

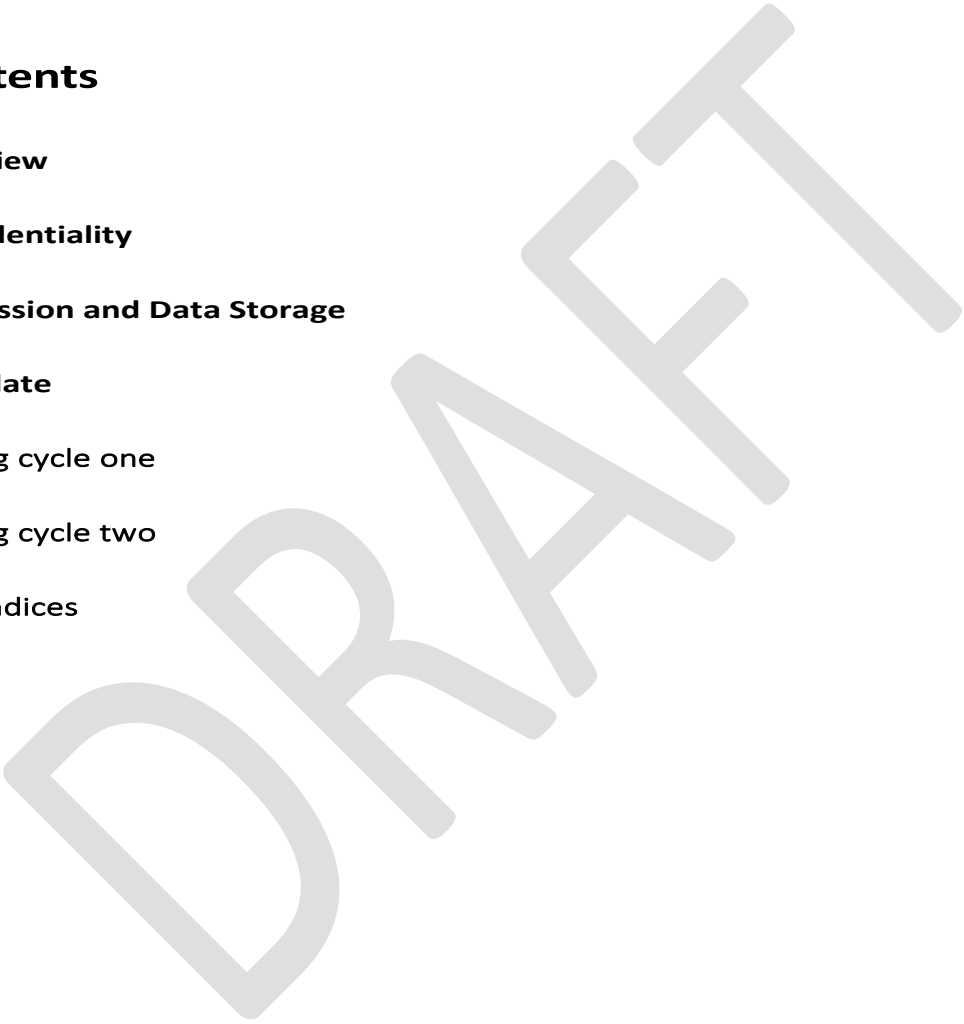
Title: Data analysis Protocol for a Joint Study into the Impacts of AI on professional Competencies of IT Professionals and Implications for Computing Students.

Author: Tony Clear

Date: May 1, 2024 Version 1, June 2 2024 Version 2, June 3 2024 Version 3, June 18 Version 4, June 19 Version 5

Contents

Overview	2
Confidentiality	3
Permission and Data Storage	3
Template	3
Coding cycle one	5
Coding cycle two	6
Appendices	8



Overview

The purpose of this protocol is to help us define a common protocol for sharing and analysing data for the ITiCSE 2024 working group: “WG02: A Multi-Institutional-Multi-National Study into the Impacts of AI on Work Practices of IT Professionals and Implications for Computing Students”. Excerpts from the working group plan to place the protocol in context (Clear et al., 2024) are given below.

Background and Related Work

As Artificial Intelligence (AI) continues to make its presence felt in transforming workplaces around the world [1,10], and the Information Technology industry in particular, it is essential to understand its impact on the work practices of IT professionals, and the implications for computing students and curricula. This research project builds on work initiated jointly, in Sweden, New Zealand and Scotland, investigating concerns about the increasing impacts of Artificial Intelligence in IT Sector workplaces for employee work engagement [11,13,1] and the implications for tertiary study, assessment and curricula in computing [4, 8, 10, 9].

