Human-centred SE & CS in the age of 'generative AI'

ARC Laureate Professor John Grundy

HCSE&CS Workshop 2023

(All Images generated by Bing image creator ; some text generated by Bing, ChatGPT and Bard)

https://www.monash.edu/it/humanise-lab

MONASH University

Australian Government
Australian Research Council

HumaniSE

Outline

- HCSE-CS key themes (from the workshop CFP :-))
- How does/might Gen AI help us achieve these?
- How does Gen AI not help us achieve?
- HCSE-CS issues with Gen AI
- Gen AI issues with HCSE-CS
- Future directions (for discuss!!)
- Summary

HCSE-CS key themes (from the AI…)

[After several goes w ChatGPT and Google doc AI… :-)]

ChatGPT: Certainly, here's a shortened summary in max 10 bullet points:

- Focus on human factors in software development.
- Emphasis on usability, accessibility, and emotions in software.
- Topics include modeling, requirements, and methodologies.
- Integration of human behavior into cybersecurity solutions.
- Evaluation of usable security and privacy.
- Study of mental models in security and privacy.
- Exploration of emerging technologies in human-centric applications.
- Case studies on insider whistleblowing and developing country cybersecurity.

Gen AI + HCSE-CS - achieving our goals…

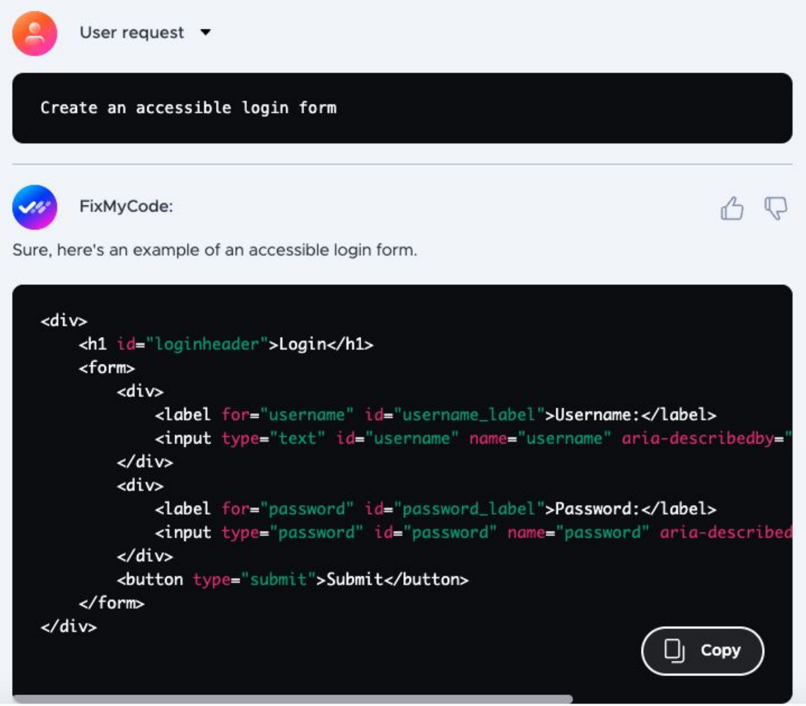
Human factors in SE, CS

Tools to help achieve e.g.:

- Accessibility - Gen AI to add to code, analyse code/UIs for issues, reproduce behaviour/errors, style transfer for inclusivity…
- Cybersecurity - Process suspicious email, generate examples of adversarial attacks, suggest steps, processes to ‘fix’
- Team work - Monitor team metrics, detect issues & suggest remediation
- Defect fixing - Human-centric defect reporting, defect analysis, fix suggest/help

Example 1: Fixmycode - “accessible code with AI”...

