Ideating XAI: An Exploration of User's Mental Models of an AI-Driven Recruitment System Using a Design Thinking Approach

HELEN SHERIDAN | DYMPNA O'SULLIVAN | EMMA MURPHY

HELEN.SHERIDAN@TUDUBLIN.IE

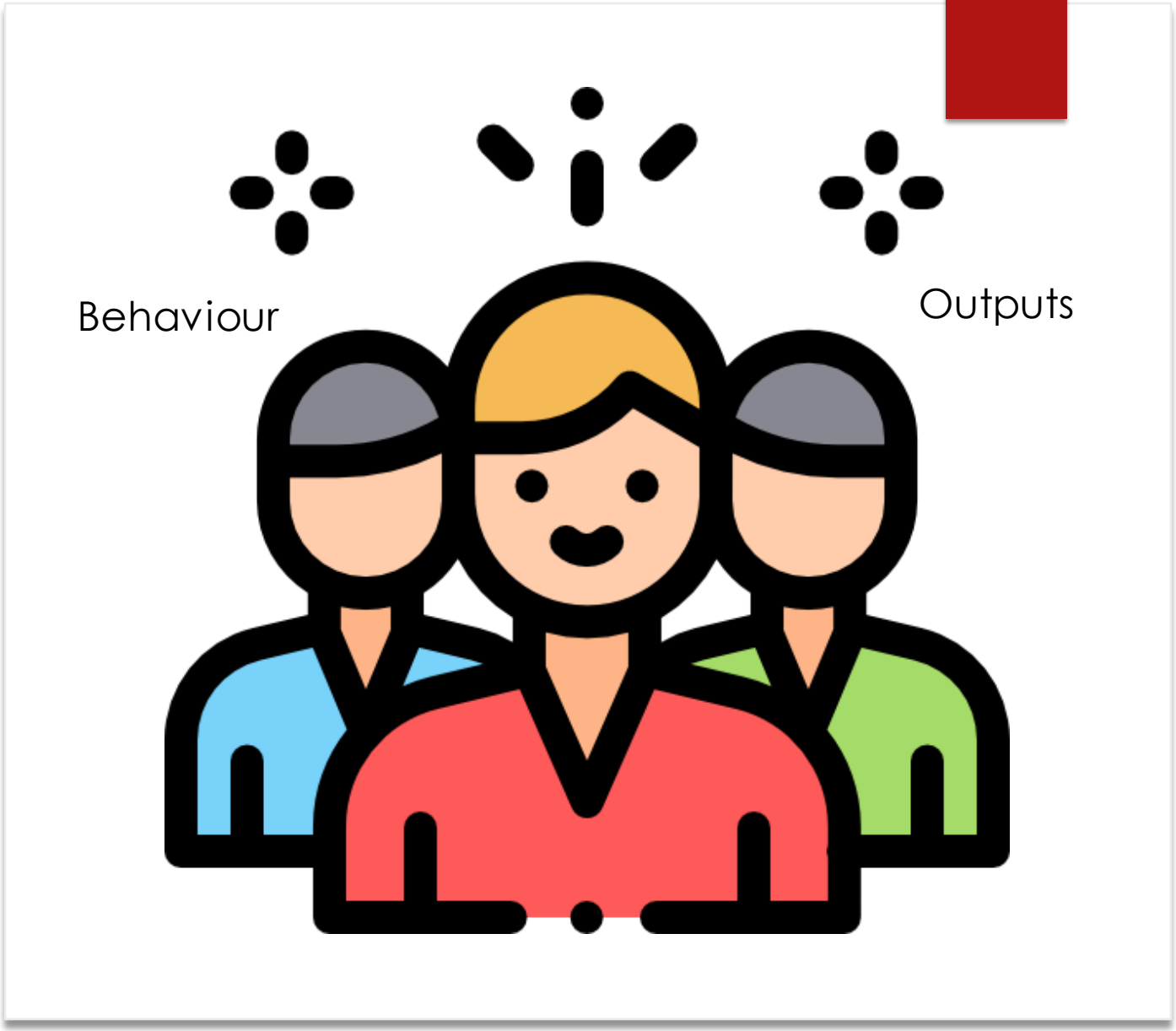
CENTRIC: International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services
Wednesday October 19th 2022