“Work engagement”, has been defined as the positive inner state where employees are fully present and engaged in their work, and is closely linked to motivation, learning, productivity, and accountability [11, 13]. Within the context of (Generative) AI at work, IT professionals have been noted as early adopters of AI [10, 1]. Their involvement in implementing and utilising AI technologies can provide valuable insights into the interplay between AI and work engagement. The implications for students are significant as future IT professionals, who must acquire and enhance competencies to adapt and thrive in digital workplaces.

2 Goals of the Working Group

By exploring the relationship between work engagement and learning, this study aims to shed light on the dynamics that drive employee engagement and its connection to the professional development of competencies. The previous study has interviewed IT professionals with the following research questions (RQ):

RQ1: How does AI influence work engagement for IT professionals?

RQ2: How does AI affect the socio-technical work dynamics for IT professionals?

RQ3: What are the implications of integrating AI on the acquisition and enhancement of professional competencies and the learning processes of IT professionals?

3 Methodology

This working group aims to analyse the corpus of interview data collected from multiple countries to better understand the implications for computing students, tertiary computing education curricula and assessment of the new professional competencies emerging from this work. This study informed by the literature on work engagement, automation and motivation for IT professionals [11, 13], will use a combination of multi-vocal literature review [7] and qualitative research methods [2, 5], including thematic analysis of the interviews, to investigate the state of the practice in and challenges IT Professionals face within their local/global work contexts. The literature on professional competencies in computing [4, 3, 6] will be drawn upon to characterise the new needs identified in this analysis. Further implications for computing curricula design and assessment will be developed from this analysis.

REFERENCES

- [1] ACM Technology Policy Council. 2023. Principles for the development, deployment, and use of generative AI technologies, ACM New York.
- [2] Braun, V. and Clarke, V. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology*, 18 (3). 328-352.
- [3] Clear, A., Clear, T., Vichare, A., Charles, T., Frezza, S., Gutica, M., Luntz, B., Maiorana, F., Pears, A. and Pitt, F. 2020. Designing Computer Science Competency Statements: A Process and Curriculum Model for the 21st Century in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, ACM, New York.

- [4] Clear, A., Parrish, A. and CC2020 Task Force. 2020. Computing Curricula 2020 - CC2020 - Paradigms for Future Computing Curricula ACM and IEEE-CS eds. *A Computing Curricula Series Report* ACM, New York.
- [5] Cruzes, D.S. and Dyba, T. 2011. Recommended steps for thematic synthesis in software engineering. in *2011 international symposium on empirical software engineering and measurement*, IEEE, 2011, 275-284.
- [6] Frezza, S., Clear, T. and Clear, A. 2020. Unpacking Dispositions in the CC2020 Computing Curriculum Overview Report in *2020 IEEE Frontiers in Education Conference (FIE)*, IEEE, Uppsala, Sweden.
- [7] Garousi, V., Felderer, M., & Mäntylä, M. V. 2019. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*, 106. 101-121
- [8] Jacques, L. 2023. Teaching CS-101 at the Dawn of ChatGPT. *ACM Inroads*, 14 (2). 40-46.
- [9] Liffiton, M., Sheese, B., Savelka, J. and Denny, P. 2023. CodeHelp: Using Large Language Models with Guardrails for Scalable Support in Programming Classes. *arXiv preprint arXiv:2308.06921*.
- [10] Prather, J., Denny, P., Leinonen, J., Becker, B.A., Albluwi, I., Craig, M., Keuning, H., Kiesler, N., Kohn, T. and Luxton-Reilly, A. 2023. The robots are here: Navigating the generative ai revolution in computing education. *arXiv preprint arXiv:2310.00658*.
- [11] Roto, V., Palanque, P. and Karvonen, H., 2019. Engaging automation at work—a literature review. in *Human Work Interaction Design. Designing Engaging Automation: 5th IFIP WG 13.6 Working Conference, HWID 2018, Espoo, Finland, August 20-21, 2018, Revised Selected Papers 5*, Springer, 158-172.
- [12] SFIA Foundation. 2023. SFIA skills aligned to EU ICT Profiles, SFIA Institute, London.
- [13] Sharp, H., Baddoo, N., Beecham, S., Hall, T. and Robinson, H. 2009. Models of motivation in software engineering. *Information and software technology*, 51 (1). 219-233.

Confidentiality

Interview recordings and transcriptions from the three countries participating in the original data collection, have been collected as a result of the respondents' participation in the study under the ethical protocols applied by each institution. For the purposes of the working group being able to analyse and aggregate the data collected, only the de-identified transcript data will be made available. Any data relating to the interviews will be securely stored.

Permission and Data Storage

The data collection for this multi-institutional multinational research project has been approved by the Auckland University of Technology and Eastern Institute of Technology, New Zealand; Robert Gordon University, Aberdeen; and Uppsala University, Sweden.

