

COMP 3647

Human-AI Interaction Design

Topic 9

***Case Studies of LLMA:
Customer Service Chatbots***

Prof. Effie L-C Law

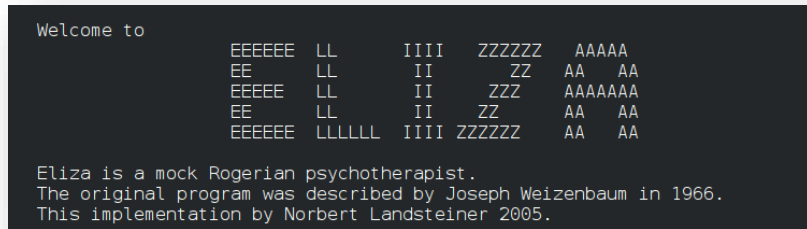
Overview

- Basics of Conversational Agents (CA)/Chatbots
- Chatbots for Customer Services
- Factors for Trust in Chatbots
- Impact of Online Banking on Older Adults
- Implications for the Design of Chatbots for Older Adults

Conversation is the Interface

As technology gets smarter, and more anticipatory
of our wants and intent, the interface gets smaller
and smaller until it disappears.

A Very Brief History of Chatbots



Keyword matching
Minimal context identification

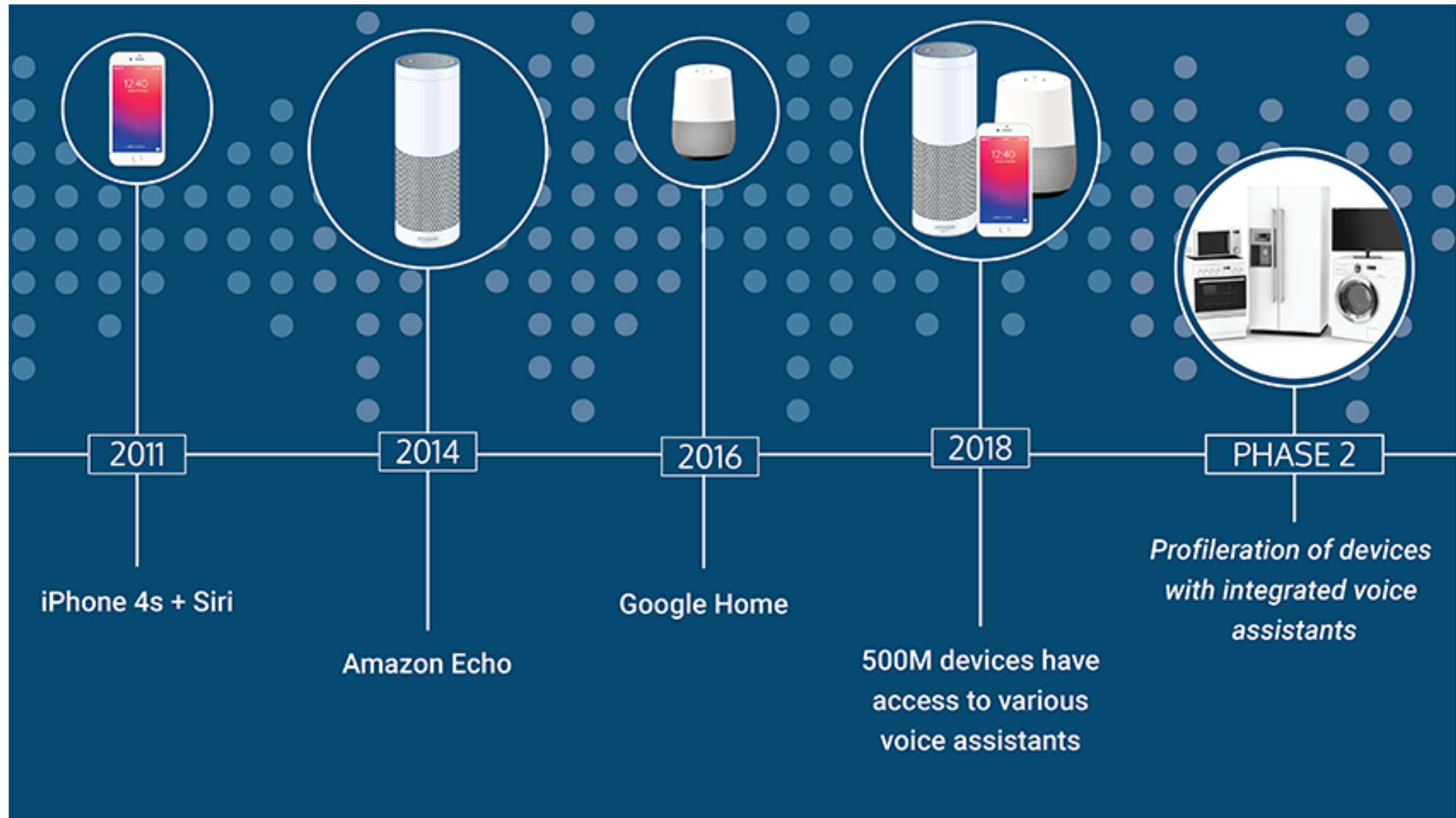


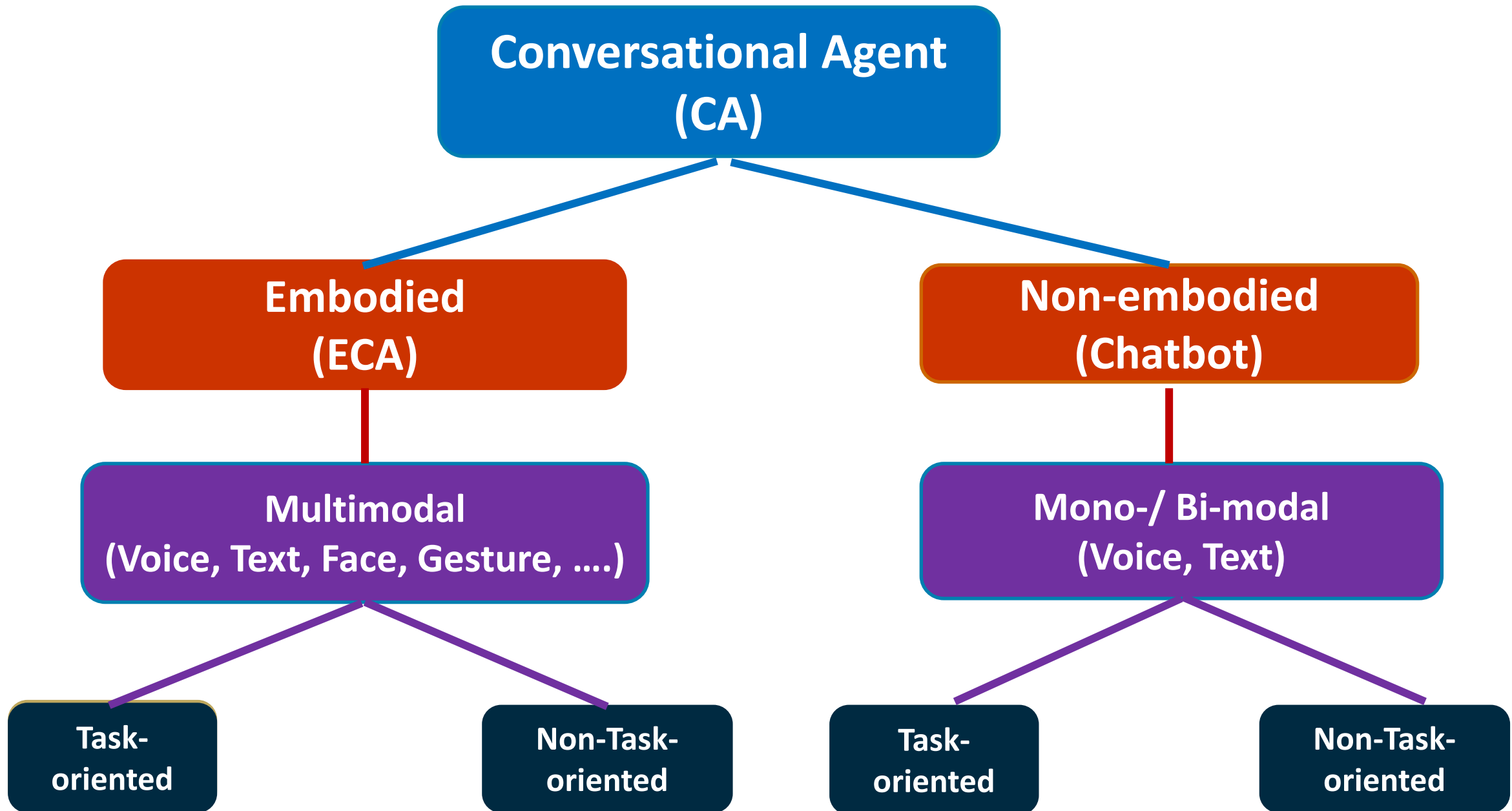
Conversational capability
Artificial Intelligence Markup Language (AIML)



Machine Learning-powered
Deep Neural Networks

Conversational/Voice User Interface (VUI)





Chatbots/ Conversational Agents

```
graph TD; Root[Chatbots/ Conversational Agents] --> IM[Interaction Modality]; Root --> KD[Knowledge Domain]; Root --> UG[User Goal]; Root --> DMA[Dialogue Management Approach]; IM --> TB[Text-based]; IM --> VB[Voice-based]; KD --> OD[Open domain]; KD --> CD[Closed domain]; UG --> TO[Task-oriented]; UG --> NTO[Non-task Oriented]; DMA --> RB[Rule-based]; DMA --> RetB[Retrieval-based]; DMA --> GB[Generative-based];
```

**Interaction
Modality**

Text-based

Voice-based

**Knowledge
Domain**

**Open
domain**

**Closed
domain**

**User
Goal**

**Task-
oriented**

**Non-task
Oriented**

**Dialogue
Management
Approach**

Rule-based

**Retrieval-
based**

**Generative-
based**

Types of CAs



(a)



(b)



(c)