About

- PhD researcher with TUDublin School of Computer Science, January 2022
- Lecturer for 18+ years in Computer Science, Visual Communications & UI/UX
- 23 years industry experience in Design, UI/UX & Film Production
- Publications with IHCI 2022, EUT+ 2022 & IARIA 2022
- IBM Enterprise Design Thinking Practitioner & Team Essentials for AI

End Users' & AI

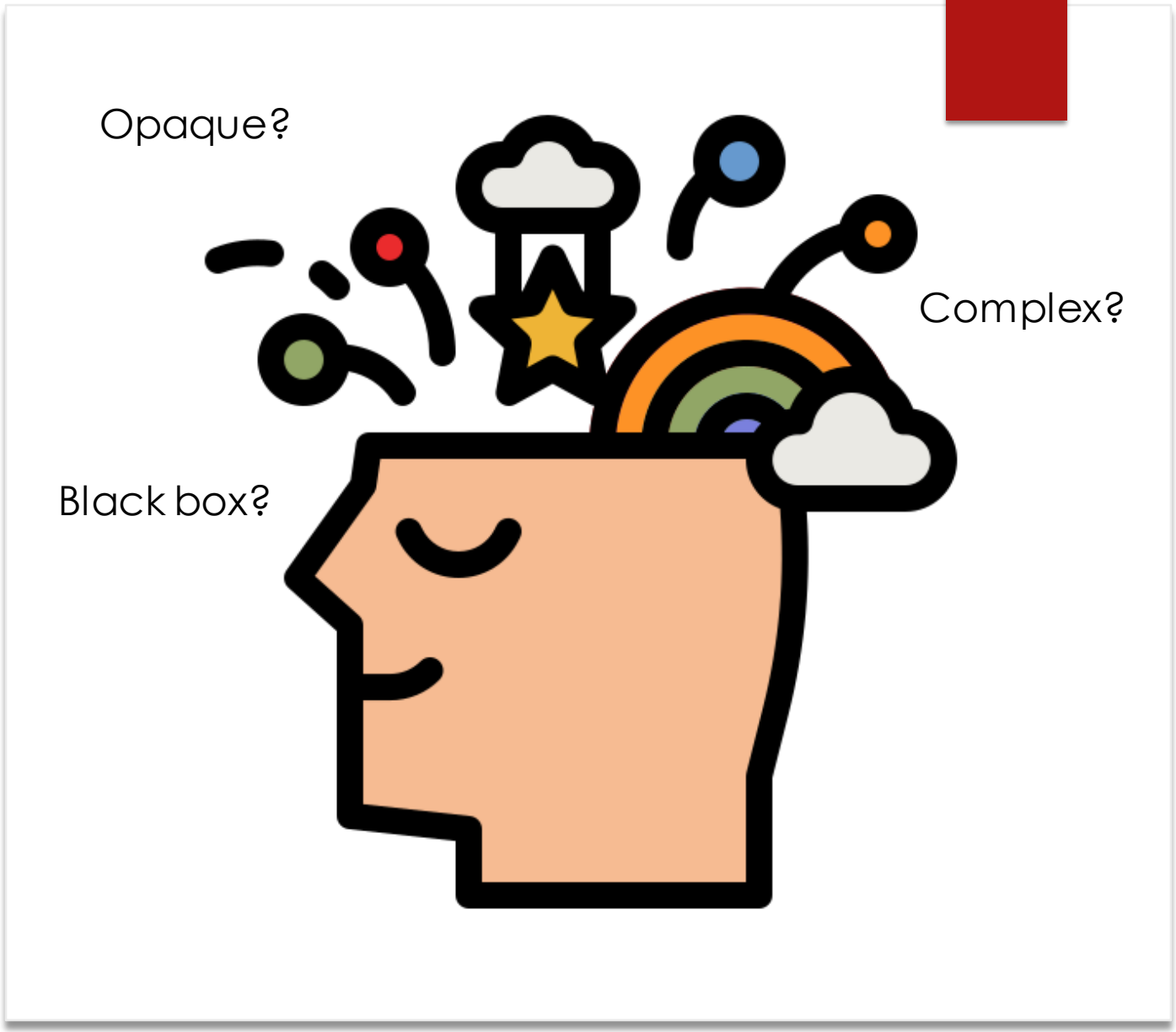
- Understanding AI behaviour
- Understanding how AI computes outputs
- Crucial in developing XAI for users