For working group participants to gain access to the de-identified transcript data, stored at Uppsala University in a secure repository, a common data sharing agreement stipulating requirements for making use of the data will need to be signed. The transcripts will be stored in the Uppsala Myfiles repository, and made available to members whose institutions have signed the data sharing agreement. For participants unable to gain institutional approval to access the full transcript data, the coded datasheets from each interview, containing only the extracted codes may be shared.

Template

In preparing this data analysis protocol for the working group, the literature on protocol templates has been consulted (Brereton et al., 2008; King et al., 2018; Stol and Fitzgerald, 2014). The excerpt below from Brereton et al., (2008), although positioned at the overall case study level, makes some highly relevant points: 1) no existing data analysis template was found to be available, 2) the two phases of "creating instruments and protocols" and "analysing data" were considered most applicable for this protocol. Therefore some relevant excerpts from the literature precede a determination in this section.

As part of our overall planning of our research program (Brereton et al., 2007) we identified the need to plan a series of case studies. To ensure a consistent planning process, we decided to develop a basic case study protocol template. The aim of the template was to provide a common structure for our case study protocols and guidance on how to construct them. We reviewed some of the well-known case study papers and text book but did not find any existing template, although Yin (2003) presented an example of a case study protocol. For this reason we constructed our own template based on basic case study methodologies described by Eisenhardt (1989), Stake (1995), and Yin (2003).

At a high level most of the approaches to case study conduct are quite similar, in spite of the very different philosophical approaches taken by different case study experts. For example Eisenhardt (1989) who is concerned with using case studies to develop theories suggests the following activities:

- **Getting started** by defining the research question and a priori questions but not defining hypotheses or theory.
- **Selecting cases** by considering a particular population and using theoretical concerns to focus on specific cases.
- **Crafting instruments and protocols** using multiple data collection methods, using qualitative and quantitative data, and preferably multiple researchers.
- **Entering the field** i.e. incorporating field notes with data analyses and using flexible and opportunistic data collection methods.
- **Analysing the data** both within case and across cases.
- **Shaping hypotheses** by iterative tabulation of evidence looking for identified constructs, replication logic across cases, and looking for evidence to explain why relationships exist.
- **Enfolding the literature** i.e. comparing with existing similar and conflicting literature.
- **Reaching closure** using the concept of “theoretical saturation” which says researchers stop looking for more cases/data when they believe more data will only give a marginal improvement to the existing results.

Stake (1995), like Eisenhardt, takes an interpretive approach to case study research, but is particularly concerned with “program evaluation” which involves evaluation of social or education policies.” (Brereton et al., 2008)

Further guiding the protocol was the work of Stol and Fitzgerald (2014),

DATA ANALYSIS AND REPORTING

The collected data were analyzed using qualitative techniques described by Seaman [34]. All interviews were transcribed, resulting in approximately 112 pages of text (A4 format, 10 points font, single line-spacing). The analysis consisted of coding the transcripts using the six themes of our framework (see Section 3) as seed categories. The transcripts were analyzed in parallel by both authors and several analytical memos were written. The memos established an audit trail of the analysis, and facilitated a process of peer debriefing for the researchers.” (Stol and Fitzgerald, 2014)

King et al., (2018) presenting the approach termed “Template Analysis” have observed that: “Thematic analysis is widely acknowledged as an accessible and useful approach to the analysis of rich and meaningful qualitative data—indeed, Clarke and Braun (2013) describe thematic analysis as the ‘basic’ method of qualitative data analysis.

...the principal focus of all thematic analysis approaches is on identifying, organizing, and interpreting themes in detailed qualitative (textual) data to highlight and convey key messages. In this chapter, our focus is on Template Analysis as a particular style of thematic analysis

They also distinguish between *Generic Template Analysis* as a *method*, and as applied within a broader research *methodology* such as grounded theory or Interpretative Phenomenological Analysis

[IPA], and further observe that issues about differing philosophical, theoretical or methodological positions can be better accommodated, as noted below:

Generic styles of thematic analysis can provide researchers more flexibility and adaptability to the particular requirements of their own work—rather than applying a methodology as a whole package.

They also distinguish between inductive and deductive forms of reasoning in template analysis.

