L-Università ta' Malta
Faculty of Information & Communication Technology

Intelligent Interfaces Weekly Tasks

Matthias Bartolo

B.Sc. It (Hons) Artificial Intelligence (Third Year)

Study Unit Code:	**ICS3211**
Study Unit:	**Intelligent Interfaces 2**
Lecturer:	**Dr Vanessa Camilleri**
Date:	**December 2023**

Contents

1 Week 2 Task - Mental Models

1.1 Mental Models and Conceptual Frameworks

Cognitive tools like mental models and conceptual frameworks aid people in comprehending, interpreting, and navigating the complexity of the outside world. Our minds construct mental models, which are condensed versions of reality, based on our experiences, assumptions, and knowledge. They act as mental short cuts that speed up decision-making and problem-solving for humans. The mental model of supply and demand, for instance, aids economists in their analysis of market dynamics. [1]

Contrarily, a conceptual framework is a well-organised network of related ideas and concepts that offers a logical approach to comprehending and evaluating a certain subject or area. It provides a thorough lens through which one may look at all facets of the topic, enabling a greater understanding of its core ideas and connections. [2]

Moreover, according to [3], *mental models* and *conceptual frameworks* can be considered staple concepts when marketing new technologies. These notions establish a theoretical foundation for comprehending user mindsets, requirements, and behaviours. Additionally, they could also guide product review, marketing, and design while providing a more detailed explanation of the links between the core elements, or building pieces, of a behavioural theory. [3]

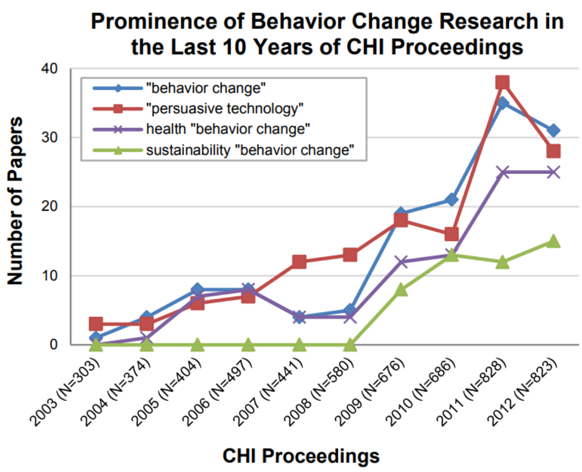

Figure 1.1 The significance of behaviour change research in the last ten years of CHI sessions. (Source: [3])

Evidently, the authors of [3] also argue that although conceptual frameworks and mental models are related, they comprise separate design aspects for technology user interfaces. People develop mental models, which are individualised cognitive representations, to comprehend their surroundings. *HCI* research is used by designers to make technology more user-friendly and captivating by coordinating interfaces with users' mental models. [3]

Nevertheless, conceptual frameworks, which have their roots in well-known theories like the trans-theoretical model or the theory of planned behaviour, are crucial in the development of *HCI* behaviour modification technologies. These frameworks provide methodical ways to analyse and shape user behaviour. They direct the development of functions that increase user satisfaction and engagement, assuring user-friendly designs and widespread technology usage. [3]

Furthermore, as mentioned in [4] different users possess various degrees of understanding of systems or processes from which they need to extract information. Mental models are quite beneficial when it is possible to identify patterns which apply to more than one individual, and system designers can employ these patterns after they have been established in order to tailor new programs to groups of individuals. Unfortunately, some of these models could not be complete; they might merely represent analogies, which frequently function rather well. [4]

1.2 Chosen Application: Facebook

Facebook is a well-known social networking site where users may connect with friends and family, exchange material, and participate in a variety of online communities. It provides tools including news feeds, chat, events, and company sites, making it a flexible platform for both private and public engagement. [5]

Figure 1.2 Facebook Logo.

