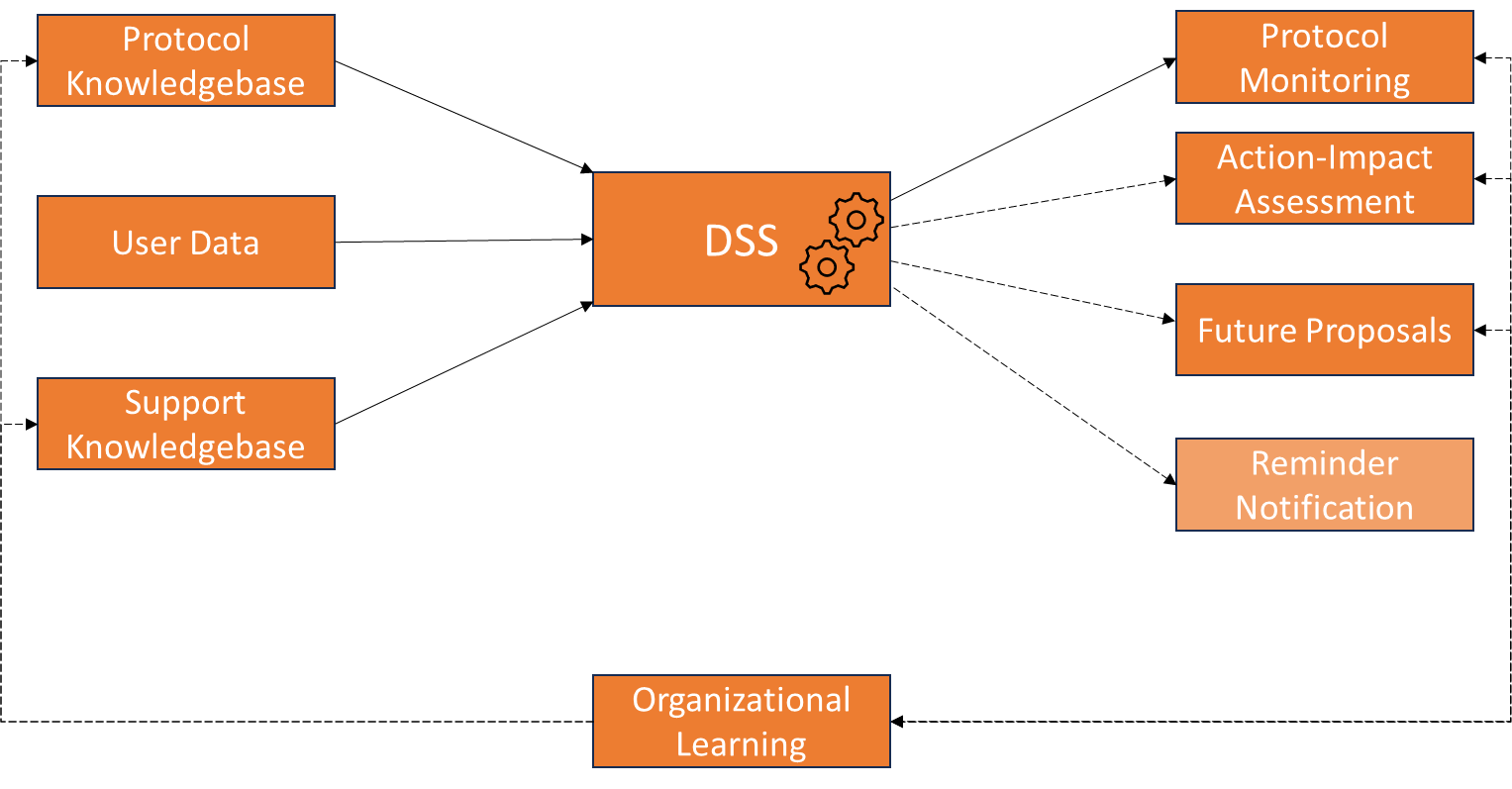


Figure clogic: Logic of AI Coach

All states were not presented at once, rather they were generated as per user input. Therefore, when a state was mark as completed/monitored, the related causal effects were estimated, using the computational logic (Treur, 2020), and the affected states were shown to the user for further monitoring. As an example, consider the example discussed earlier, i.e., development of postpartum depression (Weigl et al., 2022). For protocol monitoring, there are many factors that can cause depression e.g., deviation in normal delivery (represented by state birth\_dev), or stress (stress). To monitor depression (mppd; pppd), human coach interaction will estimates these factors. If these factors are found (Kowalski, 2011), it will communicate and provide the related support through the designed support knowledgebase. For instance, for state birth deviation (birth\_dev), AI Coach will monitor if mother is in labor (baby\_deliv) and if there was riskiness of giving birth (riskiness). If, both of the factors (states) are observed, AI Coach will enable the monitoring of deviation in birth. Therefore, a related monitoring message can be ‘Please make sure that mother is in labor and there is a risk of giving birth’. Similarly, when mental support is monitored (mental\_supp), then AI Coach ensures that the therapist has been referred (refer\_therapist). So, a related support message can be ‘Please make sure that therapist is referred’.

In other words, human coach interaction acts as a decision support system that manages and help the medical team to make decisions an follow protocol about the current situation(s) in neonatology which cannot be easily specified beforehand. These support messages are designed by the expert users. Moreover, we have implemented the functionality of ‘Protocol Monitoring’, however, further research is needed to implement the rest of the functionality of the coach, or extend it as per needs. Open source code is available online[[1]](#footnote-1). The outcome is a JSON file which contains a log as the output. A sample JSON file is shown in Figure jsonoutput, which can also be altered as per further requirements. This log can be used in various ways. For instance, it can be analyzed to assess the impacts of a particular action. Similarly, it can be used to propose a modification in the design of the coach (by editing a model).

[{"iteration":0,"date":"2023-08-28T08:38:12.575Z"},  
{"id":"X17","name":"age","user\_name":"age between 25-29","completed":true,"time\_stamp":""},{"id":"X3","name":"epidural\_used","user\_name":"Epidural used","completed":false ,"time\_stamp":""},{"id":"X1","name":"type\_mom","user\_name":"nulliparous mother","completed":false,"time\_stamp":1693211888317},{" id":"X19","name":"dilation\_mom","user\_name":"dilation of mother","completed":false,"time\_stamp":""},{"id":"X20","name ":"premature\_birth","user\_name":"premature birth","completed":false,"time\_stamp":""}]

Figure jsonoutput: JSON Output for protocol monitoring

