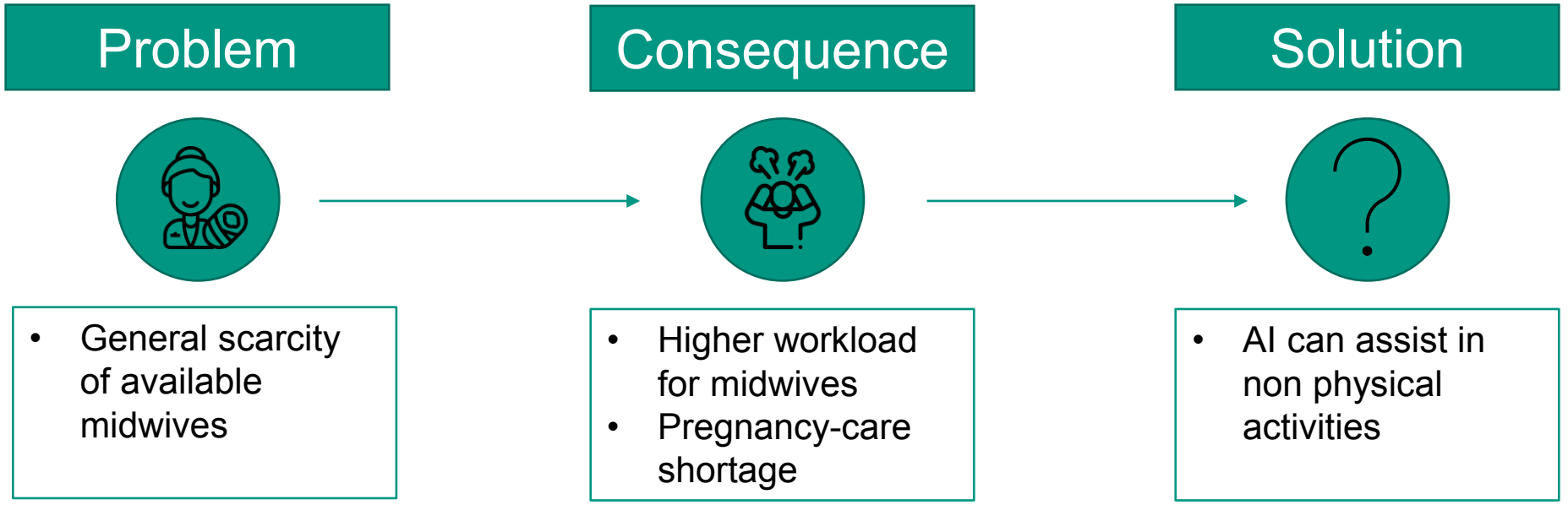
Assessment for Adoption Behavior on the Usage of Chatbots in Midwifery Consultation