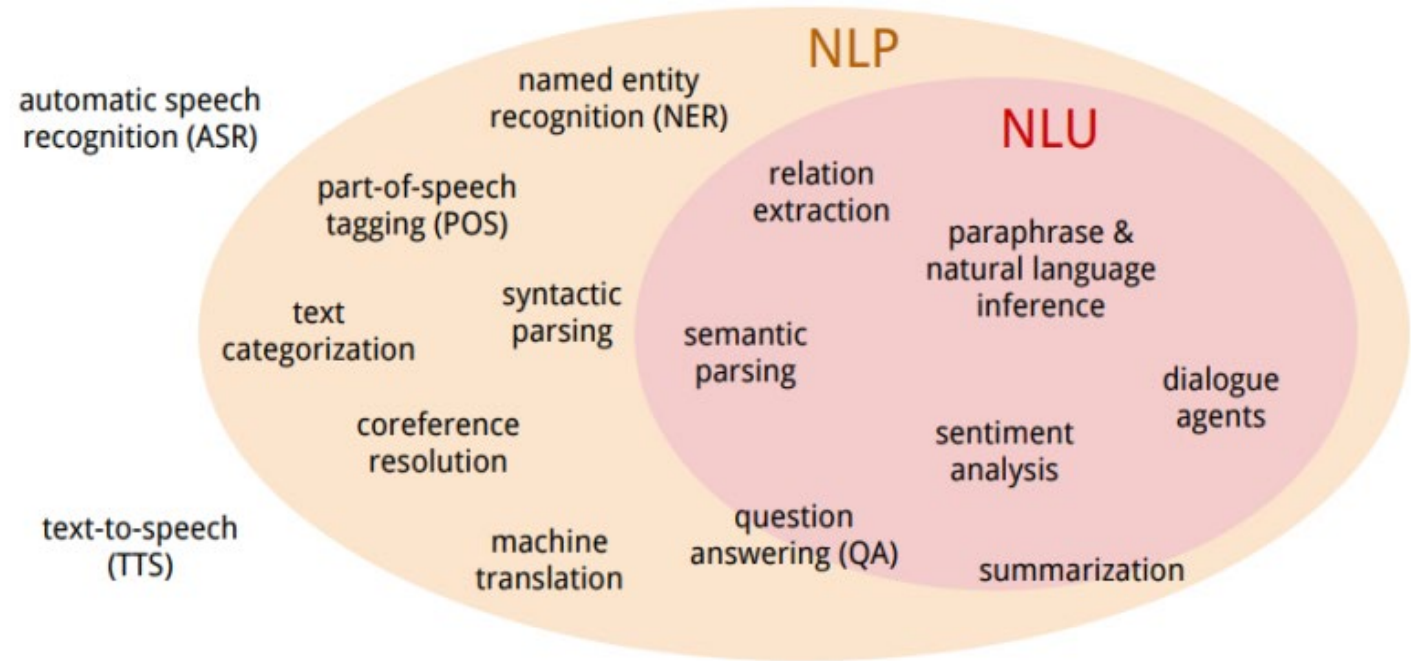
(d)

(a) Non-embodied text-based CA; (b) Non-embodied voice-based CA;
(c) Embodied CA, multimodal social robot; (d) Embodied CA, multimodal mobile CA

Natural Language (NLP/U/G)

Processing/Understanding/Generation

- **NLP**: unstructured data → a structured data format for understanding and response formulation
- **NLU**: machine reading comprehension to identify the intended meaning of a verbal expression
- **NLG**: construction of a verbal expression by a machine based on a given dataset



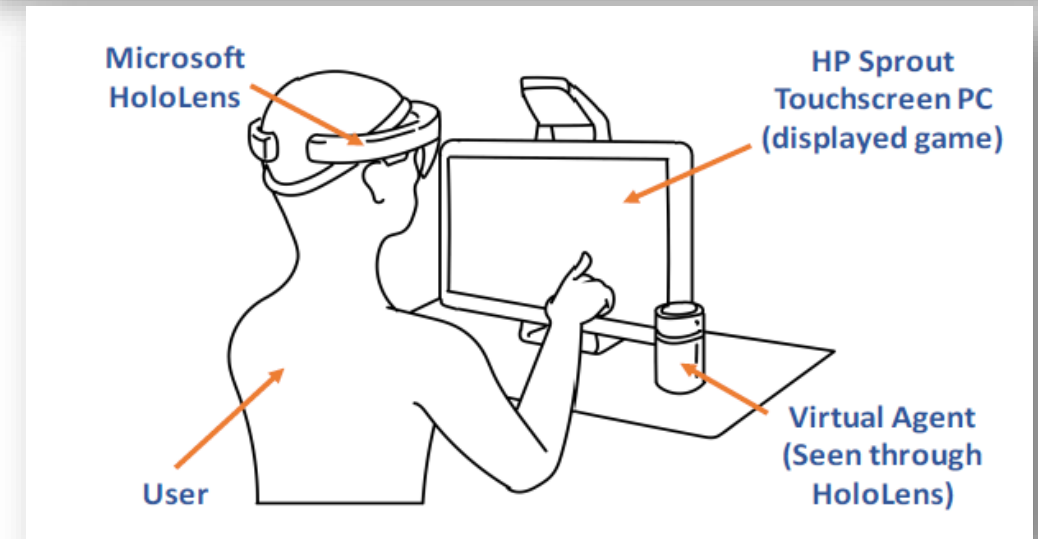
Embodied CA: Human-Robot Interaction (HRI)

➤ User (dis)engagement

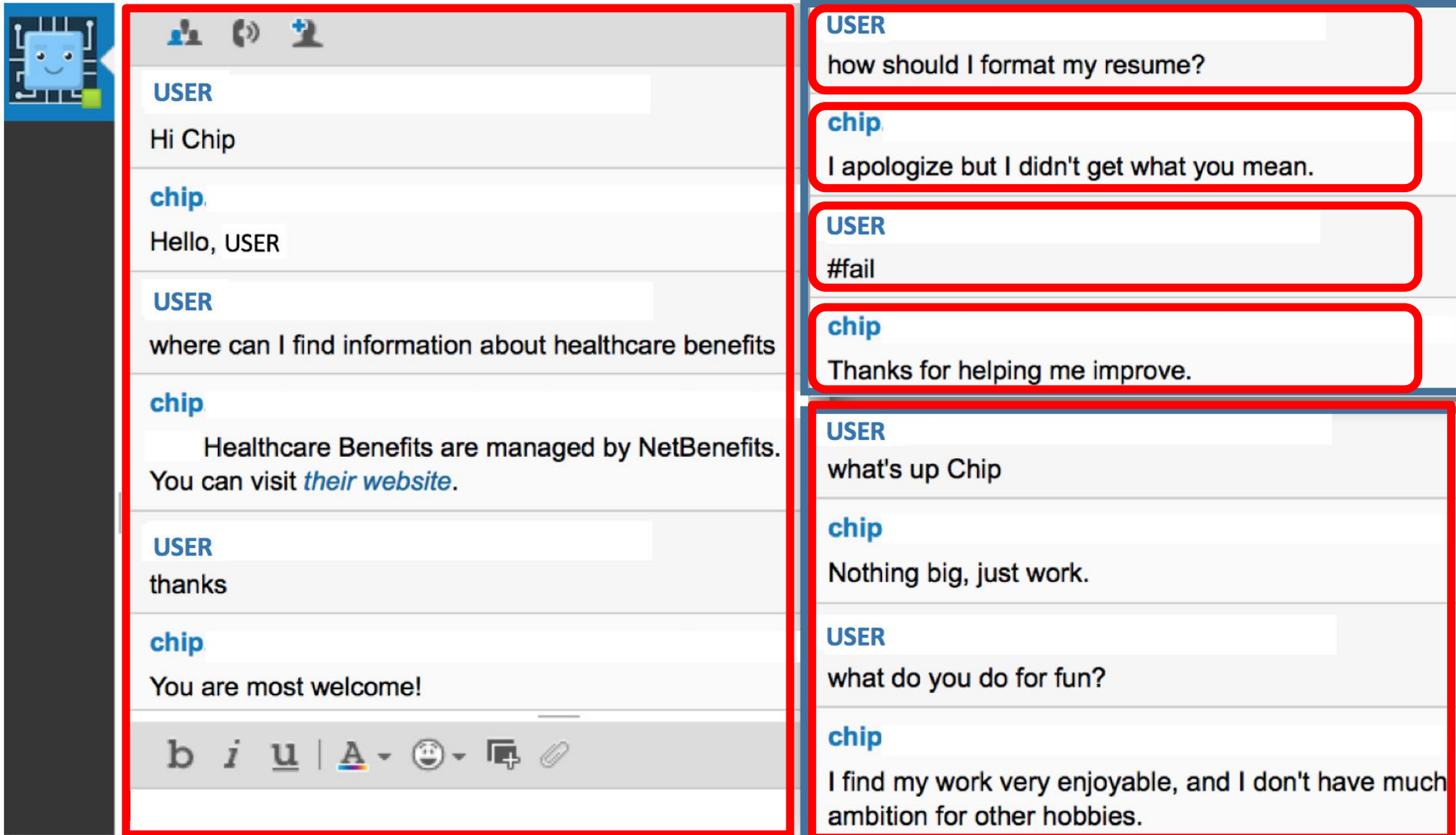
- *Behavioural signal detection*: verbal (text/voice); non-verbal (gaze, posture, gesture)
- *Explicit self-reported measures*: presence, relatability, realism, usefulness, trust
- *Implicit feedback*: dwell time, number/length of utterance

➤ Real-time algorithmic adaptations

- *User goals/needs*
- *Agent representation*
- *Utterance presentation*



Non-Embodied QA Chatbots: Watson Assistant



The image shows a chat interface with a user and a chatbot named Chip. The chatbot's responses are highlighted with red boxes. The interface includes a header with a blue robot icon and three status icons (people, speech, and a plus sign). The chat history shows the following exchanges:

- USER:** Hi Chip
- chip:** Hello, USER
- USER:** where can I find information about healthcare benefits
- chip:** Healthcare Benefits are managed by NetBenefits. You can visit [their website](#).
- USER:** thanks
- chip:** You are most welcome!

The chatbot's responses are highlighted with red boxes. The chatbot's responses are:

- chip:** I apologize but I didn't get what you mean.
- chip:** Thanks for helping me improve.

The chatbot's responses are highlighted with red boxes. The chatbot's responses are:

- chip:** I find my work very enjoyable, and I don't have much ambition for other hobbies.

Evaluation of the QA Chabot (Liao et al. CHI 2018)

- **Prediction of user satisfaction**
 - Conversational Act: Natural language classifiers
 - Lexical features: Bag-of-words (tri-gram)
- **Satisfaction with Functionality**
 - Positive:
 - Formal questions “What/Where/Which/How....”
 - Negative:
 - Sarcastic remark: “You are smart”
 - *Agent ability check*: “What can you do?” “Can you do ...?”
 - Closing: “bye”..”
- **Seeking Playfulness**
 - Chit-chat (off topic)
 - Asking about agent’s trait, asking about agent’s status, talking about oneself
 - Testing intelligence “search information about my boss”
 - Second-person pronoun “how are you?”