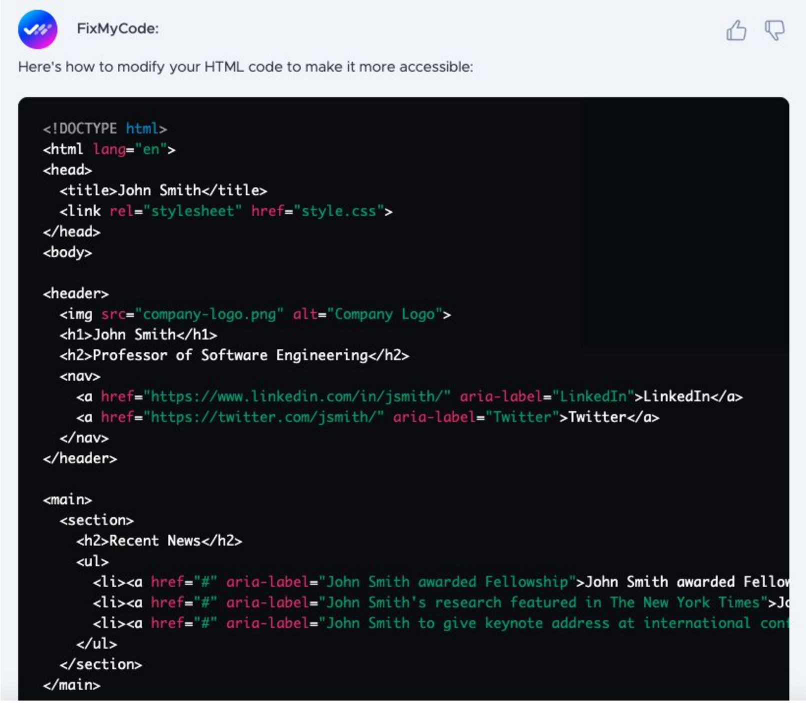
https://userway.org/fixmycode



```
Create an accessible login form
```

Sure, here's an example of an accessible login form.

```
<div>
    <h1 id="loginheader">Login</h1>
    <form>
        <div>
            <label for="username" id="username_label">Username:</label>
            <input type="text" id="username" name="username" aria-describedby="
        </div>
        <div>
            <label for="password" id="password_label">Password:</label>
            <input type="password" id="password" name="password" aria-described
        </div>
        <button type="submit">Submit</button>
    </form>
</div>
```



Here's how to modify your HTML code to make it more accessible:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <title>John Smith</title>
  <link rel="stylesheet" href="style.css">
</head>
<body>

<header>
  <img src="company-logo.png" alt="Company Logo">
  <h1>John Smith</h1>
  <h2>Professor of Software Engineering</h2>
  <nav>
    <a href="https://www.linkedin.com/in/jsmith/" aria-label="LinkedIn">LinkedIn</a>
    <a href="https://twitter.com/jsmith/" aria-label="Twitter">Twitter</a>
  </nav>
</header>

