Natural Language Processing for Maternal Healthcare: *Perspectives and Guiding Principles in the Age of LLMs*

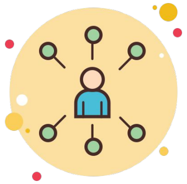
Maria Antoniak

Allen Institute for AI

Research Overview

The goal of my research is to **use NLP and LLM tools creatively to study human experiences**.

I focus my research questions on:



people (e.g., members of online communities, healthcare workers)

subjective cultural concepts (e.g., narratives, values)

person-focused healthcare (e.g., maternal health, pain)

Cultural Analytics

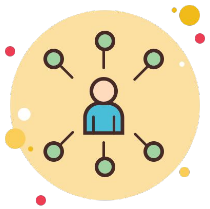
"the computational study of cultural objects, processes, agents"

(Journal of Cultural Analytics)

Traditional unsupervised NLP methods continue to be popular in cultural analytics research.

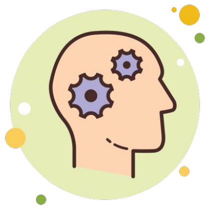
And now, LLMs are unlocking new ways to study human culture.

Computational analysis of storytelling



As a **rhetorical strategy**, storytelling can drive social movements (#MeToo, #BLM), spread misinfo (vaccine side effects), and educate (Ted Talks)

As a form of **self-disclosure**, personal storytelling can strengthen social bonds, build trust in a community, and benefit the storyteller’s wellbeing



As a **sensemaking** strategy, communal storytelling can help individuals and groups learn from their shared (healthcare) experiences

For **NLP** researchers, modeling narratives is a very challenging task that is also important for better understanding our **pretraining** and **prompting** data

Person-Focused Healthcare

Biomedical NLP: information extraction, diagnosis prediction

Instead: text analysis for **online health support communities**, **LLM-based chatbot support**, **sensemaking strategies**

Narrative medicine lies at the intersection of the humanities, clinical practice, healthcare justice: “uses patients’ narratives to promote healing”

Yang, Kraut, Smith, Mayfield, and Jurafsky. “Seekers, Providers, Welcomers, and Storytellers: Modeling Social Roles in Online Health Communities.” CHI, 2019.

Ayers, Poliak, and Dredze. “Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum.” JAMA, 2023.

Rita Charon. “Narrative Medicine: A Model for Empathy, Reflection, Profession, and Trust.”

Brief Intro to NLP

What is natural language processing?

NLP, or **computational linguistics**, uses computational methods to study human language.

This can include **analyzing** human language and **generating** human language.

Methods can include statistics, machine learning, linguistics, and programming.

Examples:

- Google Search
- Alexa
- Email spam filters
- Autocomplete
- ChatGPT

“You shall know a word by the company it keeps.”

The **distributional hypothesis**, popularized by Firth (1957).

In other words: *Which words often appear together?*

We can learn about the **meaning** of a word by studying its **usage**.

In NLP, we often use a word’s usage patterns as a **proxy** for its **semantic relationships** to other words.

What is a language model?

"a model that assigns a probability to sequences of words"

(Jurafsky & Martin, *Speech and Language Processing*)

Given a word sequence, can we **predict** the next word sequence?

What are “large language models”?

Also referred to as **pretrained models** and **foundation models**.

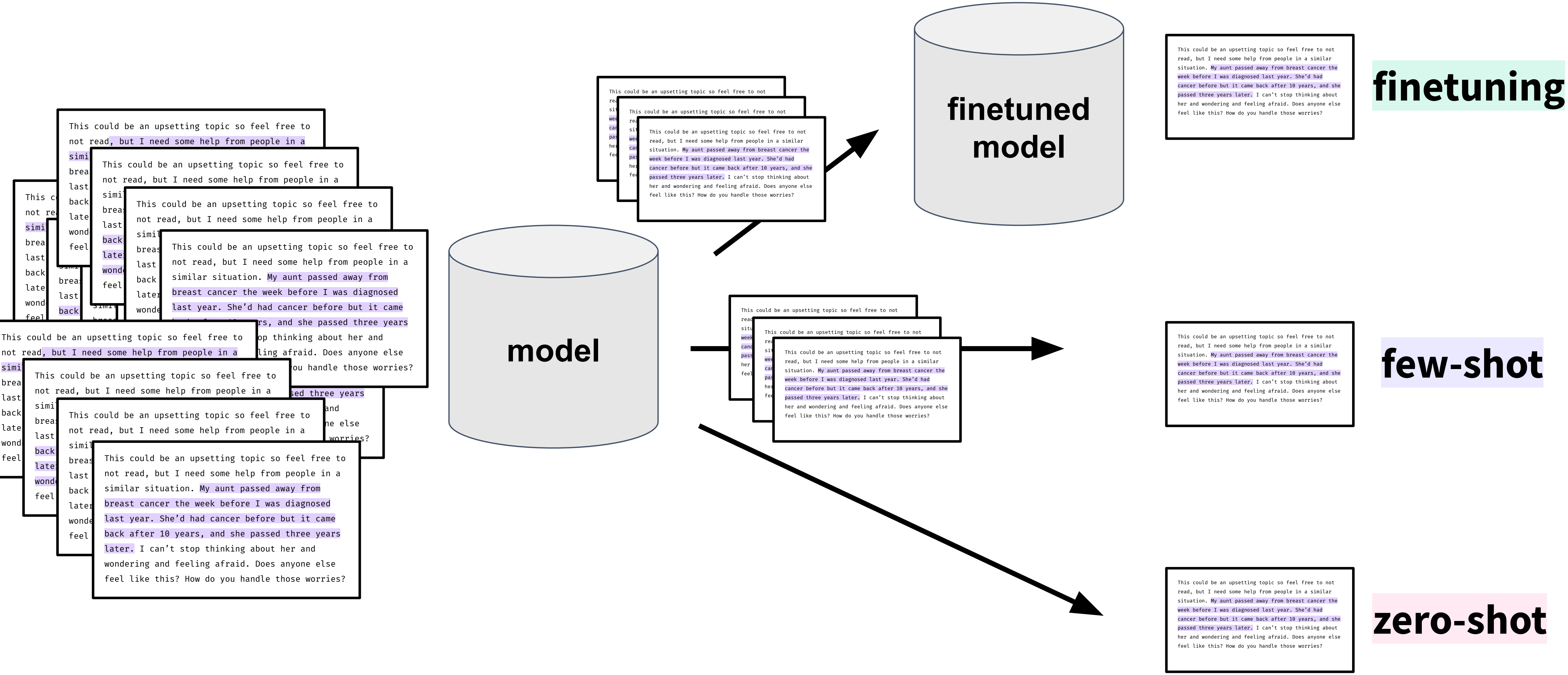
These models rely on **vast** collections of pretraining data.

Common sources include web scrapes, books, Wikipedia, Reddit, and scientific publications.

Some risks to keep in mind

1. Lack of interpretability
2. Giant datasets that are very difficult to document
3. Poor representation and quality for non-English languages
4. Toxicity/bias that is baked into the models

How do we use NLP models?



LLMs for Healthcare

Information Retrieval: Find all the EHRs with patients age 30-35

Prediction: Given this health history, how likely is postpartum depression?

Question Answering: General purpose tools for patients and providers

Linguistic Patterns: How do providers refer to different demographic groups?

→ We could use the same model for all of these tasks

→ Or we could use a smaller model customized to the individual task

Education & Tools

AI for Humanists

www.aiforhumanists.com

Stand-alone tutorials, references, and code for humanities researchers

NLP for Maternal Healthcare: Perspectives and Guiding Principles in the Age of LLMs

Maria Antoniak, Aakanksha Naik, Carla S. Alvarado, Lucy Lu Wang, Irene Y. Chen

FAccT 2024




Maternal Health Equity & NLP

This was a collaboration between the Allen Institute for AI, the University of Washington, UC Berkeley-UCSF, and the Center for Health Justice at the Association of American Medical Colleges (AAMC).

We gathered perspectives on LLMs from many different groups, and we developed guidelines for using NLP for maternal health.

Prior ethical guidelines

Chen et al and Wiens et al. overview the ML/NLP development pipeline and recommendations focused on each pipeline step. Mccradden et al. focus on ethical guidelines for ML-informed clinical decision-making.

The recommendations in these works are based on **literature reviews** and **broad sets of healthcare examples.**

Sendak et al. design guidelines based on the deployment of a specific sepsis-detection machine learning tool, and Petti et al. focused on developing ethical guidelines for the use of NLP and AI methods for early detection of Alzheimer's disease.

Guidelines for NLP for maternal health

We build on prior work that constructs ethical guidelines for machine learning practitioners by **narrowing** our focus.

NLP methods and applications, especially LLMs

Directly solicited perspectives from many affected groups

Focused on a specific healthcare topic: maternal health

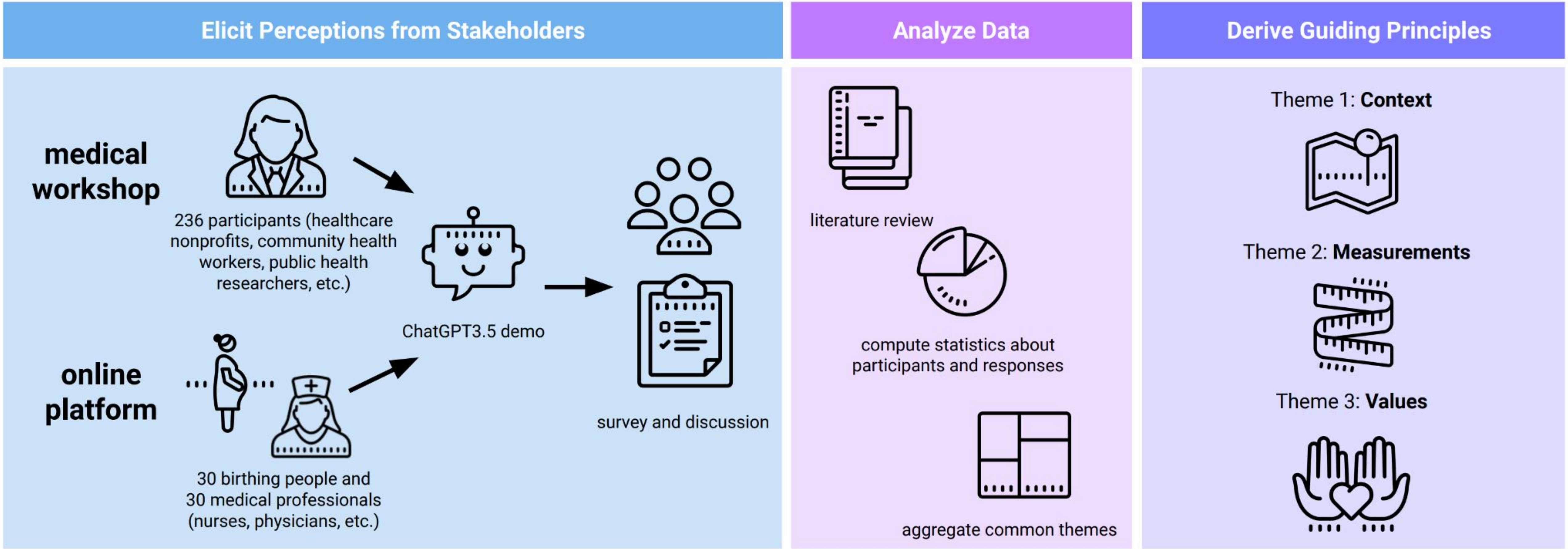
Why maternal health?

1. Many prior research studies and applications of NLP methods focused on maternal healthcare.
2. Pregnancy and childbirth are common events that often comprise a person's sole or major interaction with the healthcare system, increasing the significance and also abundance of perspectives on this topic.
3. Maternal health is a "perfect storm" of healthcare vulnerabilities, with historical biases and power dynamics influencing care.

Who should we ask?

We talked to many different stakeholders, including:

- **nurses**
- **physicians**
- **healthcare management**
- **birthing people**
- **pharmacists**
- **population/public health researchers**
- **community health workers**
- **local/state public health**



Elicit Perceptions from Stakeholders
Analyze Data
Derive Guiding Principles
medical workshop
236 participants (healthcare nonprofits, community health workers, public health researchers, etc.)
online platform
30 birthing people and 30 medical professionals (nurses, physicians, etc.)
ChatGPT3.5 demo
survey and discussion
literature review
compute statistics about participants and responses
aggregate common themes
Theme 1: Context
Theme 2: Measurements
Theme 3: Values

MATERNAL HEALTH EQUITY WORKSHOP

FROM STORY TO DATA TO ACTION | MAY 18, 2023

AAMC CENTER FOR HEALTH JUSTICE

LIVE VIRTUAL RECORDING YOLANDA LIMAN Drawing Change

Ask a question

Think of a situation when a person might have questions about maternal health.

Ask a question related to this situation. Then click Submit to generate a response.

For example:

- *What is the difference between preterm labor and Braxton Hicks contractions?*
- *What is the newest research about the relationship between air pollution and preterm labor?*

Please ask at least five new questions. You can edit your question and click Submit to generate a new response.

⚠ Don't include any private information like real names or dates. ⚠

⚠ Always check with a healthcare professional before making any healthcare decision. ⚠

Question *

What is the difference between preterm labor and Braxton Hicks contractions?

Submit

OpenAI Response:

Preterm labor and Braxton Hicks contractions are two different conditions related to contractions during pregnancy: 1. Preterm labor: This refers to the onset of regular contractions before 37 weeks of pregnancy. It is also known as premature labor or premature

Tell us what you think

What do you think about this response? Check all that apply.