Conversational
NLC1

OPENING

CLOSING

COMPLIMENT

ACKNOWLEDGE

COMPLAINTS

FEEDBACK

AGENT STATUS

CHITCHAT

AGENT TRAIT

CHITCHAT

AGENT ABIL-
ITY CHECK

OFF TOPIC RE-
QUEST

ABOUT ME

EMOTICON



Effects of Humanlikeness and Conversational Breakdown on Trust in Chatbots for Customer Service

Effie Lai-Chong Law, Durham University, UK

Asbjørn Følstad, SINTEF, Norway

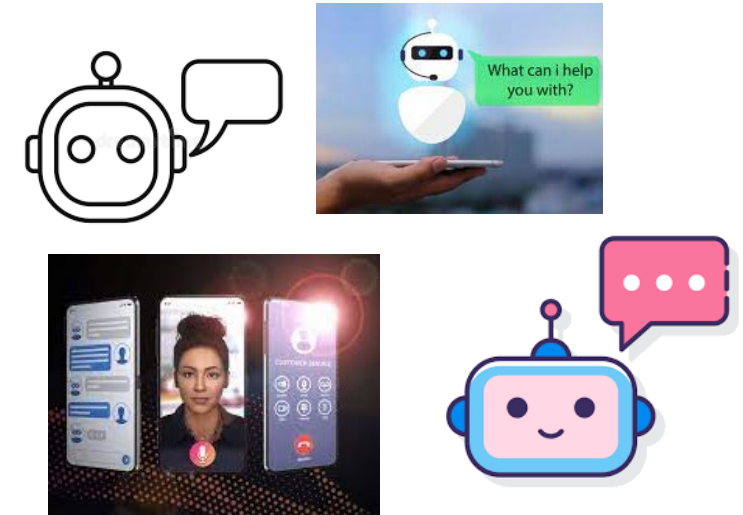
Nena van As, boost.ai, Norway



Motivation

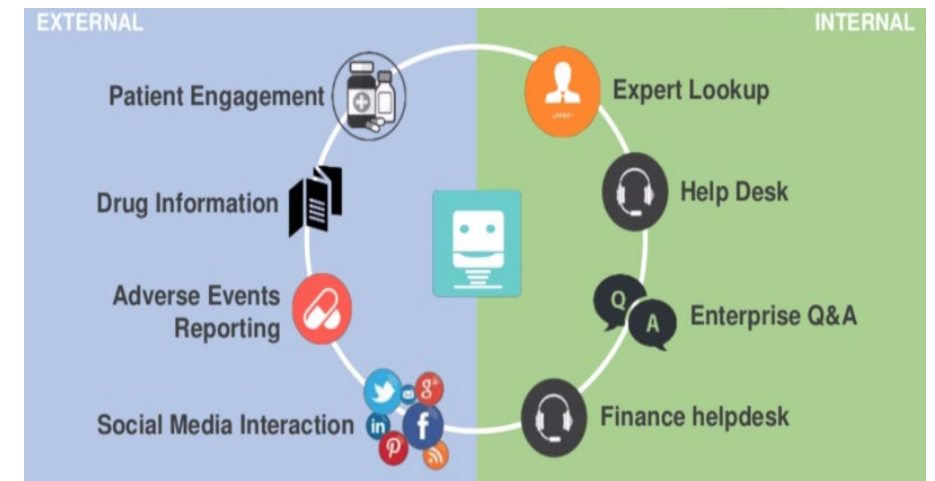
What are Customer Service Chatbots?

Chatbots are **AI-powered** software agents that interact with users through **natural language** and may, hence, be seen as a supplement to chat-based service offered by human support personnel.

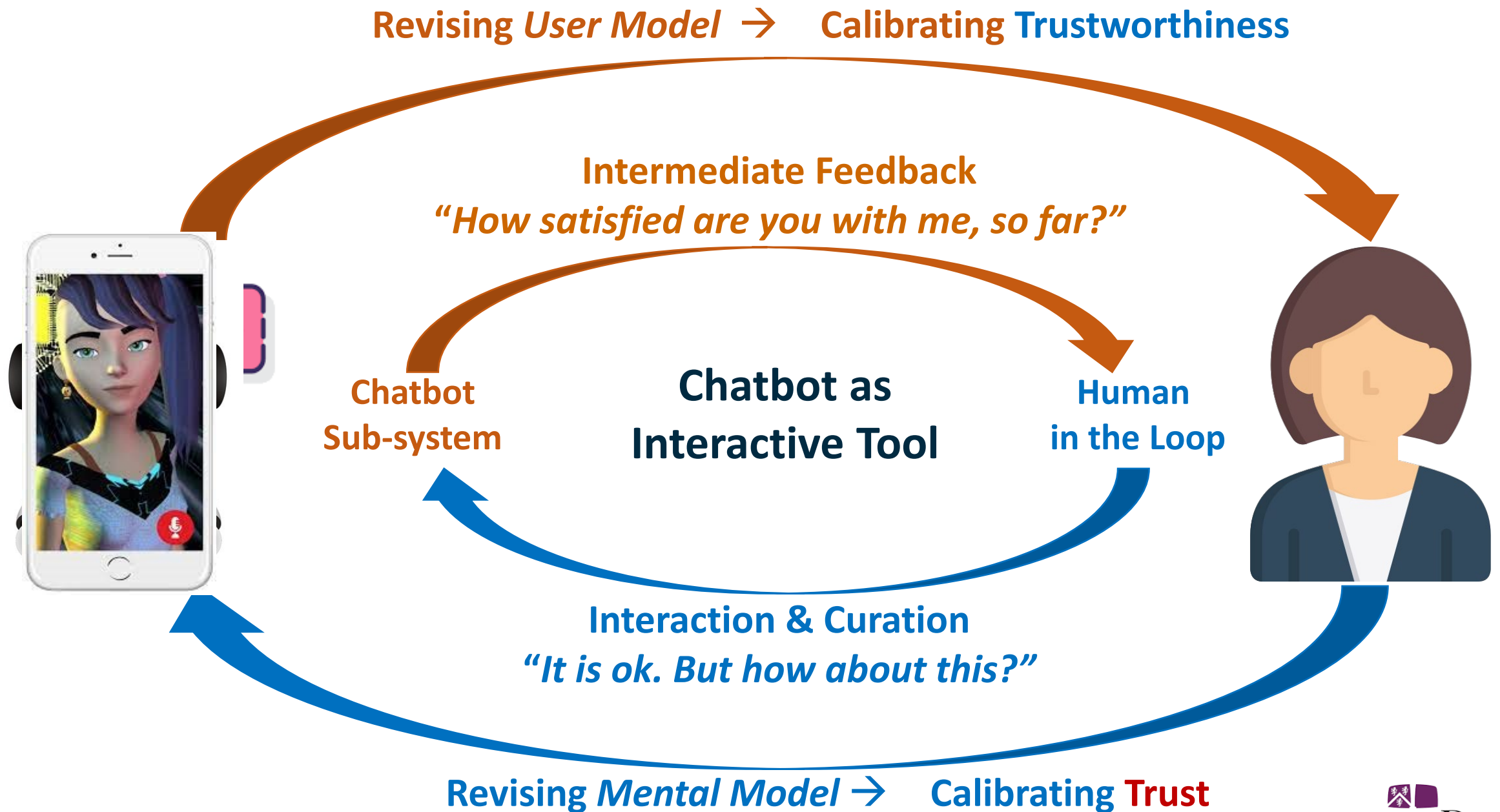


What are Applications?

- Retail and e-commerce
- Travel and hospitality
- Banking, finance, and fintech
- Healthcare
- Media and entertainment
- Education
- ...

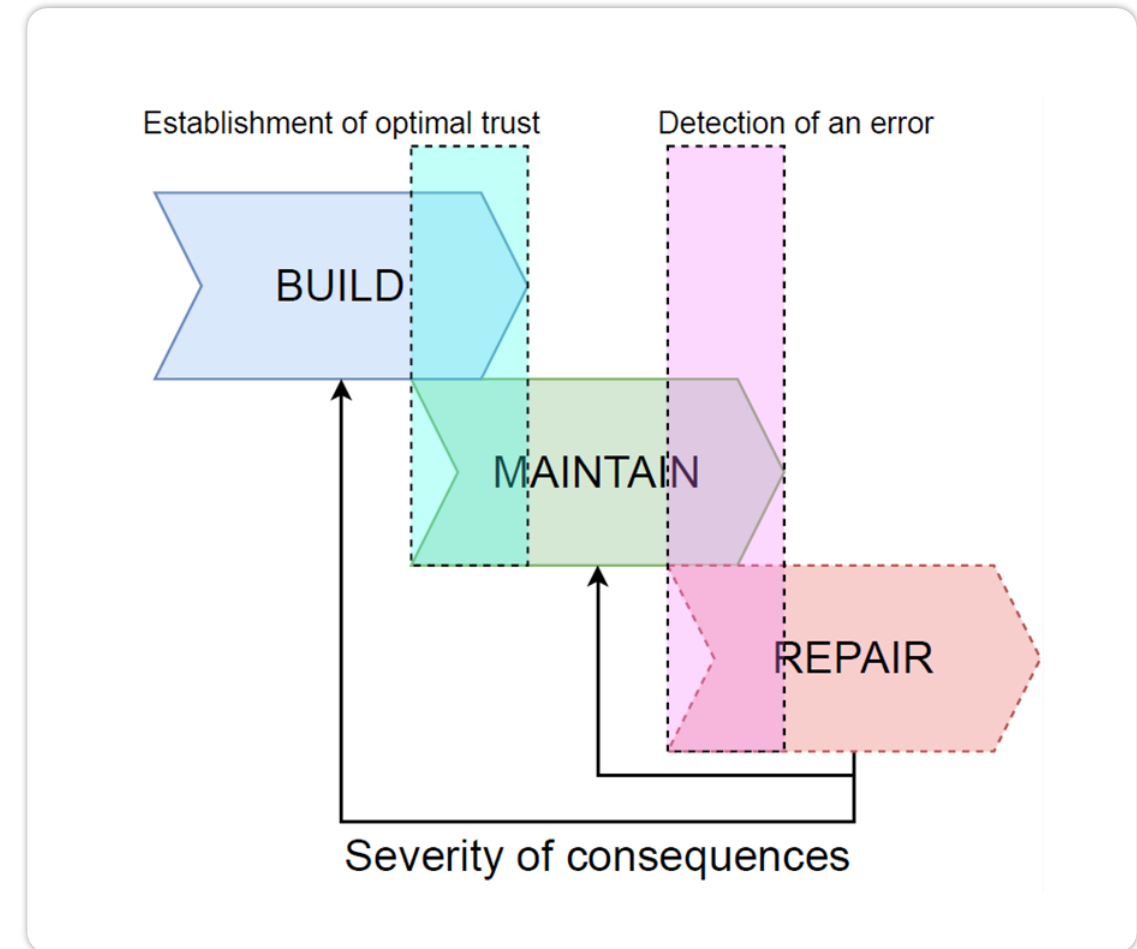


<https://research.aimultiple.com/business-chatbot/>



What is Trust?

