

Phase 1 Usability Findings

JackDaw GeoAI — UC1 & UC2 Session Analysis

9 participants · 3 evaluators · Mixed methods (observation + SUS + online form)

AURA · PoliRuralPlus · Horizon Europe Grant No. 101136910



Issues at a Glance

6 thematic clusters identified across all 9 usability sessions

Very Frequent

Map Interaction
& Area Selection

HIGH

Very Frequent

Chat Input &
Submit Mechanics

HIGH

Frequent

Confirm Dialog
& Loading States

HIGH

Very Frequent

Answer Quality
& Data Reliability

HIGH

Common

Language &
Localisation

MED

Occasional

Internal Process
Transparency

MED

Users Can't Select an Area

The most frequent and blocking usability issue across all sessions

HIGH

Zoom ≠ Selection

Every user initially assumed that zooming into a map area meant selecting it. The actual draw tool was discovered late, by accident, or not at all without assistance.

HIGH

Draw tool is invisible & hard to use

The draw button is not prominent. Once found, freehand drawing is difficult — some users made 20+ attempts before succeeding, and several needed facilitator help.

HIGH

No persistent selection feedback

After panning or zooming, users could not tell whether their drawn area was still selected. The map gave no persistent visual indication, causing repeated re-draws.

HIGH

Can't edit an existing area

Drawing a new area while one exists does not replace the old one. Users were unsure whether the old selection persisted and did not know how to update it.

Broken Conversation Flow

Users couldn't submit questions, couldn't find the confirm button, and misread loading states

Chat Input & Submit

HIGH Enter key doesn't submit

Nearly every user pressed Enter to send a message. When nothing happened they felt confused or thought the tool was broken.

HIGH Input field is hard to find / activate

Users clicked around the interface searching for where to type. Clicking the field often gave no visible active state feedback.

MED Input field jumps from top to bottom

As the conversation grew, the prompt field moved from top to bottom of the screen, confusing users who looked for it at the original position.

MED Placeholder text still shows 'Draw an area...'

After an area is already drawn, the hint text still prompts the user to draw — reinforcing false uncertainty about whether the selection persists.

Confirm Dialog & Loading States

HIGH Confirm button is off-screen

The SQL-confirm dialog appears below the visible viewport while the spinner continues — users assume the system is still processing, not waiting for their input. Appearing repeatedly (2–3×), in English only, it causes deep confusion and stalls sessions.

Users Don't Trust the Answers

Generic, incorrect or irrelevant responses eroded confidence across all use cases

HIGH

Consistently generic responses

JackDaw frequently gave answers that could apply to any European region. Users noticed they received no location-specific insight, making the tool feel no better than a basic web search.

HIGH

Tool tries to please rather than admit limits

When JackDaw lacked data it gave vague, hedging responses ('it may', 'possibly') or suggested generic alternatives — instead of clearly stating it cannot answer. Users found false confidence more damaging than an honest 'I don't know'.

HIGH

Factual errors (temperatures, dates, wrong region)

The tool stated incorrect temperatures, gave wrong weekday calculations, referred to Czech data for a query about Kilkis, and confused Veria with Vergia. Each error was immediately spotted.

HIGH

Says 'no data' for info users know exists

JackDaw denied having data on bus routes, local shops, and irrigation — all freely available. Map markers often contradicted text answers: 'I can see them on the map; why does it say there are none?'

Language Friction & Opaque Internals

Greek users in an English system, and unexplained technical noise

Language & Localisation

MED

Search in Greek, results in English

Typing Greek returns English/EU labels (e.g. 'Lagadas PPL Greece'). 'PPL' was universally misunderstood. Slight spelling variations ('Λαγκαδάς' vs 'Λαγκαδά') or Greeklish returned wrong results — users blamed themselves.

MED

Mixed-language interface

Chat replies came in Greek, but confirmations, errors and processing labels were English-only — jarring and inaccessible to non-English speakers.

