

# European Research and Innovation

## Large Language Models as Design Partners: Automating Graphical Mockups to Refine Requirements

by Giovanna Broccia, Maurice H. ter Beek (CNR-ISTI), and Alessio Ferrari (University College Dublin and CNR-ISTI)

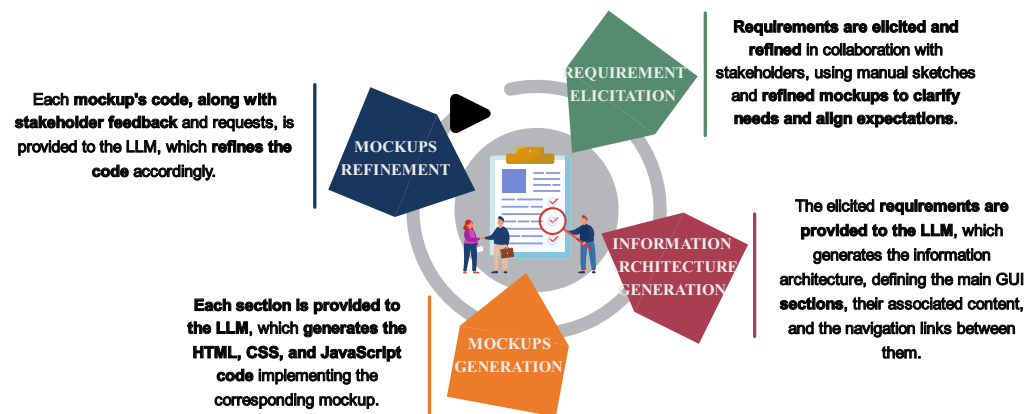
*Can large language models help designers move faster without sacrificing human centrality? Researchers from CNR-ISTI and University College Dublin are exploring how large language models can support the rapid creation and refinement of industrial graphical user interfaces through a case study involving the Italian railway operator Trenord, helping development teams move more quickly from textual requirements to interactive mockups, while keeping humans at the centre of the design process.*

Creating graphical user interface (GUI) mockups is a crucial step in software design. Mockups help designers, engineers, and end-users discuss requirements, validate ideas, and refine functionalities before implementation begins. However, the process is often time-consuming, requiring multiple design iterations and repeated stakeholder reviews. It is also technically demanding, involving manual coding and debugging activities, especially in industrial contexts where interfaces evolve continuously and many stakeholders are involved [1].

Researchers from CNR-ISTI and University College Dublin collaborated with Trenord [L1], a railway operator in Northern Italy, to investigate how large language models (LLMs) can support GUI design activities by acting as “design partners”. The work was done within Spoke 4 “Railway Transportation” of MOST – the National Centre for Sustainable Mobility [L2], which received EU funding through NextGenerationEU – and focused on the design of a predictive maintenance dashboard for railway operations. The dashboard is intended to support maintenance and engineering personnel in monitoring train fleets, analysing diagnostic information, and inspecting failures predicted by maintenance algorithms. The goal was not to replace designers, but to understand whether LLMs could automate repetitive coding activities and accelerate the transition from requirements to interactive mockups [2].

The proposed approach adopted a human-LLM co-design process consisting of four main activities (see Figure 1). In the first phase, requirements were elicited collaboratively through focus groups, analysis of existing documentation, and discussions with Trenord stakeholders. During this stage, designers also created manual sketches to clarify stakeholders’ needs and align expectations. In the second phase, the elicited requirements were provided to the LLM, which transformed textual requirements into a structured information architecture by identifying dashboard sections and navigation paths. In the third phase, the LLM generated interactive mockups in HTML, CSS, and JavaScript for each identified dashboard section. Finally, the generated mockups were refined iteratively according to

**Figure 1: Human-LLM co-design process** illustrating how requirements elicitation, information architecture generation, mockup generation, and iterative refinement are collaboratively performed by human designers, stakeholders, and LLMs to accelerate the creation of interactive GUI mockups.



stakeholder feedback. The code of each mockup, together with requests and comments from stakeholders, was provided to the LLM, which updated and refined the interfaces accordingly.

The resulting dashboard included six mockups representing different sections of the dashboard, including fleet overview pages, train configuration views, and detailed prediction panels showing the status of onboard components. Rather than static images, the generated outputs were interactive interfaces that stakeholders could immediately visualise, interact with, and discuss. This rapid feedback cycle enabled participants to identify missing requirements and propose changes early in the design process.

One of the main outcomes of the project was the observation that LLMs can significantly reduce the effort spent on repetitive technical activities, such as coding layouts and refining HTML/CSS structures. This allowed designers to focus more on stakeholder interaction, requirement clarification, and design reasoning. Importantly, the proposed methodology accelerates clerical activities such as programming and debugging while keeping humans central throughout the entire process – from requirements elicitation and refinement, via information architecture definition, to the design and evaluation of the mockups themselves. The most important contribution of the approach is therefore not full automation, but the possibility of supporting collaboration and accelerating iteration cycles.

The generated mockups were later presented and discussed in a focus group involving three members of Trenord's engineering personnel. During the session, the interfaces were shown as an operational dashboard supporting two classes of users: maintenance personnel and engineering personnel. Stakeholders were asked to provide feedback both on the generated sections and on the overall process.

The participants described the generated interfaces as more "concrete" and closer to the final product than manually created sketches. Because the interfaces could be modified rapidly, discussions became more dynamic and productive. Stakeholders were able to request immediate changes, validate functionalities earlier, and better understand how the final system could evolve. Rather than constraining creativity through predefined solutions, the approach appeared to support continuous exploration and refinement of ideas.

The experience also highlighted several practical lessons. First, requirements quality strongly influences the quality of the generated interfaces: ambiguous or incomplete requirements frequently lead to unsatisfactory outputs. Second, human involvement remains essential throughout the process: the LLM can generate and refine interfaces quickly, but humans are still responsible for defining goals, validating outputs, and guiding the design direction.

Another important lesson concerns the choice of the LLM itself. Different models showed different strengths in terms of visual quality, consistency, and adherence to requirements. We also observed that prompting style plays a key role: structured and sequential prompts significantly improved the coherence of generated interfaces and reduced the number of refinement iterations needed.

Beyond the railway domain, we believe that this approach could support many industrial design activities where rapid prototyping and continuous stakeholder feedback are essential. Future work will involve analysing how LLMs could generate multiple interface alternatives automatically and how mockups themselves could help refine requirements in future design cycles.

#### Links:

[L1] <https://www.trenord.it/en/>

[L2] <https://www.centronazionalemost.it/en/>

#### References:

- [1] T. R. Silva et al., "A Comparative Study of Milestones for Featuring GUI Prototyping Tools", *J. Softw. Eng. Appl.*, vol. 10, no. 6, 2017, 564–589. DOI: 10.4236/jsea.2017.106031
- [2] G. Broccia et al., "An experience report on leveraging LLMs for GUI generation: Automating coding to prioritise creativity," in *Joint Proc. of REFSQ 2025 Co-Located Events, CEUR Workshop Proc.*, vol. 3959, 2025. <https://ceur-ws.org/Vol-3959/CreaRE-paper2.pdf>

#### Please contact:

Giovanna Broccia  
CNR-ISTI, Italy  
[giovanna.broccia@isti.cnr.it](mailto:giovanna.broccia@isti.cnr.it)