- Trust: the willingness of a trustor to accept **vulnerability** (risk) based on positive **expectations** of the **intentions** or behaviour of the other (Rousseau et al., 1998)
- Trust depends on user perceptions of system **functionality** (competence), **helpfulness** (benevolence), and **reliability** (integrity) (Mcknight et al., 2011)
- Trust is hard earned and easily lost.



Source: Dagstuhl Seminar Trust-CA 21381

Research Questions

- Are chatbots' **humanlike** features as important for **trust** as their ability to reliably provide support?
 - Design: conversational style, visual representation, informal language, features hinting at chatbot intelligence (Go & Sundar, 2019)
 - Impact: hedonic user experience, brand perception, user sentiment, user compliance, transaction conversion, and intention to use.
- Is **trust** dominated by the chatbot's **conversational performance** rather than their humanlikeness?
 - The chatbot's ability to provide relevant and helpful responses to users' requests.
 - Chatbot productivity, efficiency and effectiveness → Chatbot user engagement (Brandtzaeg and Følstad, 2017)
- Does chatbot humanlikeness strengthen in the face of performance issues? **trust resilience**
 - Trust resilience: upkeep of trust in spite of undesirable system outcomes (De Visser et al., 2016)

Method

2 x 3 factorial experimental design with two Independent Variables (IVs):

- Humanlikeness (low vs. high)
- Conversational performance (No breakdown, Breakdown with Repair, Breakdown without Repair)

CHATBOT FOR CUSTOMER SERVICE

On this page, you find a chatbot for customer service. The chatbot represents a fictitious consumer bank called Boost Bank.

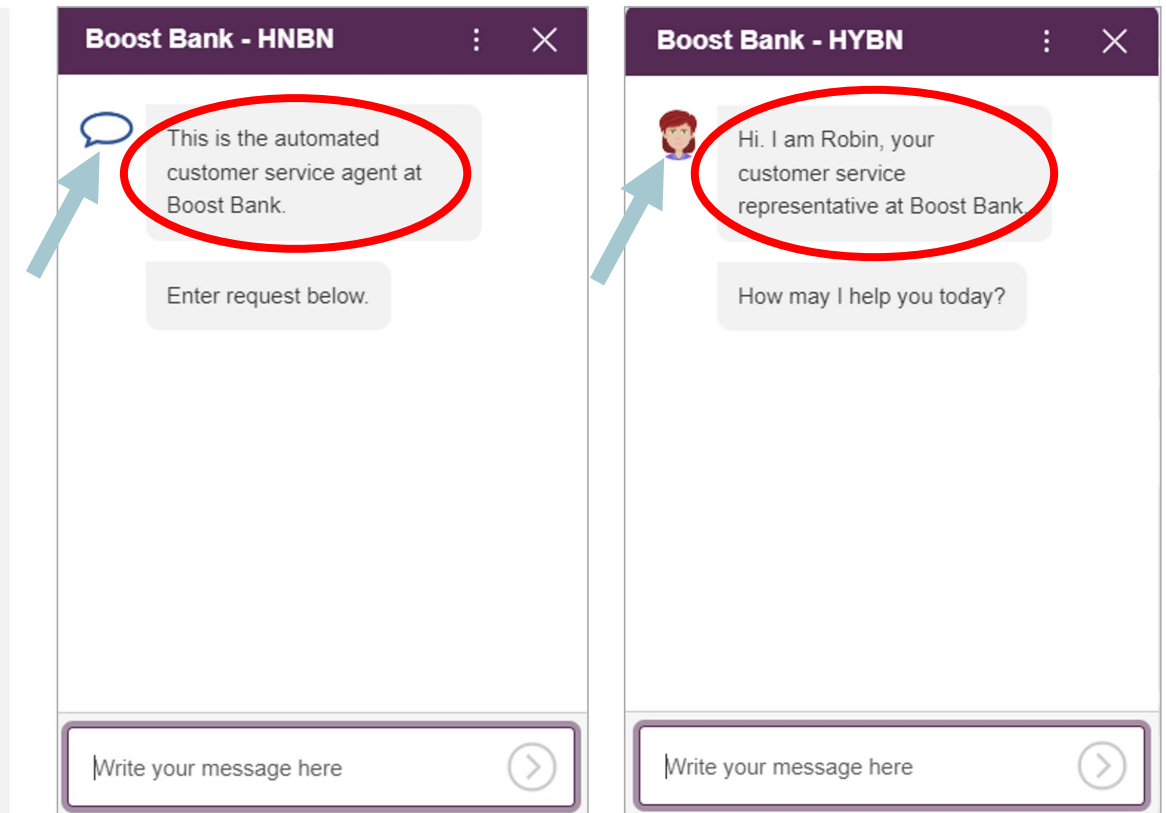
Your first task is to use the chatbot to get information about the following:

- **First**, the Boost Bank loan interest rates
- **Second**, how to apply for a loan at Boost Bank
- **Finally**, making an appointment with a bank advisor at Boost Bank

When you have found the information, the chatbot will provide you a link to a questionnaire for your feedback.

You may also at any time send the message “stop” to the chatbot to end the dialogue and move on to the questionnaire.

Open the chatbot by clicking the icon in the lower right corner.





Humanlike

Chatbot: Hi. I am Robin, your customer service representative at Boost Bank.

How may I help you today?

User: Loan interest rate

Chatbot: I am happy to help you with loan interest rates. At Boost Bank we are proud/happy to offer competitive interest rates. Which type of loan are you interested in?

- Home
- Vehicle
- Flexible

User: [Clicks button]

Chatbot: Thank you. Let me give you a brief overview of our Home / Vehicle / Flexible loan interest rates:

- If the loan is for 70% of the total value or less, you get our best interest rate of 3%
- If the loan is for more than 70% of the total value, interest rates are slightly higher at 4%

Non-Humanlike

Chatbot: This is the automated customer service agent at Boost Bank.

Enter your request below.

User: Loan interest rate

Chatbot: Boost Bank offers competitive loan interest rates. For information, select one of the following.

- Home
- Vehicle
- Flexible

User: [Clicks button]

Chatbot: Home / Vehicle / Flexible loan interest rates are as follows:

- Loans less than or equal to $\leq 70\%$ of total value: interest rate 3%
- Loans greater than 70% of the total value: interest rate 4%

If chatbot does not understand:

Chatbot: I am sorry that I was not able to understand your question.

You may try to ask again using slightly different words. Please keep questions brief and to the point.

If chatbot understands the second time (With Repair)

[the regular conversation on “Apply for a loan”]

If chatbot does not understand – second time (No repair):

Chatbot: I am sorry, but it seems that I am not able to help you with this question.

Is there anything else I can help you with?



If chatbot does not understand:

Chatbot: Request not identified.

Consider entering the request in other words. Requests should be brief and concise.

If chatbot understands the second time (With Repair)

[the regular conversation on “Apply for a loan”]

If chatbot does not understand – second time (No repair):

Chatbot: Unable to respond to request.

A new request may be entered.



Measures

Variable (nr. items)	Example item (7-point Likert scale)	Source
Trust General (3)	When in need of customer service, I feel I can depend on the chatbot	Lankton et al. (2015)
Task-specific Trust (3) TT1, TT2, TT3	Considering the chatbot's answer on [Task 1/2/3], I feel I can depend on it. I can rely on the support provided by the chatbot on [Task 1/2/3]. I feel I can count on the chatbot for questions on [Task 1/2/3]	Home-grown
Trust Belief Reliability (3)	The chatbot is a very reliable solution	Lankton et al. (2015)
Social Presence (4)	I felt like I was engaged in an active dialogue with the chatbot	Laban & Araujo (2019)
Perceived Anthropomorphism (3)	Machine like vs. Human-like (semantic differential)	Araujo (2021)
Prior Chatbot Experience	<i>Prior Use Preference (3)</i> : I frequently use chatbots for customer service <i>Prior Satisfaction (3)</i> : I usually find chatbots for customer service pleasant to use <i>Prior Use Frequency (1)</i>	Home-grown
Demographic	Gender; Age; Country of residence; Education	Home-grown

Null Hypotheses

- Overall trust (OT) is the same across the six conditions.
- Task specific trust (TT) is the same...
 - For each task **between** conditions
 - For each task **within** conditions

	No Breakdown (n=77)	Breakdown with Repair (n=85)	Breakdown without Repair (n =89)
Humanlike (n=121)	Group 1 (n =39)	Group 2 (n =40)	Group 3 (n =42)
Non-humanlike (n=130)	Group 4 (n =38)	Group 5 (n =45)	Group 6 (n =47)

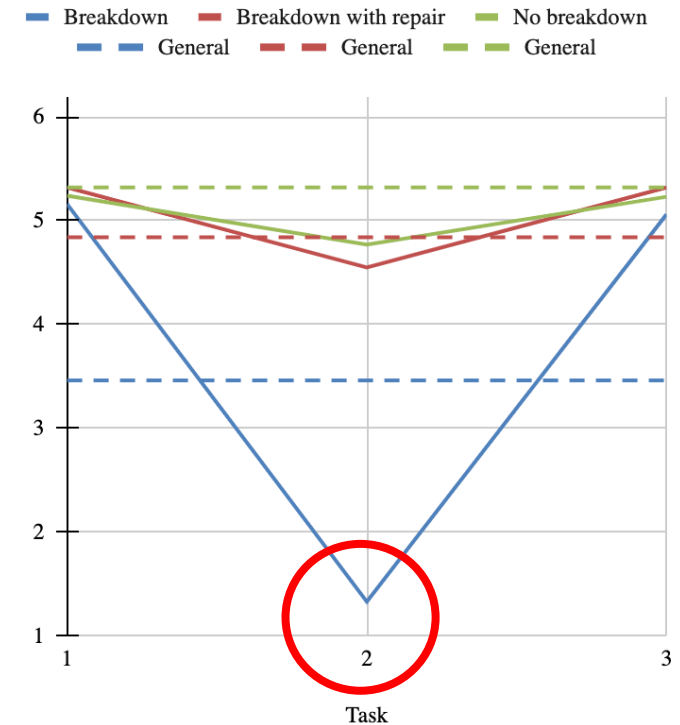
Results

- Breakdown severely impacts task-specific trust.
- Similar trend for general trust.
- Humanlike and non-humanlike conditions seem similar, especially in the breakdown conditions.