# User Interface and Interaction

This section presents the user interfaces for the AI Coach system. Preliminary, there are two type of stakeholders who may use interface of AI Coach system, i.e., the expert users who design the functionality of the coach and end users who will interact with the coach. Each user has to login to the AI Coach system to use the related functionality. The considered constraint was that the interaction was intended mainly for the baby birth. It is to be noted that the related requirements were identified by the invited interviews which involved the stakeholders from the organization, AI experts and experts from psychology.

## Expert User Interface

This interface is meant to model the behaviors of the involved stakeholders and the role of AI Coach in assisting them. Therefore, the main aim is to design and reflect the adaptive behavior of AI Coach which varies as per scenarios.

It is to be noted, that all the models are based on the shared mental models which are subjected to preliminary analysis and simulation by AI experts and researchers in the same domain, for example, in neonatal care, the work from Doornkamp et al. which specifies the communication behavior among the medical team to ensure patient safety (Doornkamp et al., 2022), or by Weigl et al. which identifies the factors that can play role in postpartum depression (Weigl et al., 2022).

When an expert user (modeler) is designing the AI Coach communication model, (s)he has to specify a logical name for the model (See Figure statenames - a). The name should be self-explanatory which can be helpful to search the model. Moreover, the modeler has to specify the number of actors who would be interacting as a team. This information helps in selecting the model for the human coach interaction. Once the users information is there, he can design the model using the states information.

|  |
| --- |
| A screenshot of a chat  Description automatically generated |
| (a) |
| A screenshot of a computer  Description automatically generated |
| (b) |

Figure statenames: a) specifies the name of the conceptual model. b) Specifying the related state names

#### Incoming Connections

Figure statenames – b indicates the incoming states at the base level. For example, here we consider that Y has two incoming connections that is from states Input X1, and Input X2 (See (Treur, 2020), and Section 3.1.3 for further details). User has to specify different details related to a state, i.e., state name, the user name or the friendly user name, the type of the state (internal or mental state/actionable state) and then the respective incoming state. An ‘internal state’ is a state which does not need user monitoring and will not be shown to the end users, but, they are part of predicting certain behavior of a state. While, an ‘actionable / observation state’ is a state that needs input from the user and thus involved in the monitoring process. First column id is auto generated, it is unique and is used for further specification of any state (see below).

