

REPORT 4

Foundations of Information Systems – Speculative Design

[ANONYMIZED REPORT – FULL IPSIS LITTERIS TRANSLATION]

Abstract

The growing population and increased life expectancy have led to a higher demand for healthcare services, surpassing the supply of healthcare professionals. This situation results in dissatisfaction among patients and healthcare professionals due to inadequate digitalization of health services and the burdensome documentation process, which can lead to professional burnout. This study investigates the application of real-time automatic speech recognition (ASR) systems for medical consultations, using advanced Natural Language Processing (NLP) and Artificial Neural Networks (ANN) technologies. A speculative design approach is employed to imagine future scenarios in which virtual health assistants improve service accessibility and quality. A hybrid telemedicine platform that combines AI efficiency with human expertise is proposed to mitigate potential negative implications. This study highlights the importance of ethical, inclusive, and effective implementation of ASR technologies in healthcare, suggesting further research in real-world contexts and stakeholder engagement to refine the proposed solutions.

1. Introduction

1.1 Identification of the Problem

As a result of factors such as population growth and increased life expectancy—which consequently lead to an observed increase in comorbidities [Malta, 2021]—but without a proportional increase in healthcare professionals to meet this greater demand for health services, healthcare systems are unable to meet the population's needs, either in quantity or in quality.

A comparison of results from the National Health Survey (PNS) conducted in 2013 and 2019 revealed a 22% increase in demand for healthcare services in the 15 days prior to the survey,

accompanied by an 11% reduction in access to services, a scenario repeated in most [Omitted for blind review] states and capitals [Silva et al., 2021].

Although the use of medical scribes enables physicians to see more patients per hour, dissatisfaction remains among patients, who consider health services to be poorly digitized and express concern about the quality of physician interaction [Fuerst, 2017].

[Omitted for blind review] guarantees universal and free access to healthcare through the Federal Constitution and the [Omitted for blind review]. However, the digitalization of healthcare has produced mixed effects. While technology optimizes certain processes, it may also contribute to professional burnout, particularly due to the use of electronic medical records [Hodgson & Coiera, 2016].

Patients still prefer face-to-face communication, valuing direct human interaction. Human speech plays a fundamental role in expressing thoughts, ideas, and emotions. Therefore, it is hypothesized that a system capable of recording, transcribing, and automatically populating medical records during consultations could improve both professional working conditions and patient satisfaction.

1.2 Challenges

Automatic Speech Recognition (ASR) faces several challenges, including:

- Scarcity of annotated data;
- Linguistic variability due to regional accents and slang;
- Spontaneous speech phenomena, such as hesitations and pauses;
- Phonetic complexity;
- High costs associated with high-quality speech data collection.

In healthcare, additional challenges include interoperability between systems, ethical issues related to patient privacy, and accessibility across different socioeconomic contexts.

1.3 Risks

The use of ASR technology in healthcare involves risks related to bias. If training datasets are not sufficiently diverse—considering age, gender, regional accents, and speech impairments—system performance may be uneven, leading to discriminatory outcomes.

1.4 Research Proposal

Due to the complexity involved in integrating multiple systems, this work focuses specifically on automatic speech recognition and transcription in Brazilian Portuguese, rather than on full system integration.

1.5 Structure of the Work

This work is structured as follows:

- Presentation of reference books and articles;
- Definition of concepts necessary to understand the technology;
- Overview of the state of the art in ASR;
- Introduction to speculative design;
- Mapping of the current scenario (*Where are we?*);
- Exploration of possible futures (*Where are we going?*);
- Projection of a desirable future (*Where do we want to go?*);
- Conclusion.

2. Presentation of the Main References

2.1 Book

Selected chapters from *Natural Language Processing: Concepts, Techniques, and Applications in Portuguese* were used, addressing core NLP concepts, speech processing, resources, and healthcare applications.

2.2 Articles

Nine articles were selected covering automatic speech recognition, healthcare speech technologies, and artificial intelligence applications in medicine. Summaries were translated using the DeepL translation tool.

3. Automatic Speech Recognition

3.1 Natural Language Processing

Natural Language Processing (NLP) is the field responsible for computational processing of human language in written, spoken, or signed form. Speech processing differs from text processing due to its acoustic, temporal, and contextual characteristics.

3.2 Concept of Automatic Speech Recognition

ASR systems convert spoken language into text by capturing audio signals, extracting features, applying acoustic and phonetic models, and using linguistic models to generate textual output.

3.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) are currently the most effective approach for ASR tasks. Architectures include DNNs, CNNs, RNNs, LSTMs, and Transformers.

3.4 State of the Art

There is no single consolidated ASR architecture, but deep neural network approaches dominate. Two notable open-source models are:

- **wav2vec 2.0**, based on convolutional neural networks;
- **Whisper**, based on Transformer architectures and trained on multilingual datasets.

4. Speculative Design

Speculative design, introduced by Dunne and Raby, creates artifacts to reflect on possible futures. It does not aim to predict the future, but to provoke discussion, reduce uncertainty, and explore preferable futures.

5. Mapping the Current Scenario

Where Are We?

During medical consultations, professionals must manually fill out clinical documentation, either digitally or on paper. This process interrupts communication, generating silences and disrupting patients' flow of thought.

Identified actors include:

1. Patients
2. Healthcare professionals
3. ASR technologies
4. Data
5. Healthcare institutions
6. Regulation and ethics
7. Society

Trend analysis (e.g., Google Trends) indicates growing interest in speech transcription technologies, particularly after the rise of generative AI.

6. Exploring Possible Futures

Where Are We Going?

Using speculative tools, future scenarios for 2030 were envisioned. These include widespread ASR adoption, virtual health assistants capable of real-time transcription and emotional analysis, and ethical dilemmas related to privacy, consent, and bias.