Humanlike



Non-humanlike



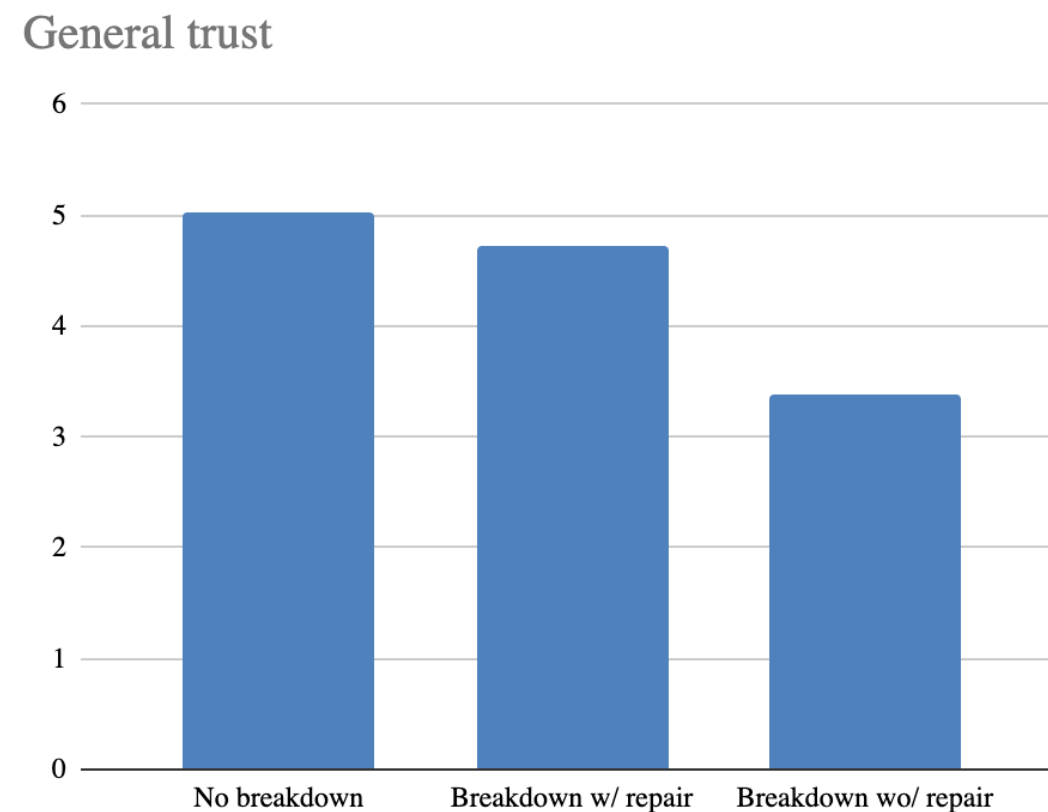
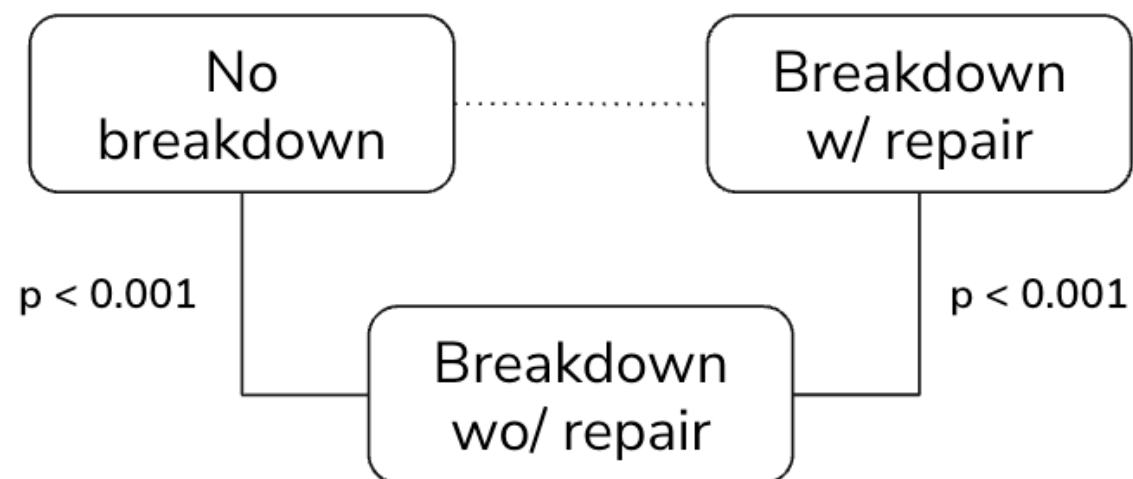
- **First**, the Boost Bank loan interest rates
- **Second**, how to apply for a loan at Boost Bank
- **Finally**, making an appointment with a bank advisor at Boost Bank

Analysis

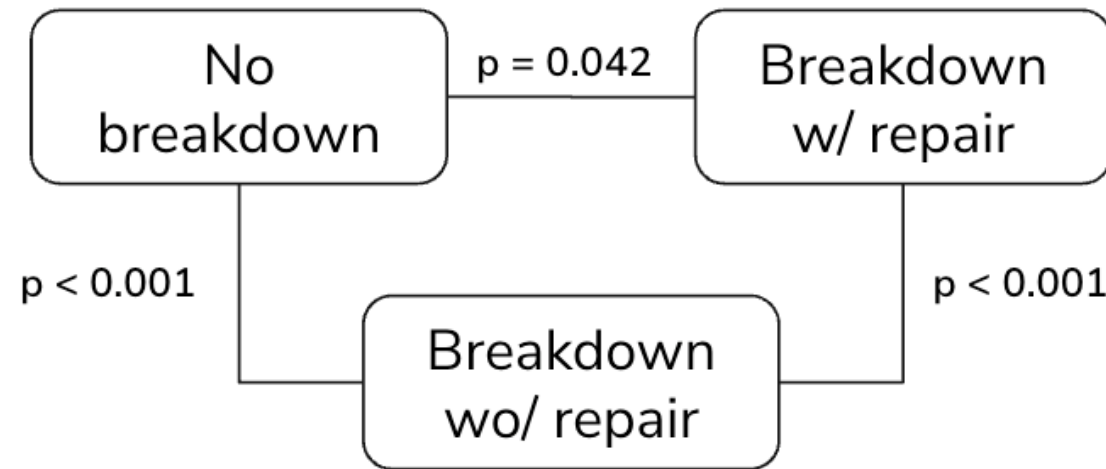
- Both *Humanlikeness* and *Conversational Performance* have an effect on General Trust (GT) and Task-specific Trust (TT)
- For Task-specific Trust:
 - *Humanlikeness* only had effect on TT1.
 - *Conversational performance* only had effect on TT2.
 - **TT3 not affected.**

		GT	TT1	TT2	TT3
Humanlikeness	F(1,245) p η	4.617 0.033 0.18	8.802 0.003 0.035	2.902 0.096 0.025	1.149 0.285 0.015
Conversational performance	F(1,245) p η	36.98 <0.001 0.23	0.87 0.422 0.007	210.91 <0.001 0.21	0.840 0.433 0.007
Interaction Effect	F(1,245) p η	0.437 0.647 0.004	1.740 0.178 0.014	1.904 0.151 0.015	0.247 0.781 0.002

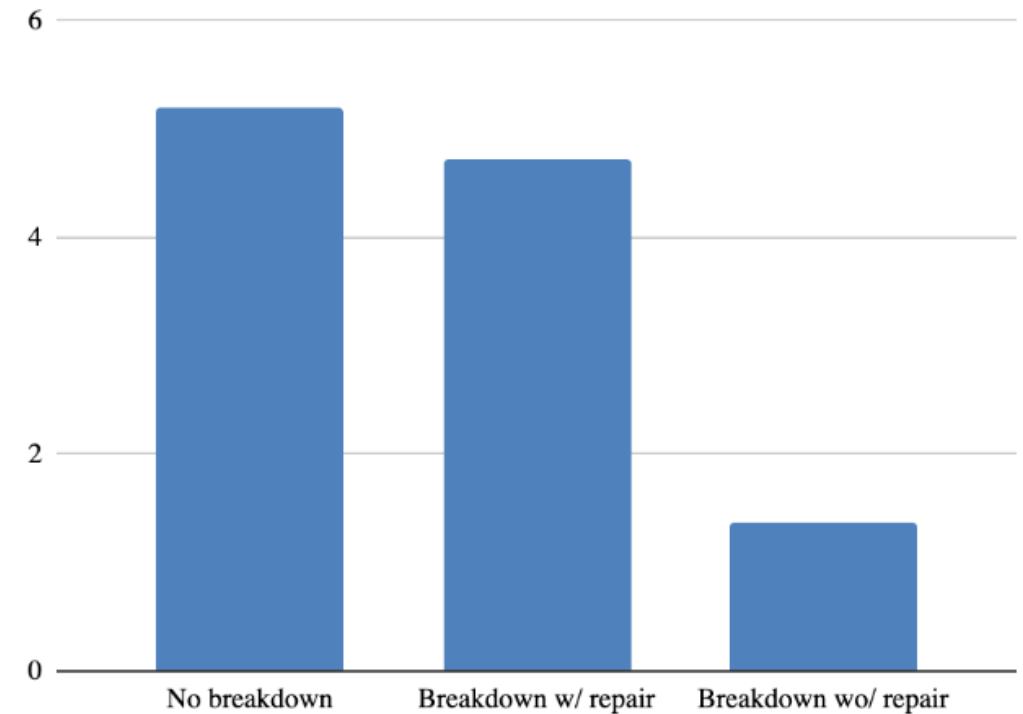
Conversational Performance on General Trust



Conversational Performance on Task-Specific Trust (Task 2)



Task specific trust for task 2



Discussion

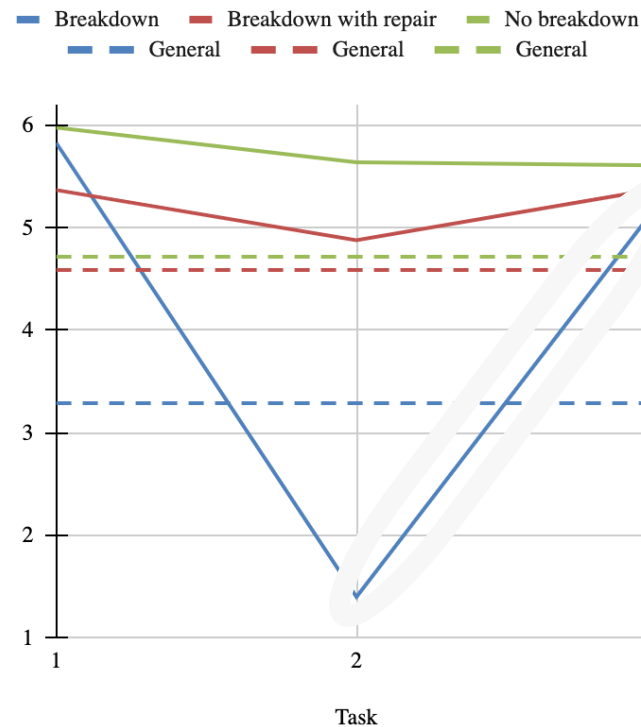
Does chatbot **humanlikeness** strengthen **trust resilience** in the face of performance issues?