1.3 Key Features

- **News Feed** - Facebook's constantly updating stream of posts from various sources.
- **Profile** - Users can create personal profiles with photos and information.
- **Friending** - Users can connect with others by sending and accepting friend requests.
- **Groups** - Users can join interest-based groups to connect with like-minded individuals.
- **Marketplace** - Users can buy and sell items locally through Facebook.
- **Pages** - Facebook enables public figures and businesses to use pages to connect with followers.
- **Events** - Users or established company groups can create and coordinate events, inviting others to participate.
- **Messaging** - Facebook also enables the feature to send text, audio, or video messages to friends and groups.

1.4 Strengths

- **Massive User Base** - With 2.8 billion monthly active users, Facebook has an extensive reach.
- **High User Engagement** - Facebook's algorithmic news feed keeps users engaged for extended periods.
- **Facebook Identity** - Facebook's real name policy adds authenticity to user profiles.

1.5 Usage

- **Staying Current** - Users rely on the news feed to stay updated on friends, news, and trends.
- **Content Sharing** - Sharing life updates, articles, photos, and videos is a common activity.
- **Messaging** - Users engage in individual and group messaging for socialising and coordination.
- **Entertainment** - Facebook provides games, quizzes, and other forms of entertainment.

1.6 Improvements

- **Reduce Misinformation** - Facebook could implement features to mitigate the spread of false or misleading content.
- **Enhance Privacy** - Facebook could also offer users more control over their data privacy and sharing settings.
- **Algorithmic Transparency** - Facebook could render the news feed algorithm’s workings more transparent.

1.7 Questionnaire

1. Participant's Background
   - **Person 1:** (age 20) is currently a Student
   - **Person 2:** (age 37) is a parent of 1 child
   - **Person 3:** (age 56) is a parent of 2 children
2. What product/online service are you using?
   - **Person 1:** Twitter/X
   - **Person 2:** Spotify
   - **Person 3:** Messaging (SMS Mobile App)
3. Why do you use it?
   - **Person 1:** I use Twitter to stay up to date with the latest news and updates whilst, allowing me to send messages to my close contacts and also post tweets occasionally.
   - **Person 2:** I use Spotify to listen to different music, and broaden my music taste.
   - **Person 3:** I use Messaging to be able to communicate with my family and close contacts.
4. How do you use it?
   - **Person 1:** I don't engage with Twitter by posting myself, but I use it to keep up to date with my close friends.
   - **Person 2:** I listen to Spotify in order to listen to music, whilst commuting.
   - **Person 3:** I use Messaging by sending regular messages to my family and close friends to update them on my current situation.
5. Do you think there are things in it which you would improve?
   - **Person 1:** I would like Twitter to have a more personalised discovery.
   - **Person 2:** I would improve Spotify by introducing a greater variety of content.
   - **Person 3:** There isn't much really to improve upon, as the simple design is sufficient to send messages, however, a better user friendly design would help.

6. If you were creating something similar how would you re-design it to suit your needs?
    - **Person 1:** I would re-design Twitter by making the interface more user friendly, whilst also changing the feed to be more personalised to my particular needs.
    - **Person 2:** I would re-design Spotify by improving the music recommendation algorithm, and playlist sorting.
    - P**erson 3:** I would improve the Messaging app, by improving upon its design to make it more user friendly and have a modern theme.

1.8 Conclusions and Analysis

From the collected feedback it was noted that mental models and conceptual frameworks are key factors in influencing how people choose to utilise social media sites. These models affect how they use the platform, as well as how useful they perceive it to be. Users' motivations for using the platform are consistent with how they perceive its main features. The subtle interaction between these models is essential for comprehending and meeting the varied expectations of users, eventually influencing the development and layout of social media experiences.

2 Week 4 Task - Evaluating Intelligence

The two papers selected for this analysis are ”See What You Want to See: Visual User-Driven Approach for Hybrid Recommendation” [6], which describes SetFusion, and ”A Cognitive Perspective on Gestures, Manipulations, and Space in Future Multi-Device Interaction” [7], which focuses on gestures, manipulations, and space.