In Figure statenames – b, it can also be seen that X1 has one incoming connection i.e., X1, so if a state (like X1) has an incoming connection from the state itself, the aim is to maintain the state value over the period of time (for further details see (Treur, 2020)). A similar pattern is developed for the adaption layers, i.e., first order level and the second order level (See Appendix A). However, the states on these levels can only be the internal states.

#### Connection weight, Speed and Initial Value of a state

Figure specs indicates the connection weights for the incoming states for the complete model specified in Figure statenames. These weights can be either a number ranging between 0 – 1 or a state specified by id. A state having parameter as a state id (e.g., X1, or X2), indicates that the behavior of the a state will vary as per values of the state with that id.

|  |  |
| --- | --- |
|  |  |
| (a) | (b) |

Figure specs: a) Connection weights b) Initial values and Speed factor for the incoming states

#### Impact Computation of a state

Figure combfunc indicates how can a modeler choose the combination function (For theory Section 2.1). By default Euclidean function is selected, however, user can modify the function by choosing it through dropdown list. After choosing the combination function, the modeler has to choose the respective parameters. Modeler can also add or remove another combination function by choosing ‘+’ or ‘-’ respectively. Once the modeler specifies these specifications, (s)he can save them. Now the designed model is ready to be used for a human coach interaction.

|  |
| --- |
|  |
| Figure combfunc: Combination Function Specification of the model states. |

## End User Interface

End users can interact with AI Coach through an interactive interface, with a visual and audio support. AI Coach selects a scenario based interaction, which depends on multiple factors, for example, number of participants, current use case and so on. As a precondition an end user should be aware of how to interact with AI Coach. Figure monitoring shows the interface of AI Coach, responsible for monitoring of neonatal care protocols, where each step is represented by a state (rectangle).

AI Coach uses network models to provide human coach interaction. For example, let’s consider the designed model by Weigel et al. The model reflect different factors, which may lead to postpartum depression (e.g., if mother is nulliparous, epidural is used). If a mother develops postpartum depression, there are higher chances for the other parent (or family member) to develop postpartum depression. Thus AI Coach needs to follow, monitor the development of postpartum depression. The main aim of this interaction is to assist the stakeholders to follow certain steps which might be skipped leading to a medical error. For example, the medical team has to refer a therapist (ref\_therapist) or parent needs to ask for support (ask\_sup) if there is stress (stress). In other words, these models provides a platform to AI Coach interact with the medical team to assist them, leaving no chance of possible error (for modeling details please see (Weigl et al., 2022)).

AI coach is a human based coach, and it interacts as per modeled behaviors of the individuals in the team. These behaviors are dynamic in nature and therefore doesn’t have predefined steps except the initial factors in the main protocol. Figure monitoring illustrates how AI Coach provides interaction based on the causal model designed by Weigel et al. The interaction starts by showing some initial states of the main protocol. Figure monitoring a indicates the initial factors (shown by grey rectangles – e.g., nulliparous mother, epidural used) that might be observed by the medical team. If these states are observed during the process, these states need to be marked as ‘monitored’ or ‘observed’ (marked as green). Marking these states will enable the end users to keep the track. Moreover, when a user clicks any state state, the coach shows further states known as ‘predictable states’ (grey rectangles – e.g., risk and mppd), which need to be monitored to follow the protocol further. If medical team somehow misses these steps, this may cause an error in the protocol. AI Coach plays its role to avoid missing this step by ensuring its monitoring.

|  |  |
| --- | --- |
|  |  |
| Figure monitoring: a) initial states of model by Weigl et al (Weigl et al., 2022). b) Observed states are shown as green, while initial/predicted states are shown as grey states | |

To maintain the rapport and to help avoid forgetfulness, the users are reminded if they are missing a task. This reminder is a highlighted state (in yellow) which allows user to ensure if they have performed the task or certain state is observed. An information slider and audio support is provided to increase the visibility of the tasks to be performed and allowing users not to miss a single task. Figure forgetfulness shows how a state is highlighted, and a related status is shown to the user(s), if they click a state. To explain it further, state MPPD can’t be observed or monitored if there the related things are not observed or performed (i.e., epidural used, cesarean, baby is born and mental support). In other words, the human coach interaction lies on how coach understands (with humanoid-reasoning) to estimate the current situation (Tauber & Ackermann, 2013). This intervention is also shown in terms of a status, which is played back by AI Coach (Figure forgetfulness – b), so that the task is not missed in the case of lesser eye contact with the coach. The user is also facilitated to terminate the process by an end button. A log is maintained to trace the actions taken over time.