- ☐ This response is accurate.
- ☐ This response is trustworthy.
- ☐ This response is useful.
- ☐ This response is up to date.
- ☐ I'm not sure what to think about this response.

Notes

Submit

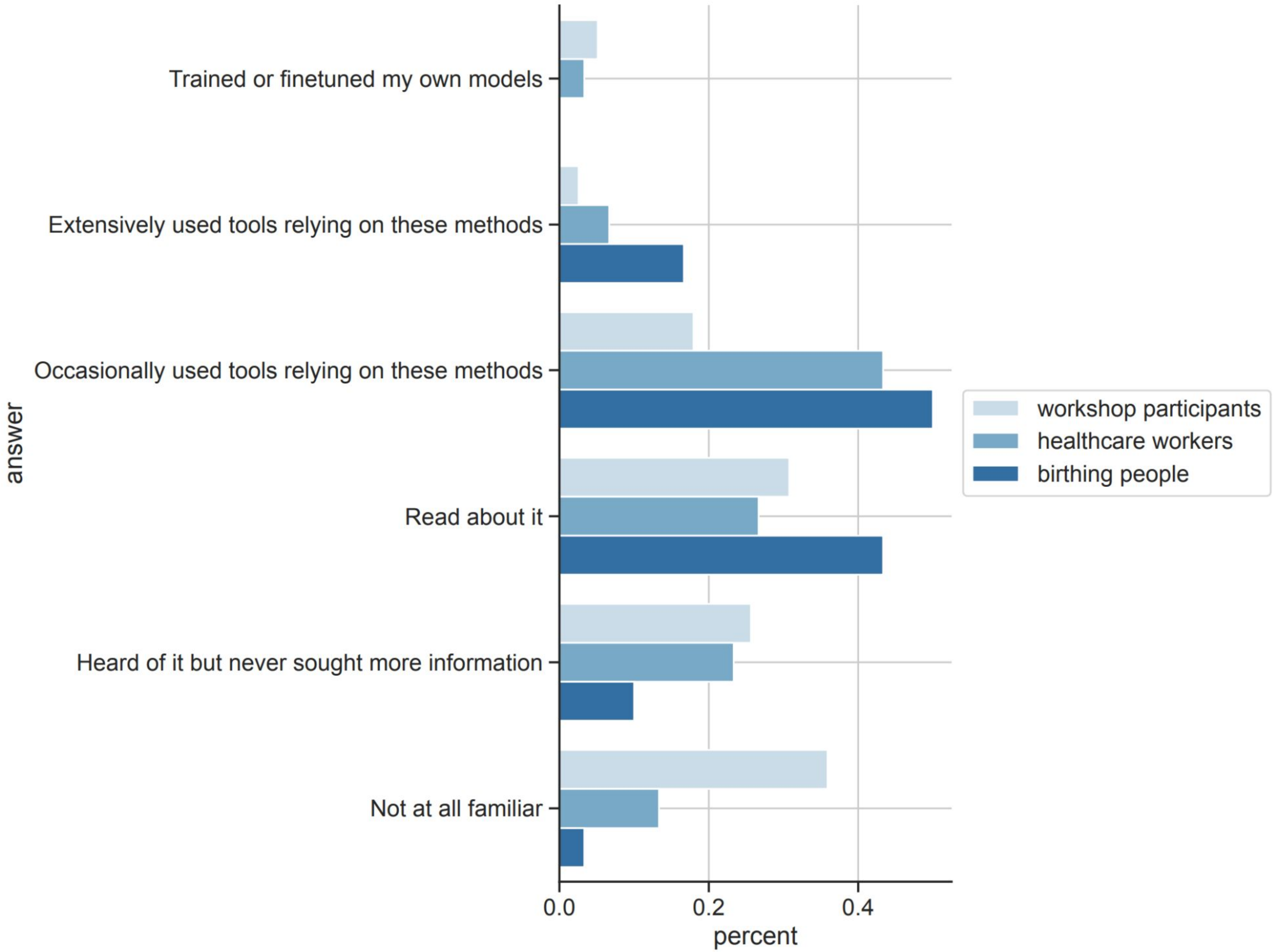
1. How was your experience with the chatbot? What stood out to you about the responses?

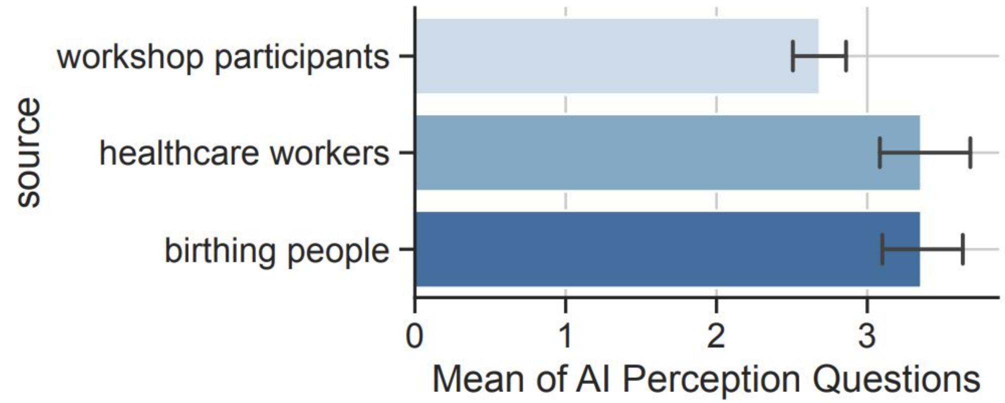
2. What are your dream NLP tools for maternal health? What tools should never be built?

3. Which maternal health stakeholders (birthing people, nurses, doulas, etc.) would benefit or be hurt by NLP tools?

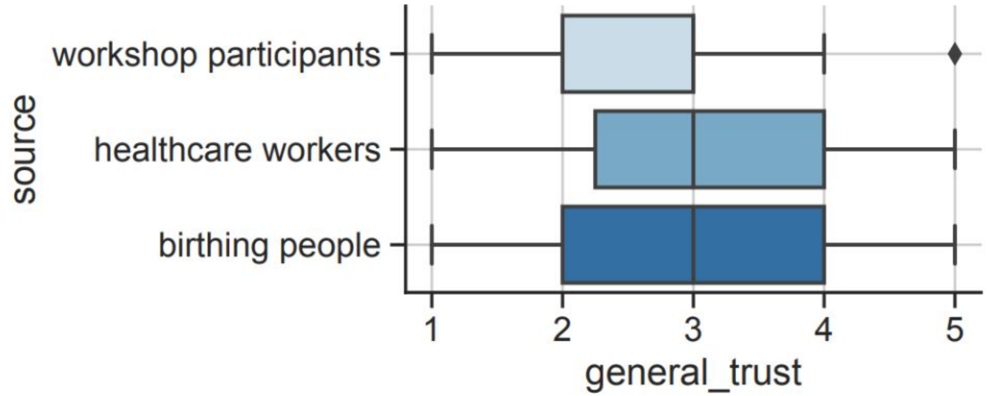
4. What principles should guide the use of NLP for maternal health? What should be the goals and guardrails?

Cohort	Race/Ethnicity	Age	Highest Education	Gender
Workshop Participants (N = 39)				
38% community nonprofits 27% pop./public health research 24% comm. health/promotara 24% local/state public health 19% healthcare management/admin 16% healthcare services researcher 11% other perinatal healthcare provider 11% other non-healthcare perinatal support 8% doula 8% non-perinatal healthcare provider 13.5% *all other groups*	41% African-American/Black 41% White 16% Hispanic/Latino/a/x 5% South Asian 19% *all other groups*	35% 35-44 30% 25-34 19% 45-54 11% 55-64 5% 65-74	38% MS, MPH, etc. 30% PhD 24% BA, BS, etc. 11% *all other groups*	92% women 5% men 3% no answer 0% non-binary
Healthcare Workers (N = 30)				
20% nurse 17% pharmacy 10% physician 10% medical tech 10% medical assistant/aide 10% research 23% *all other groups* 33% have worked in maternal/perinatal healthcare	57% White 23% African-American/Black 7% East Asian 7% Southeast Asian 9% *all other groups*	33% 35-44 30% 25-34 10% 18-24 3% 65-74 3% 55-64	50% BA, BS, etc. 17% MS, MPH, etc. 17% Trade School 10% MD, DO, etc. 7% Community College 6% *all other groups*	79% women 21% men 0% non-binary
Birthing People (N = 30)				
20% have worked in healthcare 7% have worked in maternal/perinatal healthcare	73% White 20% Hispanic/Latino/a/x 17% African-American/Black 12% *all other groups*	53% 25-34 37% 35-44 10% 65-74	33% BA, BS, etc. 30% High school or GED 13% Community College 10% MS, MPH, etc. 10% Trade School 7% PhD 3% Prof. Degree	97% women 7% non-binary 0% men

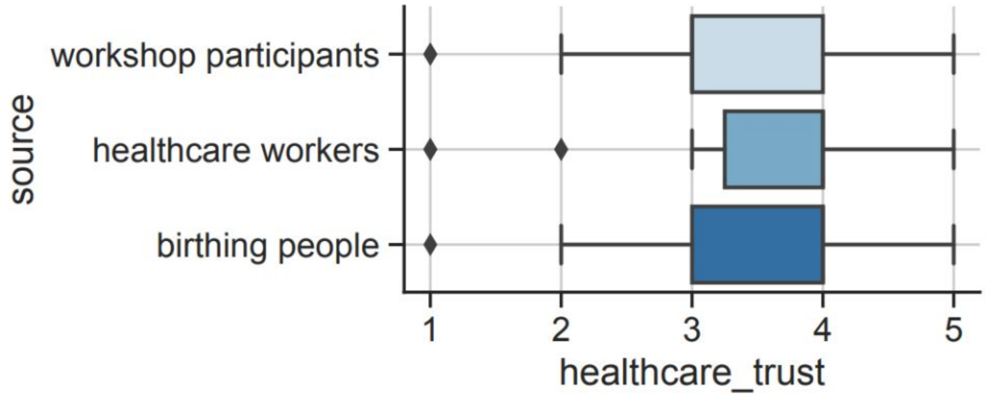

Trained or finetuned my own models
Extensively used tools relying on these methods
Occasionally used tools relying on these methods
Read about it
Heard of it but never sought more information
Not at all familiar
answer
0.0
0.2
0.4
percent
workshop participants
healthcare workers
birthing people



workshop participants
healthcare workers
birthing people
source
0
1
2
3
Mean of AI Perception Questions