[1]

Users' Mental Models

- How a user believes a system works
- Can be misaligned to how a system actually works
- Crucial in explaining AI



[2]

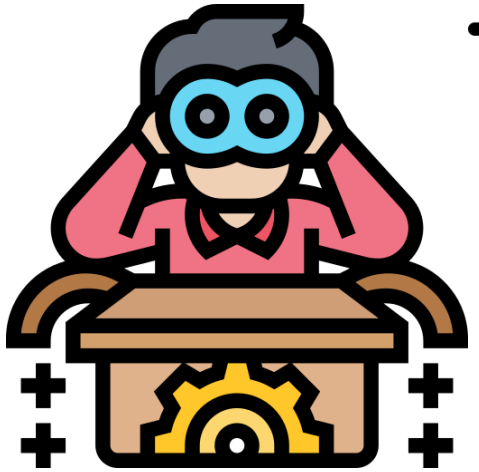
Current evaluation methods

- Primarily used for assessing interactive systems
- Evaluation gaps still around users understanding of AI
- Surveys, interviews, observations assess interaction
- How to assess cognitive perceptions & mental models?

[3] [4] [5] [6]

Surveys

Interviews

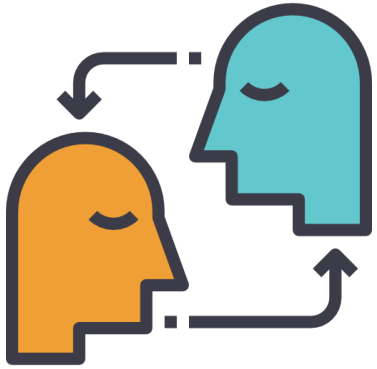
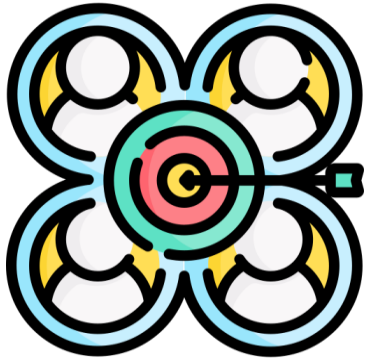


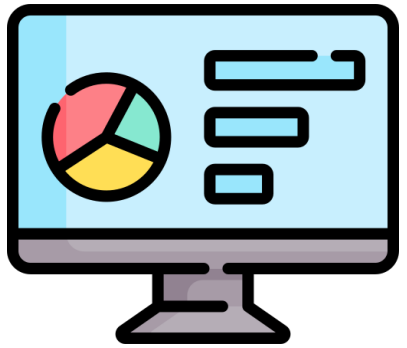
Observations

Design Thinking

- Problem solving method
- Non-linear with defined steps
- User centered at early stage
- Big ideas to explore concepts usually difficult to articulate
- Minor modifications vs novel ideas
- Pain point definition based on users' needs

[7]