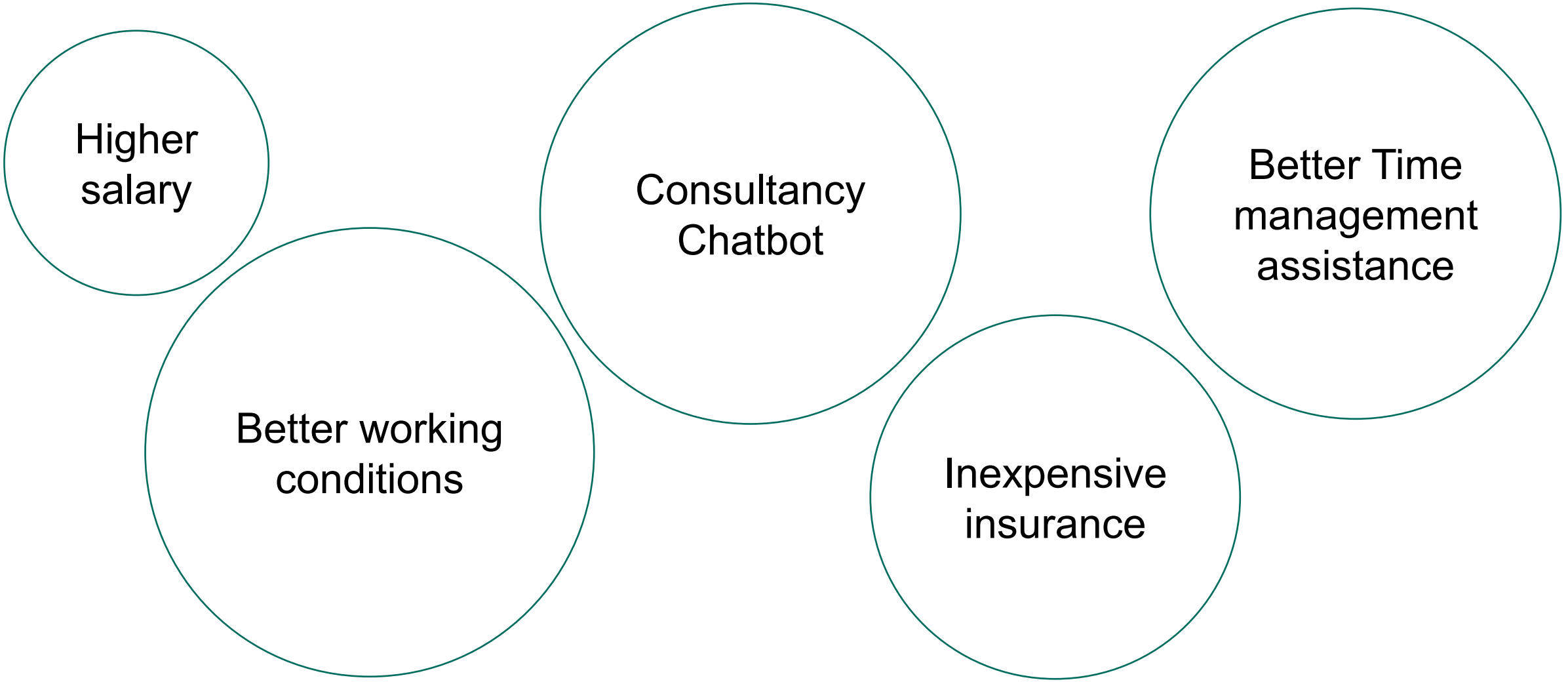
Shaping AI for Society

Kieu, Nils, Magnus

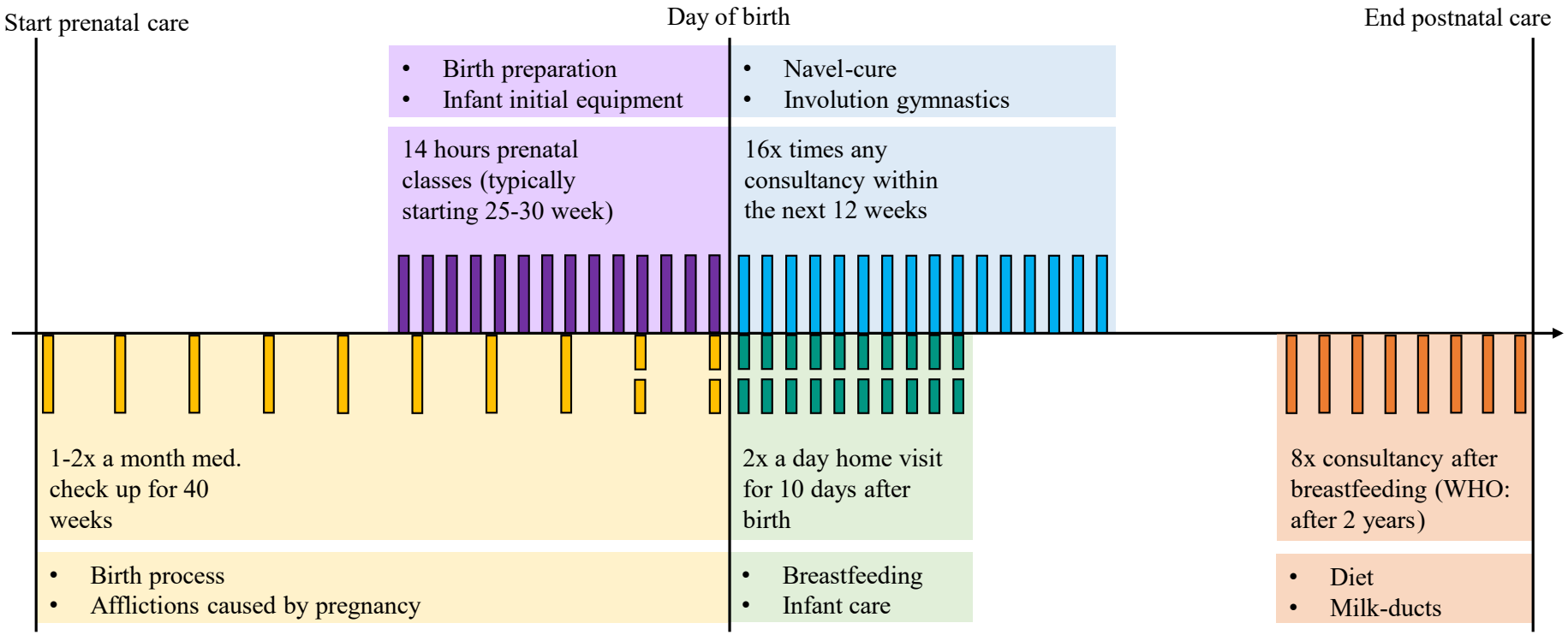
Background Information



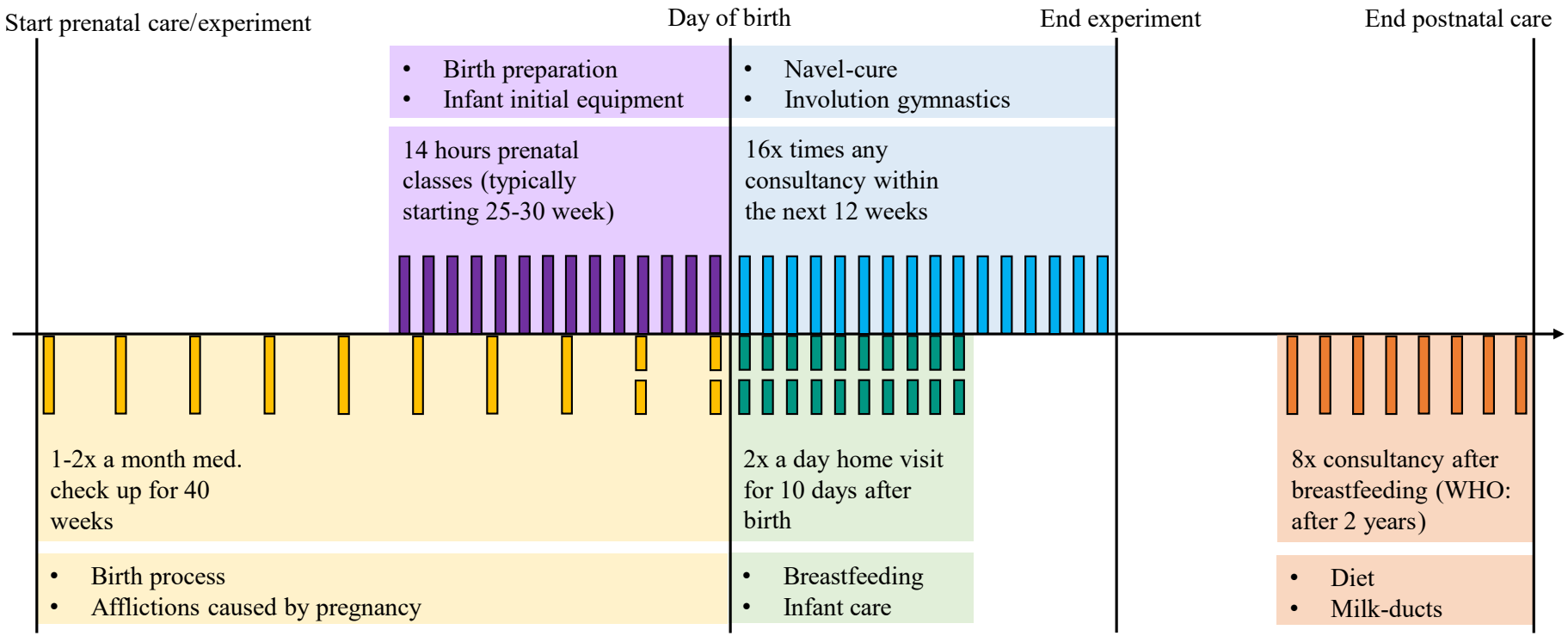
Possible Solutions/Assistances



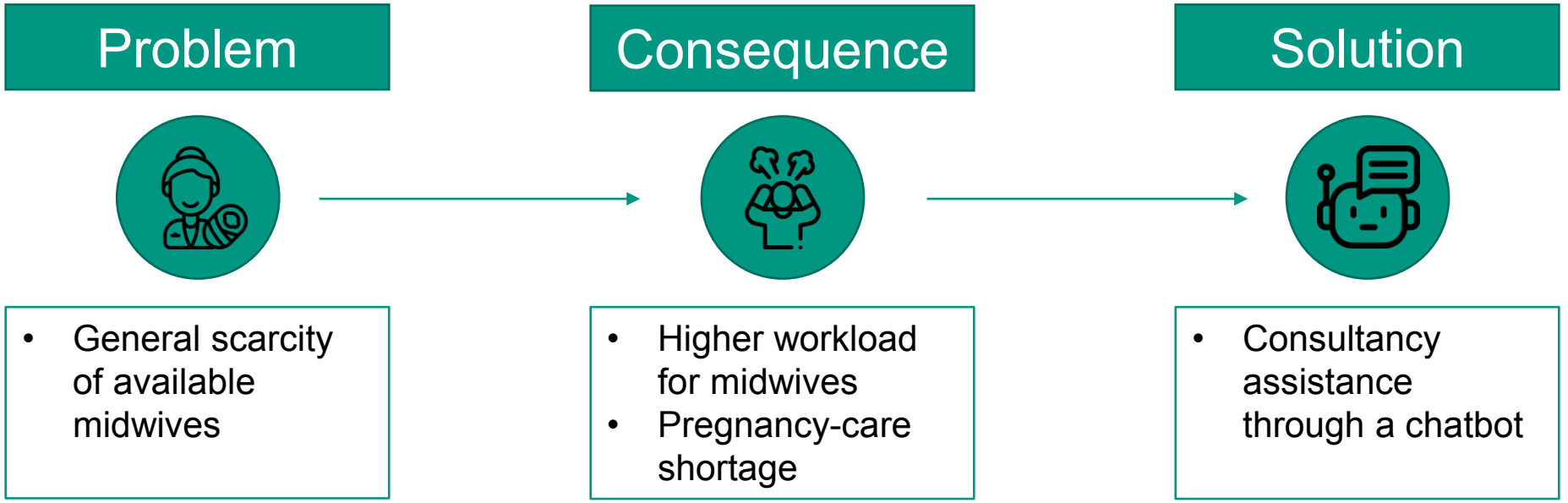