Forms of thematic analysis vary in the extent to which they use inductive or deductive reasoning: where different approaches to thematic analysis position themselves on this inductive-deductive continuum depends very much upon the methodological approach being taken...Generic approaches to thematic analysis (such as Template Analysis) can, in contrast, be used from a variety of methodological positions and do not therefore have a single fixed position on this continuum. (King et al., 2018)

The process of coding in template analysis is also not prescriptive over coding approaches and does not stipulate a “sequence of coding levels or an explicit distinction between descriptive and interpretive coding”. In template analysis the general steps in the process are defined by King et al., (2018) as below:

The procedural steps that are characteristically followed in Template Analysis are:

- Familiarization with the data
- Preliminary coding
- Clustering
- Developing the initial template
- Modifying the template
- Defining the ‘final’ template
- Using the template to interpret the data
- Writing-up

For the working group’s study, the data analysis sub-group will conduct the three steps prior to developing the initial template “Familiarization with the data, preliminary coding, clustering”, using a subset of the transcript data, to derive an initial template using a defined spreadsheet for coding each transcript.

The template will be refined as the process progresses as noted by (King et al., 2018):

Revisions might include: re-defining themes to increase or narrow their scope (shown through moving them up or down hierarchical levels), moving themes between clusters, adding new themes—or even entire new clusters—and deleting themes that have become redundant as the template has developed.

Coding cycle one

Version one of the template allowed for coding of each question. For some questions a predefined set of deductive codes may be considered as an initial set of codes for that question, for other questions a more inductive strategy may be more suitable.

Version two of the template again allows for coding of each question, and adding contextual information to support each code. This coding stage [termed coding cycle one] will typically code by phrase excerpted from the transcript relating to each question. The template has also been adapted for programmatic extraction and pooling of data, [By labelling each line of the spreadsheet with the transcript id, overall section/subsection, or interview section,/subsection and an ignore character ‘%’ for template comment data or blank lines].

Note: when inserting additional rows to the template⁵ a coder will need to include a transcript row identifier and section/subsection identifier.

The notes of the data analysis subgroup meeting on 29/05/2024, attached as an appendix, discuss a number of issues arising from a pilot coding exercise conducted by the data analysis subgroup, leading to version 2 of the template, as excerpted below.

Decisions about coding parties therefore recommend that work be distributed amongst the WG members to share the load, as the template is now considered sufficiently stable. Members of the WG will generally code whole transcripts for cycle one.

Note: *When saving the coded transcript into the template, only the first worksheet (with data rather than candidate codes) can be converted into a .csv file for further processing. The first worksheet can be selected for this purpose when saving as a .csv file.*

A master spreadsheet with allocations of transcripts to WG members has been developed and shared.

A supporting pair coding strategy may be adopted, first individually coded, then by comparison between pairs, with the data analysis subgroup aiming to independently code a third of the transcripts if time allows.

Note: *In a meeting of three WG leaders plus one WG member – 03/06/2024, we decided against this supporting pair coding strategy, given remaining available time and preference to focus energies on cycle two coding and broader thematic analysis.*

Coding cycle two

For coding cycle two when the scripts have been developed and the pooled data for each question is considered for coding and extracting themes, it may be preferable to assign questions to sub-groups, to enable greater topic focus, e.g. question 4.3 by the competencies sub-group?

The transcripts will be stored in the Uppsala Myfiles repository, and made available to members whose institutions have signed the data sharing agreement.

The process of thematic analysis will broadly follow the stages recommended by Braun & Clarke (2006) and Cruzes and Dybá (2011), as depicted below, [although the development of a 'model' as proposed by Cruzes and Dybá (2011) may be too ambitious], but tempered by the needs of the working group to adopt a form of "Qualitative pragmatism" (Braun & Clarke, 2021), and apply that within its two cycles of coding and theme development .

Step 1: Data familiarisation.

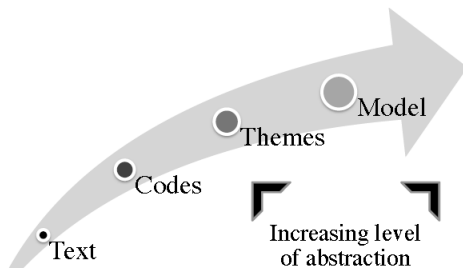
Step 2: Obtain codes.

Step 3: Searching for themes.

Step 4: Review themes.

Step 5: Defining and naming themes.

Step 6: Reporting. (Braun, & Clarke, 2006).



Ex. Cruzes, & Dyba, (2011)

For coding cycle two, the individual coded files uploaded to the Uppsala repository, will be concatenated by a simple script into a common file, to enable further analysis across sections and subsections, when developing codes and themes. The use of filtering and sorting, grouping or extracting within excel files should provide a relatively low-tech common solution giving flexibility for individuals and across groups. The coding process could then be viewed as equivalent to laying the codes out on a large table for sorting and grouping. As proposed to the data analysis subgroup (14/06/2024):

Just thinking ahead to how to analyse the data for each question once pooled, maybe borrowing some basic ideas of 'open card sorting' may provide a pragmatic approach to a method?