2.1 Key similarities:

These two publications are fundamentally similar in a number of ways. First of all, the two articles have the same objective of improving the user experience and creating more efficient and intuitive interfaces through investigating novel interface designs. Second, in an effort to increase the range of interaction options, both papers concentrate on pushing the limits of conventional input techniques by experimenting with touch, gestures, and spatial manipulations. Finally, by attempting to create more seamless links between user input and system output, both articles aim to close the gaps between interface design execution and assessment, thus guaranteeing a more effective and user-friendly interaction experience.

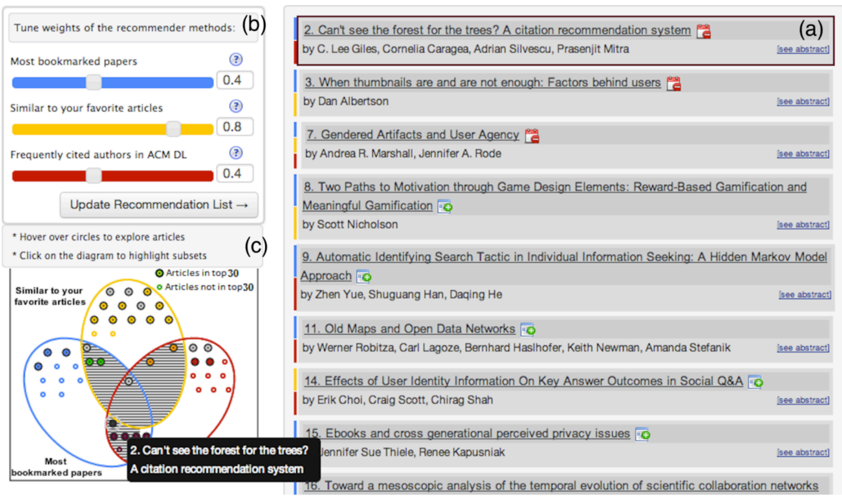

Figure 2.1 Screenshot of SetFusion presenting (a) a filtered list of recommended papers, (b) sliders, and (c) the Venn diagram. (Source: [6])

2.2 Key differences:

In order to enhance user cognition and spatial memory, Jetter's study [7] focuses on the investigation of gestures and spatial manipulations as input modalities, stressing the importance of blended interactions. In contrast, Parra et al. [6] focus on innovative visualization approaches, especially for recommender system output. Its main goal is to provide consumers with more influence over suggestions by means of an interactive visualization. Specifically, Jetter [7] explores cognitive perspectives on interaction through a more conceptual approach, whereas Parra et al. [6] offer a more practical and tangible approach to their research by presenting a tangible system implementation alongside user studies.

2.3 Usability Characteristics:

2.3.1 Learnability:

In terms of learnability, Jetter [7] observes that gesture sets are an artificial language that, with careful design, most users should be able to adopt, rather than something that is inherently ”natural.” He also points out that examples of the huge potential of Application-specific gesture sets are demonstrated by apps like the remarkable tabletop system for document editing with pen and touch gestures by Matulic & Norrie.

Moreover, Para et al. [6] note that SetFusion included fusing sliders to enhance user control over source integration, and instead of employing clustermaps as in TalkExplorer, more basic Venn diagrams were utilized to depict set intersections. Ultimately, the use of these Venn diagrams presents an easier-to-understand visual paradigm for set presentation, thus enabling faster learnability.

2.3.2 Efficiency:

A sensation of flow, control, and directness that is unparalleled is achieved through the use of such gesture sets, as demonstrated in Jetter's study [7] in the best instance. Assuming that a gesture is quicker than choosing a menu item or remembering keyboard shortcuts, in a poor scenario, users only save a few seconds, and they are not disturbed (assuming that switching between a mouse, touchpad, and keyboard is no longer necessary).

