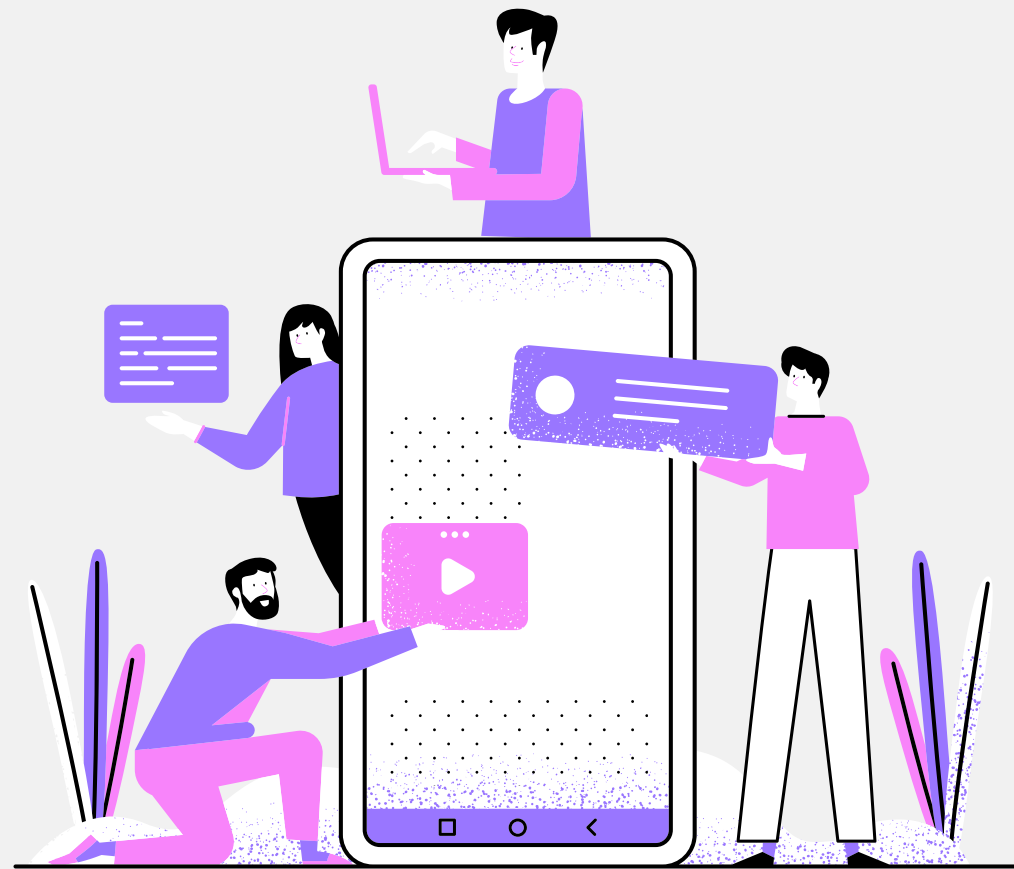# Machine-Assisted Topic Analysis

A human-AI collaboration framework for qualitative research in healthcare

Tris Papakonstantinou

UCL

Evaluation gap

Digital health and public health services

Available apps/services not evaluated

Evaluated apps/services not available

Unregulated industry

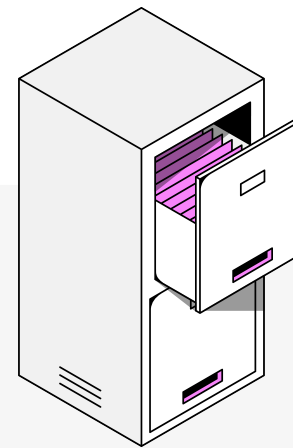User feedback not integrated in time

Adverse effects

Qualitative data in healthcare

Surveys and feedback

Public health
Digital applications
User feedback

File drawer

Vast volumes
Resource-intensive
Only subsets

Too late

Time-sensitive
Dynamic
Adverse effects

Artificial Inteligence

Human Inteligence

Artificial Inteligence	Human Inteligence
Speed	Synthesis
Accuracy	Meaning-making
Cost-effectiveness	Oversight