|  |  |
| --- | --- |
|  |  |
| Figure forgetfulness: If mppd is observed, then a) a visual interaction for the related unobserved states are highlighted. b) status displayed (along with audio support) to provide a better understanding of the missing tasks. | |

The interface of the coach is kept simple, and is designed to handle time constraint situations and to perform their role with the minimum involvement of AI Coach. An audio support is also added, so that if a user is not able to track a predictable state, the coach assists the user to ensure that the particular state is also monitored. For example, if the user wants to make risk as ‘observed’, AI coach will ensure that riskiness is observed only if age of the mother is between 25 – 29. Therefore, an audio is played (e.g., Mark age as monitored) to ensure is risk is there. We also added a preliminary tone support for audio. To do this, we performed sentiment analysis of the text (e.g., state name, or related message) by huggingface library[[2]](#footnote-2). For instance, if a negative state (e.g., stress) is observed by user, the audio tone of coach is high (pitch = 0.9) and rate = 0.7), while a positive tone is bit low (pitch = 0.8) and rate = 0.85). The implementation details of the coach can be found in appendices. When the process is completed or is terminated by the user, a log is generated for analysis (for example, see in Figure jsonoutput). This log can be used for further analysis.

# System Evaluation and Discussion

Preliminary discussions helped us to clarify the system boundaries and the basic objectives of AI Coach (See Section 3). We considered them as the basis for system evaluation. Therefore, system evaluation was based on, e.g., a) how well AI Coach interaction can be designed by the system, b) how far it can meet the consumers’ needs, and c) what is the usability of the system and AI Coach in terms of its assisting the medical team. Table feedback mentions the feedback and related score of the questions that were asked from 5 participants (applied psychologist = 1; software developers = 2; AI researchers = 2) to evaluate the system.

It is to be noted that, the knowledgebase (models used by AI Coach) is to be designed, simulated and verified mathematically before using AI Coach. However, to give an essence of how it can be created, we explained some test models that can be designed and used dynamically by the AI Coach system. For this we gave them a brief walk-through tutorial which included a) the goals of the AI Coach system, b) the knowledgebase creation of AI Coach, c) interaction with the AI-Coach, d) feedback questions, and e) an informal session about the coach. This session not only helped us to study the look and feel of the users (expert + end users) for human coach interaction, but it also helped the participant to understand how AI Coach will interact with the end users. Following we explain the results of using the AI Coach system.

## Test Model

A sample model was designed to see how far it is easy to design the model using the AI Coach system. This model is inspired from newborn life support (Madar, Roehr, Ainsworth, Ersdal, et al., 2021), and uses adaptive network modelling technique (Treur, 2020). The model designed in this phase is used for protocol monitoring and the results mentioned in Section 4.2 are compiled accordingly.

Therefore, the participants of the study had to play the role of a) an expert users and b) an end users. This is because, our focus is investigate how easy is to use the system to design the coach and using it. Moreover, following this will develop a deeper understanding of the system and algorithm for protocol monitoring. The example model is mentioned in Figure samplemodel, where each state specifies actions to be monitored by AI Coach. These actions are similar to the actions, which are to be performed by medical staff. Here state XXX

Figure samplemodel: Sample model inspired from life support neonatal care (Madar, Roehr, Ainsworth, Ersdal, et al., 2021)

## Results and Discussion

All the participants were explained the purpose to design a model, and all the participants were successful to design the model (in average time = 25 mins). Participants found the framework quite easy and useful to develop a human coach interaction, which will serve in a dynamic manner. All users commented that the interactive visualization could help and save time to develop such coaches, which are not only capable of providing human computer interaction to reflect human behaviors, but also can be customized and used for different purposes.

A participant also mentioned that he used to develop rule based chatbots, which rely on structures like decision trees, however, this system can enable users to develop a human-bot interaction to address graph based structures. Moreover, writing sample messages were also found to provide human-bot interaction. Overall, positive user experience was noticed for the coach design.