An analysis of the research findings presented in [6] revealed that SetFusion operates much more significantly on user motivation, performance, and attitude when used in natural suggestion mode. First, a number of factors showed that SetFusion is more user-friendly when used in its natural state. Twice as many users used SetFusion, bookmarked chats with it, and used it often, compared to a comparable number of people exposed to the systems. Additionally, users spent a significant amount more time on average on the system and bookmarked nearly four times as many discussions. Secondly, the data presented compelling evidence that users utilized SetFusion in its native form more productively, reducing the ratio of support actions, the overall time spent to process yield, or the quantity of chats bookmarked.

2.3.3 Memorability:

In the worst scenario, learning and remembering gestures becomes as challenging as using keyboard shortcuts or command-line languages, and using them could have a detrimental impact on memorability. This is especially true if gestures are not used frequently. For many application contexts, it is just not possible to create a mapping between instructions and gestures that is distinct, obvious, and simple to remember. [7]

Regarding SetFusion, while it isn't said explicitly, it may be inferred that users can quickly regain skill because of the increased number of visual interface components between the UMAP research and the 2013 pilot trial. In contrast, the latter witnessed a rise in overall user engagement, which eventually resulted in increased application competency. [6]

Figure 2.2 Usability Characteristics. (Source: [8])

2.3.4 Errors:

Errors can occur in gesture-based interactions, particularly when dealing with intricate sequences of motions. This is due to the fact that understanding the functions of each relevant gesture require users to make initial guesses and memorise enormous sets of gestures. [7]

In the case of SetFusion, user faults were not recorded as the focus of the article was on assessing user engagement with the revised system. However, when it came to the suggestions it made, the system itself was prone to inaccuracies. Fortunately, these could be easily fixed with sliders and other simple tweaks, allowing users to get the right result. The fact that users modified their list of recommendations 4.36 times on average after adjusting the sliders lends more credence to this. demonstrating the ease of use and potency of such a basic function. [6]

2.3.5 Satisfaction:

It can be noted that the goal of using gestures as user input in HCI is to meet the demands of users in contexts that are either context- or domain-specific. For instance, using gestures instead of touch controls or other kinds of interaction might be helpful in hygienic environments like hospitals. [7]

Furthermore, Para et al. [6] note that SetFusion's visual interface, which gives users control over the fusion process and enhances controllability through enhanced visualization of the process, may be more conducive to user satisfaction and success than traditional "black box" system-driven hybridization in terms of satisfaction.

2.4 Elements indicating intelligence in Parra et al.'s work:

The interface design by Parra et al. incorporates a number of clever components. They include interactive visualizations, which greatly increase system transparency by giving users the power to exercise control and a clear knowledge of the suggestion process. Furthermore, users may customize their experience by customizing suggestions to suit their interests thanks to the integration of numerous recommendation sources that can be adjusted using sliders. Furthermore, by using Venn diagrams to show the intersections between different suggestion lists, complicated information is made easier to understand and comprehend for consumers, resulting in an interface that is easier to use.

[6]

2.5 Elements indicating intelligence in Jetter's work:

Jetter's research is notable due to its emphasis on using spatial cognition and the proprioceptive feedback that is present in movements and spatial adjustments. This strategy attempts to improve learning and memory retention while lowering the cognitive strain on users. Moreover, a more intuitive user experience is achieved through Jetter's emphasis on body-based interactions and users' natural spatial abilities rather than abstract dialogue metaphors. Because of its intuitive design, which mimics how people naturally interact with their actual surroundings, it is easier to use and more accessible.

[7]

2.6 Summary

Essentially, the goal of Parra et al.'s research on SetFusion and Jetter's work on gestures and spatial manipulations is to create interfaces that are transparent and intuitive by design. These interfaces close the gap between execution and assessment by using users' pre-existing cognitive and physical capacities, thereby accelerating the transition to more intelligent, user-centered systems. Their different approaches—one focusing on cognition and input techniques, the other on visualization and control—taken together help to shape the development of user interfaces that are more in line with human capabilities and provide a more intuitive and intelligent interaction experience.