Health insurance benefits for women with statutory health insurance



Health insurance benefits for women with statutory health insurance



Background Information

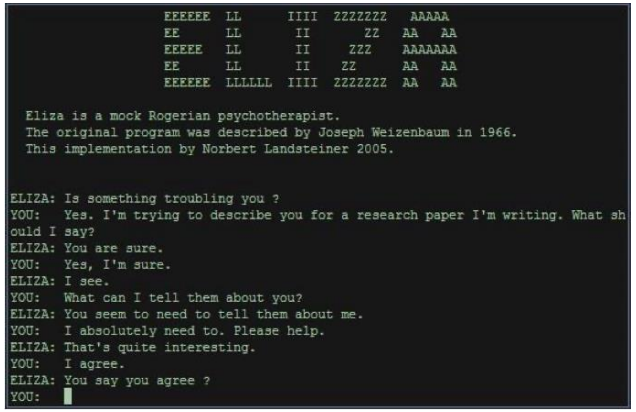


State of the Art Chatbots in Medical Care

PARADISE	Technical Performance in Health Care	General Performance using AI
Ease of usage Clarity Natural Friendliness Robustness regarding misunderstanding Willingness to use the system	1. Usability 2. Classifier Performance: AUC, accuracy, Recall, Precision 3. Assessment of intelligence	Efficiency Effectiveness Satisfaction

Why not use AI performance metrics exclusively?