As a next step, participants were briefed, that the models they designed so far can also be used for monitoring a neonate. For simplicity, a pre-designed model was considered for monitoring through human coach interaction.

Table feedback: Participants feedback on usefulness and usability of AI Coach (5 point Likert Scale)

|  |  |  |
| --- | --- | --- |
| **Statement** | **Mean** | **S.D** |
| The system is easy to use | 3.8 | 0.81 |
| I find the system design is helpful to develop the human coach interaction | 2.9 | 0.58 |
| The system helped me to understand the flow of AI Coach which will interact with the people | 3.8 | 0.5 |
| AI Coach can help to achieve the goal, i.e., reduction of medical errors | 3.9 | 0.68 |
| The interaction of AI Coach was useful | 3.1 | 0.48 |
| The reminder prompts dynamically generated are useful. | 4.0 | 0.88 |
| Audio utterances were helpful to keep the flow of AI Coach | 2.9 | 0.66 |
| I enjoyed real-time interaction with AI Coach | 3.2 | 0.40 |
| I am satisfied with my experience using the system and its coach | 3 | 0.7 |
| I would use the system to design my coach. | 3.9 | 0.73 |

Table feedback mentions the results of the experience of participants about the human coach interaction. Overall, all the participants were quite impressed how AI Coach system can be used to generate a human coach interaction. One of the participants liked the possibility of using such complex phenomenon in practical fields like healthcare. Also, a participant mentioned ‘it was a nice experience, coach especially when seeing how the model came to an end to ensure a good operation’.

However, a participant mentioned that it can indeed provide a real time interaction in neonatal care, but how the users will be trained to use this type of agent in stressful situations. Similarly, another concern was that the steps were not shown earlier, so how could a user know what to click. But, this was an intentional choice, so that the users don’t click a step mistakenly or unintentionally, but to ensure that (s)he has performed/observed that particular task. A participant mentioned that user engagement has been always very challenging, and users might hesitate to use it, so it needs to be more compelling and valuable enough to encourage regular usage. Moreover, another participant mentioned that how can an end-user interact with the coach in such a limited time, and how voice interaction is possible under stress.

# Limitations and Future Work

In this section, we would like to mention the limitations of the study. Initially, we will explain the limitations in the design of AI Coach, then we will explain the limitation in analysis and results of the study.

While discussing the limitations of the design of the framework, we would like to mention that the basic goals of the framework were to provide a real-time interaction among a team, with the aim to assist healthcare professionals if they forget a step in a medical protocol. The user needs and goals were taken as a primary objective, however, some requirements still need to be incorporated. Few examples are, rapport creation (e.g., a message during a conversation ‘as a coach, I can help you, do you want me to remind you’). Currently tone regarding a state is incorporated and the coach can change its tone if a state like stress in monitored, however, adopting an attachment style to make humane still needs to be considered (Roelofsma, 2013). Moreover, the current design of AI Coach totally rely on how the coach is modeled to interact with the end-users. This means that the knowledge has to be pre-programmed by the domain experts, however, introducing adaptive behavior and learning can be helpful in decision making at the runtime.

The framework has ability to integrate different features in future. For example, the domain knowledge should be increased by incorporating more mental models. This will help to improve domain knowledge and thus AI Coach can interact in a better way. The proposed framework enables handling of complex tasks based on human behaviors, however, it lacks the human emotions and personalization. Adding an avatar with emotions can be helpful to overcome this (Zmugg et al., 2015).

Unlike many other chatbots, AI Coach takes user input through button click, however, framework also allows to integrate sensors, which can improve the human experience by better understanding of environmental factors to predict an action. For example, one example can be integrating sensors for taking baby vitals’ data, to predict if the baby is breathing normally, and predicting the related outcomes which can be helpful in assisting the healthcare professionals working on the floor. Another example, can be predicting stress from sensors. As a result, stress can be reduced by music, so AI Coach interprets the stress levels through the environmental sensors, and may propose users to play a comforting music (Trappe, 2012). Similarly, adding few question answer sessions via sensors can improve the human-level understanding and intuition, which can improve human coach interaction.

Turing test

Evaluation and Validation in real-time still needs to be done <https://journal.laurea.fi/living-labs-from-abstraction-to-implementation-case-circ4life-living-labs/#800cd72b>

Still needs to be done A graph of blue and orange bars

Description automatically generated

# Conclusion

# Acknowledgements

Doctors

We will use the **Connected Living Lab approach** (WP1,3,5), where five European Hospitals (UK, NL, SL, IL, PL) will connect in Living Labs specialised in user-centred care and will co-design, test and validate the SMILE4SAFETY AI coach platform.