3 Week 6 Task - Future Trends in IUIs & the role of ASR

Intelligent User Interfaces (IUIs) comprise systems that employ AI and machine learning to improve user interactions with digital interfaces [9]. These interfaces respond dynamically to the demands, preferences, and behaviors of users, resulting in a more intuitive and personalized experience. A basic example of IUI can be demonstrated by the implementation of natural language processing in chatbots or voice-activated systems, which enhances the way individuals engage with technology.

3.1 Application of IUI in Healthcare

IUIs have the potential to improve healthcare by providing more personalized and intuitive interactions between patients, physicians, and health information systems [10]. Voice-based IUIs, for example, could enable doctors to enter patient data and access medical information conversationally, decreasing paperwork loads. Additionally, chatbot IUIs could provide patients with personalized education and assistance outside of doctor appointments [11]. Another application for IUI in healthcare can be attributed to augmented reality, whereby IUIs may help doctors navigate complicated procedures or observe the anatomy of patients [12]. However, as IUIs get more complex, ethical considerations concerning user privacy become more pressing. This is evident in IUIs that use data mining or machine learning to adapt to individual users, which risk exposing sensitive health data if appropriate privacy controls are not in place [13]. Strict access restrictions, encryption, and anonymization techniques, as well as upfront permission procedures that clarify how user data will be used, can all help preserve patient privacy. Furthermore, in a high-risk setting, IUIs should likewise avoid going beyond supporting doctors to completely replacing human judgment. Continuous monitoring and iterative design that incorporates physician and patient input will be required to ensure that these technologies supplement rather than detract from treatment. Additionally, to safeguard the health of the patients who use healthcare systems, regulations such as the Health Insurance Portability and Accountability Act (HIPAA) [14] in the United States must be enforced. Overall, properly designed IUIs offer enormous potential to improve healthcare experiences and results for all parties involved, but they must be adopted with extreme caution in terms of privacy and clinical appropriateness.

3.2 The evolution of ASR technology and its impact on the usability of IUIs

Speech can be considered to be a more natural form of human contact as opposed to text-based communication. Over the past few years, Automatic Speech Recognition (ASR) technology has advanced significantly, revolutionizing the usability of IUIs [15]. Voice-activated systems such as Siri, Alexa, and Google Assistant have swiftly gained popularity by allowing users to communicate with gadgets and applications using voice recognition. Moreover, this significant leap can be attributed to the rise of deep learning and the creation of large datasets, resulting in considerable advances in speech recognition accuracy [15]. However, bias concerns persist, as demonstrated by a case study of gender bias in Google's speech recognition algorithm [16]. Researchers discovered that women's word error rates were greater than men's, with the system more likely to incorrectly transcribe women's speech [17]. The reason for this bias lies in the fact that women are more likely to be marginalized in datasets or corpora. To ensure justice and inclusion, ASR systems should be trained and tested on varied speech data that equally reflect men and women, and participatory design techniques that directly engage women users to find pain points should be used. Designers should check for disparities in mistake rates across genders. As speech becomes a more common input modality, it will be vital to ensure that ASR operates equally effectively for people of both genders. This necessitates a greater understanding of how design decisions affect inclusion as well as a dedication to maximizing accessibility and ease of use for users of all genders.