That would view each code as a "card" equivalent which could be grouped into semantic clusters by the assigned team for each question or question group. They could also be arranged into hierarchies e.g. "challenges" including at a lower-level "security".

That should give us a workable set of codes and themes for each question derived both inductively and deductively,

refs below and attached:

Rugg, G., & McGeorge, P. (1997). The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems*, 14(2), 80-93.

Fincher, S., & Tenenberg, J. (2005). Making sense of card sorting data. *Expert Systems*, 22(3), 89-93.

As noted in the attached Appendix three the use of memoing is a further recommended practice for groups to share insights.

A draft allocation of sections for analysis is given below:

Subgroup	Section_Id	
Data Analysis	DI-0	Demographic Information:
Data Analysis	S1-0	Section 1: Understanding AI
Practitioners	S2-0	Section 2: Work Engagement and AI
Practitioners	S3-0	Section 3: Socio-Technical Work Dynamics
Competencies	S4-0	Section 4: Professional Development and Competencies
Lit Review	S5-0	5. Closing:

18/06/2024

Appendices

The appendices follow below.

Template Version 2: 02/06/2024

Appendix one

Over
all
secti
on/
subs
cript
ion
fil
e
no.

AI For IT Professionals: Interviews and Transcription

Transformed
Transcript
Questions
Mapped to
Codes

%

%

Test Example [02/06/2024]: Replace bold yellow cells with valid data from transcript

%

%

IA

00

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

Transcript Identifiers	Transcript Identifiers
Participant ID:	A000 [dummy interviewee only]
Interview Name:	AI for IT Professionals
Interviewee Subgroup #	NZ Interviewee
Site:	AUT
Date of Interview:	15/02/2024
Interviewer ID:	IA000
Transcriber:	Tony Clear

Note: use this as filename: e.g. IA000-Coded

Introduction:

1. Welcome and introductions

2. Have you read the information?

a. OK, before we begin the interview itself, I'd like to confirm that you have read and signed the informed consent form, that you understand that your participation in this study is entirely voluntary, that you may refuse to answer any questions, and that you may withdraw from the study at anytime.

3. Briefly explain the purpose of the interview and its significance for the study

a. This study aims to investigate how AI influences the work engagement of IT professionals. By exploring the relationship between work engagement and learning, the study will shed light on the dynamics that drive employee engagement and its connection to the professional development of competencies. The findings will contribute to the design of targeted interventions and strategies that enhance work engagement.

%

%

IA

00

0

0

0

0

0

0

0

0

0

0

0

0

0

Demographic Information

DI-0

Complete Cells From Transcript Data

IA
00
0 S1-
1_0

1. How would you define AI in your professional context?

*consider AI
defs
worksheet
tab - for
second
coding cycle*

IA
00
0 S1-
1_1

IA
00
0 S1-
1_2

IA
00
0 S1-
1_3

*Comment:
might be
two levels,
one
connected
to specific
processes
with specific
tools, like
for A001
NLP versus
more
general
work
processes,
like
summarizin
g or
improving
language*

IA
00
0 S1-
2-0

2. Could you describe any specific AI technologies or tools currently implemented or utilised in your work?

IA
00
0 S1-
2-1

IA
00
0 S1-
2-2

IA
00
0 S1-
2-3

IA
00
0 S1-
3-0

3. Are you aware of any specific technologies or tools that have AI functionality invisibly embedded which are currently implemented or utilised in your work?

IA
00
0 S1-
3-1

IA
00
0 S1-
3-2

IA
00
0 S1-
3-3

IA
00
0 S1-
4-0

4. What do you perceive are the main reasons or motivations behind adopting AI in your organisation or industry?

IA
00
0 S1-
4-1

IA
00
0 S1-
4-2

IA
00
0 S1-
4-3

IA
00
0 S1-
5-0

5. What do you perceive to be the main reasons or motivations behind adopting AI in your practice?

IA
00
0 S1-
5-1

IA
00
0 S1-
5-2

IA
00
0 S1-
5-3

IA
00
0 S1-
6-0

6. How has AI impacted your work processes, tasks, and responsibilities?

IA
00
0 S1-
6-1

--	--

IA
00
0

S1-
6-2

IA
00
0

S1-
6-3

IA
00
0

S1-
7-0

7. What challenges or opportunities related to integrating AI in your role have you experienced? Please elaborate.

IA
00
0

S1-
7-1

IA
00
0

S1-
7-2

IA
00
0

S1-
7-3

%
IA
00
0

S2-0

Section 2: Work Engagement and AI

IA
00
0

S2-
1-0

1. How do you perceive the concept of work engagement in using AI technologies or tools in your role?

Consider worksheet tabs motivation and engagement for second coding cycle

%
IA
00
0

S2-
1-1

IA
00
0

S2-
1-2

IA
00
0

S2-
1-3

IA
00
0

S2-
2-0

2. From your perspective, how has AI influenced your engagement at work?

IA
00
0

S2-
2-1

IA
00
0

S2-
2-2

IA
00
0

S2-
2-3

IA
00
0

S2-
3-0

3. What aspects of AI have contributed to changes in your work engagement? (e.g., autonomy, decision-making, collaboration)

IA
00
0

S2-
3-1

IA
00
0

S2-
3-2

IA
00
0

S2-
3-3

IA
00
0

S2-
4-0

4. What differences in work engagement among colleagues or teams have you observed due to the implementation of AI? Please elaborate.