<main>
  <section>
    <h2>Recent News</h2>
    <ul>
      <li><a href="#" aria-label="John Smith awarded Fellowship">John Smith awarded Fello
      <li><a href="#" aria-label="John Smith's research featured in The New York Times">J
      <li><a href="#" aria-label="John Smith to give keynote address at international con
    </ul>
  </section>
</main>
```

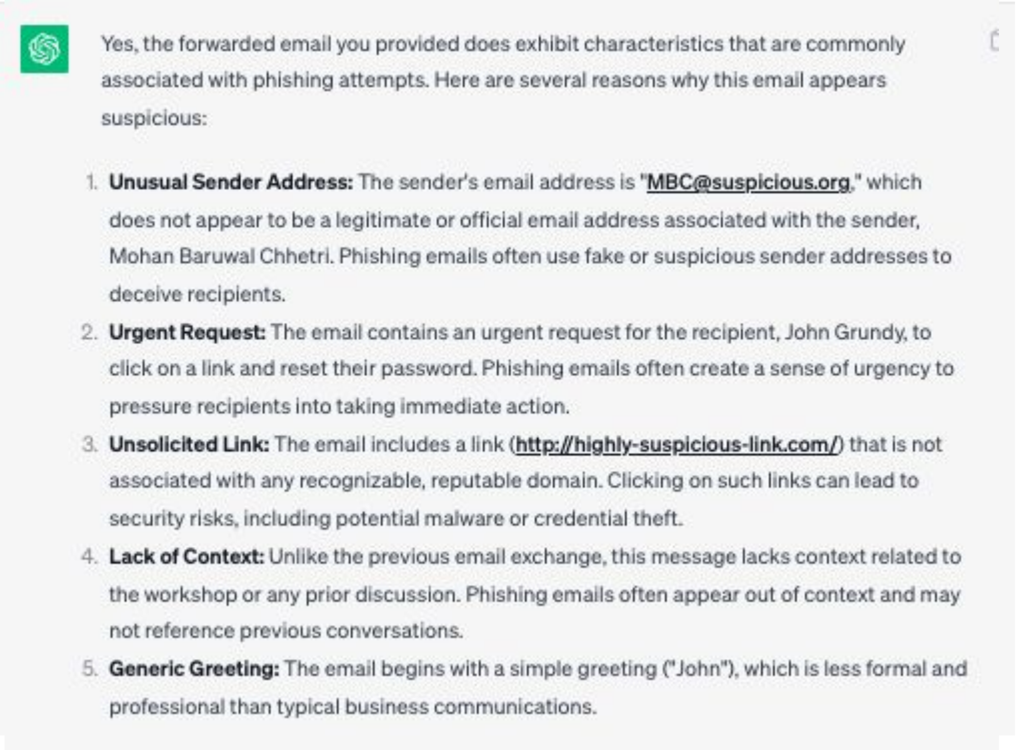
Example 2: Data analysis

Analysis of data - in this case discuss between Mohan and me of this talk focus… :-D

Email 1…

Email 2 …

I decided not to click on the link btw…



Yes, the forwarded email you provided does exhibit characteristics that are commonly associated with phishing attempts. Here are several reasons why this email appears suspicious:

1. **Unusual Sender Address:** The sender's email address is "MBC@suspicious.org," which does not appear to be a legitimate or official email address associated with the sender, Mohan Baruwal Chhetri. Phishing emails often use fake or suspicious sender addresses to deceive recipients.
2. **Urgent Request:** The email contains an urgent request for the recipient, John Grundy, to click on a link and reset their password. Phishing emails often create a sense of urgency to pressure recipients into taking immediate action.
3. **Unsolicited Link:** The email includes a link (http://highly-suspicious-link.com/) that is not associated with any recognizable, reputable domain. Clicking on such links can lead to security risks, including potential malware or credential theft.
4. **Lack of Context:** Unlike the previous email exchange, this message lacks context related to the workshop or any prior discussion. Phishing emails often appear out of context and may not reference previous conversations.