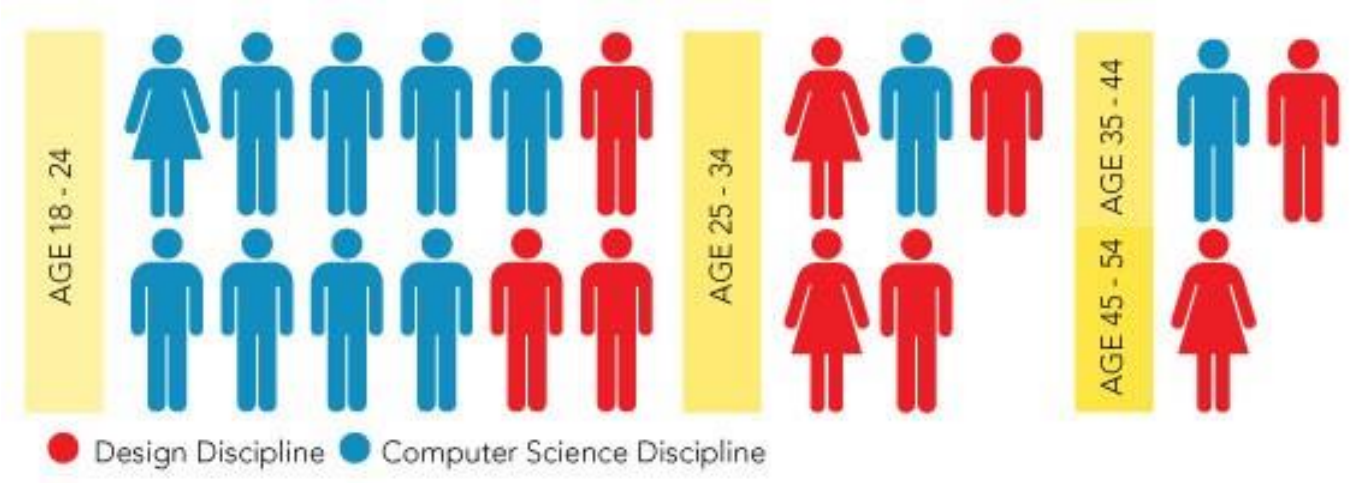






Empathise | Define | Ideate | Prototype | Test

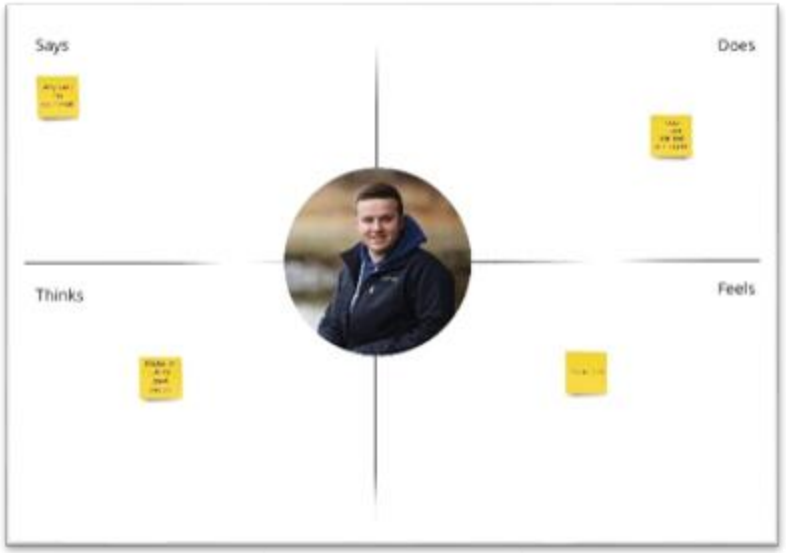
Methodology: Design Thinking

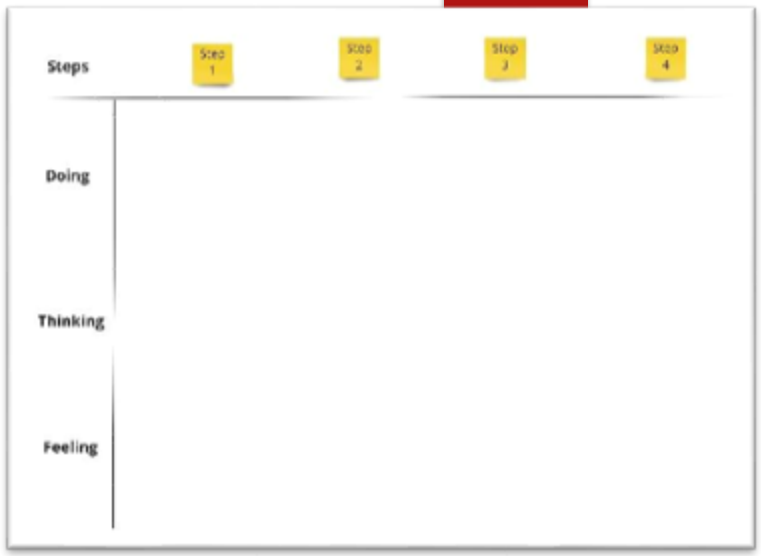
- Design thinking workshop
- 20 participants
- Multidisciplinary: Computer Science & Design Undergraduates

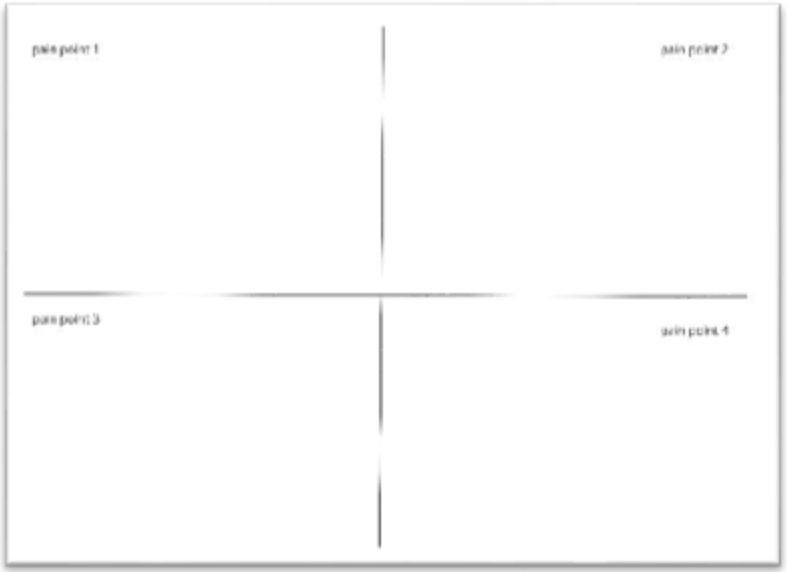


Methodology: Design Thinking

- **Empathise:** empathy mapping & as is scenario
- **Define:** pain point definition
- **Ideate:** Big ideas & prioritisation







Methodology: Personas

- Recruitment domain
- Personas: Maria Atkins a recruitment specialists and Andrew Wilson a recent graduate looking for work
- 2 different but typical users of an AI driven recruitment system
- Scenario focused on personas' frustrations with the AI system

Problem statement

How can we explain AI systems decisions, making them more transparent and understandable to users?