# References

Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M. A. A., & Dwivedi, Y. K. (2023). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge*, *8*(1), 100333. https://doi.org/10.1016/j.jik.2023.100333

Almirall, E., Lee, M., & Wareham, J. (2012). Mapping Living Labs in the Landscape of Innovation Methodologies. *Technology Innovation Management Review*, *2*(9), 12–18. https://doi.org/10.22215/timreview/603

Bardelli, S., Del Corso, G., Ciantelli, M., Del Pistoia, M., Lorenzoni, F., Fossati, N., Scaramuzzo, R. T., & Cuttano, A. (2022). Improving Pediatric/Neonatology Residents’ Newborn Resuscitation Skills With a Digital Serious Game: DIANA. *Frontiers in Pediatrics*, *10*, 842302. https://doi.org/10.3389/fped.2022.842302

Björkstén, K. S., Bergqvist, M., Andersén-Karlsson, E., Benson, L., & Ulfvarson, J. (2016). Medication errors as malpractice-a qualitative content analysis of 585 medication errors by nurses in Sweden. *BMC Health Services Research*, *16*(1), 431. https://doi.org/10.1186/s12913-016-1695-9

Clark, I. A., & Maguire, E. A. (2020). Do questionnaires reflect their purported cognitive functions? *Cognition*, *195*, 104114. https://doi.org/10.1016/j.cognition.2019.104114

Doornkamp, S., Jabeen, F., Treur, J., Taal, H. R., & Roelofsma, P. H. M. P. (2022, September 13). A Controlled Adaptive Network Model of a Virtual Coach Supporting Speaking Up by Healthcare Professionals to Optimise Patient Safety. *Proc. of the 13th Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence, BICA\*AI’22*.

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Giacomin, J. (2014). What Is Human Centred Design? *The Design Journal*, *17*(4), 606–623. https://doi.org/10.2752/175630614X14056185480186

Hammersley, M., & Atkinson, P. (2007). Ethnography: Principles in Practice, 3rd edn Routledge. *London.[Google Scholar]*.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2008). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 6.

Hu, L., Zhang, L., Yin, R., Li, Z., Shen, J., Tan, H., Wu, J., & Zhou, W. (2021). NEOGAMES: A Serious Computer Game That Improves Long-Term Knowledge Retention of Neonatal Resuscitation in Undergraduate Medical Students. *Frontiers in Pediatrics*, *9*, 645776. https://doi.org/10.3389/fped.2021.645776

Kowalski, R. (2011). *Computational Logic and Human Thinking: How to Be Artificially Intelligent* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511984747

Leminen, S., Westerlund, M., & Nyström, A.-G. (2012). Living Labs as Open-Innovation Networks. *Technology Innovation Management Review*, *2*(9), 6–11. https://doi.org/10.22215/timreview/602

Madar, J., Roehr, C. C., Ainsworth, S., Ersda, H., Morley, C., Ruediger, M., Skåre, C., Szczapa, T., Te Pas, A., Trevisanuto, D., & others. (2021). Newborn resuscitation and support of transition of infants at birth. *Notfall & Rettungsmedizin*, *24*(4), 603–649.

Madar, J., Roehr, C. C., Ainsworth, S., Ersdal, H., Morley, C., Rüdiger, M., Skåre, C., Szczapa, T., Te Pas, A., Trevisanuto, D., Urlesberger, B., Wilkinson, D., & Wyllie, J. P. (2021). European Resuscitation Council Guidelines 2021: Newborn resuscitation and support of transition of infants at birth. *Resuscitation*, *161*, 291–326. https://doi.org/10.1016/j.resuscitation.2021.02.014

Manifesto, A. (2001). *Manifesto for agile software development*.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Panagos, P. G., & Pearlman, S. A. (2017). Creating a Highly Reliable Neonatal Intensive Care Unit Through Safer Systems of Care. *Clinics in Perinatology*, *44*(3), 645–662. https://doi.org/10.1016/j.clp.2017.05.006

Petrovic, M., Nicholls, J., & Siassakos, D. (2022). Proceed with reasonable care: When legal principles inform training to prevent harm during childbirth. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *80*, 105–113. https://doi.org/10.1016/j.bpobgyn.2021.12.006