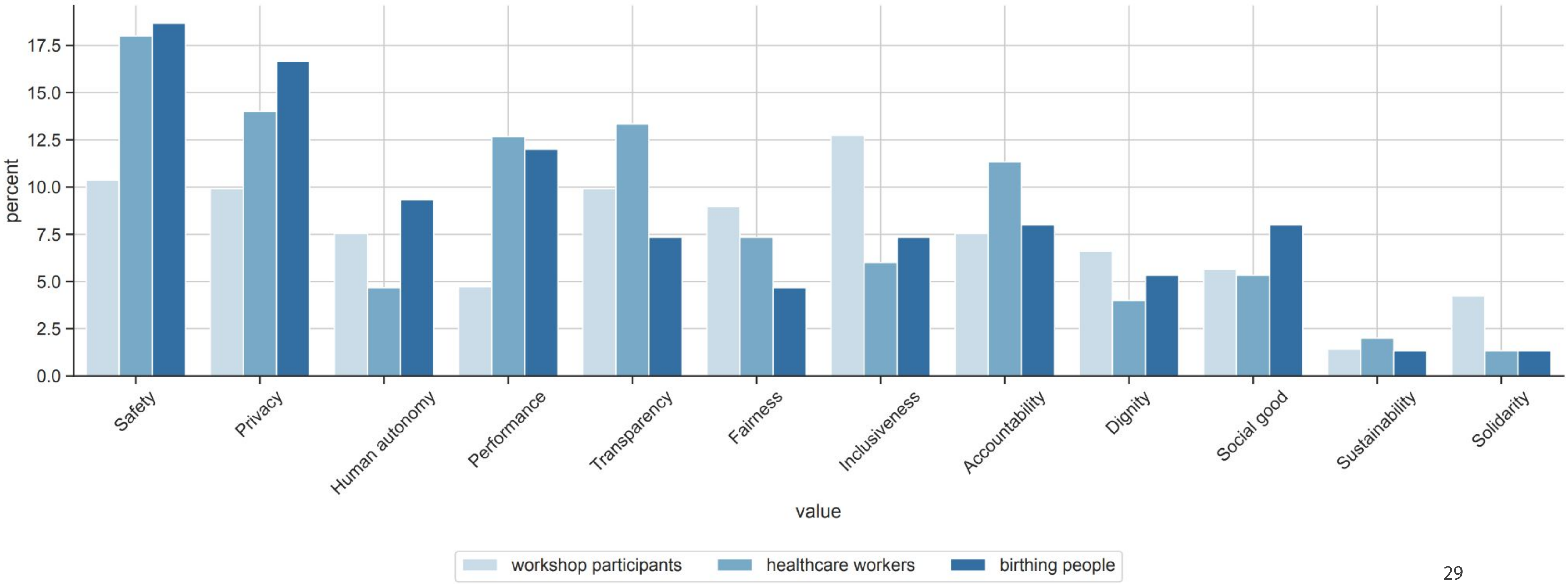


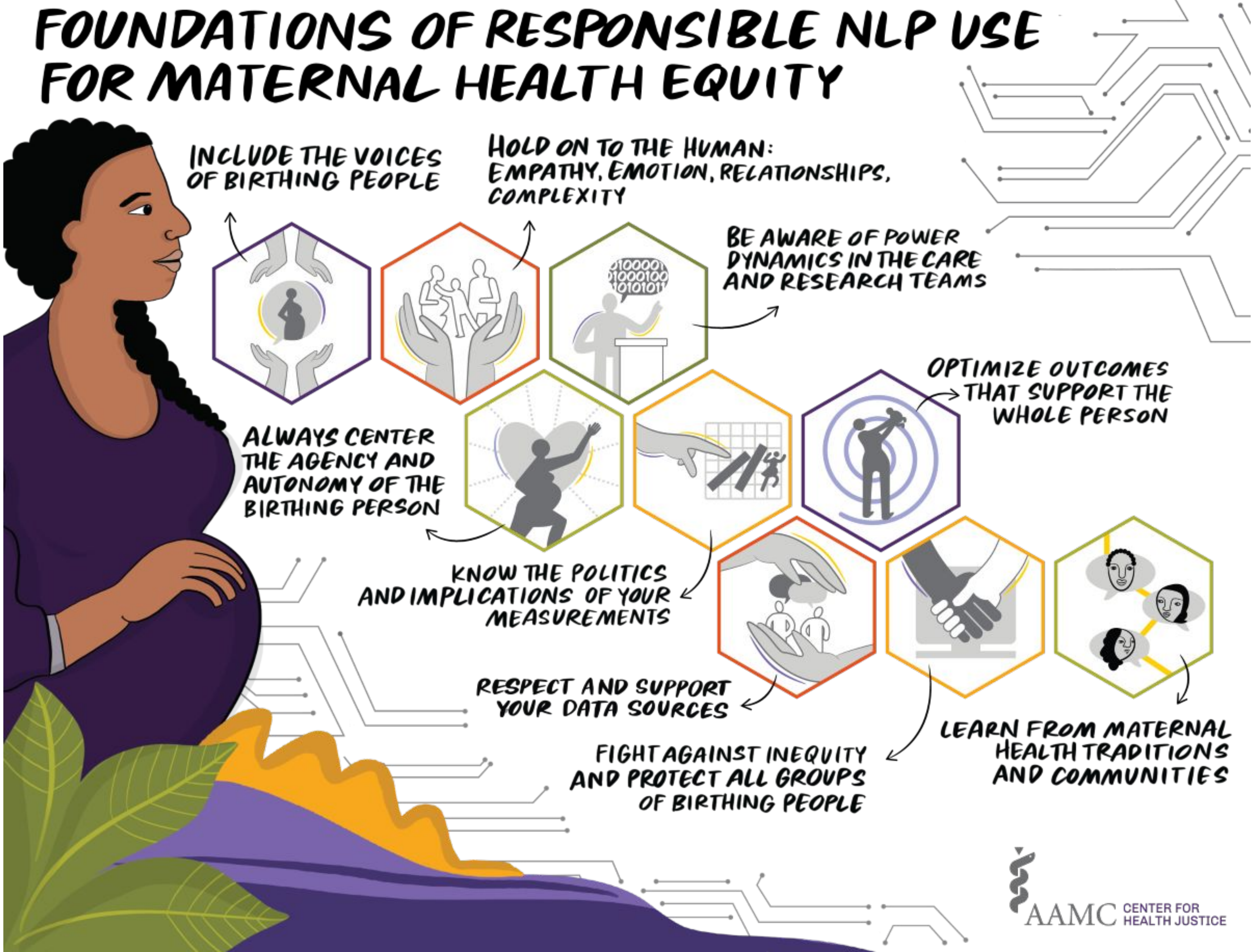
workshop participants
healthcare workers
birthing people
source
1
2
3
4
5
general_trust



workshop participants
healthcare workers
birthing people
source
1
2
3
4
5
healthcare_trust

Different groups prioritize different values




FOUNDATIONS OF RESPONSIBLE NLP USE FOR MATERNAL HEALTH EQUITY
INCLUDE THE VOICES OF BIRTHING PEOPLE
HOLD ON TO THE HUMAN: EMPATHY, EMOTION, RELATIONSHIPS, COMPLEXITY
BE AWARE OF POWER DYNAMICS IN THE CARE AND RESEARCH TEAMS
ALWAYS CENTER THE AGENCY AND AUTONOMY OF THE BIRTHING PERSON
OPTIMIZE OUTCOMES THAT SUPPORT THE WHOLE PERSON
KNOW THE POLITICS AND IMPLICATIONS OF YOUR MEASUREMENTS
RESPECT AND SUPPORT YOUR DATA SOURCES
FIGHT AGAINST INEQUITY AND PROTECT ALL GROUPS OF BIRTHING PEOPLE
LEARN FROM MATERNAL HEALTH TRADITIONS AND COMMUNITIES
AAMC CENTER FOR HEALTH JUSTICE

Guidelines: Recognizing contextual significance

Be aware of power dynamics in the care team.

“If mothers relied too heavily on AI instead of seeking professional help then the nurses and doulas may see fewer people seeking care” – *Birthing Person 11*

Know the politics and implications of your measurements.

“I get worried about what’ll happen when insurance companies think there’s cost savings to using these tools … they can cut corners, have more profit … given the incentives in … the healthcare system” – *Workshop Participant 1*

Learn from maternal health traditions and communities.

“It [(AI)] should follow the same guidelines that medical professionals do: ‘Do no harm.’ ” – *Healthcare Worker 6*

Guidelines: Holistic measurements

Optimize for outcomes that support the whole person.

“It’s not just about the outcome, right? It’s about the whole experience” – *Workshop Participant 3*

Protect all groups of birthing people.

“Everything that has the potential to benefit also has the potential to hurt” – *Workshop Participant 4*

Hold onto the human: empathy, emotion, relationships, complexity

“your own judgment and/or human compassion components, wisdom, experience [are a] part of the care” – *Workshop Participant 7*

Guidelines: Who and what is valued

Include the voices of those seeking care.

“there needs to be community input, there needs to be representation in creating these tools” – *Workshop Participant 8*

Always center the agency and autonomy of the birthing person.

“[Disclose explanations so that] the person using it knows what kind of information or advice it can and cannot give” – *Birthing Person 2*

Respect and support your data sources.

“The principle that should guide these tools is to… have transparency for its sourcing of data” – *Healthcare Worker 2*

Perceptions of risks and benefits of LLMs

"I wish it [(AI)] had been around when my son was a newborn so I could interact with it during late night feedings. One, to give me something to do, and two, **to make me feel like I wasn't alone"** – *Birthing Person 12*

"Just to have something there to ask questions to when I am not sure as to what is happening or when I need a quick answer" – *Birthing Person 9*

LLMs as part of an information ecosystem

“Often times people will **google questions and try to sift through all the search results** to find the applicable information. AI could make that a much more efficient process” – *Healthcare Worker 9*

“It would be nice to be able to type in **worries and fears** to an AI bot and get accurate answers **instead of going down rabbit trails on search engines** that leave you more concerned” – *Birthing Person 11*

“People already diagnose themselves on WebMD. Providing more tools can be **dangerous”** – *Birthing Person 4*

Next Steps

Examining our collected query data!

Small but high quality query dataset written by diverse stakeholders and professionals about a specific healthcare topic.

Expand this data? Combine with another dataset? Measure empathy in responses?

THEME	COUNT
pregnancy	45
best practices	43
symptom interpretation	38
labor	34
risks	26
pain	21
weighing options	19
definitions and information	15
finding resources	15
inequity	15
dismissal	14
care team	13
race	13
what to expect	12
is this normal	11
breastfeeding	10
fear	10
postpartum	9
planning communication with providers	8
finances	7
first time pregnancy	7
blood pressure	5

Themes: *Seeking definitions and information*

SITUATION	QUESTION
A patient is approaching her due date and her doctors has recommended an induction because of her age if she does not go into labor spontaneously. She has had 3 vaginal deliveries with no complications.	**What is** a labor induction?
An analyst is interested in learning about different maternal health indicators.	**How** do Nulliparous, Term, Singleton, and Vertex Cesarean Birth Rates inform maternal health?
A patient is worried about not knowing the difference between preterm labor and or Braxton Hicks contractions. They're experiencing some contractions but they aren't sure what to do next.	**What is the difference** between preterm labor and Braxton Hicks contractions?
Patient doesn't know whether they want a natural birth or induced	**How often** is pitocin used
A researcher is interested in learning more about racial inequities in maternal health.	**What are** racial inequities?
Women entering prenatal care late	How do we **define** late entry to prenatal care for a pregnant patient?
entering late into prenatal care	how do you **define** late entry to prenatal care?
A pregnant patient is concerned about back pain she is experiencing and wants to know if it is sciatic pain or contractions.	**What is the difference** between sciatic back pain and contractions?
When they are having vaginal discharge. Often women are concerned about whether it is was normal or not.	**What is the difference** between vaginal discharge in pregnancy and rupture of membrane?

Themes: *Planning communication with provider*

SITUATION	QUESTION
A first time mother is worried about pelvic pain but has been told that this is just something all pregnant mothers experience. It has become difficult for her to sleep at night and feel rested for work the next day. **She is at her appointment with a new provider and considering sharing what she has been feeling but is hesitant because she has been ignored by providers in the past.**	Is pelvic pain normal for first time mothers?
A patient is expressing concerns about a procedure and its associated risks. They have had negative experiences before and **wants to make sure she is making an informed decision about her options. During this process, she feels like she is not being listened to.**	**What are the pros and cons of each of the c-section option that you are presenting to me?**
A middle eastern pregnant patient is nervous because her doctor is dismissive of her feeling and symptoms.	I am middle eastern and my doctor keeps dismissing my symptoms, **how do I make my doctor take me seriously?**
Karen is 24 weeks gestation, and reports to her OB/GYN that her mouth is bitter and she is not able to eat. While the OB/Gyn is concerned about Karen's not gaining weight according to her gestational age, he is not addressing Karen's concerns.	**How can I elevate my concerns as priorities?**
Patient is 39 weeks pregnant and provider wants to induce labor. Patient wants to wait until 41 weeks for spontaneous labor.	I want to wait for spontaneous labor until 41 weeks and my provider wants to induce me at 39 weeks. I do not want an induction. **What should I say to my provider.**

Narrative Paths and Negotiation of Power in Birth Stories

Maria Antoniak, Karen Levy, David Mimno

Computer-Supported Cooperative Work (CSCW), 2019

A maternal mortality crisis in the U.S.

In the U.S., rates of pregnancy-related deaths and complications are rising but many of these are potentially **preventable.**

Berg et al. “Preventability of pregnancy-related deaths: results of a state-wide review.” *Obstetrics & Gynecology*, 2005.
Creanga et al. “Pregnancy-related mortality in the United States, 2006-2010.” *Obstetrics & Gynecology*, 2015.

Postpartum depression affects 6-13% of people after childbirth, but it can be **prevented or mitigated.**

Lavender & Walkinshaw. “Can midwives reduce postpartum psychological morbidity? A randomized trial.” *Birth*, 1998.
Stewart & Vigod. “Postpartum Depression.” *New England Journal Medicine*, 2016.

Research Questions

Negative birthing experiences are associated with a **loss of control**.

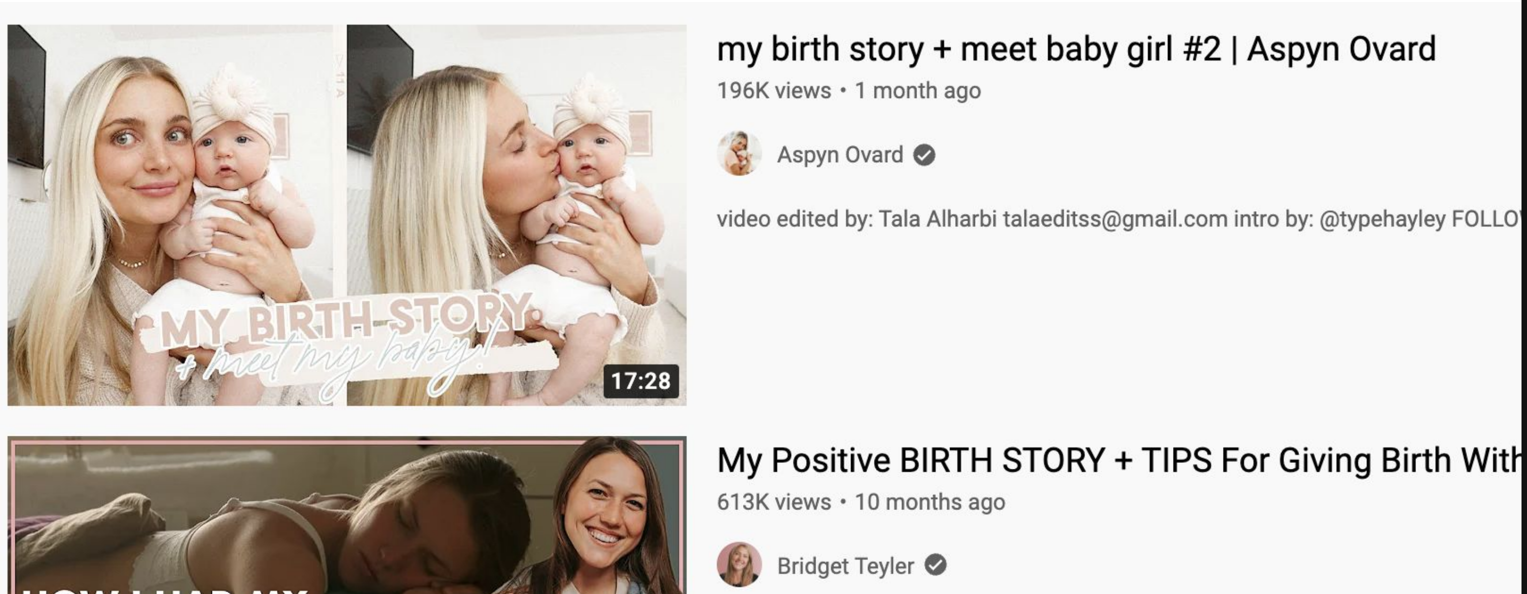
NLP methods let us explore birthing people's voices **at scale**.