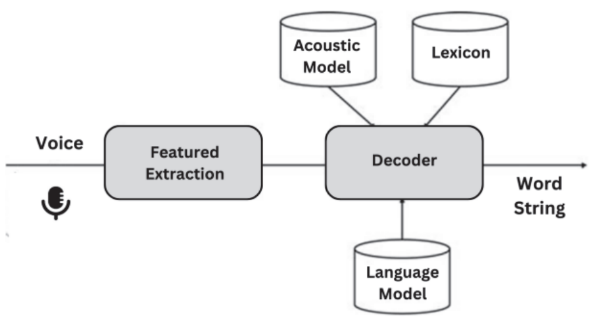

Figure 3.1 Working Speech Recognition Model. (Source: [15])

4 Week 8 Task - Multimodal Interfaces

4.1 Facilitating Multiparty Dialog with Gaze, Gesture, and Speech

The study describes a computer model for coordinating turn-taking in multiparty talks using an embodied conversational agent. The main issues addressed include observing conversational dynamics in real time, determining appropriate turn-taking judgements, and conveying those decisions through synchronised gaze, gesture, and voice behaviours to control the flow of discussion.

The solution consists of representations and algorithms for sensing the speech source and targets, inferring floor status and intentions, determining when to contribute, and creating verbal and nonverbal behaviours to convey turn-taking intents. The results show that by manipulating these cues, the agent may successfully participate in and influence multiparty interactions. In two-party interactions, for example, the addressee being glanced at answers 86 percent of the time.

These ideas might be used to improve the seamless conversational interactions between people and intelligent technologies in real-world contexts such as schools. The models can enable smart boards to track students' willingness to contribute and coordinate taking turns. Alternatively, intelligent tutoring systems might use these behaviours to increase student interest and involvement. Aside from education, corporations might use similar embodied AI interfaces to engage consumers or deliver training while avoiding inconvenient overlaps. As technology advances, any organisation that wishes to promote inclusive conversation may furnish facilities with intelligently regulated multimodal smart boards or virtual assistants. Their ability to discern underlying dynamics inside groups and gently mould interaction patterns indicates a paradigm shift in how we create venues to foster human relationships and egalitarian principles on a large scale.

The capacity of the conversational agent to orchestrate smooth encounters might also help older people, particularly those with deteriorating cognitive or physical capacities. The system's ability to manage turn-taking gracefully while modifying behaviors to effectively indicate intentions may make digital assistants more accessible for senior citizens. Those with visual or hearing impairments might use the synchronised gaze signals and voice changes to enhance genuine discourse. Tailored embodiments that utilise personal context might stimulate positive interaction in elderly people suffering from dementia or memory difficulties. As our population ages, building AI interfaces that accommodate aging users' specialised demands will be critical for inclusion.

However, there are still challenges in effectively perceiving loud multiparty situations and dealing with interruptions or simultaneous speech. These issues might be addressed by advances in speaker diarization and sound source localization, as well as reinforcement learning to optimise turn-taking rules depending on dialogue context.