Profit, J., Sharek, P., Kan, P., Rigdon, J., Desai, M., Nisbet, C., Tawfik, D., Thomas, E., Lee, H., & Sexton, J. (2017). Teamwork in the NICU Setting and Its Association with Health Care–Associated Infections in Very Low-Birth-Weight Infants. *American Journal of Perinatology*, *34*(10), 1032–1040. https://doi.org/10.1055/s-0037-1601563

Redlich, B., Siemon, D., Lattemann, C., & Robra-Bissantz, S. (2017). Shared Mental Models in Creative Virtual Teamwork. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. https://doi.org/10.24251/hicss.2017.057

Roelofsma, P. H. M. P. (2013, November). The Ambient Intelligent Environment as an Attachment Figure: How to Create Sensitive Coaching Systems for Human Care. *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. https://doi.org/10.1109/wi-iat.2013.180

Rosson, M. B., & Carroll, J. M. (2002). *Usability engineering: Scenario-based development of human-computer interaction*. Morgan Kaufmann.

Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Stavroudis, T. A., Miller, M. R., & Lehmann, C. U. (2008). Medication Errors in Neonates. *Clinics in Perinatology*, *35*(1), 141–161. https://doi.org/10.1016/j.clp.2007.11.010

Stepke, F. L. (2012). WORLD HEALTH ORGANIZATION Standards and operational guidance for ethics review of health-related research with human participants. *Acta Bioethica*, *18*(1), 129–132. https://doi.org/10.4067/s1726-569x2012000100014

Tauber, M. G., & Ackermann, D. (2013). *Mental models and human-computer interaction*. Elsevier.

Trappe, H.-J. (2012). Role of music in intensive care medicine. *International Journal of Critical Illness and Injury Science*, *2*(1), 27. https://doi.org/10.4103/2229-5151.94893

Treur, J. (2020). *Network-Oriented Modeling for Adaptive Networks: Designing Higher-Order Adaptive Biological, Mental and Social Network Models* (p. XVII, 412). Springer Nature Publishing. https://doi.org/10.1007/978-3-030-31445-3

Tsiouris, K. M., Tsakanikas, V. D., Gatsios, D., & Fotiadis, D. I. (2020). A Review of Virtual Coaching Systems in Healthcare: Closing the Loop With Real-Time Feedback. *Frontiers in Digital Health*, *2*, 567502. https://doi.org/10.3389/fdgth.2020.567502

van Ments, L., Treur, J., Klein, J., & Roelofsma, P. (2021). A second-order adaptive network model for shared mental models in hospital teamwork. *Computational Collective Intelligence: 13th International Conference, ICCCI 2021, Rhodes, Greece, September 29–October 1, 2021, Proceedings 13*, 126–140.

Weigl, L. M., Jabeen, F., Treur, J., Taal, H. R., & Roelofsma, P. H. M. P. (2022). Modeling Learning for a Better Safety Culture within an Organisation Using a Virtual AI Coach: Reducing the Risk of Postpartum Depression by More Communication with Parents. *Cognitive Systems Research*, *80*, 1–36.

Weizenbaum, J. (1983). ELIZA — a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *26*(1), 23–28. https://doi.org/10.1145/357980.357991

Xu, Y., Jabeen, F., Treur, J., Taal, H. R., & Roelofsma, P. H. M. P. (2022, November 7). Adaptive Agent Network Models with Internal Mental Models Supporting Patient Safety. *Proc. of the 15th International Conference on Social Computing and Networking, SocialCom’22*.

Zmugg, R., Braun, A., Roelofsma, P., Thaller, W., Moeskops, L., Havemann, S., Reljic, G., & Fellner, D. W. (2015). Personalization of Virtual Coaching Applications using Procedural Modeling. *Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and E-Health*. https://doi.org/10.5220/0005435600370044

# Appendices

1. **Interface of AI Coach**

**Creating Model**

While considering work from Weigel et al., this interface shows how can a model be specified

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Moreover, the modeler has to specify the number of actors who would be interacting as a team. This in

1. **Implementing AI Coach**

1. https://github.com/MsFakhra/AI\_in\_Healthcare [↑](#footnote-ref-1)
2. <https://huggingface.co/transformers/usage.html> [↑](#footnote-ref-2)