Inclusivity Evaluation of LLM generated code

Mandatory questions are marked with an asterisk (*)

Survey Statement

We are conducting this survey as part of a study on evaluating a security mechanism (for inclusivity) generated using Large Language Models (LLMs). The following questions guide the research.

Effectiveness: How effectively can LLMs generate code that implements security and inclusivity requirements, particularly for neurodivergent users?

Influencing Factors: What factors (e.g., prompt design, model choice, and specification level) influence the quality of LLM-generated code with respect to security and inclusivity?

Evaluation: How does LLM-generated code perform when reviewed by humans and LLMs in terms of identifying security vulnerabilities and inclusivity gaps?

In this study, we used prompts with varying levels of detail regarding inclusivity considerations in the development of security mechanisms—from minimal to highly detailed specifications (three cases in total).

You are requested to evaluate the output from each case for inclusivity. The survey will take approximately 10 minutes to complete.

We sincerely appreciate your time and participation.

Participation consent *

☐ I agree to participate in the survey

Current role

Number of years of experience

Case I: No Inclusivity Specification

User Story
Helena is a 67 year patient with chronic heart disease. She calls the hospital to book an appointment to review her medication dosage. Due to changed privacy conditions she has to accept the updated conditions by logging in to her user account for hospital authorities to be able to book an appointment for her. The authentication to her personal healthcare account requires a strong password which she has forgot, therefore she needs to reset the password to be able access the account and accept the privacy statement.

Please access the application here: https://health-c1.scientificresearch.xyz/

Email: helena@example.test

Inclusivity Evaluation criteria

Scale (1-5)
1: Strongly disagree, 2: Disagree, 3: Neutral, 4: Agree, 5: Strongly Agree

Attention *

	1	2	3	4	5
User interface is clean and does not have moving images or pop-ups. *	○	○	○	○	○
Important items clearly stand out. *	○	○	○	○	○
It is easy for to find out on the screen what she needs quickly. *	○	○	○	○	○

Memory *

	1	2	3	4	5
The user does not need to remember many rules or steps. *	○	○	○	○	○
Options like Password manager, autofill, or “remember me” are present *	○	○	○	○	○

	1	2	3	4	5
The user can easily recognize saved or familiar items. *	○	○	○	○	○

Comprehension *

	1	2	3	4	5
Text is short, clear, and easy to read. *	○	○	○	○	○
Error messages describe the problem and the way to fix it. *	○	○	○	○	○
Examples show what to type or how to do a step. *	○	○	○	○	○

Decision Making *

	1	2	3	4	5
Choices (like password, code, or face login) are clear. *	○	○	○	○	○
The user is not provided with too many choices at one time. *	○	○	○	○	○
The system clearly informs the outcome of each action. *	○	○	○	○	○

Learning *

	1	2	3	4	5
The user gets required help or short tips for new steps. *	○	○	○	○	○
The user can easily try again if she makes a mistake. *	○	○	○	○	○
The design stays consistent so that the user can learn it over time. *	○	○	○	○	○

Case II: Moderate Inclusivity Specification

User Story

Helena is a 67 year patient with chronic heart disease and cognitive condition ADHD. She calls the hospital to book an appointment to review her medication dosage. Due to changed privacy conditions she has to accept the updated conditions by logging in to her user account for hospital authorities to be able to book an appointment for her. The authentication to her personal healthcare account requires a strong password which she has forgot, therefore she needs to reset the password to be able access the account and accept the privacy statement.