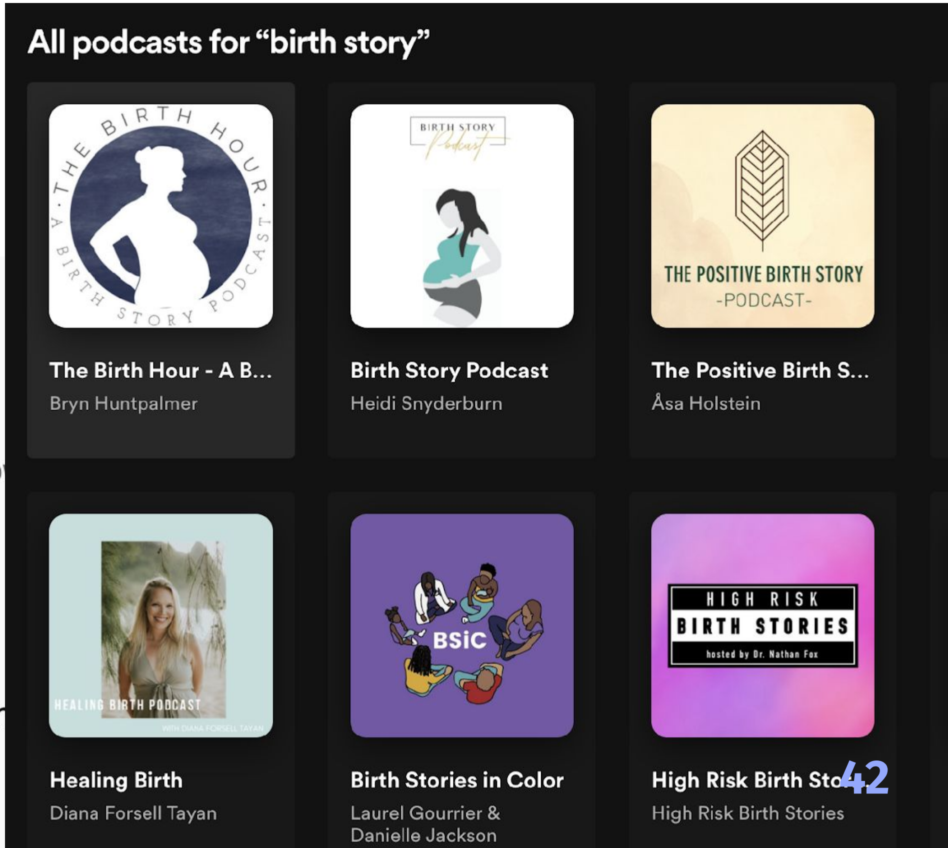
We use **quantitative measurements** to add **qualitative detail** to prior work on birthing experiences and decision making.

Carma L Bylund. "Mothers' involvement in decision making during the birthing process: a quantitative analysis of women's online birth stories." *Health Communication*, 2005.

Online Birth Stories

Detailed personal narratives of giving birth that have been shared throughout history and are very popular (but under-studied) online.





Research Questions

What narrative **pathways** are described in birth stories?
How are **expected** or **unexpected** event sequences **framed**?

Whom do the authors frame as holding **power**?
What **actions** do these actors take that makes them powerful?

r/BabyBumps: Post Titles

- Fiona's **Birth Story**: The best laid plans go awry? Long read
- Thomas Berry **Birth Story** (Planned C-Section w/Complications)
- John’s **Birth Story** - (planned for natural birth, got an ER c-section)
- Emily Rose's **Birth Story**
- My **birth story**! [home, unmedicated, midwife-assisted water-birth]
- Damian's Home**birth Story**!
- Alice's **birth story**. (long, didn't go as planned)

“I’ll begin just by saying that this subreddit has taught me so much in preparation for my first birth. I gained a lot of insight into other people’s experiences, and I loved reading all the stories, whether positive or negative. I hope that sharing my story helps someone in the same way.

“It was 1AM, April 5th. I was experiencing some light cramping and couldn’t fall asleep. This was pretty normal since 39 weeks, so I wasn’t worried. My husband woke up and was rubbing my back when I felt a pop. I told my husband, and he thought it was just my back cracking, but I realized that it was my water breaking. I told my husband that it looks like it’s time to go to the hospital! It was 3:30AM at that point.

“We headed to the hospital and got admitted. The nurse checked me, which was extremely painful, to the point that I wanted to cry. She said that I wasn’t even one centimeter dilated, which didn’t surprise me because I was supposed to be induced on April 8th. They got me started on pitocin to help start contractions and moved me to the delivery room at 4:45AM... "

r/BabyBumps: Ethical Tensions

Need to consider:

- Privacy tradeoff
- Setting and context
- Who is most vulnerable

Our decisions for this dataset:

- Paraphrase all quotes
- No data release
- Share our findings on r/BabyBumps

Narrative Patterns: Topics Over Time

We trained a latent Dirichlet allocation (LDA) **topic model** on the stories.

LDA is an unsupervised, generative model that produces a **probability distribution over words** for each topic and a **probability distribution over topics** for each document.

LDA continues to beat other topic modeling methods (including LLM-based methods!) in human coherence tests.

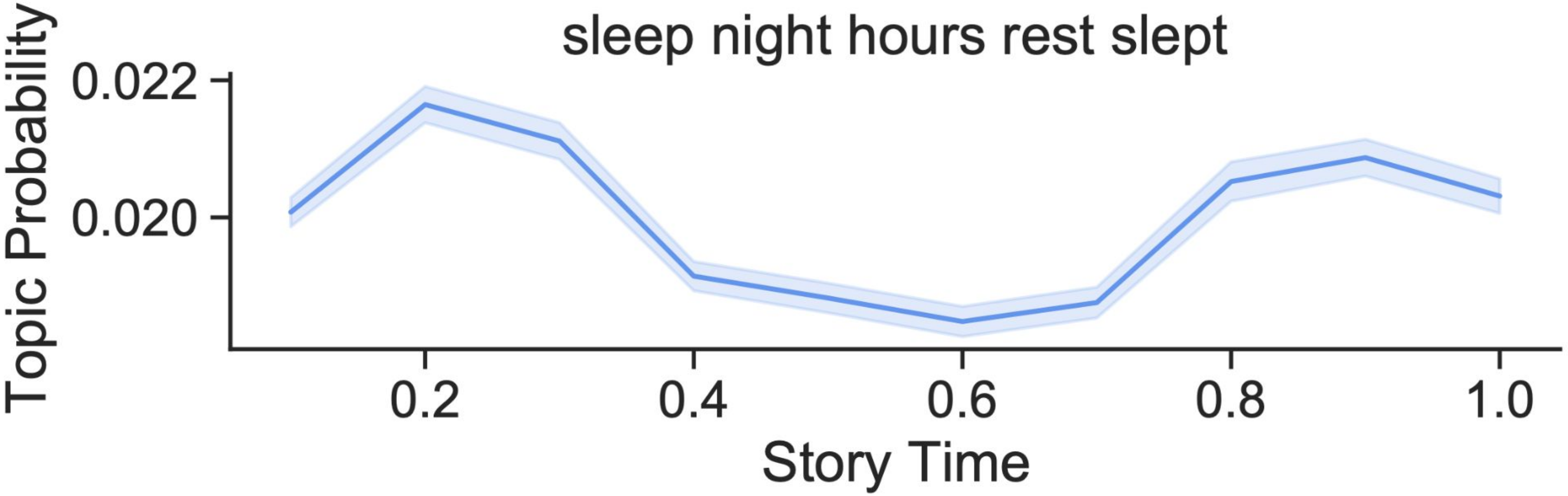
Blei et al. "Latent Dirichlet allocation." *JMLR*, 2003.

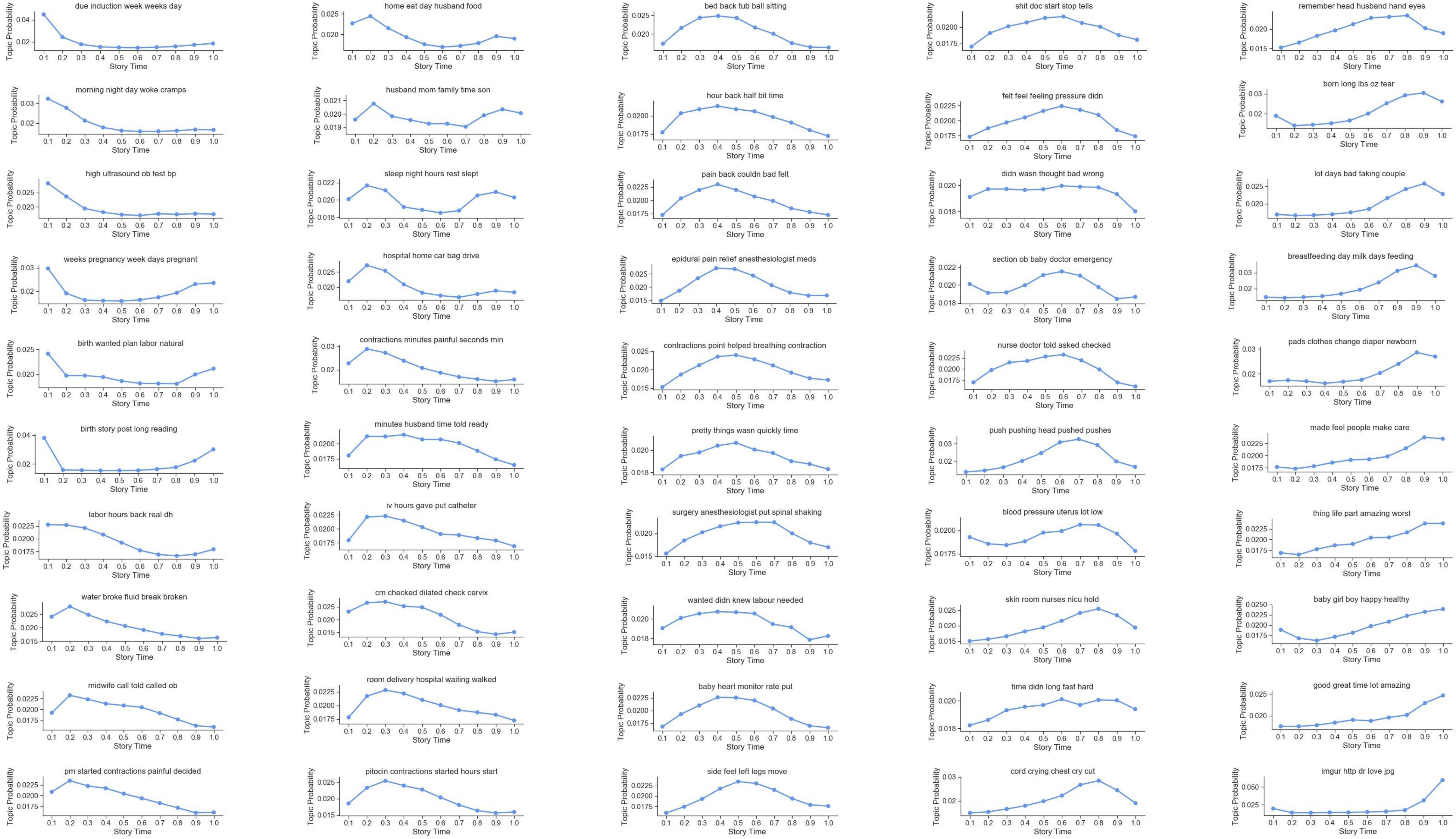
Harrando et al. "Apples to Apples: A Systematic Evaluation of Topic Models." RANLP, 2021.

Hoyle et al. "Are Neural Topic Models Broken?" Findings of EMNLP, 2022.