Says
Does
Thinks
Feels
Who's Active
Does Degree = Job?
Get a Job
Asks for Advice
Searching Job Sites
What am I doing wrong?
Why me!
??? Confused
Angry
Is it my CV?
Tired
Depressed
Hopeful
Freelance
I need the money
Excluded
Ignored
Angry
Upset
Is the system correct.
Become more unique
Talk with engineers about system.
Understand the problem better
Talk to experts
Top talent not being gained
HELP
THINK
Research
Get HELP
Ask for help
Let's talk together
I want more clarity
Hmmm.
Put data in Excel
Confused
Not helpful
Overwhelmed
Should I change anything
I don't want to fail
Double verification system
Get second opinion
Automate the process
How to improve
Steps
Doing
Review Applicants
Reading
Is it a fair system
Search
Apply
Revise CV
Call
Indeed
Send Letters
Analyse replies
Never Get Job
Lack of feedback
Dejected
Hopeful
Excited
Stressed
Second eyes
Proof Read
Hire Billboard
Emigrate
Research CEO Refine CV
Frustrated
Depressed
Hopeful
Login to system
Calm Self
Update System?
Confused
Anxiety
Guilty
Ambitious
Unsure
Nervous
Skeptical
Unfair
HR
Research
Meditate
Hidden agenda
Big Bets
No Brainers
Utilities
Unwise
Talk to a Robot
Questions the System
Change input criteria
Ask the experts
Compare to the old system
Share her tracking
Virtual help
Dash base
Importance to user
Feasibility for us
High
Low
Difficult / Expensive
easy / cheap
Make it a game
Make it fun to Apply
Analytics
Email Sent
RATING
CEO
Employer can send feedback
Creates Personal Website
Re apply Button
CLIPPY
Buys illegal firearm
Highlight problem in CV
Score in categories
Create Dashboard of Applicants
Show success
Escalate Rejection
Hire Hitman
Allow Cover Letters
Template

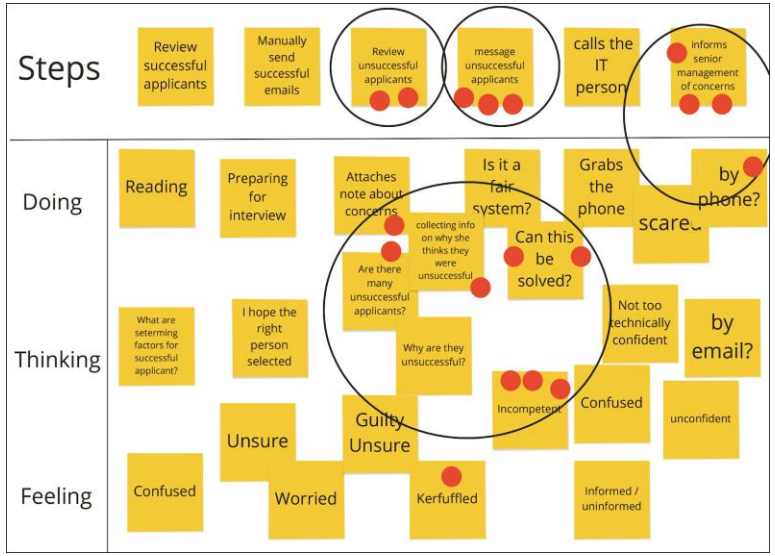
Results: Empathising & Definition

- Group 4: Maria Atkins Persona
- Opaque & Confusing
- Powerless, Out of control and has a sense of guilt

- As is scenario steps
- Reviewing & messaging unsuccessful and successful applicants
- Informing management

- Voting on pain points, 5 votes each
- Clustered around areas
- 4 pain points identified

[7]



Graphical Representation of Empathy Map, As is scenario and Pain Points Group 4

Results: Empathising & Definition

- Group 1: Andrew Wilson Persona
- Negative & Confusing
- Upset, Angry and unmotivated

- Searching & applying, waiting, receiving reply, updating CV & reapplying
- Repeating steps with no feedback

- Voting on pain points, 5 votes each
- Clustered around areas
- 4 pain points identified

[7]