IA
00
0

S2-
4-1

IA
00
0

S2-
4-2

IA
00
0

S2-
4-3

%
IA
00
0

S3-0

Section 3: Socio-Technical Work Dynamics

%
IA
00
0

IA
00
0 S3-
1-0 1. In your experience, how have AI technologies or tools affected the overall dynamics (e.g., interactions between colleagues, team structures, communication patterns) in your workplace?

IA
00
0 S3-
1-1

IA
00
0 S3-
1-2

IA
00
0 S3-
1-3

IA
00
0 S3-
2-0 2. Have you noticed any changes in the division of tasks or responsibilities within your team or organisation due to AI? If so, please provide examples.

IA
00
0 S3-
2-1

IA
00
0 S3-
2-2

IA
00
0 S3-
2-3

IA
00
0 S3-
3-0 3. How has AI influenced the interaction and collaboration between humans and systems or machines in your work environment?

IA
00
0 S3-
3-1

IA
00
0 S3-
3-2

IA
00
0 S3-
3-3

%
IA
00
0 S4-0 **Section 4: Professional Development and Competencies**

%

%

IA
00
0 S4-
1-0 1. Has the integration of AI impacted a need for professional development?

IA
00
0 S4-
1-1

IA
00
0 S4-
1-2

IA
00
0 S4-
1-3

IA
00
0 S4-
2-0 2. What specific skills or knowledge areas (competencies) have become more important or in demand due to AI? Please elaborate.

IA
00
0 S4-
2-1

IA
00
0 S4-
2-2

IA
00
0 S4-
2-3

IA
00
0 S4-
3-0 3. What training or support have you received in developing AI-related competencies or adapting to the changes induced by AI?

--	--

consider for competencies worksheet tabs: Knowledge, skills, dispositions in task? And vocabularies ? in second coding cycle

IA
00
0 S4-
3-1

Data Analysis Subgroup Notes - Appendix Two

ITiCSE 2024 Working Group 2 – Data Analysis Subgroup meeting notes - 29/05/2024

Attendees: Tony Clear, Monica Divitini, Maria Kasinidou via MS-Teams

Meeting Aim: To confirm data analysis strategy and viability of draft protocol and coding template

Prior Activity Completed: Two coded spreadsheets (by one member) and their tidied transcripts have been shared and reviewed by subgroup members.

Issues Raised in the meeting:

Issue: Consistency of coding the data will be dependent in some cases on technical expertise and/or methodological ability to draw codes from the data.

Resolution: We will need to allow for coders to highlight areas of doubt, and confer with a colleague in the WG with more specific expertise.

Issue: The coding template has maybe focused too much on discrete codes rather than setting them in their context of the interview.

Resolution: Coding at this stage based on phrases [as largely verbatim excerpts] seems to be a partial resolution. [More specific codes e.g. by relating to vocabularies in the extra worksheets in the template will need to be developed as we progress to a further level of abstraction when grouping codes and themes].

Action TC:

But we agreed to add an extra column in the template spreadsheet to add more context excerpts to support the codes, and provide examples of verbatim quotes for later use in publication.

An updated protocol and template will be provided.

Issue: the need to concurrently refer to transcript and template when coding

Resolution: While the transcripts could be extracted into the same spreadsheet as the coding template, this may raise issues for programmatically grouping the codes for each question – with each line sequentially extracted into a question pool, for review and group coding, and developing themes, where that suits, [e.g. professional competencies and student competencies to be coded and thematically grouped by relevant sub groups].

Action TC:

To discuss with Niklas Humble over how best to set up the template and script for extracting codes.

Note: the possibility of using dedicated software such as NVivo was discussed for organising the data, and while desirable, the inability of all team members to participate would limit the applicability of

this approach. So a simple spreadsheet based template and a script to extract question content seemed the most practical option for now.

Issue: the volume of data to be analysed.

Resolution: allocate four transcripts per WG member to code (approx. $4 \times 15 = 60$ transcripts). We moved away from the idea of a more concentrated question group based coding strategy as too complex to coordinate, but could work once coded questions have been collated into a pool by the scripts, as noted above re. concurrent transcript and question coding.