5. **Generic Greeting:** The email begins with a simple greeting ("John"), which is less formal and professional than typical business communications.

Example 3: Team ‘management’ w AI help

MotiveMetrics tool - capture motivation, personality, satisfaction, performance metrics during agile SD

MotiveMetrics++

- Emotions, task performance metrics
- Detect team member “under stress”
- Flag to manager
- Suggest some mitigations
- Generate task plans for diff SEs

Figure 4: Example card (A)

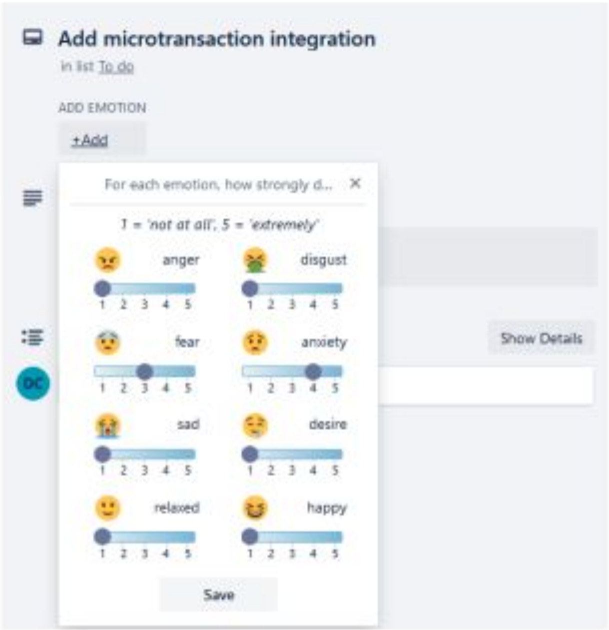

Add microtransaction integration
in list To do
ADD EMOTION
+Add
For each emotion, how strongly d...
1 = 'not at all', 5 = 'extremely'
anger
disgust
fear
anxiety
sad
desire
relaxed
happy
1 2 3 4 5
Save
Show Details
DC

Fig. 9: Performance Ratings shown on pie charts.

Example 4: Human-centred defect reporting,fixing

Some defects we define as “human-centric” issues or “human-centric” defects

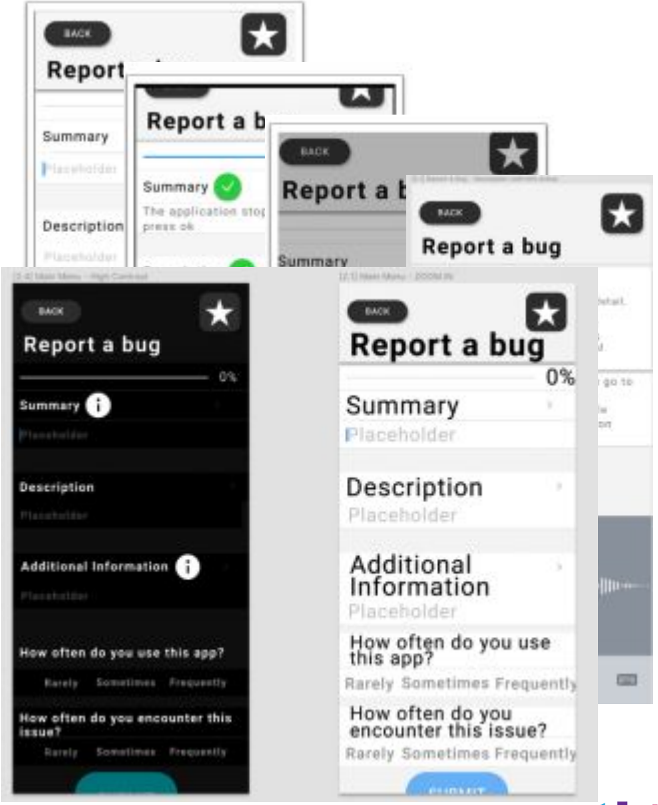
“In the eye of the beholder” - some users think a defect, some not

Manifests differently to different users, in different usage scenarios

Want improved defect reporting UIs

Use Gen AI to determine kind of defect, recommend possible fixes

Use gen AI to help craft, manage human-centric defect reports


BACK
Report a bug
0%
Summary
Placeholder
Description
Placeholder
Additional Information
Placeholder
How often do you use this app?
Rarely Sometimes Frequently
How often do you encounter this issue?
Rarely Sometimes Frequently
SUBMIT
The application stop press ok

https://www.monash.edu/it/humanise-lab

Gen AI + HCSE-CS - maybe NOT achieving our goals…

I will use the last example to start:

- What if over-loads already stressed dev with too much, inappropriate work? Mis-leads them? Forgets critical path work items?
- What if our AI suggests e.g. fire this person they are unproductive / not “good fit “ in the team…?
- What if the data for training is biased e.g. female SEs, neurodiverse SEs, older SEs…?
- What if team/manager misses gen AI mistakes (“hallucinations”)?

HCSE-CS issues with Gen AI

This is a really interesting area IMO e.g.

- Is ChatGPT/Copilot type ‘dialogue’ best approach to interact with?
- How do different developers/users find it?
- What are key issues with using - good and bad?
- How do diff users use e.g. personality, motivation, emotions etc + Gen AI?
- How do teams best leverage Gen AI - mix of human and AI?
- What are key cybersecurity issues e.g. IP leakage, privacy, etc
- How integrate with e.g. low code/no code so end users can create/modify code?
- Open datasets vs closed?
- Where are key improvements needed?

Example - code gen & privacy (code)

We recently did a study of developer perceptions of privacy related issues in AI generated code

Showed a bunch of good/bad gen code examples

Asked to identify privacy-related issues with the Gen code, how would fix etc

Asked perceptions on using gen AI in their dev work, whether they identify such issues in practice

Paper being finalised… :-)

Gen AI issues with HCSE-CS

“Personality” of AI/generative AI/conversational agent interfaces?

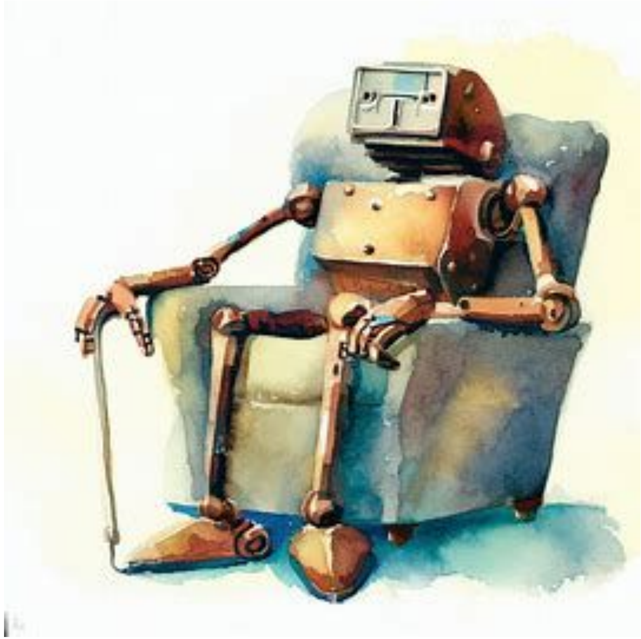
How modify the (gen AI) interface to better suit user, context of use, etc

How monitor the gen AI to detect e.g. bias, inaccuracy, performance, dift, etc

How protect from adversarial attacks - gen AI and human users