- These Metrics do not represent the humanness of a Chatbot
  - Especially in medicine and health care the Usability seems to be more important



Current Research on Chatbots for Pregnancy care and Midwifery

Mental Health Care	Usability-test for Perinatal Women and Partners	1
	Perinatal health care Chatbot to assess mental state	2
Pregnancy education	Experiment on the usage of a pre-visit education chatbot on genetic testing	3
Breastfeeding Consultancy	Interviews with mothers on the adoption rating	4

Experimental Design examples in the Medical and Pregnancy field

Mental Health Care

Qualitative Usability test

13 women; 2 men

Giving instant feedback to the Chatbots answers

Asking at least 3 questions per day

Final questionnaire showed that the users could be encouraged to use medical chatbots

[1]

Breastfeeding Consultancy

Quantitative observational Design

125 Participants

Telephonic Interview

Adoption rating on the usage of a Chatbot for Breastfeeding Consultancy

[2]

Respondent City	Positive	Neutral	Negative
Amritsar	34	27	2
Jaipur	61	2	0

Research Question:

Is the provision of a midwifery chatbot adopted by pregnancy stakeholders*?

* mothers, parents, parental authority

Hypothesis

H1.1: It is assumed that chatbots in midwifery consultation will be adopted by pregnancy stakeholders (without any incentives).

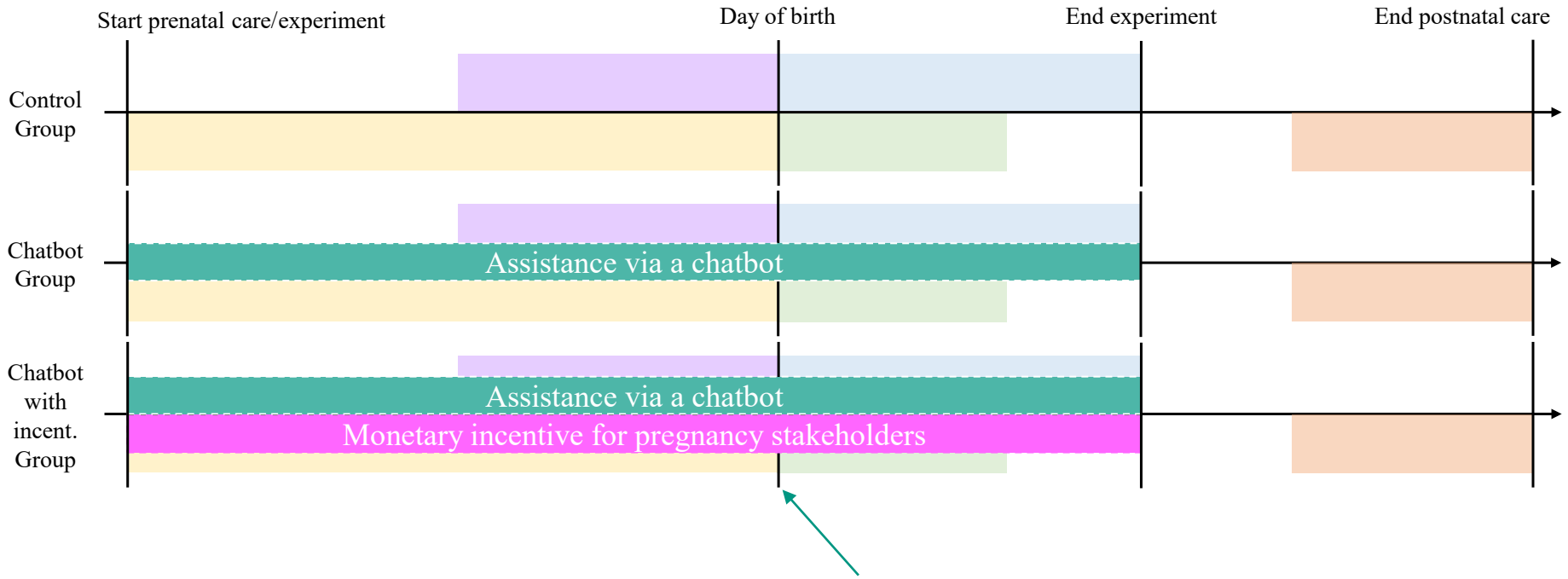
H1.2: It is assumed that chatbots in midwifery consultation will be fully adopted by pregnancy stakeholders (without any incentives).