Note: Transcripts will be stored in the Uppsala myfiles directories and access can be arranged.

Action TC:

To set up master sheet with transcript codes and allocations.

Action Transcribers:

To upload transcripts to central directories.

Action Åsa or Niklas:

To arrange access to central directories.

Issue: the quality and consistency of coding.

Resolution: Partly addressed by consistent protocol and spreadsheet template, and pragmatic approach to the analysis, also by a whole group coding demonstration to give a sense for the process.

Action TC:

To coordinate full group coding demo.

A further strategy is for each of the data analysis subgroup members to aim to code 20 transcripts themselves as a cross validation strategy. Even a partial completion of this [ambitious?] target will enable some validation of consistency.

Action Data Analysis Subgroup members:

To code 20 transcripts each.

Action TC: to allocate transcript codes

Note: Decided against this approach in the interests of saving time and focusing energies - 03/06/2024

Issue: Monica's issues [cf. appendix] over the quality and consistency of coding.

Resolution: All good points, partly addressed by improving and making consistent the protocol and spreadsheet template, and a pragmatic approach to the analysis, as noted above. Will also review in more detail and update the protocol e.g. "how to deal with levels of use of technology", or coding

“what others are doing versus what one is reporting doing”. Some of these, or similar or new questions may arise as we progress anyway,

Action TC: to update protocol and template.

DRAFT

Appendix

some notes about the coding

Monica Divitini <divitini@ntnu.no>

Wed 29/05/2024 6:41 PM

Tony Clear <tony.clear@aut.ac.nz>; Maria Kasinidou <maria.kasinidou@ouc.ac.cy>

Dear Tony and Maria

Sorry my comments come very late. I do not expect you to read them before the meeting, we can go through them during the meeting. Just as a reference.

Talk to you soon

Monica

Coding approach

- I looked at the examples sent and I normally use a different coding style, at least it seems, if I understand correctly the examples. I record text that is relevant together with a short code. In this way, it is easier to bring together pieces of text from different parts of the interview that talk about the same thing. For example, for IA001, at the beginning when they talk about the technology they use, it is also discussing their understanding of AI and how they use it.
- The “Complete Cells From Transcript Data” does not seem to collect excerpts, but summaries/interpretations? The risk is that we miss the connection to the text and it is difficult to detect interpretation “mistakes”, especially if we distribute the work.
- The “excerpts” are decontextualized – it might be difficult to get an understanding depending on how people code. This, added with no use of specific codes, leads to text that is difficult to analyze, for example Q7 challenges and opportunities, where it is difficult to understand if they are talking about challenges or opportunities. I would suggest using two codes to specify if we are talking about OPP or CHAL

Process ahead

- A bit cumbersome to work with 2 files
- Too much material to do with Excel?! Is there any tool that we could use?
- The quality of the transcription at points makes it rather cumbersome to do the coding. Should we do something with it?
- We need to “operationalize” more the models...or do we plan to use the models also later in the analysis? If we want to share the coding process with others in the working group, we might need a codebook explaining explicating what is relevant to look for.

- The protocol is very detailed, and it contains a lot of information. I think we should add a detailed instruction guide with practical information, e.g. convention for naming of the files, etc.

More specifically about the coding:

- Should we avoid naming the exact location and rather code at the country level? For this study, it should be enough and we reduce the risk to infringe the anonymity of the respondents. For example, for IA001 I would use NZ rather than Hamilton
- Do we need education level? For example, for IA001 it would be postdoc...but it might be an information that we do not have for all the respondents
- Employment status: for example, IA001 is looking for a job. Is this relevant?
- Experience might be a range rather than an exact number – how do we want to capture it?

About the codes

- When asked about technology, A001 actually seems to talk about use?
- There might be two levels, one connected to specific processes with specific tools, like for A001 NLP versus more general work processes, like summarizing or improving language
- There might be something interesting with what others are doing versus what one is reporting doing. This might tell something about what is perceived as an “acceptable” usage versus something that is a bit borderline.
- It is not very clear the distinction between these two – I intuitively seem to get it, but then in the coding I am getting confused.

4. What do you perceive are the main reasons or motivations behind adopting AI in your organisation or industry?			
	QUALITY	I pass the text to API then API will generate different kinds of embedding to represent this text.	find basic function
	TIME	I haven't used this function very often. Because I think I can read you all the summary by myself, but some people develop a tool so you can upload like 100 PDF. Then he can summarise up to this tool and this AI large language model can summarise each of the PDF and give you very short sentence of each PDF.	summarise text
5. What do you perceive to be the main reasons or motivations behind adopting AI in your practice?			
		review work	review work
		check expression	check expression
		train and fine tune AI models	train and fine tune AI models

Memoing Example - Appendix Three

ITiCSE 2024 Working Group 2 – Memoing example – 18/06/2024

AI Impact and Competencies - Memos:

Consider “memoing” as a process supporting thematic analysis and developing codes and themes, condensing data as below:

Cruzes, D. S., & Dyba, T. (2011). Recommended steps for thematic synthesis in software engineering. In *2011 international symposium on empirical software engineering and measurement* (pp. 275-284). IEEE.