Graphical Representation of Empathy Map, As is scenario and Pain Points Group 1

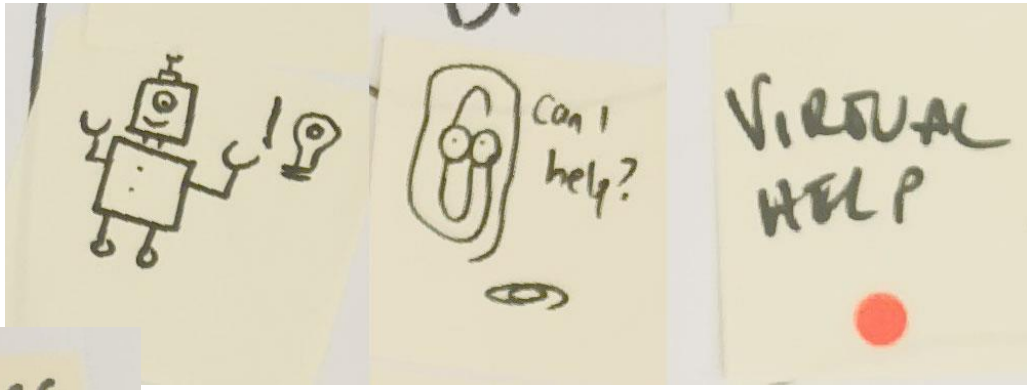
Results: Categories

- **Data:** Pre-workshop survey, audio recording during workshop, photographs of worksheets, post workshop interview
- Consolidated findings
- Categorisation to group findings into topic areas
- 2 common categories: Visual feedback & analytics and Visual Comparisons

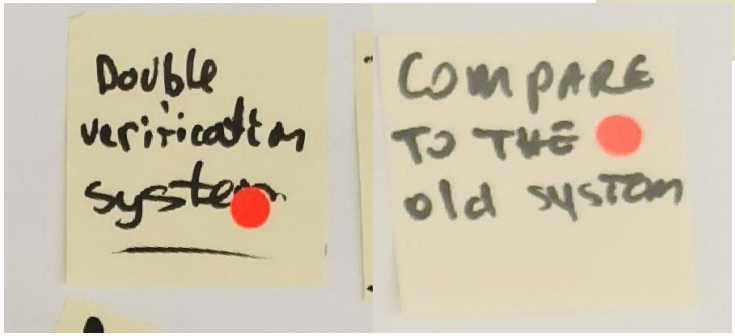
Results: Pain point to big ideas

- Pain points reflect opaque areas of the AI system for users
- Big ideas reflect solutions to provide explanations in order to enhance end-users' understanding of AI system and potentially explain AI systems' behaviour
- Pain points mapped to big ideas for Maria Atkins

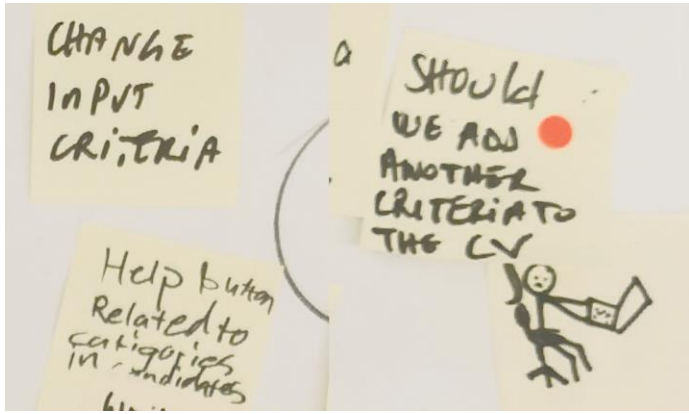
Pain points for Maria Atkins Group 4



Participants visualisation of visual feedback & analytics



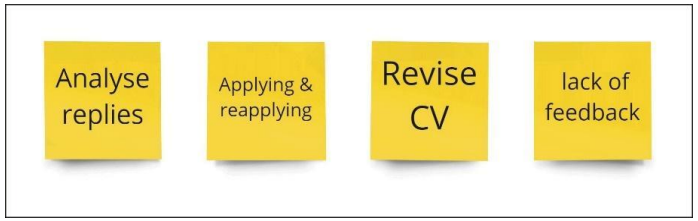
Participants visualisation of visual comparisons