H2: It is assumed that monetary incentives will motivate pregnancy stakeholders to use chatbots in midwifery consultation.

The experimental design

General specifications	
Type of experiment:	Field experiment
Location of experiment:	One midwifery unit in Germany
Relevant parties:	Pregnancy stakeholders, midwives, healthcare insurer, software developers
Period of experiment	
Start:	Prenatal Care
End:	12 weeks after birth
Total duration:	Max. 52 weeks
Chatbot application	
Language:	German
Availability:	24/7
Initial registration by the pregnancy stakeholders to the application is required	
Measured variables	
Variable 1:	Time spent at pregnancy stakeholder visit
Variable 2:	Frequency of addressed topics by the pregnancy stakeholders towards the midwife
Variable 3:	Frequency of addressed topics by the pregnancy stakeholders towards the chatbot

The experimental design – three groups



Day of birth of all participants have to be within a 8 weeks timeframe

The experimental design – three groups

- Total number of participants (pregnancy stakeholders): ~30

	Control Group	Chatbot Group	Chatbot with incentive Group
Number of participants	~10	~10	~10
Physical assistance by midwife	Yes	Yes	Yes
Access to chatbot application	No	Yes	Yes
Additional monetary incentive (5 € per week)	No	No	Yes

The experimental design – Procedure

Preparation

- Explanation of rules and purpose of experiment
- Guidance for use of chatbot
- Instruction for midwives

Execution of experiment

- One midwife per experiment
- Groups are assigned randomly
- Participants of chatbot group and chatbot group with incentive have the option to choose between midwife or chatbot
- Midwives responsible for documentation

Analysis and evaluation

- Statistical analysis of time spent at visits and frequency of addressed topics

The experimental design – Analysis example

NOT real NUMBERS

Example (mean)	Control Group (#1)	Chatbot Group (#2)	Chatbot w. incentive Group (#3)
Variable 1 (in hours)	50	40	45
Variable 2 (topics adressed)	15	10	17
Variable 3 (topics adressed)	0	4	8
Variable 2 (norm) = (Variable 2/Variable 1)	0.3	0.25	0.378
Variable 3 (norm) = (Variable 3/Variable 1)	0	0.1	0.178

$$H1.1\ accepted: Variable\ 3_{norm}(\#2) > 0$$
$$H1.2\ accepted: Variable\ 3_{norm}(\#2) \geq Variable\ 2_{norm}(\#1)$$

$$H2\ accepted: \quad Variable\ 3_{norm}(\#3) > Variable\ 3_{norm}(\#2)$$

Expected results

H1.1: It is assumed that chatbots in midwifery care will be adopted by pregnancy stakeholders (without any incentives). ✓

H1.2: It is assumed that chatbots in midwifery care will be fully adopted by pregnancy stakeholders (without any incentives). ✗

H2: It is assumed that monetary incentives will motivate pregnancy stakeholders to use chatbots in midwifery care. ✓

Limitations

Different personalities of midwifes may have an effect on experiment

Chatbots in healthcare are not fully developed yet

Less training data for German language

Data can differ from since participants know that they participate in an experiment

Digital competency of participants