Please access the application here: https://health-c2.scientificresearch.xyz/

Email: helena@example.com OR helena67

Inclusivity Evaluation criteria

Scale (1-5)
1: Strongly disagree, 2: Disagree, 3: Neutral, 4: Agree, 5: Strongly Agree

Attention *

	1	2	3	4	5
User interface is clean and does not have moving images or pop-ups. *	○	○	○	○	○
Important items clearly stand out. *	○	○	○	○	○
It is easy for to find out on the screen what she needs quickly. *	○	○	○	○	○

Memory *

	1	2	3	4	5
The user does not need to remember many rules or steps. *	○	○	○	○	○
Options like Password manager, autofill, or “remember me” are present *	○	○	○	○	○
The user can easily recognize saved or familiar items. *	○	○	○	○	○

Comprehension *

	1	2	3	4	5
Text is short, clear, and easy to read. *	○	○	○	○	○
Error messages describe the problem and the way to fix it. *	○	○	○	○	○
Examples show what to type or how to do a step. *	○	○	○	○	○

Decision Making *

	1	2	3	4	5
Choices (like password, code, or face login) are clear. *	○	○	○	○	○
The user is not provided with too many choices at one time. *	○	○	○	○	○
The system clearly informs the outcome of each action. *	○	○	○	○	○

Learning *

	1	2	3	4	5
The user gets required help or short tips for new steps. *	○	○	○	○	○
The user can easily try again if she makes a mistake. *	○	○	○	○	○
The design stays consistent so that the user can learn it over time. *	○	○	○	○	○

Case III: Detailed Inclusivity Specification

User Story

Helena is a 67 year patient with chronic heart disease and cognitive condition ADHD. She calls the hospital to book an appointment to review her medication dosage. Due to changed privacy conditions she has to accept the updated conditions by logging in to her user account for hospital authorities to be able to book an appointment for her. The authentication to her personal

healthcare account requires a strong password which she has forgot, therefore she needs to reset the password to be able access the account and accept the privacy statement.

Inclusivity description: People with ADHD can lose track of multi-step tasks, get easily distracted, and feel overwhelmed by unclear instructions or time pressure. The login and password reset flow must be forgiving, clear, and low-stress. Avoid dense instructions, timeouts, or unexpected page changes. Support memory and focus by showing visible progress and offering reminders of the next step.

Additional requirements:
• Keep the process simple and structured with clear, step-by-step guidance.
• Use straightforward and consistent language without unnecessary information.
• Provide clear feedback at each step so the user knows what has happened and what to do next.
• Allow enough time for actions without rushing or session timeouts.
• Reduce visual and cognitive distractions in the interface.
• Ensure the user can pause and return to the task without losing progress.
• Make help options easy to find and access at any stage.

Please access the application here: https://health-c3.scientificresearch.xyz/

Email: helena.patient@examplehospital.org

Inclusivity Evaluation criteria

Scale (1-5)
1: Strongly disagree, 2: Disagree, 3: Neutral, 4: Agree, 5: Strongly Agree

Attention *

	1	2	3	4	5
User interface is clean and does not have moving images or pop-ups. *	○	○	○	○	○
Important items clearly stand out. *	○	○	○	○	○
It is easy for to find out on the screen what she needs quickly. *	○	○	○	○	○

Memory *

	1	2	3	4	5
The user does not need to remember many rules or steps. *	○	○	○	○	○
Options like Password manager, autofill, or “remember me” are present *	○	○	○	○	○
The user can easily recognize saved or familiar items. *	○	○	○	○	○

Comprehension *

	1	2	3	4	5
Text is short, clear, and easy to read. *	○	○	○	○	○
Error messages describe the problem and the way to fix it. *	○	○	○	○	○
Examples show what to type or how to do a step. *	○	○	○	○	○

Decision Making *

	1	2	3	4	5
Choices (like password, code, or face login) are clear. *	○	○	○	○	○
The user is not provided with too many choices at one time. *	○	○	○	○	○
The system clearly informs the outcome of each action. *	○	○	○	○	○

Learning *

	1	2	3	4	5
The user gets required help or short tips for new steps. *	○	○	○	○	○
The user can easily try again if she makes a mistake. *	○	○	○	○	○

	1	2	3	4	5
The design stays consistent so that the user can learn it over time. *	○	○	○	○	○