Human-AI teaming - who does what

Can I retire or will I be retired by Gen AI…???



Example 1 - gen AI “personality”

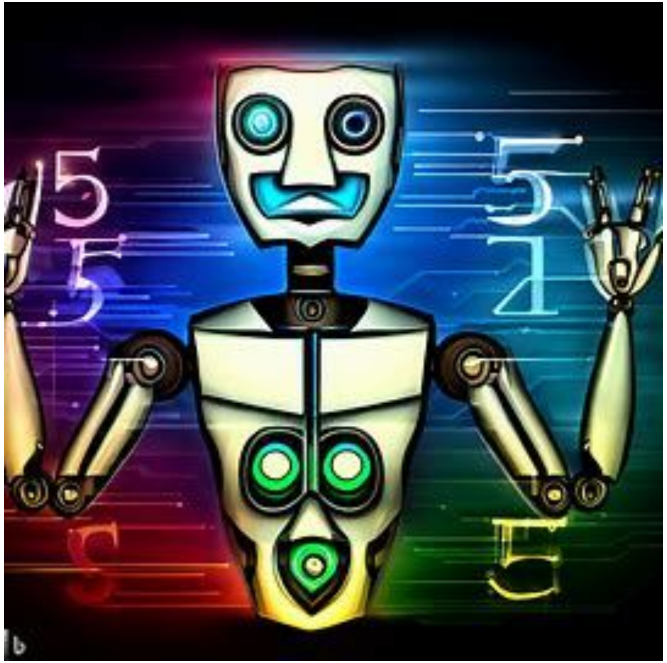
Conducting a study of the ‘personality’ of Gen AI

Building on our earlier work of developer personality, motivation, satisfaction work

How does gen AI do on a personality test?

Can gen AI be “prompted” with different Five-factor style ‘personality’ in its interactions with user?

Does this improve user/gen AI interaction? Or make it worse, or no difference??



Example 2 - Monitoring (Gen) AI components

Range of human values important to ensure ML/AI meets

Incl bias, fairness, privacy, honesty, transparency, engagement, …

How do we specify these (in Human-centric way!)

How do we generate “monitors” to observe the AI behaviour @ runtime

How do we mitigate detected human value violations in the AI @ runtime?

(this image was not really like the opposite topics - but I asked for ‘in the style of Van Gough’ and kind of liked the result - very human-centric :-))

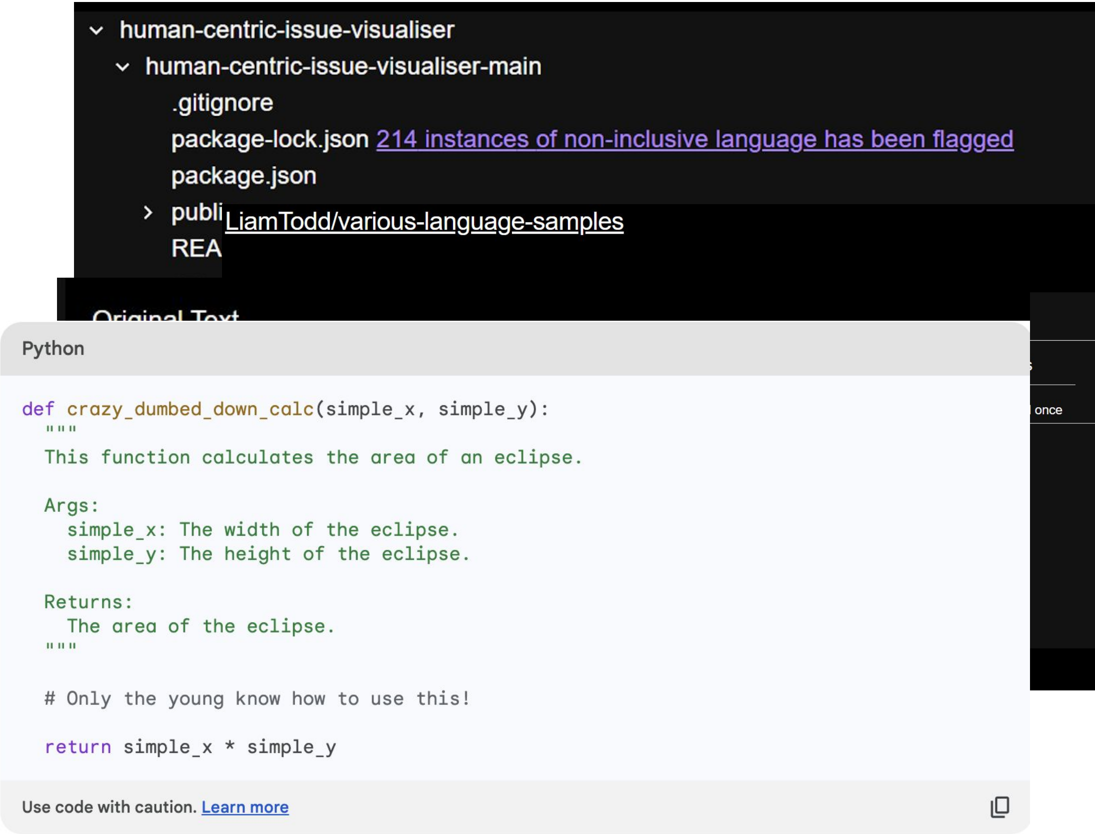
Example 3 - Inclusifying (Gen AI) components

We did analysis of language in GitHub and mobile app descriptions for “non-inclusive” terms

We have been experimenting with using LLMs to aid “inclusifying” terms - incl app descriptions, app content, documentation, code comments, code elements…

Can Gen AI be used to inclusify its own language…???

[It was nice of bard to give me the code caution warning…] ->



human-centric-issue-visualiser
human-centric-issue-visualiser-main
.gitignore
package-lock.json 214 instances of non-inclusive language has been flagged
package.json
publi
REA

LiamTodd/various-language-samples

Original Text

once

Python