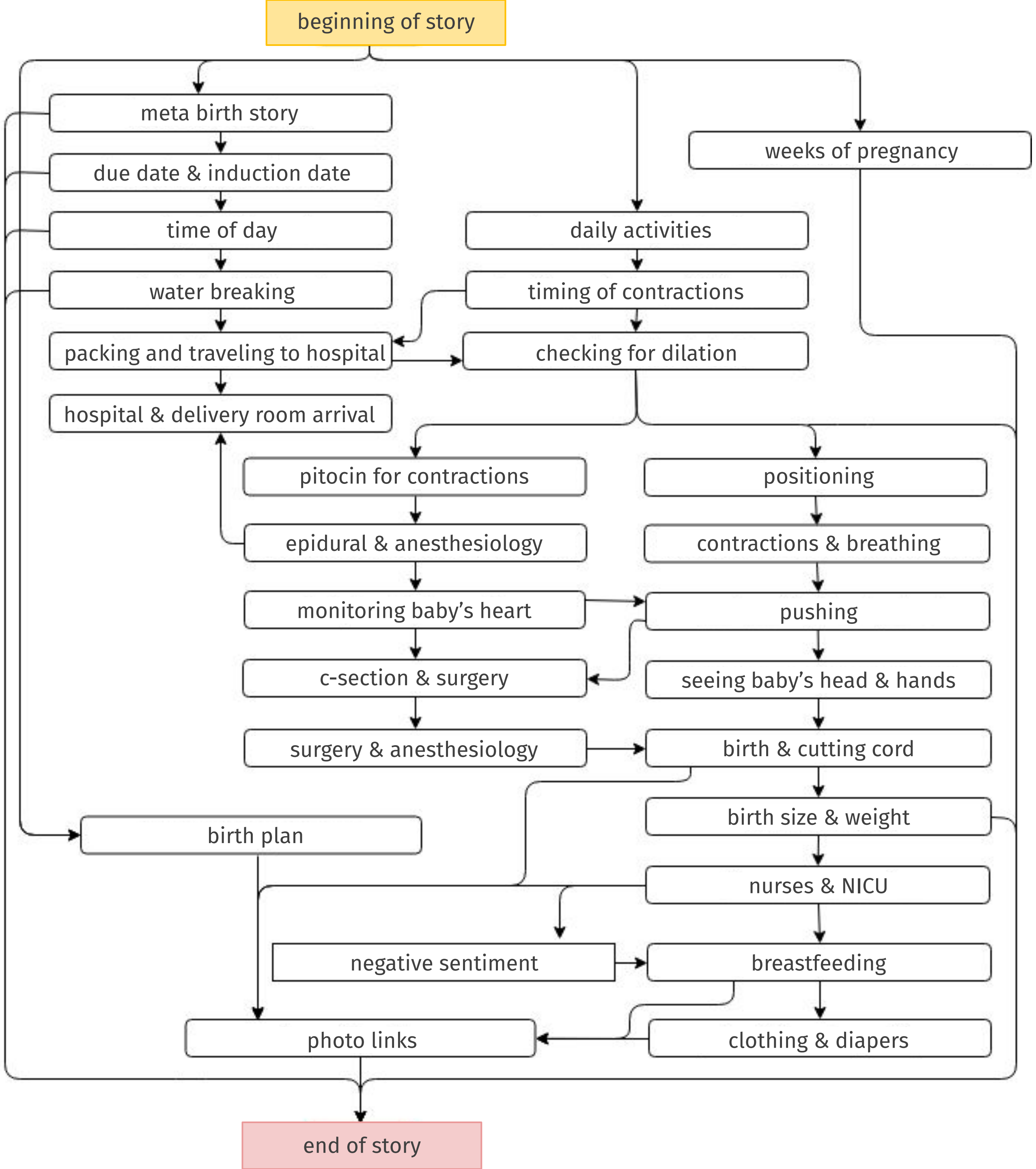
Story Time		Topic 1	Topic 2	Topic 3	Topic 4	
0.10	I'll begin just by saying that this subreddit has taught me so much…	[0.01,	0.20,	0.03,	0.56,	…]
0.20	It was 1am, April 5th. I was experiencing some light cramping…	[0.23,	0.11,	0.02,	0.01,	…]
0.30	We headed to the hospital and got admitted. The nurse checked me…	[0.03,	0.01,	0.32,	0.04,	…]
…	…	…				

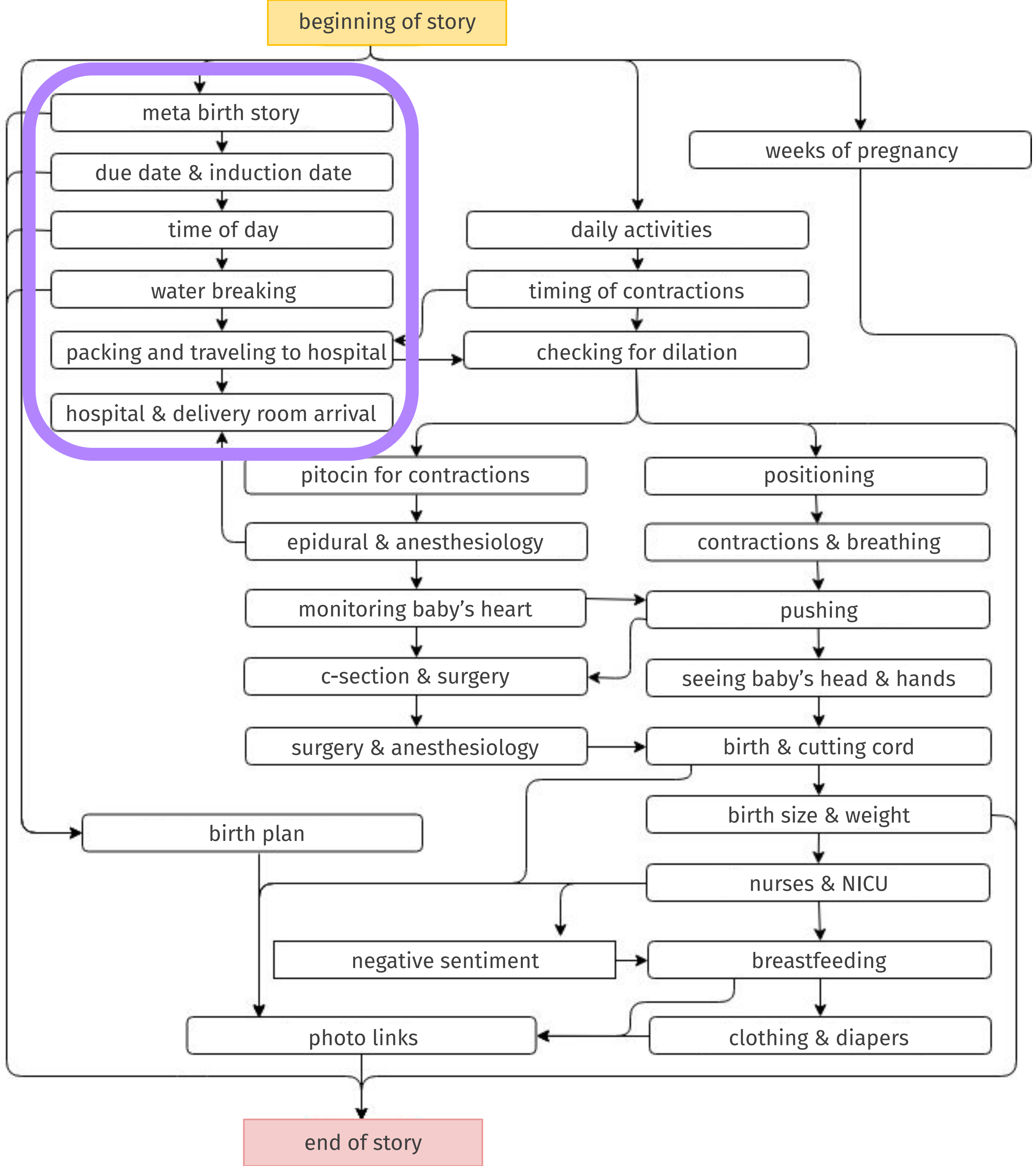
Narrative Patterns: Topics Over Time

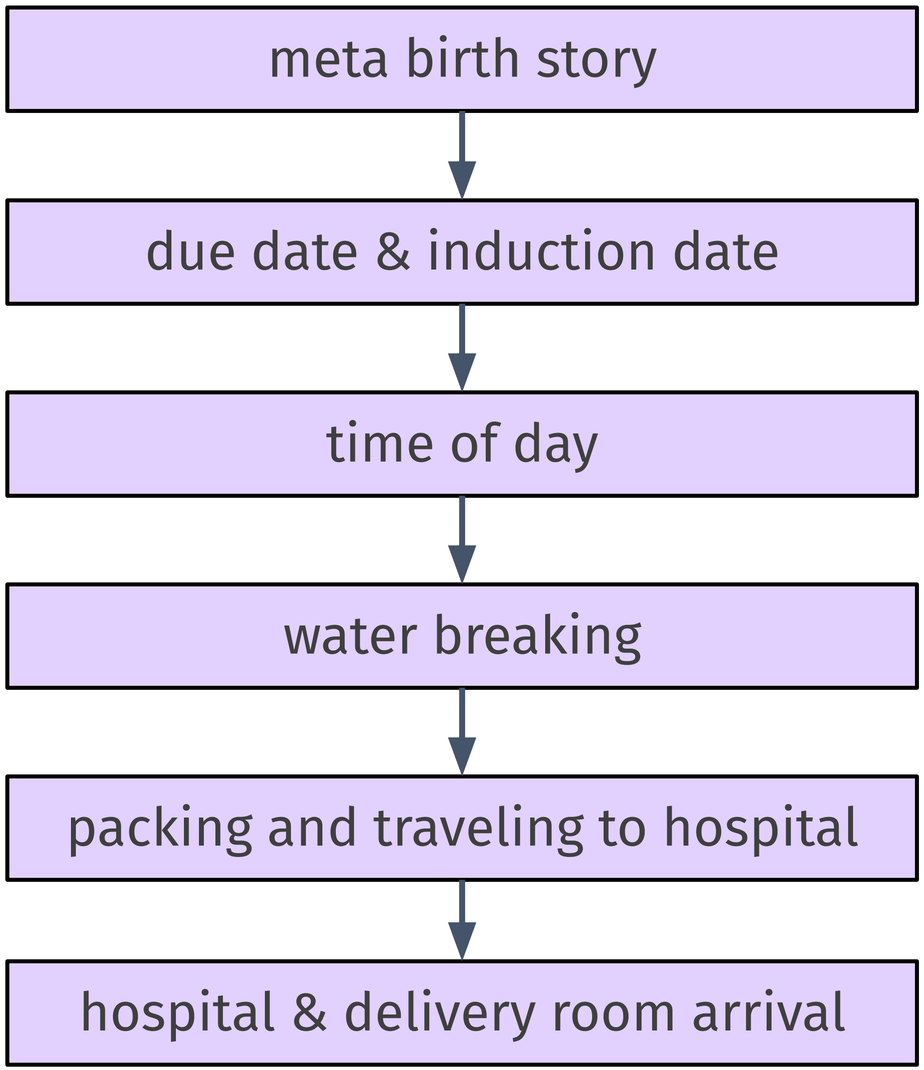


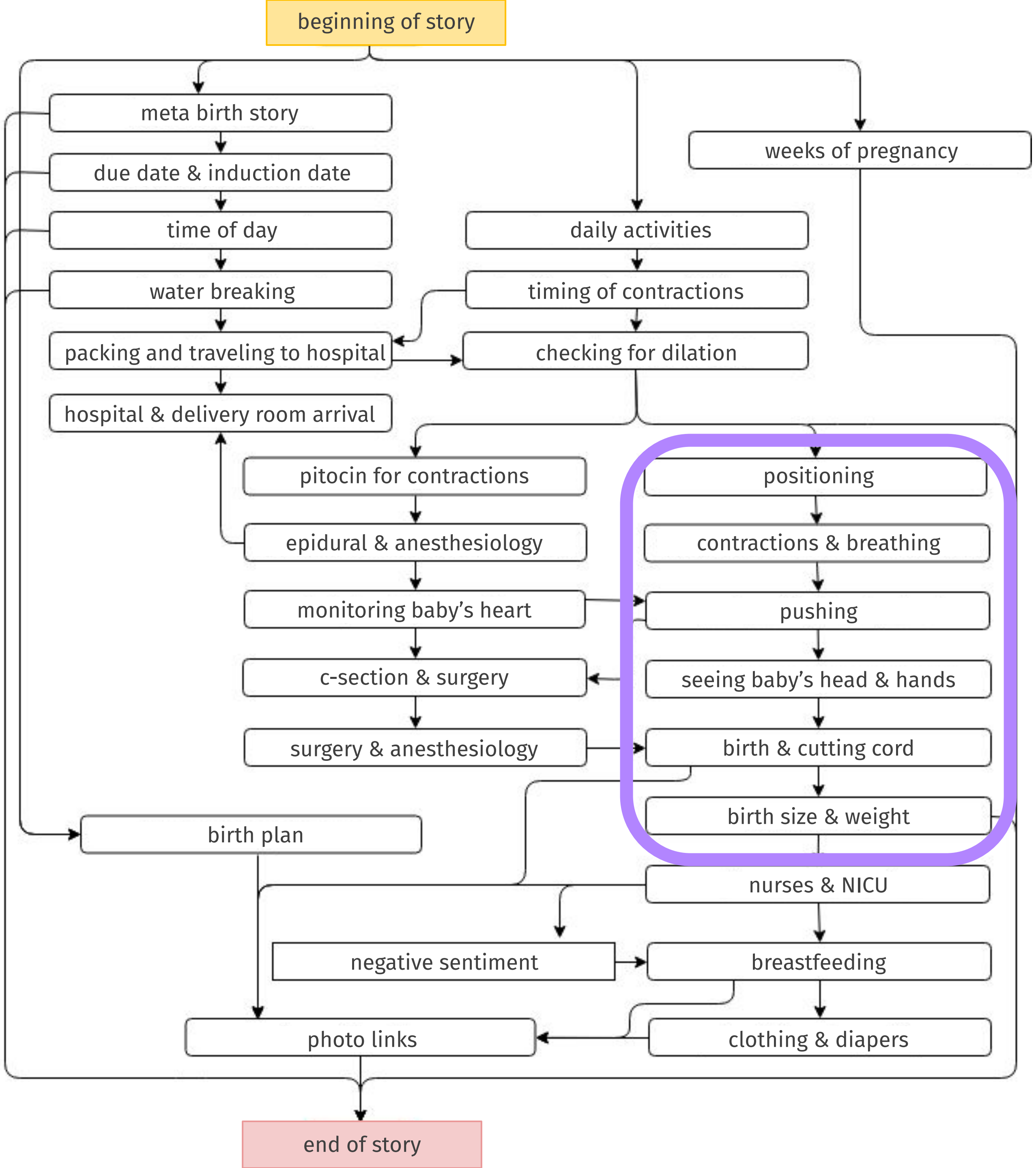

due induction week weeks day
home eat day husband food
bed back tub ball sitting
shit doc start stop tells
remember head husband hand eyes
morning night day woke cramps
husband mom family time son
hour back half bit time
felt feel feeling pressure didn
born long lbs oz tear
high ultrasound ob test bp
sleep night hours rest slept
pain back couldn bad felt
didn wasn thought bad wrong
lot days bad taking couple
weeks pregnancy week days pregnant
hospital home car bag drive
epidural pain relief anesthesiologist meds
section ob baby doctor emergency
breastfeeding day milk days feeding
birth wanted plan labor natural
contractions minutes painful seconds min
contractions point helped breathing contraction
nurse doctor told asked checked
pads clothes change diaper newborn
birth story post long reading
minutes husband time told ready
pretty things wasn quickly time
push pushing head pushed pushes
made feel people make care
labor hours back real dh
iv hours gave put catheter
surgery anesthesiologist put spinal shaking
blood pressure uterus lot low
thing life part amazing worst
water broke fluid break broken
cm checked dilated check cervix
wanted didn knew labour needed
skin room nurses nicu hold
baby girl boy happy healthy
midwife call told called ob
room delivery hospital waiting walked
baby heart monitor rate put
time didn long fast hard
good great time lot amazing
pm started contractions painful decided
pitocin contractions started hours start
side feel left legs move
cord crying chest cry cut
imgur http dr love jpg
Topic Probability
Story Time