Participants visualisation of criteria manipulation or tracking

Results: Pain point to big ideas

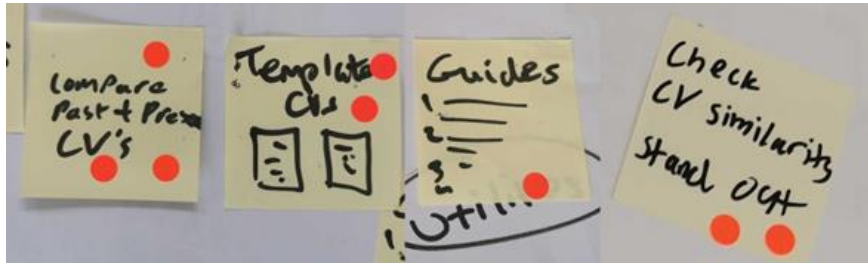
- Pain points reflect opaque areas of the AI system for users
- Big ideas reflect solutions to provide explanations in order to enhance end-users' understanding of AI system and potentially explain AI systems' behaviour
- Pain points mapped to big ideas for Andrew Wilson



Pain points for Andrew Wilson Group 1

Participants visualisation of visual feedback & analytics

Participants visualisation of offering chances to rectify and reapply



Participants visualisation of visual comparisons

Discussion: Explanations

- **Factual Explanations:**
- Useful when system output is as expected

- **Counter Factual Explanations:**
- Useful especially when system output isn't met

- **Principal Reason Explanations:**
- Allowing for criteria manipulation & chance to achieve different result

Implementing Design Thinking

- A Design Plan for Ideating AI Using a Design Thinking Approach

1. Persona & Scenario	2. Participants	3. Size	4. Multimodal	5. Playbacks	6. Embrace the absurd	7. Data Collection
Design persona on **real use cases** & devise scenario to establish users' frustrations for better pain point identification.	**Interdisciplinary participants** are favoured preferably with domian stakeholders represented.	Min workshop: **4 participants**	Encourage **drawing & writing**.	Hold playbacks at **critical moments**.	Keep groups **on task and aligned** to the problem.	**Photograph** worksheets regularly.
		Max workshop: **Dependant on facilities** such as location size or number of facilitators.	**One idea** per sticky note.	Allow all team members to contribute to **ensure alignment.**	**Embrace the absurd**, no idea is rejected initially.	**Number worksheets** to identify groups.
			Quantity over quality.	Focus on **solutions to pain points.**		**Audio record** playbacks and closing reflections.
			Engage designers as participants.			

Thank You

- Questions

References

- [1] Shin, D., 2021,The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 146, p.102551.
- [2] Kaur, H., Williams, A. and Lasecki, W.S., 2019. Building shared mental models between humans and ai for effective collaboration. CHI'19, May 2019, Glasgow, Scotland.
- [3] Brennen, A., 2020, April. What Do People Really Want When They Say They Want" Explainable AI?" We Asked 60 Stakeholders. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems(pp. 1-7)
- [4] Chazette, L. and Schneider, K., 2020,Explainability as a non-functional requirement: challenges and recommendations. Requirements Engineering, 25(4), pp.493-514.
- [5] Wang, D., Yang, Q., Abdul, A. and Lim, B.Y., 2019, May. Designing theory-driven user-centric explainable AI. In Proceedings of the 2019 CHI conference on human factors in computing systems(pp. 1-15)
- [6] Sperrle, F., El-Assady, M., Guo, G., Borgo, R., Chau, D.H., Endert, A. and Keim, D., 2021, June. A Survey of Human-Centered Evaluations in Human-Centered Machine Learning. In Computer Graphics Forum(Vol. 40, No. 3, pp. 543-568)
- [7] Sheridan, H, O'Sullivan, D and Murphy, E, 2022, Ideating XAI: An Exploration of User's Mental Models of an AI-Driven Recruitment System Using a Design Thinking Approach, In Proceedings IARIA, SoftNet 2022 Congress, Lisbon, October 2022
- Photographs from www.pexels.com, icons from www.flaticon.com