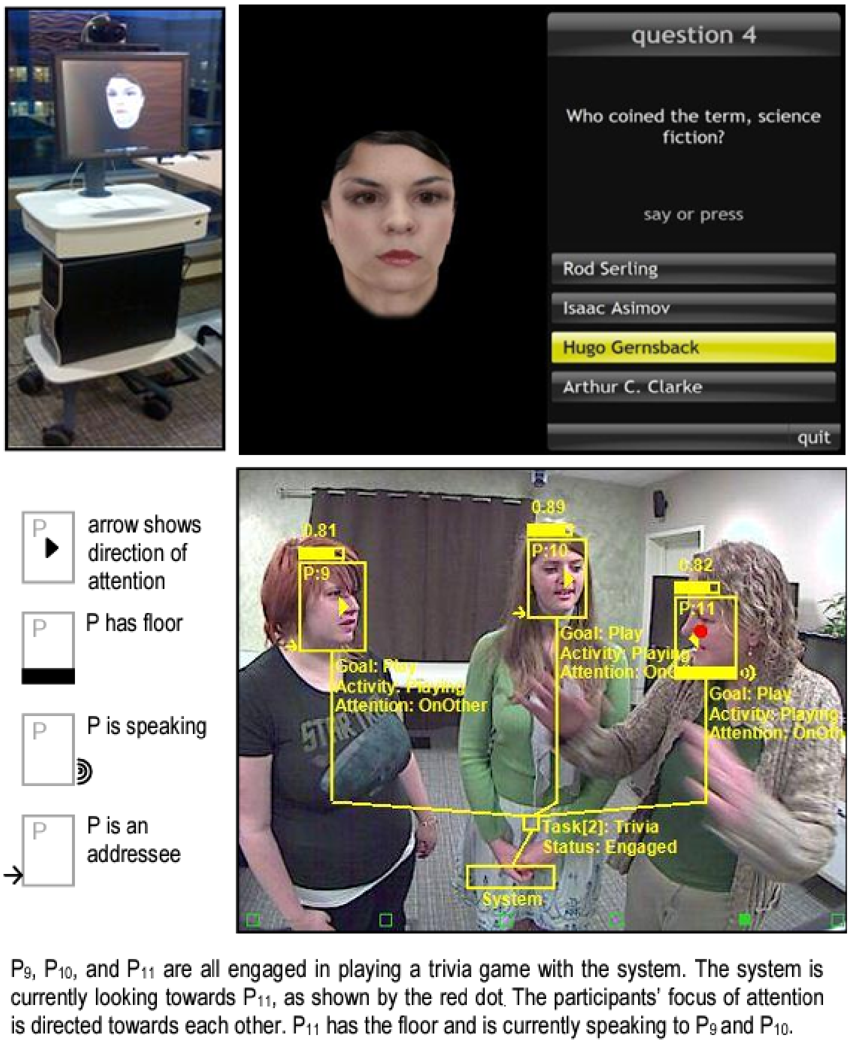

Figure 4.1 System with real-time scene analysis and the question game application operating on it. (Source: [18])

Looking ahead, the given architecture provides a basis for building fluid conversational skills in intelligent assistants, cars, household robots, or virtual reality avatars. The capacity to engage elegantly in sophisticated, coordinated human discussions has the potential to alter how humans interact with technology. As artificial intelligence affects more elements of our professional, private, and social lives, providing pleasant user experiences through natural interface behaviours will become increasingly important. Research on this crucial frontier is still in its early stages.

This field of study has the potential to have far-reaching societal implications. We can favourably affect emerging communication abilities by defining how future generations interact with AI through multimodal technologies. In general, building these technologies with human wants and behaviours in mind will be critical to establishing symbiotic interactions between man and machine.

[18]

References

[1] D. Gentner and A. L. Stevens, *Mental Models*, 1st. Psychology Press, 1983. DOI: 10.4324/9781315802725.

[2] J. A. Luft, S. Jeong, R. Idsardi, and G. Gardner, "Literature reviews, theoretical frameworks, and conceptual frameworks: An introduction for new biology education researchers," *CBE—Life Sciences Education*, vol. 21, no. 3, rm33, 2022, PMID: 35759629. DOI: 10.1187/cbe.21-05-0134. eprint: https://doi.org/10.1187/cbe.21-05-0134. [Online]. Available: https://doi.org/10.1187/cbe.21-05-0134.

[3] E. B. Hekler, P. Klasnja, J. E. Froehlich, and M. P. Buman, "Mind the theoretical gap: Interpreting, using, and developing behavioral theory in hci research," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13, Paris, France: Association for Computing Machinery, 2013, pp. 3307–3316, ISBN: 9781450318990. DOI: 10.1145/2470654.2466452. [Online]. Available: https://doi.org/10.1145/2470654.2466452.

[4] L. Westbrook, "Mental models: A theoretical overview and preliminary study," *J. Information Science*, vol. 32, pp. 563–579, Dec. 2006. DOI: 10.1177/0165551506068134.

[5] Facebook Inc., *Facebook*, Feb. 29, 2020. [Online]. Available: https://www.facebook.com/.

