Politecnico di Torino

e-Lite

DAUIN

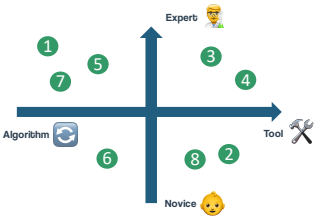
Dottorato di Interesse Nazionale in Intelligenza Artificiale – Industria – (XXXVII cycle)

Ph.D. Candidate: Tommaso Calò
Supervisors: Luigi De Russis, Fulvio Corno

Personal Website | e-Lite Lab @ PoliTO | National PhD in AI | Politecnico di Torino

From Human Representations to AI Realization: Algorithms and Tools for Creating and Refining Interactive Systems

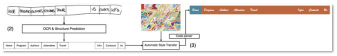
Motivations



Humans convey ideas through rich, natural representations such as **text**, **sketches**, **illustrations** and **visual cues**. Recent AI systems have the ability to recognize the complex nuances of human intent and map them to procedural artifacts, such as code or domain-specific languages. In this thesis, we explore how these human representations can be leveraged to create and refine interactive systems along two fundamental dimensions: tools and algorithms (from novel AI methods to practical implementations) and user expertise (from novice to expert users). Through crontributions spanning both dimensions, we investigate how these capabilities can benefit users across various domains, including **UI design**, **education**, and **smart home automation**.

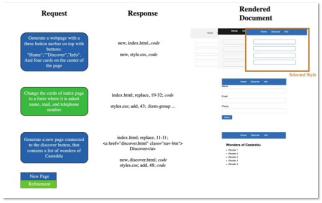
Sketches and Text to Web Uis

1



We introduce a style-aware sketch-to-code conversion method for web interfaces that not only translates hand-drawn sketches into functional HTML/CSS code, but also automatically applies visual styles from reference images using machine learning techniques for color extraction and font matching.

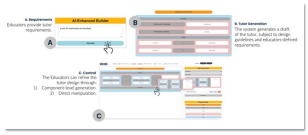
2



We introduce a natural language-based website generator leveraging Large Language Models that enables non-technical users to create and iteratively refine complete websites through simple text descriptions, making web development more accessible while maintaining user control over the final output.

Pedagogical Goals to Tutors' UIs

3



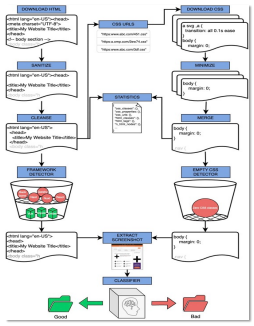
We introduce an approach for democratizing Intelligent Tutoring System creation by leveraging Generative AI to help educators design tutor interfaces without specialized programming skills. The system translates educators' high-level requirements into interface designs through prompt engineering and DSL, demonstrating time reductions compared to traditional drag-and-drop methods in preliminary tests.

4

Building on previous work, we present a an AI-assisted system for ITS interface design that introduces pedagogical step decomposition and preference-driven UI refinement. Through a user study with 8 educators, we demonstrate that the system achieved 33% higher interface quality while maintaining comparable design time to traditional methods.

Design2Code model and dataset advancement

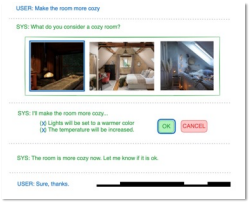
5



This research advances code generation from visual web designs by introducing: 1) WebUI2Code - a specialized pipeline and dataset that processes real-world website code and screenshots to produce cleaner training data, 2) a synthetic dataset of Bootstrap websites and their sketched variations that better captures modern web component diversity, and 3) an evaluation demonstrating that transformer-based architectures outperform traditional RNN models on code generation benchmarks, while identifying challenges in handling real-world website complexity.

Multimodal Disambiguation in Smart Homes

6



In this work we introduce a system that combines LLMs with visual and textual cues to resolve ambiguous user commands in smart home environments. The approach advances smart home interaction by providing a more intuitive way to handle ambiguous requests like "make the room cozier".

Representation Learning of UIs

7

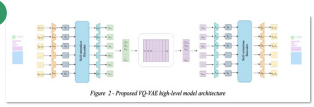
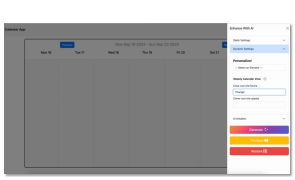


Figure 2 - Proposed VQ-VAE high-level model architecture

This research introduces a novel architecture for learning meaningful representations of graphical user interfaces, combining a Transformer-based VQ-VAE model with a Figma-compatible dataset to bridge the gap between AI research and real-world design tools.

Integrating GenAI into Dynamic UIs

8



We introduce a framework which enables applications to leverage Large Language Models for real-time interface personalization through natural language interactions, addressing the limitations of traditional static customization approaches.

Calò, T., & De Russis, L. (2022). Style-Aware Sketch-to-Code Conversion for the Web. In Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '22 Companion) (pp. 44-47). Association for Computing Machinery. https://doi.org/10.1145/3531706.3536462
Calò, T., & De Russis, L. (2023). Leveraging Large Language Models for End-User Website Generation. In End-User Development: 9th International Symposium, IS-EUD 2023 (pp. 52-61). Springer. https://doi.org/10.1007/978-3-031-34433-6_4
Calò, T., & De Russis, L. (2024). Enhancing smart home interaction through multimodal command disambiguation. Personal and Ubiquitous Computing, 1-16. https://doi.org/10.1007/s00779-024-01827-3
Calò, T., & Maclellan, C. (2024). Towards Educator-Driven Tutor Authoring: Generative AI Approaches for Creating Intelligent Tutor Interfaces. In Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S '24) (pp. 305-309). Association for Computing Machinery. https://doi.org/10.1145/3657604.3664694