

Can AI Build a Usable Website?

Evaluating AI's Ability to Create a Usable Website Prototype

Research Plan



By Theresa Wilkinson

Study Objectives

The objective of this study is to evaluate AI website builders' ability to create a complete, working home-remodeling website for use as a usability-testing prototype.

Specifically, the study examines each builder's ability to:

- Follow the supplied prompt and requirements.
- Generate realistic fictional content, images, and design.
- Create working navigation, links, and interactions.
- Meet the requested WCAG 2.1 accessibility requirements.
- Support participants in completing usability-testing tasks.
- Produce a practical prototype requiring minimal editing.

Study Design

AI website builders will be evaluated using a standardized prompt requesting a complete home-remodeling website suitable for usability testing. Each builder independently generates the prototype.

Outputs will be compared using a predefined evaluation framework that assesses completeness, accuracy, quality, practicality, and the editing required. Build attempts, problems, and researcher modifications will be documented.

Selected working prototypes will undergo unmoderated usability testing using the same tasks and follow-up questions. Participant behavior and feedback provide evidence of how well the prototypes support the intended tasks.

Study Approach

- Request a website prototype for the fictional company Leonard Remodel & Build in Columbus, Ohio.
- Ask the AI to create the design, content, photos, and interactions.
- Provide navigation and tasks covering services, past projects, reputation, contact, process, and company background.
- Document generated outputs, access problems, design observations, and troubleshooting.

Study Prompt

The first prompt given to the AI website builders was very general.

Prompt: Build a complete clickable home remodeling website prototype for usability testing. Leonard Remodel & Build, Columbus OH. Include working navigation, clickable pages, buttons, and contact form (no real submissions). Include Services, Process, Portfolio, About Us, and Contact with realistic content, photos, reviews, team information, and responsive design.

Evaluation Framework

Each AI-generated website is evaluated using the same five criteria:

- Completeness: Includes the requested pages, content, images, and features.
- Accuracy: Follows the supplied company details and requirements; content is internally consistent.
- Quality: Provides clear navigation, readable content, appropriate images, and accessible interactions.
- Practicality: Works as a prototype that participants can access and use to complete the testing tasks.
- Editing Required: Amount and type of researcher intervention needed before testing.

Researcher checks and usability findings support the evaluations. Changes made after generation are documented separately from the original output.

Research Questions

The study seeks to answer the following questions:

- Can AI website builders create a complete, working usability-testing prototype from the supplied prompt?
- Can they generate realistic fictional content and appropriate images?
- How well do they follow navigation, interaction, and accessibility requirements?
- Can participants find the information needed to complete the usability tasks?
- How much editing or technical setup is required before testing?
- Which builder produces the most complete, usable, and practical prototype?
- Which AI model produces the highest-quality analytical output?

Analysis Materials

The evaluation uses the following materials:

- Standardized website-building prompt and any follow-up prompts.
- Original AI-generated websites and available code or downloads.
- Records of build attempts, publishing problems, and required setup.
- Researcher checks the functionality and accessibility of the prototype.

- Records of prototype changes made before testing.
- Unmoderated usability-testing results and participant responses.
- Evaluation criteria: Completeness, Accuracy, Quality, Practicality, and Editing Required.

Analysis Tasks

The researcher will:

1. Review each builder's output against the supplied prompt.
2. Evaluate completeness, accuracy, quality, practicality, and editing required.
3. Document build problems, setup requirements, and prototype changes.
4. Analyze usability results for the tested prototypes.
5. Compare findings across builders, distinguishing original output from researcher modifications.
6. Develop evidence-based conclusions and recommendations.
7. Discuss implications for UX researchers and identify study limitations.

Limitations

1. This study evaluates selected AI website builders using one home-remodeling website prompt. Results reflect the tools, versions, and account access available at the time of testing.
2. Usability findings apply only to the prototypes tested and participants recruited. Researcher modifications may affect results and will be documented.
3. Accessibility checks do not establish full WCAG 2.1 compliance. Publishing or access problems may arise from the builder, hosting service, or local internet connection.

4. Conclusions are limited to the evidence collected and should not be generalized to all AI-generated websites.