No!

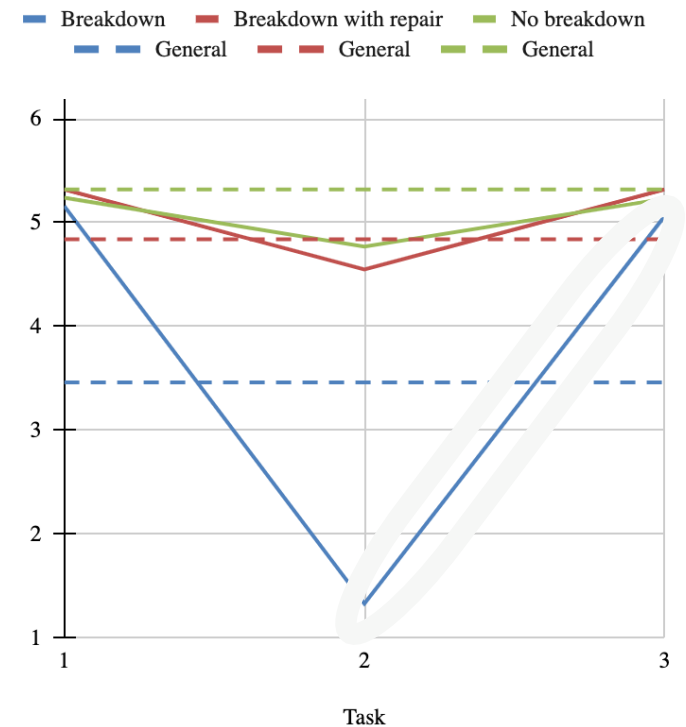
Humanlikeness and Conversational Performance exerted their effect independently of each other.

Recovery of trust from Task 2 to Task 3 is **NOT** stronger for the humanlike condition.

Humanlike



Non-humanlike



Discussion

Is **trust** dominated by the chatbot's **conversational performance** rather than their humanlikeness?

Yes!

Humanlikeness only impacted Task-specific Trust upon first encounter. After experience, no longer significant.

		GT	TT1	TT2	TT3
Humanlikeness	η	0.18	0.035		
Conversational performance	η	0.23		0.21	
Level 1 : 2	t(160) p			2.12 0.042	
Level 2 : 3	t(172) p	6.85 <0.001		17.81 <0.001	
Level 1 : 3	t(164) p	7.71 <0.001		19.22 <0.001	

Practical Importance of Repair

Strength of repair:
as if no breakdown occurred at all!

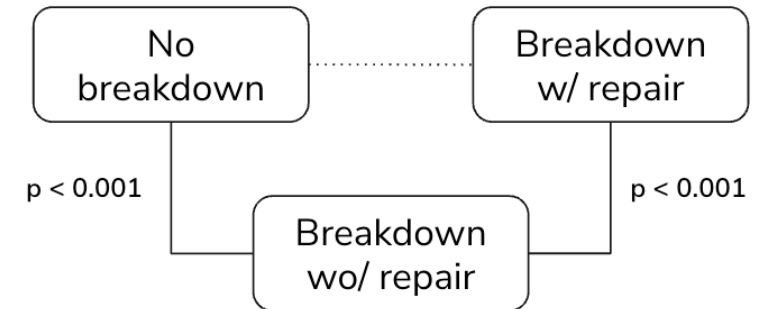
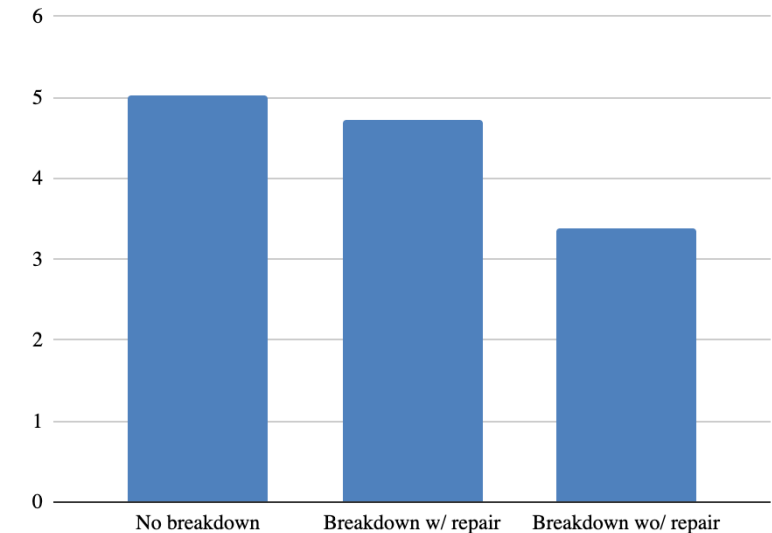
Trust resilience

Prioritisation:

- Anticipate on failure
- Design for repair
 - with/without apology
 - type of apology



General trust



Limitations & Implications

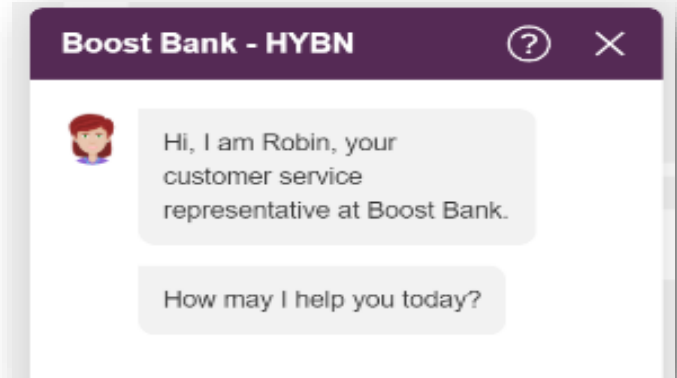
➤ *Humanlikeness*

- Appearance: Modality, Embodiment, Realism
- Humanlike chatbot conversations (Jain et al. 2017)
 - Clarify the chatbot **capabilities** (disclosure)
 - Sustain the **context** of the talks (context-awareness)
 - Acknowledge dialog **failures** (with/without apology)
 - End the conversations with **grace** (Peak-End effect)

➤ *Task criticality*

- Perceived relevance and riskiness
- Order effect

- **First**, the Boost Bank loan interest rates
- **Second**, how to apply for a loan at Boost Bank
- **Finally**, making an appointment with a bank advisor at Boost Bank





Effects of Prior Experience, Gender, and Age on Trust in a Banking Chatbot with(out) Breakdown and Repair

Effie Lai-Chong Law, Durham University, UK

Nena van As, boost.ai, Norway

Asbjørn Følstad, SINTEF, Norway



Research Question

What is the respective effect of

(a) prior experience

(b) gender

(c) age

on the perceived **trust** and interaction qualities of the chatbots characterised by humanlikeness and conversational performance?

Prior Experience on Trust

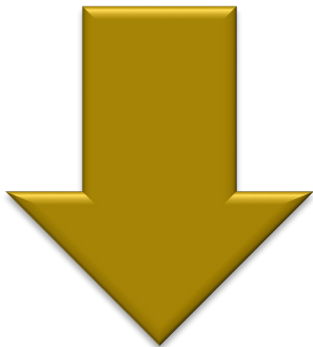
Impressions formed in **previous interactions** with humans/ products/ services can shape people's attitudes and behaviours in subsequent encounters with entities having some **similar traits**: **Cognitive Bias**



Halo Effect

- Good impression
- Positive transfer
- More trust and better value

Beauty and Usability in HCI
Marc Hassenzahl (2004)



Horn effect

- Bad impression
- Negative transfer
- Less trust and lower value

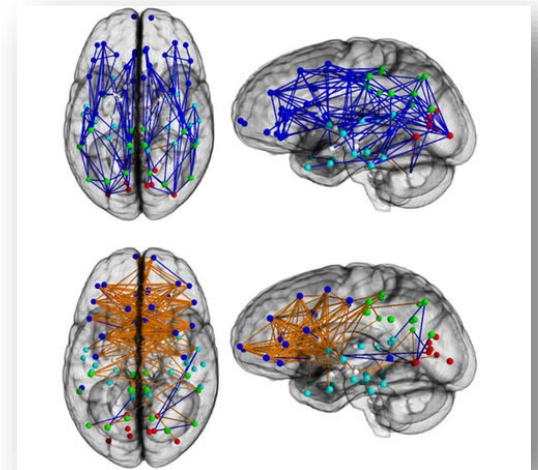
Marketing research
Marie-Sophie Schönitz (2019)

Gender Effect on Trust

- Men tend to trust interaction objects (human, non-human) *more* than women (Buchan et al., 2008; Wu et al., 2020)
- Women more sensitive to *social risk* while trusting (Wu et al., 2020)
- Gender differences in *predisposition* to trust (Rachid, 2020) and *trust dynamics* (Haselhuhn et al 2015)
 - following a trust violation, women less likely to lose trust and more likely to restore trust in a transgressor than men
- Gender of chatbot avatars
 - female avatar more likely forgiven for error (Toader et al 2019)
 - online shopping: no gender effect (Kasilingam 2020)
 - Student support chatbot: no gender effect (Pesonen 2021)