Machine-Assisted Topic Analysis

1

2

3

4

5

Data collection

Surveys / Reviews
Social media data

Pre-processing

Text cleaning
Tokenization
Stopwords
Stemming

Topic Model

STM
Token distribution
Quotes
Covariates

Qualitative

Topic
Interpretation
Thematic analysis

Service improvement

High-level
Recommendation

NHS Test and Trace

Study 1: Proof of concept

37,914 free text responses of dissatisfied users

Service improvement suggestions

CTAS text analytics

1. Too many contacts and incompatible contact method

- Too many calls received (topic 14)
- Multiple calls per household received (topic 20)
- Multiple calls for a house called reiterating the same process (topic 23)
- Multiple accounts generated for one user resulting in multiple contact attempts through different channels (topic 25)
- Too much information collected through various channels; concern about data privacy (topic 9)
- Contact method incompatible with user's digital skills and technology access (topic 19)

2. Too late

- Conflicting information about isolation dates from different sources of communication (topic 15)
- Users were already self isolating and hence suggested isolation period perceived as incorrect (topic 10)
- Users perceive they were contacted to late and expressed the need to speak to a contact tracer for clarification and more information (topic 4)
- Users contacted too late as they were already isolating since they were contacted personally by the case and hence the service is extending their isolation period (topic 2)
- User contacted by the case personally before contacted to the service, the user expressed their concern that the service is too slow and consequently they could have been infecting others (topic 18)

3. Too complex and rigid

Compexity

- Password setup perceived as excessively complex (topic 12)
- Difficulty completing the form in the context of experiencing Covid-19 symptoms; expressed perceived lack of financial and mental health support (topic 16)

Rigidity

- Lack of option to communicate that the users contracted Covid-19 whilst inpatient in a hospital (Topic 7)
- Covid-19 symptoms list perceived as not exhaustive (topic 21)
- Unable to go back, adjust, and add responses; unable to add information that is not on the form; form perceived as repetitive (topic 13)
- Issues with entering postcodes; ambiguity between home and current address; invalid response when address outside the UK; ethnic backgrounds options not represented (topic 17)
- Unable to input certain information related to specific context: coming into contact with people but wearing PPE (e.g., key worker); unable to convey information that the users perceived as of importance (being on a busy flight, in a shopping centre) (topic 8)
- Dissatisfaction with questions: repeated, irrelevant, and complexly phrased that are difficult to respond when experiencing Covid-19 symptoms (topic 24)

Germ Defence

Study 2: Validation

1472 free text responses

Parallel human-only and MATA analysis

Triangulation

Triangulation results

Human-only themes	Human-only codes	Triangulation with MATA codes	
		Agreement	Complementary
Layout and language style	Clear and simple	A1, A3, A5, A22	
	Not enough information	B9, B11, B14	
	Not streamlined or sophisticated	B5, B2, B6, B8	
	Too repetitive	B3, B5, B10	
	Too simplistic/patronising	B3, B5, B11, B14	
Confidence in how to perform the behaviours	Clear practical advice and troubleshooting is helpful	A2, A6, A9, A10, A12, A13, A20, A24, A7, A21, A14, A18	B12, B9
	Feeling informed and reinforced by reliable sources is empowering	A12, A13, A16, A20, A7, A14, A25, A19, A17	A15
	Inconsistencies undermine confidence		A20, A17
Reducing all or nothing thinking	Trying to perform all the behaviours is exhausting		B12, B6
	Understanding that small changes matter is motivating		A8, A21, A19
	We should act according to risk	B1	A16, A19