7. Projecting a Desirable Future

Where Do We Want to Go?

A hybrid telemedicine platform is proposed in which:

- AI performs transcription and preliminary analysis;
- Human professionals validate diagnoses;
- Complex cases are escalated to direct human care;
- Continuous professional training is emphasized.

This approach seeks to balance efficiency and empathy.

8. Conclusion

This work explored the application of real-time automatic speech transcription for healthcare consultations, integrating NLP, ANN technologies, and speculative design. While promising, ethical, cultural, and institutional challenges remain. Future work should involve real-world testing and broader stakeholder participation.

Future work should involve field studies in simulated and real clinical environments, broader stakeholder engagement, and regional evaluations to ensure ethical, inclusive, and effective implementation.

References

(Some of the references were suppressed for blind review purposes.)

BANK, Alan J.; GAGE, Ryan M. Annual impact of scribes on physician productivity and revenue in a cardiology clinic. **ClinicoEconomics and outcomes research**, p. 489-495, 2015. Doi: <https://doi.org/10.2147/ceor.s89329>

CASELI, Helena de Medeiros; NUNES, Maria das Graças Volpe. Processamento de linguagem natural: conceitos, técnicas e aplicações em português. 2024. Disponível em https://brasileiraspln.com/livro-pln/2a-edicao/Livro_PLN_BPLN_2ed.pdf

DHANJAL, Amandeep Singh; SINGH, Williamjeet. A comprehensive survey on automatic speech recognition using neural networks. **Multimedia Tools and Applications**, v. 83, n. 8, p. 23367-23412, 2024. DOI: <https://doi.org/10.1007/s11042-023-16438-y>

ESPÍNDOLA, Fábio Pereira. A teoria do Affordance no design especulativo e transumanismo do filme A Vigilante Do Amanhã: Ghost In The Shell. **Revista de Design, Tecnologia e Sociedade**, v. 11, n. 1, 2024. Disponível em <https://periodicos.unb.br/index.php/design-tecnologia-sociedade/article/view/51626>

FUERST, Mark L. Patients prefer face-to-face communications with doctors. 2017. DOI: <https://doi.org/10.1097/01.COT.0000527379.44018.e1>

GRIS, Lucas Rafael Stefanel et al. Desenvolvimento de um modelo de reconhecimento de voz para o Português Brasileiro com poucos dados utilizando o Wav2vec 2.0. In: **Anais do XV Brazilian e-Science Workshop**. SBC, 2021. p. 129-136. DOI: <https://doi.org/10.5753/bresci.2021.15798>

HODGSON, Tobias; COIERA, Enrico. Risks and benefits of speech recognition for clinical documentation: a systematic review. **Journal of the american medical informatics association**, v. 23, n. e1, p. e169-e179, 2016. DOI: <https://doi.org/10.1093/jamia/ocv152>

KUMAR, Vijay; SINGH, Hemant; MOHANTY, Animesh. Real-Time Speech-To-Text/Text-To-Speech Converter with Automatic Text Summarizer Using Natural Language Generation and Abstract Meaning Representation. **International Journal of Engineering and Advanced Technology (IJEAT)**, v. 9, p. 2361-2365, 2020. Disponível em <https://ssrn.com/abstract=3925035>

LATIF, Siddique et al. Speech technology for healthcare: Opportunities, challenges, and state of the art. **IEEE Reviews in Biomedical Engineering**, v. 14, p. 342-356, 2020. DOI: <https://doi.org/10.1109/rbme.2020.3006860>

LI, Calandra et al. Electronic medical record-related burnout in healthcare providers: a scoping review of outcomes and interventions. **BMJ open**, v. 12, n. 8, p. e060865, 2022. DOI: <https://doi.org/10.1136/bmjopen-2022-060865>

MALTA, Deborah Carvalho et al. Analysis of demand and access to services in the last two weeks previous to the National Health Survey 2013 and 2019. **Revista Brasileira de Epidemiologia**, v. 24, p. e210002, 2021. DOI: <https://doi.org/10.1590/1980-549720210002.supl.2>

RADFORD, Alec et al. Robust speech recognition via large-scale weak supervision. In: **International conference on machine learning**. PMLR, 2023. p. 28492-28518. DOI <https://doi.org/10.48550/arXiv.2212.04356>

SANJEEV, Santosh; PONNEKANTI, Gowtham Sai; REDDY, G. Pradeep. Advanced healthcare system using artificial intelligence. In: **2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)**. IEEE, 2021. p. 76-81. DOI: <http://dx.doi.org/10.1109/Confluence51648.2021.9377084>

SCHNEIDER, Steffen et al. wav2vec: Unsupervised pre-training for speech recognition. **arXiv preprint arXiv:1904.05862**, 2019. DOI <http://dx.doi.org/10.21437/Interspeech.2019-187>

SILVA, Lívia Silveira et al. Demanda espontânea e acesso no Sistema Único de Saúde: vivências de usuários da atenção primária. **Avances en Enfermería**, v. 39, n. 1, p. 30-39, 2021. Disponível em <https://fi-admin.bvsalud.org/document/view/mqye8>

DE TOLEDO, Thiago Ferreira et al. Web System Prototype based on speech recognition to construct medical reports in Brazilian Portuguese. **International journal of medical informatics**, v. 121, p. 39-52, 2019. DOI: <https://doi.org/10.1016/j.ijmedinf.2018.10.010>

ZAMITH, Isabela et al. Design Especulativo como ferramenta de reflexão sobre futuros possíveis. **Blucher Design Proceedings**, v. 10, n. 5, p. 4994-5020, 2022.