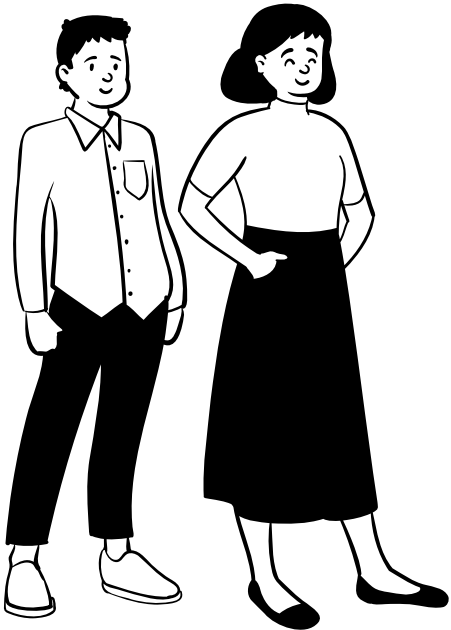
The Neurobiology of Trust. Scientific American (2008)



<https://blog.culturaldetective.com/2013/12/17/new-brain-study-illustrates-gender-differences/>

Age Effect on Trust

- Chatbot usefulness
- Chatbot enjoyment
- Use chatbot to avoid human contact
- Effort invested in using chatbot
- Appreciate chatbot with human touch (Goot & Pilgrim 2019)



Method

- Between-subject experimental studies with six conditions
- A set of measures on trust and demographic variables

	No Breakdown (n=77)	Breakdown with Repair (n=85)	Breakdown without Repair (n =89)
Humanlike (n=121)	Group 1 (n =39)	Group 2 (n =40)	Group 3 (n =42)
Non-humanlike (n=130)	Group 4 (n =38)	Group 5 (n =45)	Group 6 (n =47)

Variable (nr. items)	Example item (7-point Likert scale)	Source
Trust General (3)	When in need of customer service, I feel I can depend on the chatbot	Lankton et al. (2015)
Task-specific Trust (3) TT1, TT2, TT3	Considering the chatbot's answer on [Task 1/2/3], I feel I can depend on it. I can rely on the support provided by the chatbot on [Task 1/2/3]. I feel I can count on the chatbot for questions on [Task 1/2/3]	Home-grown
Demographic	Gender; Age; Country of residence; Education	Home-grown

Demographic variables

- Prior Use Preference
 - when choices available, chatbot is preferred
- Prior Chatbot Satisfaction
 - perceived efficiency, effectiveness and pleasure of using the chatbot for the service

7-point Likert Scale
(1: Strongly Disagreed,
7 Strongly Agreed)

Prior Chatbot Experience

Prior Use Preference (PF)

PF1: I frequently use chatbots for customer service
PF2: I use chatbots for customer service when this is provided as a service alternative
PF3: I have used chatbots for customer service for a long time

Prior Chatbot Satisfaction (SAT)

SAT1: Chatbots for customer service typically provide good help
SAT2: In general, chatbots for customer service are an efficient way to get support
SAT3: I usually find chatbots for customer service pleasant to use.

Prior Use Frequency (FQ)

Five options:

- More than 10 times
- 5-10 times
- 3-4 times
- 1-2 times
- Never

Demographic

Gender (female, male, prefer not to say)

Age (free text)

Country of residence (free text)

Education (three options)

Results: Prior Use Preference

3*2*3 ANOVAs ([Prior variables]*Humanlikeness*Conversational Performance)

Group	Range	n	Overall Trust
Low	1.00-3.67	91	3.71 (.16)
Middle	3.68-5.33	74	4.08 (.17)
High	5.34-7.00	86	5.13 (.14)

General Trust: $F_{(2,233)} = 21.920, p < .001, \eta^2 = .158$
Significant

Results: Prior Use Frequency

3*2*3 ANOVAs ([Prior variables]*Humanlikeness*Conversational Performance)

Group	Range	n	Overall Trust
Low	<5 times	64	3.95 (0.19)
Middle	5-10 times	90	4.44 (0.16)
High	>10 times	97	4.44 (0.16)

General Trust: $F_{(2,233)} = 1.917$, $p = .149$, $\eta^2 = .016$
Non-significant

Results: Prior Chatbot Satisfaction

3*2*3 ANOVAs ([Prior variables]*Humanlikeness*Conversational Performance)

Group	Range	n	Overall Trust
Low	1.00 -3.67	91	3.28 (0.14)
Middle	3.68 -5.33	74	4.32 (0.13)
High	5.34 -7,00	86	5.41 (0.14)

General Trust: $F_{(2,233)} = 65.456, p < .001, \eta^2 = .360$
Significant

Correlations among Prior Variables

- People preferred using chatbots for customer service tended to be satisfied with them.
- For Female and Older users, the more they used them, the more unsatisfied they were with them.

	All N = 251	Female N=178	Male N=69	Younger N=112	Middle N=72	Older N=67
Satisfaction vs. Preference	0.64 $p<.001$	0.640 $p<.001$	0.624 $p<.001$	0.677 $p<.001$	0.607 $p<.001$	0.611 $p<.001$
Satisfaction vs. Frequency	-0.195 $p=.002$	-0.230 $p=.002$	-0.083 $p=.496$	-0.168 $p=.076$	-0.210 $p=.076$	-0.251 $p=.041$

Results: Gender

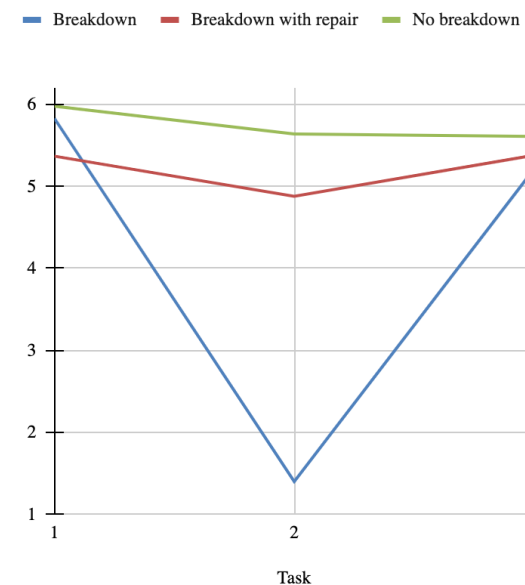
2*2*3 ANOVA (Gender * Humanlikeness * Conversational Performance)

- **General trust:** No significant difference between female and male
- **Trust Resilience:** No significant changes in *Diff(TT1-TT2)* or *Diff(TT2-TT3)* for male and female, both trust resilient.

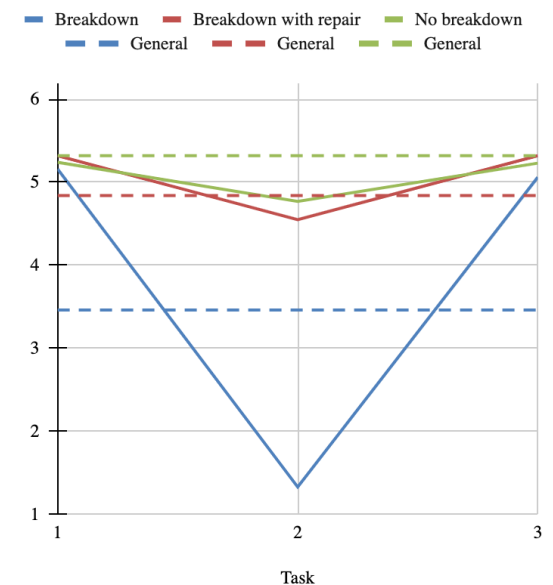
Female						
	Humanlike			Non-humanlike		
Conversational Performance	TT1	TT2	TT3	TT1	TT2	TT3
No Breakdown	6.01	5.77	5.74	5.15	4.91	5.35
Breakdown with Repair	5.27	4.76	5.31	5.37	4.66	5.32
Breakdown without Repair	6.08	1.45	5.22	5.06	1.1	4.94

Male						
	Humanlike			Non-humanlike		
Conversational Performance	TT1	TT2	TT3	TT1	TT2	TT3
No Breakdown	5.86	5.05	5	5.4	4.53	5.2
Breakdown with Repair	5.69	5.21	5.59	5.21	4.31	0.31
Breakdown without Repair	5.5	1.38	5.07	5.52	2.09	5.45

Humanlike



Non-humanlike



Results: Age & General Trust

3*2*3 ANOVAs (Age*Humanlikeness*Conversational Performance)

Group	Range (years)	n	Overall Trust
Younger	18 - 30	112	4.21 (1.61)
Middle	31 - 40	72	4.39 (1.50)
Older	41 - 68	67	4.41 (1.55)

General Trust: $F_{(2,231)} = .759, p = .469$
non-significant

Results: Age & Trust Resilience

Three-way **interaction effects**
(*Age*Conversational Performance**
Humanlikeness) for trust differences were significant:

- $Diff(TT1-TT2)$ ($F_{(4,233)} = 3.57, p = 0.008$)
- $Diff(TT2-TT3)$ ($F_{(4,233)} = 2.49, p = .044$)

Younger						
Conversational Performance	Humanlike			Non-humanlike		
	TT1	TT2	TT3	TT1	TT2	TT3
No Breakdown	5.87	5.56	5.69	4.88	4.59	4.90
Breakdown with Repair	5.15	4.75	5.32	5.20	3.8	4.98
Breakdown without Repair	5.96	1.23	5.32	4.85	1.24	4.85
Middle						
Conversational Performance	Humanlike			Non-humanlike		
	TT1	TT2	TT3	TT1	TT2	TT3
No Breakdown	6.03	5.93	5.73	5.34	4.79	5.49
Breakdown with Repair	5.89	4.83	5.39	5.54	5.28	5.72
Breakdown without Repair	6.00	1.17	5.64	5.25	1.52	5.00
Older						
Conversational Performance	Humanlike			Non-humanlike		
	TT1	TT2	TT3	TT1	TT2	TT3
No Breakdown	6.13	5.47	5.33	5.83	5.13	5.5
Breakdown with Repair	5.48	5.10	5.48	5.24	4.97	5.42
Breakdown without Repair	5.42	1.97	4.57	5.49	1.23	5.44

The three age groups changed the level of trust from task to task **significantly** under different chatbot conditions.

Implications

Prior experience: The halo or horn effect on trust (of positive/negative transfer) from previous chatbot experiences

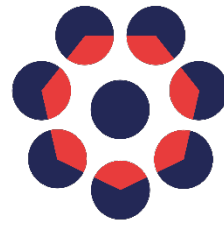
- To understand the experience users bring with them
- To design prompts or marketing strategies to address these earlier experiences.