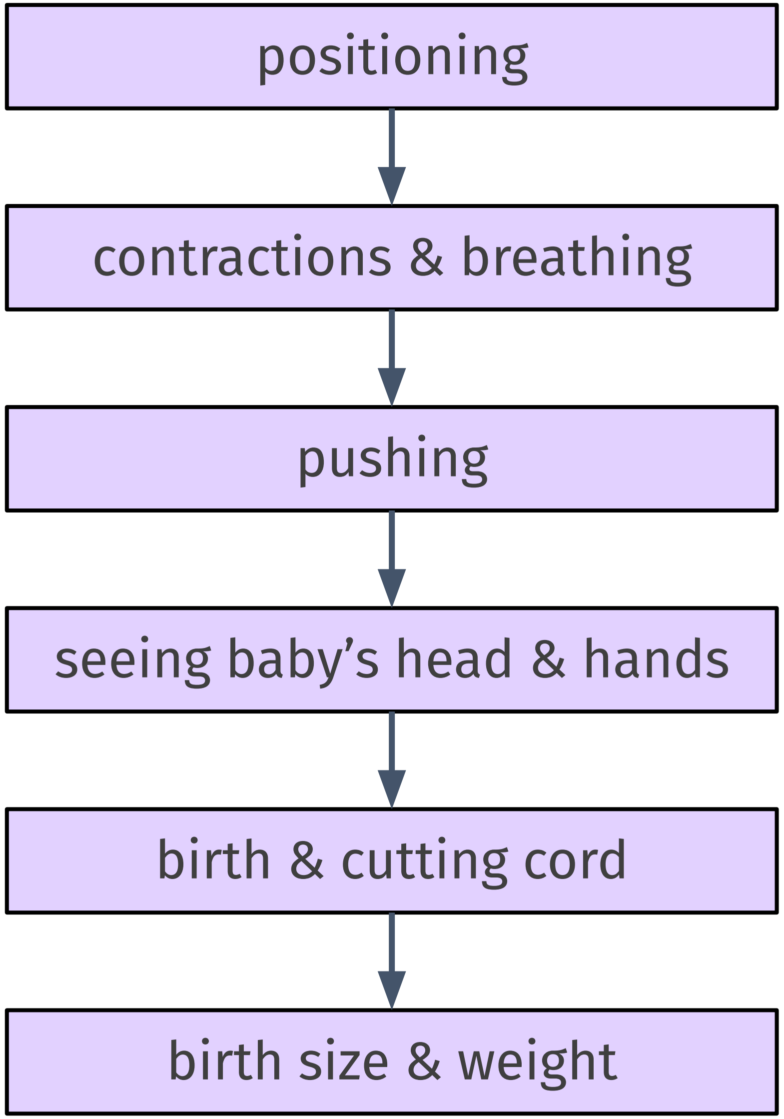
Story Time		Topic 1, Topic 2, Topic 3, Topic 4
0.10	I’ll begin just by saying that this subreddit has taught me so much …	[0.01, 0.20, 0.03, **0.56**, …]
0.20	It was 1am, April 5th. I was experiencing some light cramping …	[**0.23**, 0.11, 0.02, 0.01, …]
0.30	We headed to the hospital and got admitted. The nurse checked me …	[0.03, 0.01, **0.32**, 0.04, …]
…	…	…

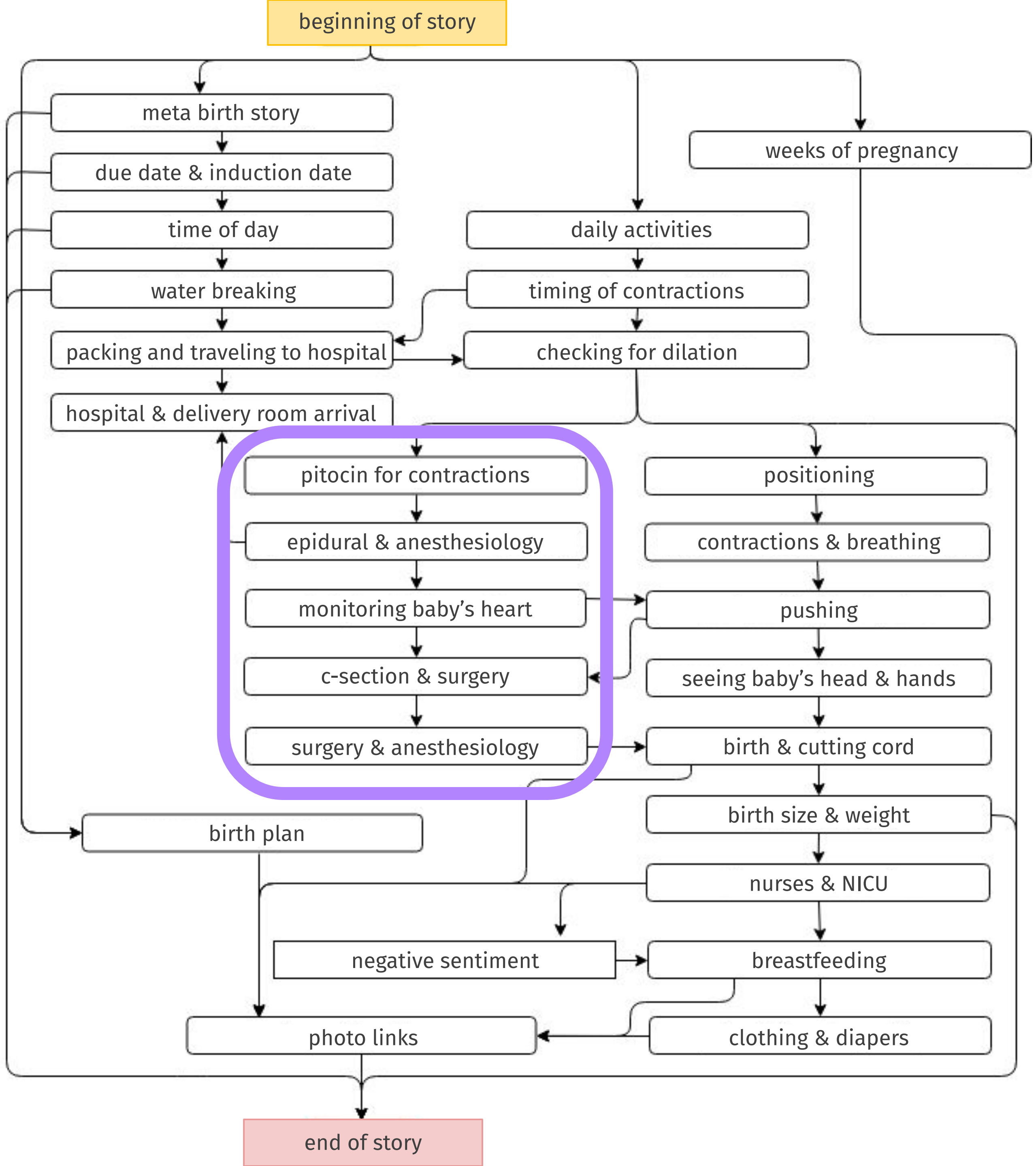

beginning of story
meta birth story
due date & induction date
time of day
water breaking
packing and traveling to hospital
hospital & delivery room arrival
weeks of pregnancy
daily activities
timing of contractions
checking for dilation
pitocin for contractions
epidural & anesthesiology
monitoring baby's heart
c-section & surgery
surgery & anesthesiology
positioning
contractions & breathing
pushing
seeing baby's head & hands
birth & cutting cord
birth size & weight
nurses & NICU
breastfeeding
clothing & diapers
birth plan
negative sentiment
photo links
end of story


beginning of story
meta birth story
due date & induction date
time of day
water breaking
packing and traveling to hospital
hospital & delivery room arrival
weeks of pregnancy
daily activities
timing of contractions
checking for dilation
pitocin for contractions
epidural & anesthesiology
monitoring baby's heart
c-section & surgery
surgery & anesthesiology
positioning
contractions & breathing
pushing
seeing baby's head & hands
birth & cutting cord
birth size & weight
nurses & NICU
birth plan
negative sentiment
breastfeeding
photo links
clothing & diapers
end of story

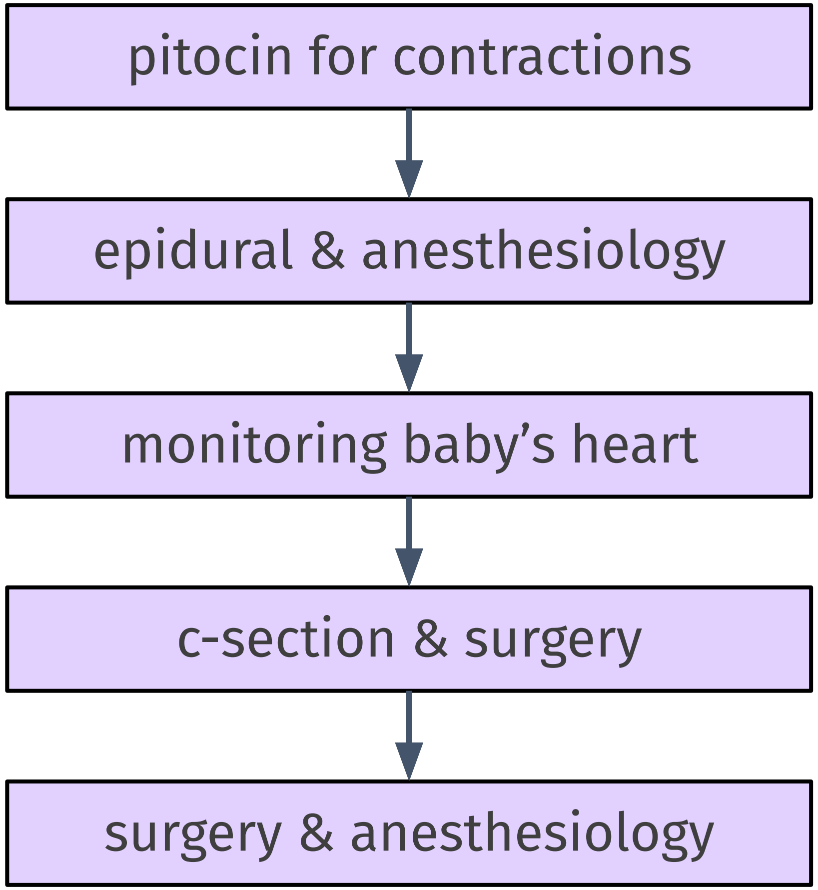

meta birth story
due date & induction date
time of day
water breaking
packing and traveling to hospital
hospital & delivery room arrival


beginning of story
meta birth story
due date & induction date
time of day
water breaking
packing and traveling to hospital
hospital & delivery room arrival
weeks of pregnancy
daily activities
timing of contractions
checking for dilation
pitocin for contractions
epidural & anesthesiology
monitoring baby's heart
c-section & surgery
surgery & anesthesiology
positioning
contractions & breathing
pushing
seeing baby's head & hands
birth & cutting cord
birth size & weight
birth plan
nurses & NICU
negative sentiment
breastfeeding
photo links
clothing & diapers
end of story


positioning
contractions & breathing
pushing
seeing baby's head & hands
birth & cutting cord
birth size & weight



beginning of story
meta birth story
due date & induction date
time of day
water breaking
packing and traveling to hospital
hospital & delivery room arrival
weeks of pregnancy
daily activities
timing of contractions
checking for dilation
pitocin for contractions
epidural & anesthesiology
monitoring baby’s heart
c-section & surgery
surgery & anesthesiology
positioning
contractions & breathing
pushing
seeing baby’s head & hands
birth & cutting cord
birth size & weight
nurses & NICU
birth plan
negative sentiment
breastfeeding
photo links
clothing & diapers
end of story



pitocin for contractions
epidural & anesthesiology
monitoring baby’s heart
c-section & surgery
surgery & anesthesiology

