HOW CAN ARTIFICIAL INTELLIGENCE (AI) BE APPLIED TO DESIGN TESTING FOR HUMAN COMPUTER INTERACTION (HCI)?

Dr Anthony Basiel and Dr Mike Howarth
Anthony.Basiel@solent.ac.uk
abasiel@gmail.com | https://abasiel.uk

OUTLINE

- Core problem
- Key terms
- Synthetic UX Personas – Dis/advantages
- AI policies and regulations
- UX/AI Persona workflow
- Case Study 1: AI video summary
- Case Study 2: AI v. Human survey creation
- AI & Human synthesis
- Conclusions

TAG CLOUD

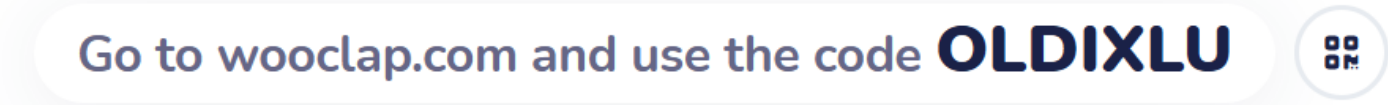
Post words for the methods / tools you use to collect primary research data.
You can do more than once.

Go to wooclap.com and use the code **OLDIXLU**

TAG CLOUD

See the results at

https://app.wooclap.com/events/OLDIXLU/live-session



Post words for the methods / tools you use to collect primary data. You can do more than once.

Interview Questionnaire Survey
Focus group

THE CORE PROBLEM

The Scale vs. Quality Trade-Off:

Traditional user research is limited by resource constraints

The Principles:

"Don’t start with AI, start with the problem"

Objective:

Establish a common language

DEFINING KEY TERMS

Artificial Intelligence (AI):

Systems characterised by adaptability and autonomy, capable of perceiving environments, interpreting data to mimic cognitive skills to make predictions.

HCI Personas:

Archetypal or 'median' representations of target user groups derived from real user data.

Purpose of Personas:

Promote perspective-taking, foster user empathy, prioritise features, and guide UX decisions.

PROS AND CONS OF SYNTHETIC USERS (1 OF 2)

- **Potential Advantages:**
- Rapid generation of UX survey questions.
- Drastic time savings when collecting mock responses to test scripts
  .

PROS AND CONS OF SYNTHETIC USERS (2 OF 2)

Potential Disadvantages:

Artificial Optimism:

Synthetic personas are frequently 'too agreeable' (sycophantic) and lack human emotional depth.

Hallucinations:

AI may invent data if prompted outside its core dataset

Mimicked Observer Effect:

AI datasets are human-derived, which can mirror the Hawthorne effect where responses alter due to perceived observation.

ACADEMIC POLICIES & LEGAL REGULATIONS

Institutional Guardrails:

Never upload confidential, sensitive, or unpublished data into third-party AI tools.

UK Regulatory Approach:

Focuses on the contextual *use* of AI rather than policing the technology itself; remains under constant review.

ACADEMIC POLICIES & LEGAL REGULATIONS

EU AI Act Framework:

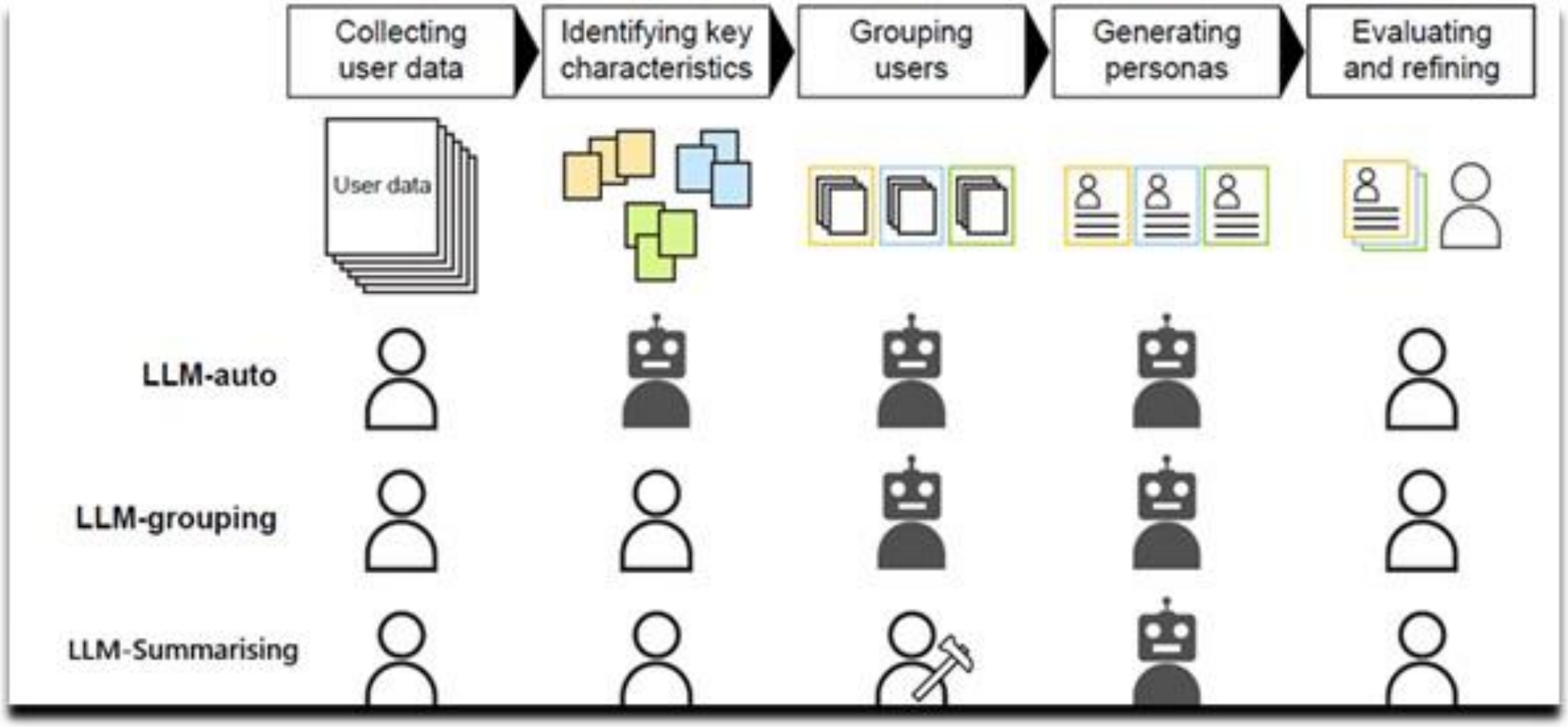
Risk-based categorisation:

- *Unacceptable Risk:* Social scoring systems (banned).
- *High-Risk:* Tools like CV scanners or automated ranking software (strictly regulated) .
- *Unregulated:* Basic applications not explicitly flagged .

THE 5-STEP HUMAN-AI PERSONA WORKFLOW

Step 1: Collecting Data – Executed by humans from stakeholders
Step 2: Identifying Key Characteristics – Collating user themes (Human).
Step 3: Grouping Users – Categorising users into low, medium, and high types (Human).
Step 4: Generating Personas – Drafting sample personas (AI).
Step 5: Evaluating and Refining – Extensively reviewing and adjusting outputs.
Optimal Model: The LLM-Summarising Model, which preserves human input across all steps except generation.

THE 5-STEP HUMAN-AI PERSONA WORKFLOW



CASE STUDY 1 – VIDEO SUMMARISATION

Task: Using Otter.ai to summarise a 9-minute academic video transcript (1,722 words) into brief formats.
AI Output Formats: Text-only Summary, Action items, and a structural Outline.
The Failure: The AI provided an adequate but bland summary that over-weighted minor details and completely lacked tone, personal witness, and nuance.
The Human Solution: Redrafting using a 'Clean English' professional style, re-ordering importance, and adding interactive user challenges.

CASE STUDY 2 – AI VS. HUMAN SURVEY CREATION

Methodology: 44 undergraduate students split into two teams to create a website UX persona survey in 15 minutes.

Team A (AI): Wrote prompts for Microsoft Co-pilot to generate draft questions.

Team B (Human): Brainstormed manually using only standard search engines.

Quantitative Findings: AI scored higher on overall clarity (4.06 vs. 3.82), structural logic, and professional readability.

Qualitative Findings: Human-generated questions scored significantly higher on feeling thoughtful, empathetic, and capturing nuanced human experiences (4.0 vs. 3.59).

SYNTHESIS:
WHERE AI EXCELS VS. WHERE HUMANS EXCEL

AI Strengths:

Establishing clean structural layout.

Massive data organisation and thematic parsing.

Standardised, highly efficient text output.

Human Strengths:

Deep empathy and contextual nuance.

Dynamic, emotionally engaging language.

Complex UX understanding and logical academic arguments.

CONCLUSIONS

The Hybrid Ideal: AI-generated foundations combined with extensive human refinement represents the current gold standard.

The "Thought Partner" Persona: AI should be applied as a conversational assistant or 'colleague' to stress-test drafts, rather than fully replacing human judgment.

Future Outlook: There will be a booming commercial demand for highly skilled text reviewers and editors capable of refining raw AI outputs.

See the research paper at:

https://shorturl.at/JGJGE

THANKS FOR YOUR TIME. QUESTIONS?

HAPPY TO COLLABORATE.

You can email me at:

abasiel@gmail.com

See my research at:

https://abasiel.uk