Person hours results

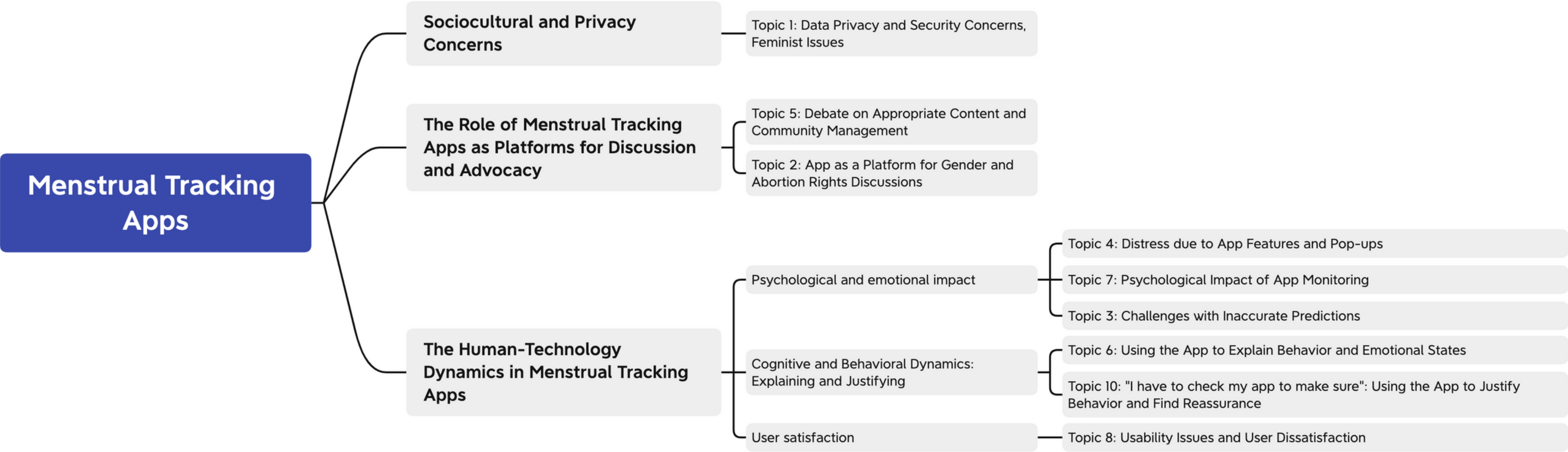|	Human-only	MATA
Preparation	25	8
Coding	95 (13.6 per coder)	0
Validation	14	4
Interpretation	13.5	28 (9 per coder)
Total person hours	**147.5**	**40**

Menstrual Tracking Apps

Sociocultural and Privacy Concerns
- Topic 1: Data Privacy and Security Concerns, Feminist Issues

The Role of Menstrual Tracking Apps as Platforms for Discussion and Advocacy
- Topic 5: Debate on Appropriate Content and Community Management
- Topic 2: App as a Platform for Gender and Abortion Rights Discussions

The Human-Technology Dynamics in Menstrual Tracking Apps

Psychological and emotional impact
- Topic 4: Distress due to App Features and Pop-ups
- Topic 7: Psychological Impact of App Monitoring
- Topic 3: Challenges with Inaccurate Predictions

Cognitive and Behavioral Dynamics: Explaining and Justifying
- Topic 6: Using the App to Explain Behavior and Emotional States
- Topic 10: "I have to check my app to make sure": Using the App to Justify Behavior and Find Reassurance