Gender: No significant difference in trust as observed in human-human interaction

- To reconsider the assumption that models on human trust in AI-powered systems can be grounded in their counterparts on interpersonal trust

Age: No significant main effect, but significant interaction effects

- Older adults may appreciate more emotional than practical value from chatbot interactions



UKRI
**Trustworthy
Autonomous
Systems Hub**

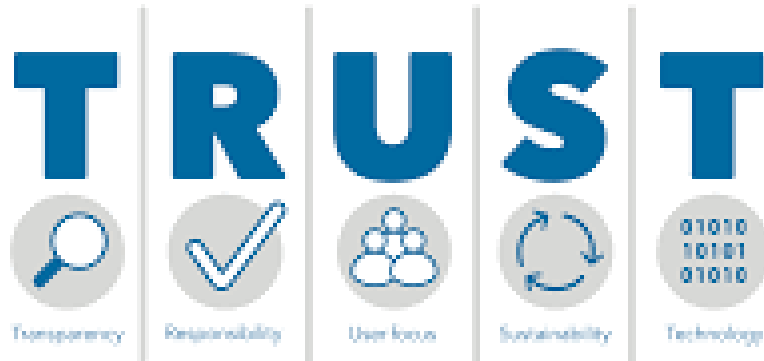
Conversational Agents for Older Adults (CA4OA)

Co-designing *Inclusive* and *Trustworthy* Conversational Agents
on Basic Services with Older Adults

April 2023 – June 2024

Motivation

- Chatbots for customer service: Healthcare, Banking, Entertainment...
- Low adoption among older adults: Trust in emerging technology
- Distrust → Underuse → Digital Marginalisation
- Mistrust → Overuse → Digital Exploitation



Transparency
Responsibility
User-focus
Sustainability
Technology

<https://www.rd-alliance.org/trust-principles-rda-community-effort>

Online Banking

How many bank branches have closed since 2015?

5,579, at a rate of around **54 per month**.

NEWS

[Home](#) | [Cost of Living](#) | [War in Ukraine](#) | [Climate](#) | [UK](#) | [World](#) | [Business](#) | [Politics](#) | [Culture](#) | [Tech](#)

[England](#) | [Local News](#) | [Regions](#)

Lloyds Bank branch closure will hit 'elderly and vulnerable'

10 August 2022



BBC Breaking News

@BBCBreaking

Banking giant HSBC to close 114 branches in UK from April 2023 as customers increasingly bank online



Business ...



Full list of 23 NatWest branch closures happenin...

Visit

Images may be subject to copyright. [Learn More](#)

Mind the Gap

FCA: Basic financial services are available to **vulnerable, digitally excluded parts of society** in physical branches rather than online only.

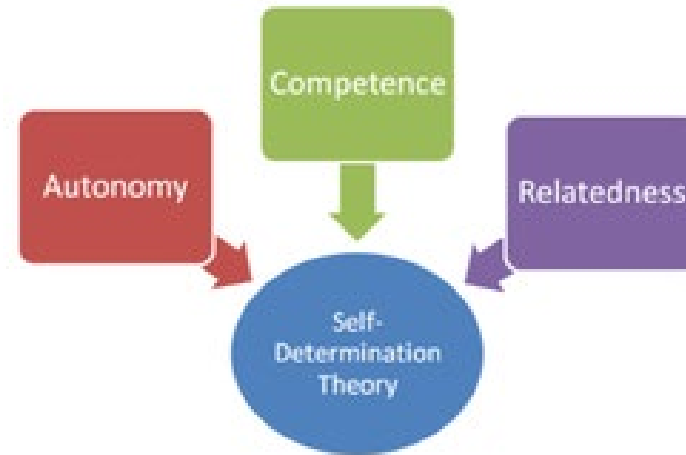
Financial Conduct Authority (FCA) (2022). Branch and ATM closures or conversions. *Finalised Guidance, FG 22/6*, October 2022.

The critical gap between online banking services provision and user adoption can be bridged by improving **customers' trust in digital tools.**



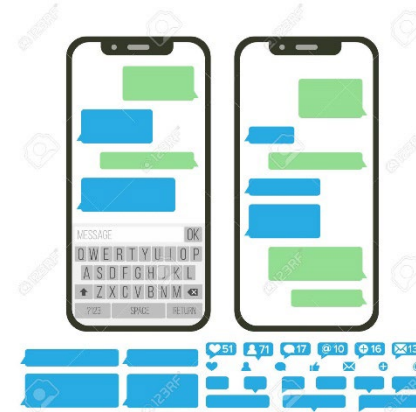
Trust & Older Adults' Attributes

- Three basic needs (SDT)
 - Autonomy
 - Competence
 - Relatedness
- Loneliness (Lieberz et al., 2021)
 - Interpersonal trust
 - Human-robot trust
- Attitudes
 - Confidence, prior experience in technology
 - Prejudice towards technology: Intergroup Contact Theory (Gordon Allport, 1954; Haggadone et al. 2021)

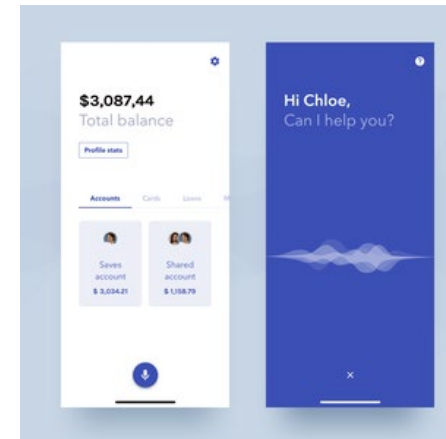


Trust by Design for Conversational Agent (CA)

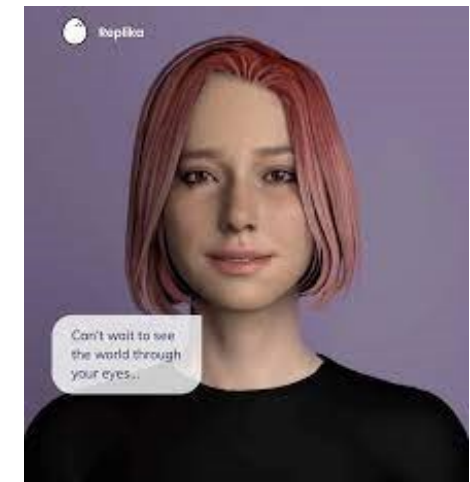
- Varied extents of CA **modality, embodiment and anthropomorphism**: 2D/3D, unimodal/multimodal, and realism (low/high avatar fidelity)
- Conversation flows to match older adults' **mental models** (i.e., human-AI metaphors) of CAs; the transferability of human-human to human-agent communication.
- Conversation patterns of CAs to match older adults' **age, gender, physical & psychological conditions**
- Varied presentations of **alerts to risks and breakdown repairs** on older adults' trust in CAs.



<https://www.allerin.com/blog/what-will-happen-when-ar-vr-and-chatbots-all-converge>



https://dribbble.com/tags/bank_bot



<https://www.superworldapp.com/content-hub/the-future-of-ai-chatbots-is-in-ar>

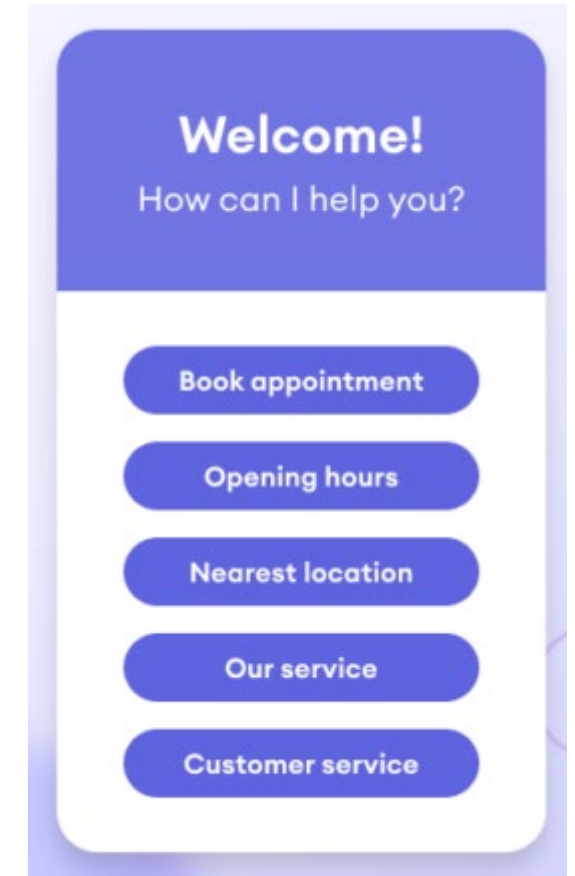
Chatbot Design Requirements (Grudin & Jacques, 2019)

Human-centred Design Approach

- 1) investigating the strengths of different platforms and tools, and identifying **user needs**;
- 2) setting a **personality**;
- 3) determining **task difficulty**;
- 4) understanding the **access environment** and users' tendency to be **forgiving of errors**;
- 5) making **amusing** scripted responses;
- 6) making chatbots reveal their **identity** (disclosure);
- 7) iterating through small-scale **user testing**

Text- vs. Button-based Interaction

- Custom buttons allow only a limited range of inputs, facilitating a **decision-making** process.
- Limiting the options, using easily navigable buttons rather than mobile keyboard makes conversations with chatbots **more accessible**. (Ryu et al. 2020)



Method: Older-Adult-centred Design

- Enabling older adults of heterogeneous backgrounds to state their needs, preferences, and experiences of using CAs
- Qualitative Method: Contextual Inquiry and Participatory Design
 - Interviews
 - Focus groups
 - in situ observations
- Quantitative Method: Questionnaire to measure
 - Loneliness
 - Trust
 - UX
 - Attitudes toward CAs
 - Intention to use
- Online banking services: issuing bank statements, checking account balance, transfer, bill payment, and cheque deposit

Summary

- Chatbots are increasingly impactful **at all levels!**
- Trust in chatbots for customer service mainly determined by their **performance** rather than humanlikeness.
- Trust in chatbots for customer service seems not gender-sensitive, but **age-sensitive**.
- Older adults can become **digitally marginalised** when more and more essential services moving online.
- Make chatbots **trustworthy**; enable older adults to trust and use them.
- **Ethics** is of utmost importance
- **Usable, useful, and safe ICT for older adults' wellbeing!**

Questions?