```python
def crazy_dumbed_down_calc(simple_x, simple_y):
  """
  This function calculates the area of an eclipse.

  Args:
    simple_x: The width of the eclipse.
    simple_y: The height of the eclipse.

  Returns:
    The area of the eclipse.
  """

  # Only the young know how to use this!

  return simple_x * simple_y
```

Use code with caution. Learn more

Summary

Generative AI has and will have a major impact on (HC) SE & CS

Interesting issues of when to use, how to use, how to integrate into workflow, how to ensure ‘correct’ usage etc

Interesting HC SE & CS issues with gen AI itself

What are the _human_ aspects that impact gen AI usage?

Can gen AI be more ‘humanised’??

A few of our related works…

(image of john grundy, very young, hot, water color)

Solanki, P., Hussain, W., Grundy, J.C. Operationalising ethics in artificial intelligence for healthcare: a framework for AI developers, AI and Ethics, Springer, July 2022 PDF

Ahmad, K., Abdelrazek, M., Arora, C., Baniya, A.A., Bano, M., Grundy, J.C., Requirements Engineering Framework for Human-centered Artificial Intelligence Software Systems, Applied Soft Computing, vol. 143, August 2023, Elsevier PDF

Gao, Z., Xia, X., Grundy, J.C. , Lo, D., Li, Y.Y-F. Generating Question Titles for Stack Overflow from Mined Code Snippets, ACM Transactions on Software Engineering and Methodology, Vol. 29, No. 4, PDF

Graetsch, U.M., Khalajzadeh, H., Shahin, M., Hoda, R., Grundy, J.C. Dealing with Data Challenges when Delivering Data-Intensive Software Solutions, to appear in IEEE Transactions on Software Engineering PDF

Pathmabandu, C., Grundy, J.C., Chhetri, M.B., Baig, Z. Privacy for IoT: Informed consent management in Smart Buildings, Future Generation of Computing Systems, vol. 145, August 2023 PDF

Obie, H., Du, H., Madampe, K., Shahin, M., Ilekura, I., Grundy, J.C., Li, L., Whittle, J., Turhan, B., Khalajzadeh, H. Automated Detection, Categorisation and Developers' Experience with the Violations of Honesty in Mobile Apps, to appear in Empirical Software Engineering PDF

Liu, P., Zhao, Y., Fazzini, M., Cai, H., Grundy, J.C., Li, L. Automatically Detecting Incompatible Android APIs, to appear in ACM Transactions on Software Engineering and Methodology PDF

Law, C.Y., Cain, A., Vasa, R., Von Baggo, K., Grundy, J.C. Case Study of Designing and Evaluating an Independent Open Learner Model Tool, Higher Education Pedagogies, vol. 8, no. 1, 2023 PDF

Khalajzadeh, H., Shahin, M., Obie, H., Agrawal, P., Grundy, J.C. Supporting Developers in Addressing Human-centric Issues in Mobile Apps, IEEE Transactions on Software Engineering, vol. 49, no. 4, PDF

Where to with Gen AI & HCSE-CS ???

Your views ???

Top priority areas to focus on?

Can Gen AI aid us in addressing some of those key HCSE-CS workshop topics I highlighted earlier?

How do we ensure HC Gen AI from SE & CS perspectives??