LOW

Sample questions fail silently without area

Tapping a sample filled the chat field but submission failed if no area was drawn. Users didn't understand why nothing happened.

Internal Process Transparency

MED

Technical noise during processing

Messages like 'Loading dataset...' and 'Generating SQL plan...' appeared while waiting. Users didn't understand these labels — they added anxiety rather than reassurance. References to internal component names (e.g. 'raven') compounded this.

LOW

Irrelevant climatology tables appear repeatedly

Climate tables and charts were appended to multiple responses regardless of relevance — users took these as evidence the tool was off-topic.

Questionnaire Summary

Phase 1 results: usability scores and scenario-level patterns

System Usability
Scale (SUS)

58.9

/ 100

Below average usability

Range: 25 – 80
Strong variation between users
indicating inconsistent experience

Below 'Good' threshold (72)

SUS Item Breakdown

% of users who agree or strongly agree

Would use frequently

44%

Easy to use

56%

Functions work harmoniously

44%

Most people learn quickly

67%

Felt confident using it

22%

Unnecessarily complex

44%

Very cumbersome to use

44%

Scenario Evaluation Key Patterns



Users can learn the workflow once they understand the interaction model.



Information quality rated insufficient or overly generic across most scenarios.



Users felt spending more time would NOT improve answer quality — suggesting a ceiling on perceived capability.



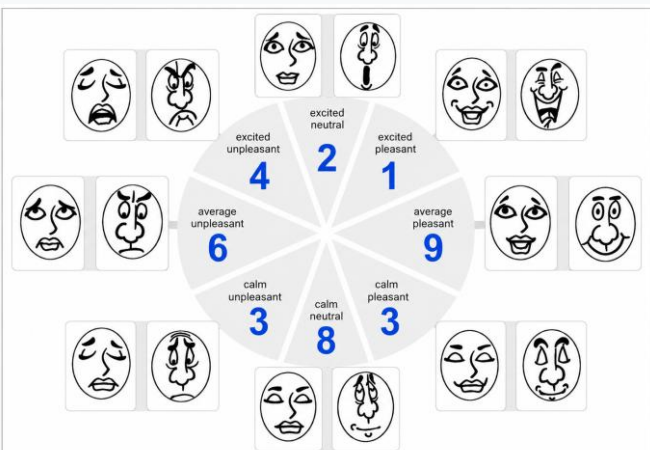
Perceived decision-support value was mixed: some found it helpful, many did not.

Takeaway: Users can learn the interface — but confidence in answer quality remains low.

What Users Said

Emotional trajectory across scenarios · Likes · Dislikes · Quotes

Emotional Response by Scenario



✓ What Users Liked

- Ability to target a specific geographic area
- System sometimes anticipated follow-up questions
- Interface perceived as simple once understood
- Geospatial focus seen as unique vs generic AI

✗ What Users Disliked

- Draw tool — hard to find & use
- Answers too generic
- Asks users for info it should retrieve
- Confirm workflow
- Missing
- Greek/English UI & labels

Representative User Quotes

"I would prefer a conversation without having to use the map."

"The draw icon is not visible."

"The answers were too general."

"It asks me for information it should already know."

"Until now it hasn't answered anything — it always says it can't."

Priority Improvements

Ranked by frequency & impact across all 9 sessions

R1 Redesign area selection

R1

Replace freehand draw with a clearer, more guided interaction (e.g. click-to-place polygon, visible 'Select Area' button with label). Add persistent selected-area indicator.

R2 Enable Enter to submit; fix input visibility

R2

Support Enter key for submission. Keep the input field in a fixed, visible position. Show a clear active state when the field is focused.

R3 Bring confirm button into view & simplify it

R3

Auto-scroll to the confirm action or keep it in a sticky footer. Reduce confirm frequency. Translate all system prompts to Greek.

R4 Say no I can't

R4

It is better to inform the users that it cannot answer, instead of trying to "please them"