My **positive medicated** birth story

Jonah’s birth story: **positive, medicated**, c-section

My **positive medicated** birth story

Topic 11

Topic 32

Topic 5

Jonah's birth story: **positive**, **medicated**, c-section

Topic 47

Topic 18

Topic 5

My **positive medicated** birth story

Topic 11

$p = 0.20$

Topic 32

$p = 0.01$

Topic 5

Jonah's birth story: **positive**, **medicated**, c-section

Topic 47

$p = 0.06$

Topic 18

$p = 0.14$

Topic 5

My **positive medicated** birth story

Topic 11

Topic 32

Topic 5

p = 0.20 → log(p) = -1.61

p = 0.01 → log(p) = -4.60

-6.21

Jonah's birth story: **positive**, **medicated**, c-section

Topic 47

Topic 18

Topic 5

p = 0.06 → log(p) = -2.81

p = 0.14 → log(p) = -1.97

-4.78

Likely Topic Transitions

Story Log Prob.	Bigram from Story Title
-34.19	positive medicated
-34.27	positive hospital
-34.30	med free
-34.52	positive induction
-34.53	story ftm
-34.73	vaginal delivery
-34.77	story hospital
-34.83	weeks pp
-34.85	hour labor
-34.88	super long
-34.92	failed induction
-34.95	super positive
-34.95	late birth
-35.01	story positive
-35.06	hospital birth
-35.07	story finally
-35.07	water birth
-35.12	vaginal birth
-35.27	line jumper
-35.31	story born

Unlikely Topic Transitions

Story Log Prob.	Bigram from Story Title
-35.79	traumatic birth
-35.82	story unmedicated
-35.93	story baby
-35.94	post partum
-35.95	story plus
-35.95	due date
-35.99	pp advice
-36.02	baby birth
-36.03	home birth
-36.04	c section
-36.05	story warning
-36.11	unplanned c
-36.13	slightly traumatic
-36.27	natural birth
-36.40	belated birth
-36.42	positive unmedicated
-36.42	emergency c
-36.53	trigger warning
-36.60	induction epidural
-36.84	happy ending

	Story Log Prob.	Bigram from Story Title
	-34.19	positive medicated
	-34.27	positive hospital
	-34.30	med free
	-34.52	positive induction
	-34.53	story ftm
	-34.73	vaginal delivery
	-34.77	story hospital
	-34.83	weeks pp
	-34.85	hour labor
Likely Topic Transitions	-34.88	super long
	-34.92	failed induction
	-34.95	super positive
	-34.95	late birth
	-35.01	story positive
	-35.06	hospital birth
	-35.07	story finally
	-35.07	water birth
	-35.12	vaginal birth
	-35.27	line jumper
	-35.31	story born

	Story Log Prob.	Bigram from Story Title
	-35.79	traumatic birth
	-35.82	story unmedicated
	-35.93	story baby
	-35.94	post partum
	-35.95	story plus
	-35.95	due date
	-35.99	pp advice
	-36.02	baby birth
	-36.03	home birth
Unlikely Topic Transitions	-36.04	c section
	-36.05	story warning
	-36.11	unplanned c
	-36.13	slightly traumatic
	-36.27	natural birth
	-36.40	belated birth
	-36.42	positive unmedicated
	-36.42	emergency c
	-36.53	trigger warning
	-36.60	induction epidural
	-36.84	happy ending

Likely Topic Transitions

Story Log Prob.	Bigram from Story Title
-34.19	positive medicated
-34.27	positive hospital
-34.30	med free
-34.52	positive induction
-34.53	story ftm
-34.73	vaginal delivery
-34.77	story hospital
-34.83	weeks pp
-34.85	hour labor
-34.88	super long
-34.92	failed induction
-34.95	super positive
-34.95	late birth
-35.01	story positive
-35.06	hospital birth
-35.07	story finally
-35.07	water birth
-35.12	vaginal birth
-35.27	line jumper
-35.31	story born

Unlikely Topic Transitions

Story Log Prob.	Bigram from Story Title
-35.79	traumatic birth
-35.82	story unmedicated
-35.93	story baby
-35.94	post partum
-35.95	story plus
-35.95	due date
-35.99	pp advice
-36.02	baby birth
-36.03	home birth
-36.04	c section
-36.05	story warning
-36.11	unplanned c
-36.13	slightly traumatic
-36.27	natural birth
-36.40	belated birth
-36.42	positive unmedicated
-36.42	emergency c
-36.53	trigger warning
-36.60	induction epidural
-36.84	happy ending

Story Log Prob.	Bigram from Story Title
-34.19	positive medicated
-34.27	positive hospital
-34.30	med free
-34.52	positive induction
-34.53	story ftm
-34.73	vaginal delivery
-34.77	story hospital
-34.83	weeks pp
-34.85	hour labor
-34.88	super long
-34.92	failed induction
-34.95	super positive
-34.95	late birth
-35.01	story positive
-35.06	hospital birth
-35.07	story finally
-35.07	water birth
-35.12	vaginal birth
-35.27	line jumper
-35.31	story born

Likely Topic Transitions

Story Log Prob.	Bigram from Story Title
-35.79	traumatic birth
-35.82	story unmedicated
-35.93	story baby
-35.94	post partum
-35.95	story plus
-35.95	due date
-35.99	pp advice
-36.02	baby birth
-36.03	home birth
-36.04	c section
-36.05	story warning
-36.11	unplanned c
-36.13	slightly traumatic
-36.27	natural birth
-36.40	belated birth
-36.42	positive unmedicated
-36.42	emergency c
-36.53	trigger warning
-36.60	induction epidural
-36.84	happy ending

Unlikely Topic Transitions

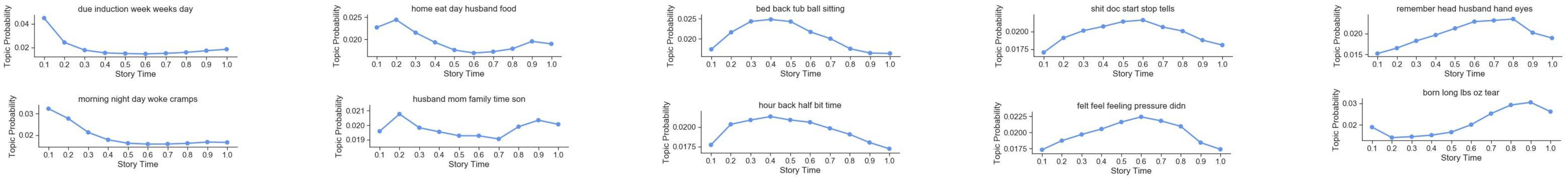
r/BabyBumps: Limitations

- Our birth stories only include “**happy endings**” with no lost pregnancies
- Skew towards **home** births and **unmedicated** births
- No explicit/verified demographic information…
- …but the authors generally
  - write in English
  - describe experiences in the U.S., Canada, and the U.K.
  - have access to Reddit

Takeaways from birth stories

We can use computational tools to model the shared **narrative patterns** and **power framing** in a community's set of birth stories.

We discovered sets of **diverging pathways**, found **outlier** stories labeled with terms of surprisal, trauma, and **happy endings**.



Sensemaking About Contraceptive Methods Across Online Platforms

LeAnn McDowall, Maria Antoniak, David Mimno

International Conference on Web & Social Media (ICWSM), 2024

Difficult healthcare choices

Birth control plays an important role beyond contraception and can be used to treat and manage medical conditions.

But birth control methods are not one-size fits all, and the choices of whether to use birth control and which method to select are complicated by **personal beliefs**, **cost and accessibility**, and a wide array of **side effects** that are difficult to predict and identify.

When navigating this decision, birth control users face a **sensemaking** challenge.

Research Questions

Which birth control methods and side effects are more likely to be discussed **on different online platforms**?

What kinds of **sensemaking activities** that birth control discussants engage in online and how these differ by platform?

Data Collection: Reddit and Twitter

Community	# of Posts	Vocab Size	Mean Tokens	Year Range	Posts Dist. (2007-2020)	Moderation	Structure
Reddit Posts	68,958	49,088	79	2010-2020		user moderators	forum posts
Reddit Comments	264,912	67,837	32	2010-2020		user moderators	replies to forum posts
Twitter Posts	499,796	398,910	12	2006-2020		company	tweets (no retweets)
Twitter Replies	211,896	73,896	12	2007-2020		company	replies to tweets

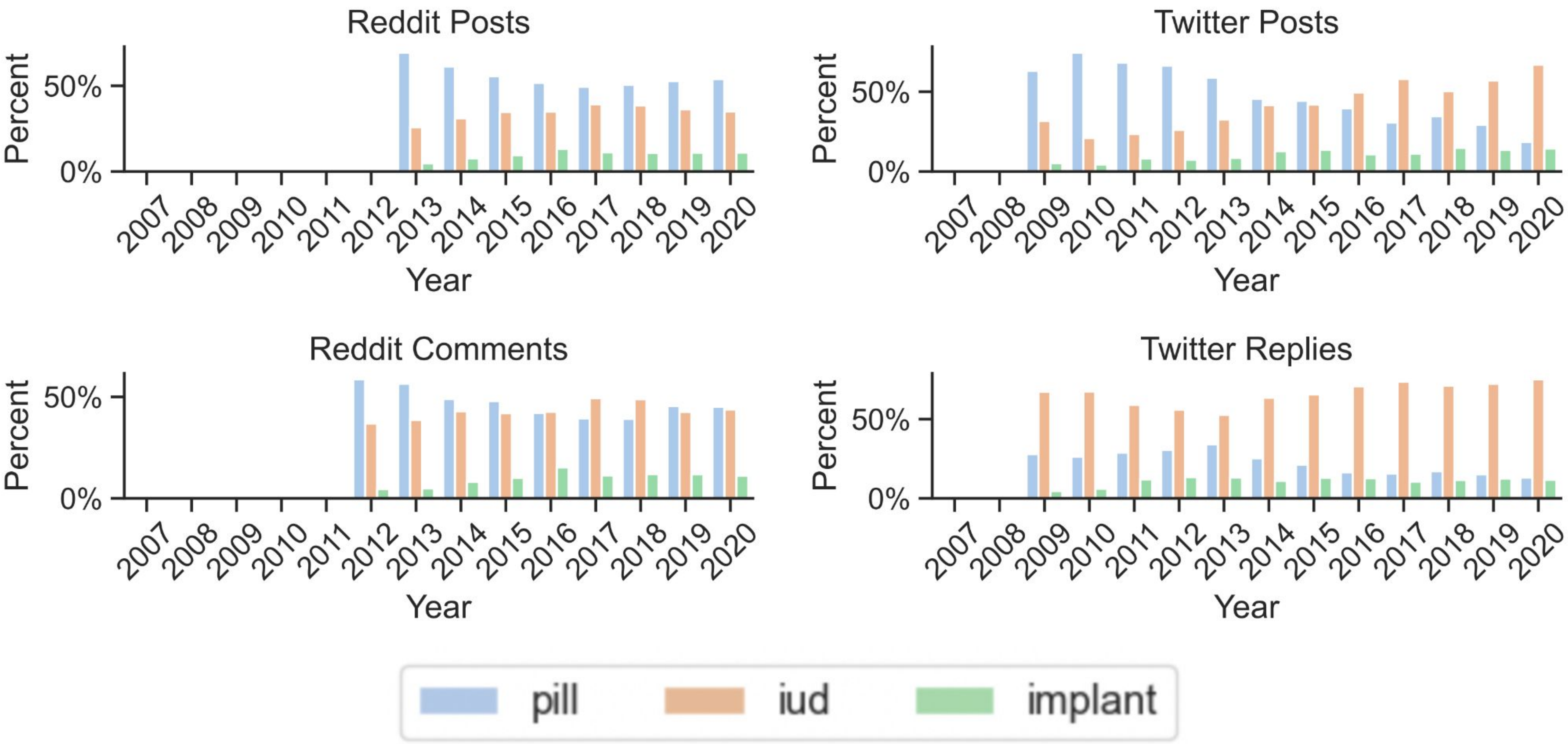
Table 1: Overview of the two datasets, including only texts mentioning our target birth control methods.