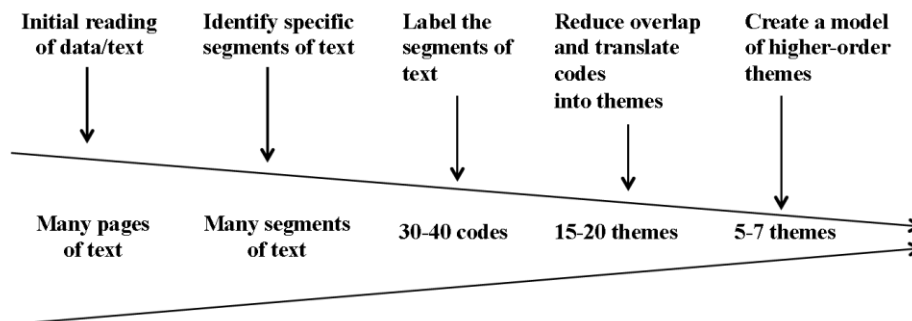


Figure 1. Thematic synthesis process (adapted from Creswell [14]).

Reference for memoing: Birks, M., Chapman, Y., & Francis, K. (2008). Memoing in qualitative research: Probing data and processes. *Journal of research in nursing*, 13(1), 68-75.

Use memos to consolidate thinking and share insights within your subgroup.

Example memo:

Date: 18/06/2024

IA003 S2-3-1

Code: Collaboration [a professional and foundational knowledge area] : Context: I think collaboration among my peers is something that has influenced my engagement because we are talking about it

IA004 S2-2-4

Code: it will isolate the people into working inside their own zone : Context: That interaction will reduce and it will isolate the people into working inside their own zone, and that you know, collaboration with other people. Having a chat about something, discussions and everything that will be reduced in future, I believe that.

Possible contrasting themes - collaboration in use of AI - increasing or reducing?:

IA003 S2-3-5

collaboration in use of AI: for us, it's been like together? Yeah, comparing comparing tools and what you use for what

IA004 S2-2-4

That **interaction will reduce** and it will isolate the people into working inside their own zone, and that you know, **collaboration** with other people. Having a chat about something, discussions and everything that **will be reduced in future**, I believe that.

DRAFT

Protocol References:

- Brereton, P., Kitchenham, B., Budgen, D., & Li, Z. (2008). Using a protocol template for case study planning. *Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering*.
- Kasinidou, M., Kleanthous, S., & Otterbacher, J. (2023). Artificial Intelligence in Everyday Life: Educating the Public Through an Open, Distance-learning Course. *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*.
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology, 18*(3), 328-352.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.
- Clear, T., Cajander, Å., Clear, A., McDermott, R., Bergqvist, A., Daniels, M., Divitini, M., Forshaw, M., Humble, N., Kasinidou, M., Kleanthous, S., Kultur, C., Parvini, G., Polash, M., & Zhu, T. (2024). A Plan for a Joint Study into the Impacts of AI on Professional Competencies of IT Professionals and Implications for Computing Students. In *Proceedings of the 2024 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 2). ACM. <https://doi.org/https://doi.org/10.1145/3649405.3659527>
- Cruzes, D. S., & Dyba, T. (2011). Recommended steps for thematic synthesis in software engineering. In *2011 international symposium on empirical software engineering and measurement* (pp. 275-284). IEEE.
- King, N., Brooks, J., & Tabari, S. (2018). Template analysis in business and management research. *Qualitative methodologies in Organization studies: volume II: methods and possibilities*, 179-206.
- Roto, V., Palanque, P., & Karvonen, H. (2019). Engaging automation at work—a literature review. *Human Work Interaction Design. Designing Engaging Automation: 5th IFIP WG 13.6 Working Conference, HWID 2018, Espoo, Finland, August 20-21, 2018, Revised Selected Papers 5*.
- Sharp, H., Baddoo, N., Beecham, S., Hall, T., & Robinson, H. (2009). Models of motivation in software engineering. *Information and Software Technology, 51*(1), 219-233.
- Stol, K.-J., & Fitzgerald, B. (2014). *Research protocol for a case study of crowdsourcing software development* (Lero Technical Report -TR_2014_03). <https://cora.ucc.ie/handle/10468/7039>