User satisfaction
- Topic 8: Usability Issues and User Dissatisfaction
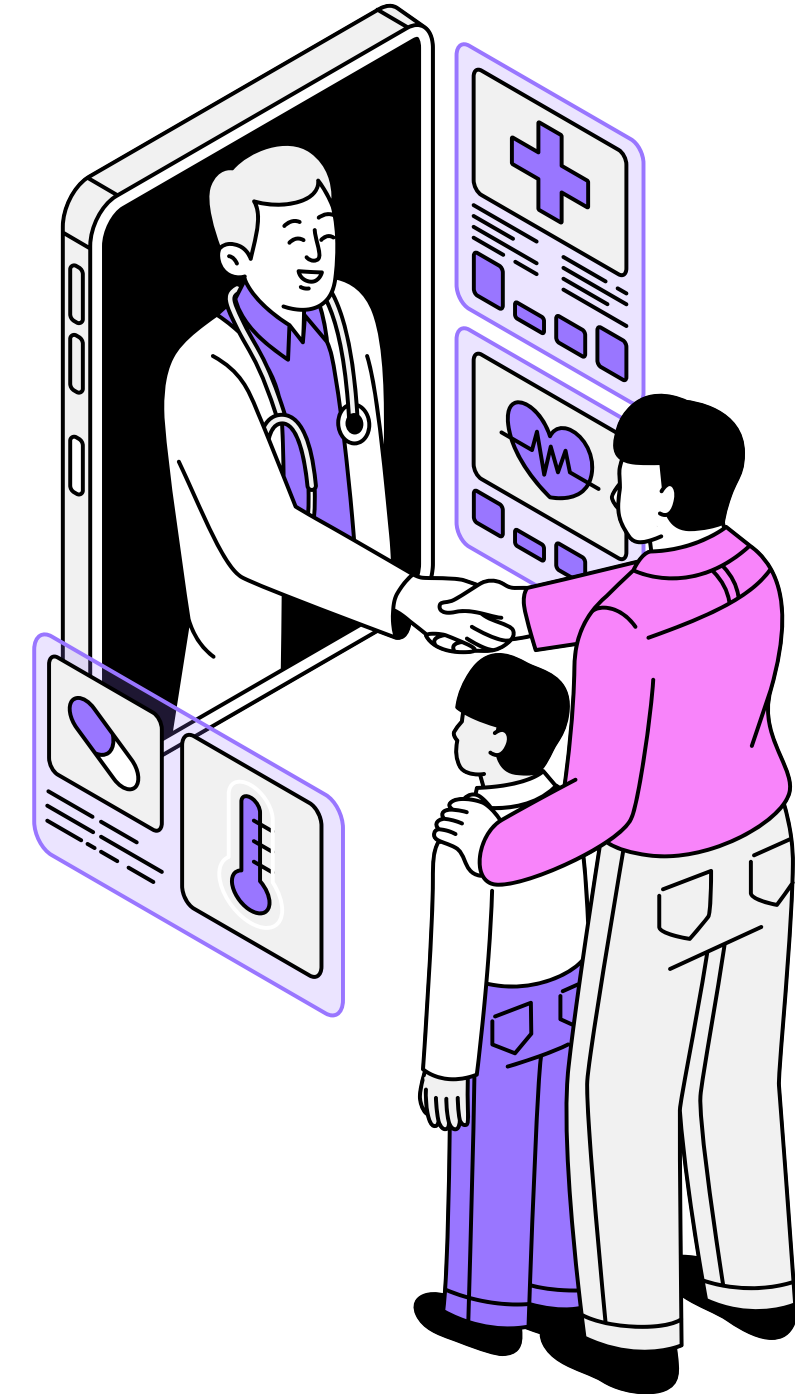
Limitations

Lack of depth and nuance

Individual perspectives get lost

Context comes later

Implications

Enables timely analysis

Low hanging fruit

Generalisable framework

Thank you

Dr Paulina Bondaronek

Collaborators

Dr Chryssa Stefanidou

Dr Tim Chadborn

Dr Lauren Towler

Dr Ben Ainsworth

Prof Lucy Yardley

Prof Richard Amlôt

Jack Bolter

Maria Ceka

	Procedure	Hours (total person-hours)
Preparation	Each of the 7 coders were assigned ~210 participants, whose responses were transferred to the NVivo software package. LT set up the initial coding template based on a codebook developed and validated during previous analyses of Germ Defence data (Morton et al., 2021), previous survey data gathered from website users, and some initial data familiarisation. Six voluntary research assistants (VRAs) were trained by LT in qualitative coding and using NVivo. This involved giving the VRAs an overview of the qualitative process and its aims, the coding process and the meaning of inductive and deductive coding, and previous qualitative analyses from the Germ Defence project.	25
Coding	Analysed using a codebook TA approach, template analysis (Brooks et al., 2015). The data were coded deductively onto the thematic codebook, though some inductive codes were integrated into the codebook upon discussion with the team.	95 (13.6 hours per coder)
Validity checks	The first 50 survey respondents allocated to each trainee coder (23.81% of average total respondents per coder) were cross-checked, and any discrepancies were discussed in subgroups until agreement was reached, under supervision of LT.	14
Interpretation	LT interpreted the findings and created themes from the coding and discussed with the team. LT presented the results to the wider team, and made any adjustments based on discussion with the coders and wider team.	13.5
Total person hours		147.5

	Procedure	Hours
Preparation	Data cleaning and conversion of data to STM format	8
Coding	The structural topic model is run. The model infers the topics from the corpus of text and maps them back to individual documents, which are now assigned topics and represented as a distribution of them.	0
Validity checks	Diagnostic analysis and evaluation of models with 5-40 topics	4
Interpretation	Interpretation of model by describing the topics (stage 1) and creation of broader themes to create the final template (stage 2)	28 (9 hours per coder)
Total person hours		40