Methods

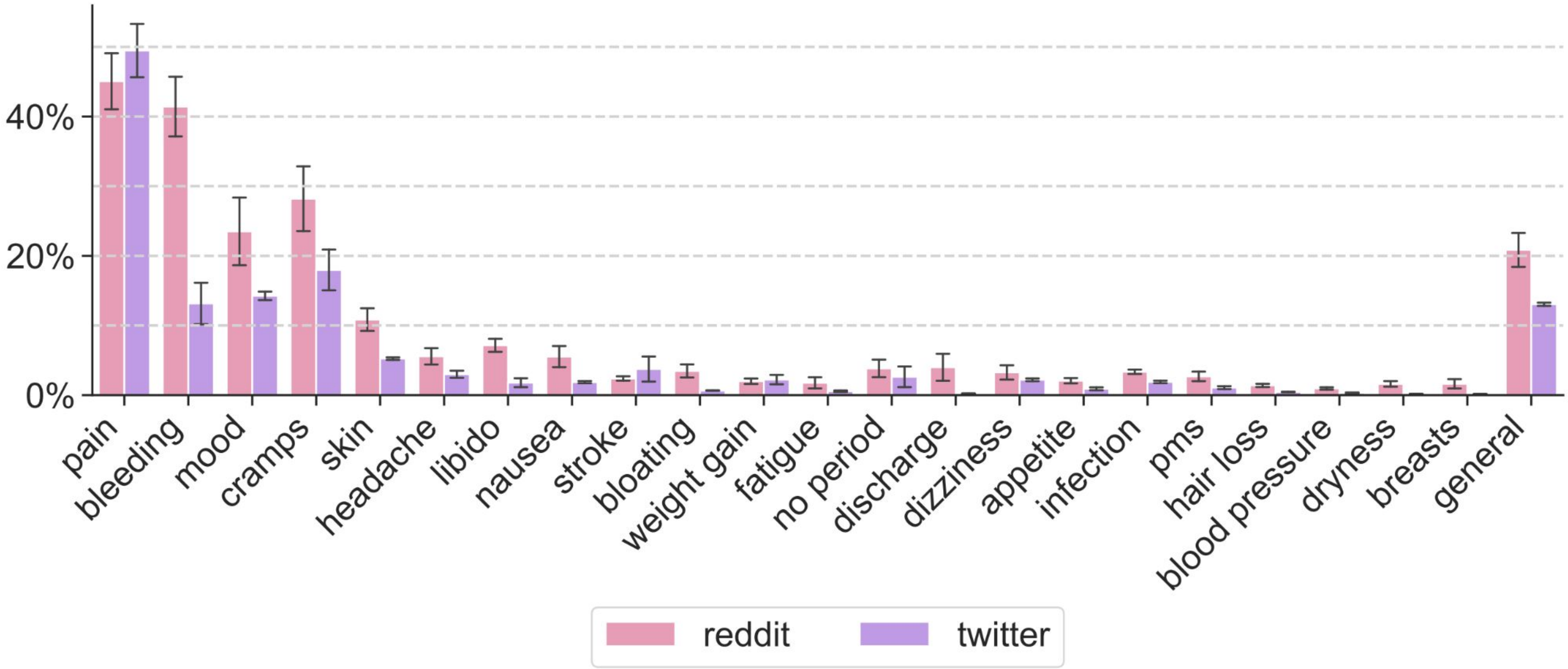
Custom Lexicon: Birth Control Methods

Custom Lexicon: Side Effects

Topic Model: Latent Dirichlet Allocation (LDA)

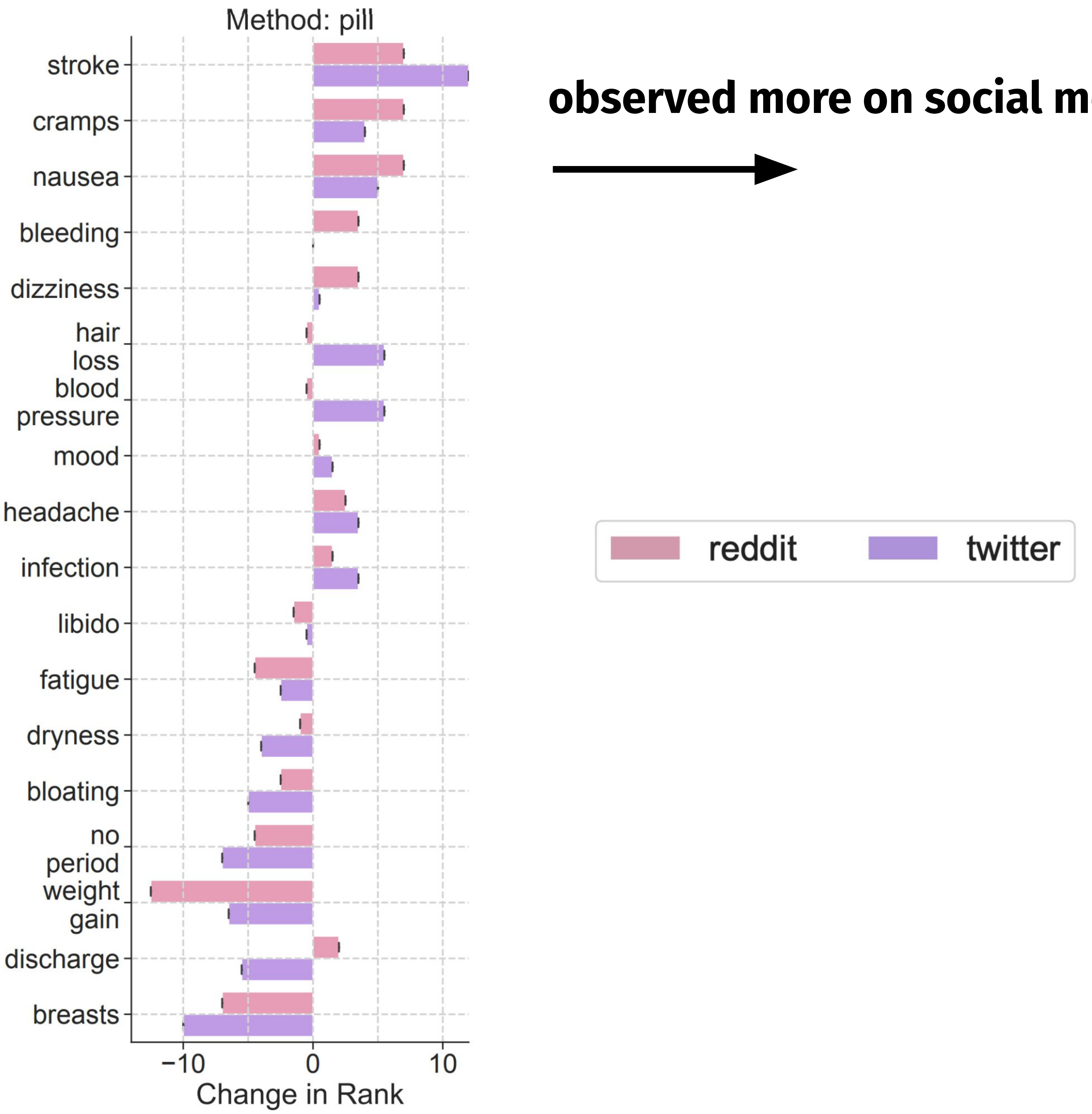


Reddit Posts
Twitter Posts
Reddit Comments
Twitter Replies
Percent
50%
0%
2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
Year
pill
iud
implant



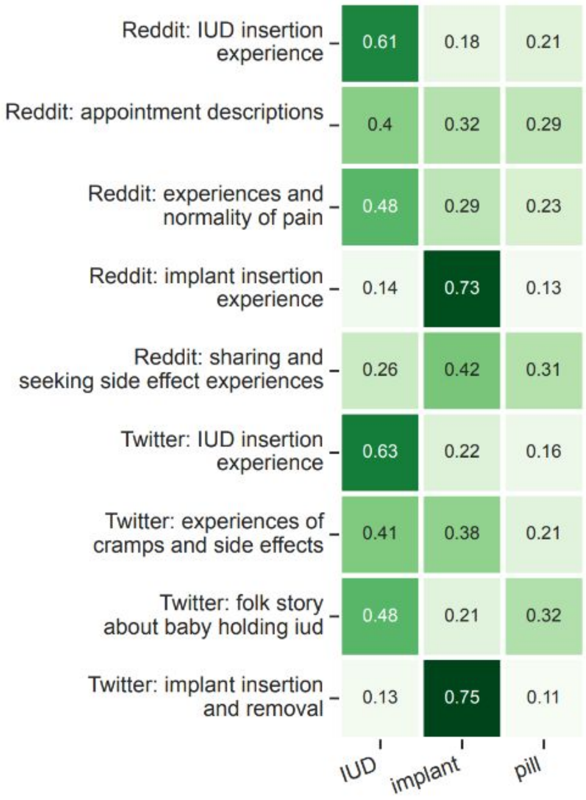
40%
20%
0%
pain
bleeding
mood
cramps
skin
headache
libido
nausea
stroke
bloating
weight gain
fatigue
no period
discharge
dizziness
appetite
infection
pms
hair loss
blood pressure
dryness
breasts
general
reddit
twitter

Comparison between self-reported survey data and observed frequencies on social media

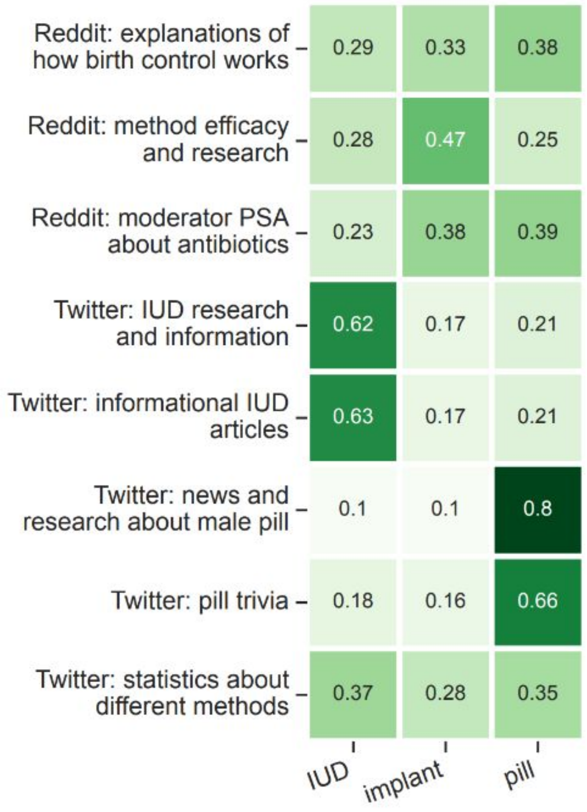


observed more on social media →

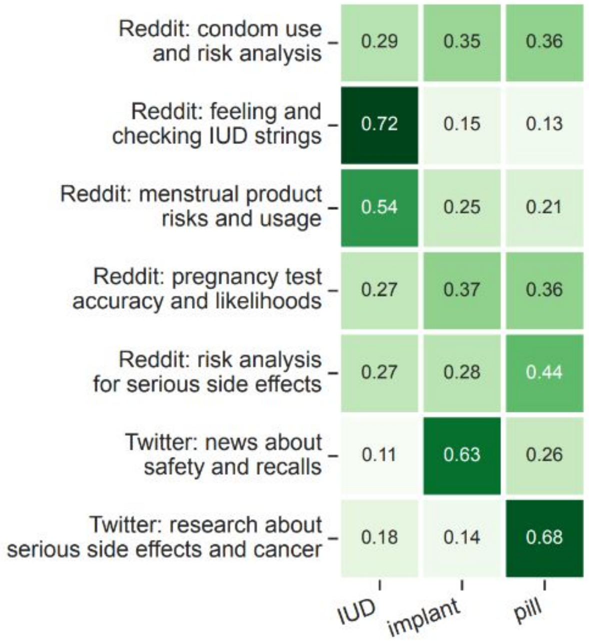
← **reported more in surveys**



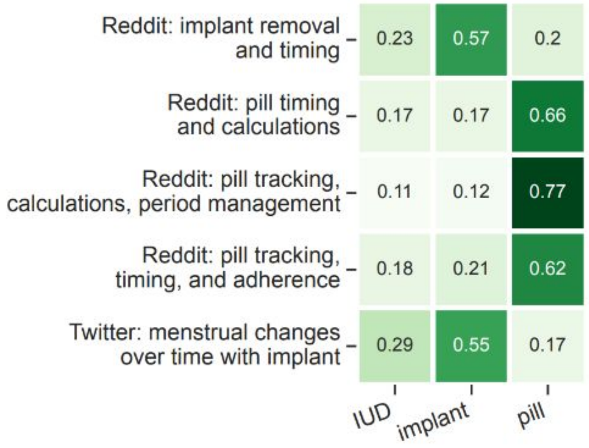
(a) Storytelling



(b) Information & Explanations



(c) Risk Analysis



(d) Timing & Calculations

Contributions

- We identify a unique combination of sensemaking strategies including storytelling, risk analysis, timing and calculations, causal reasoning, method and hormone comparison, and information and explanations.
- Across platforms, we find that storytelling is used to prepare for and overcome painful insertion experiences.
- Twitter users are more likely to discuss the IUD and severe side effects, while Reddit users frequently discuss both the IUD and the pill as well as sensitive side effects like bleeding.
- We compare our results to self-reported survey data, highlighting side effects for which Twitter and Reddit users discuss more than expected, perhaps indicating increased interest and needs met by the platforms

Future Work: Modeling Narratives

Storytelling is a powerful driver of community **sensemaking** processes that can also spread **misinformation.**

Need for more studies on new online communities like TikTok where storytelling is part of a longer **folkloric** and **memetic** tradition.

Narrative medicine to empower care seekers and motivate empathy in caregivers.

Collaborators: narratologists, NLP researchers, web & social media researchers

Future Work: LLMs and Healthcare

We need **new solutions** to inaccessibility, physician burnout, etc.

We’re **already** observing care-seekers using LLMs to script communication with their healthcare providers, seek information, etc.

But there are many **risks** involved in using NLP to address those challenges: power consolidation, privacy implications, biases and errors

Future Work: Probing LLMs

Stability evaluations when using LLMs for **small, socially-specific** datasets and tasks (or in cases where "tasks" aren't clear and data exploration is the goal).

Benchmarking model performance across **domains** and **communities**: **measuring variation** in and across specific **cultural contexts**.

Thank you! Questions?

maria-antoniak.github.io