[6] D. Parra, P. Brusilovsky, and C. Trattner, "See what you want to see: Visual user-driven approach for hybrid recommendation," Feb. 2014. DOI: 10.1145/2557500.2557542.

[7] H.-C. Jetter, "A cognitive perspective on gestures, manipulations, and space in future multi-device interaction," Apr. 2014.

[8] M. Jahn and A. Frank, "How to increase usability of spatial data by finding a link between user and data," Jan. 2004.

[9] F. Geyer, T. Memmel, and H. Reiterer, "Interactive user interface specifications : Supporting collaboration and creativity with simulation-driven modeling tools," *IADIS International Journal on WWW/Internet*, vol. 7, no. 2, pp. 63–79, 2009, ISSN: 1645-7641.

[10] R. Sumarlin, "The review of user experience and user interface design of hospital information system to improve health care service," Jan. 2018. DOI: 10.2991/icobest-18.2018.39.

[11] L. Wilson and M. Marasoiu, "The development and use of chatbots in public health: Scoping review," *JMIR Hum Factors*, vol. 9, no. 4, e35882, Oct. 2022, ISSN: 2292-9495. DOI: 10.2196/35882. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/36197708.

[12] M. Eckert, J. Volmerg, and C. Friedrich, *Systematic review of augmented reality in healthcare (preprint)*, May 2018. DOI: 10.2196/preprints.10967.

[13] S. Amer, "Ethical concerns regarding the use of intelligent user interfaces," *Journal of Human-Computer Interaction Ethics*, 2023.

[14] Centers for Medicare & Medicaid Services, *The Health Insurance Portability and Accountability Act of 1996 (HIPAA)*, Online at http://www.cms.hhs.gov/hipaa/, 1996.

[15] A. Thakur, L. Ahuja, R. Vashisth, and R. Simon, "Nlp & ai speech recognition: An analytical review," in *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2023, pp. 1390–1396.

[16] J. Bajorek, "Voice recognition still has significant race and gender biases," *Harvard business review*, May 2019.

[17] M. Garnerin, S. Rossato, and L. Besacier, "Investigating the impact of gender representation in ASR training data: A case study on librispeech," in *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing*, M. Costa-jussa, H. Gonen, C. Hardmeier, and K. Webster, Eds., Online: Association for Computational Linguistics, Aug. 2021, pp. 86–92. DOI: 10.18653/v1/2021.gebnlp-1.10. [Online]. Available: https://aclanthology.org/2021.gebnlp-1.10.

[18] D. Bohus and E. Horvitz, "Facilitating multiparty dialog with gaze, gesture, and speech," in *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, ser. ICMI-MLMI '10, Beijing, China: Association for Computing Machinery, 2010, ISBN: 9781450304146. DOI: 10.1145/1891903.1891910. [Online]. Available: https://doi.org/10.1145/1891903.1891910.

Plagiarism Form

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Declaration

Plagiarism is defined as “the unacknowledged use, as one’s own work, of work of another person, whether or not such work has been published” (Regulations Governing Conduct at Examinations, 1997, Regulation 1 (viii), University of Malta).

I /~~We~~* , the undersigned, declare that the [assignment / Assigned Practical Task report / Final Year Project report] submitted is my /~~our~~* work, except where acknowledged and referenced.

I /~~We~~* understand that the penalties for making a false declaration may include, but are not limited to, loss of marks; cancellation of examination results; enforced suspension of studies; or expulsion from the degree programme.

Work submitted without this signed declaration will not be corrected, and will be given zero marks.

* Delete as appropriate.

(N.B. If the assignment is meant to be submitted anonymously, please sign this form and submit it to the Departmental Officer separately from the assignment).

StudentName	Signature
StudentName	Signature
StudentName	Signature
Matthias Bartolo	
StudentName	Signature

ICS3211	Intelligent Interfaces Weekly Tasks
CourseCode	Title of work submitted

December 2023
Date