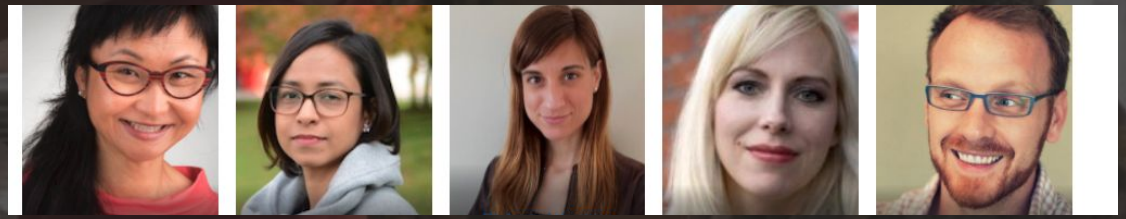
Media Understanding Workshop II

Relationships & Interactions



USC Viterbi School of Engineering

Google Research

Today's Agenda

10:00am PT Introductory remarks
Kree Cole-McLaughlin, Shri Narayanan

10:10am PT Taxonomies for Relationships & Interactions

10:30am PT Character Networks for Interaction Modeling

10:50am PT The Ethics of Measuring Body Size

11:10am PT Break

11:20am PT Understanding Conflict and Abusive Language

11:40am PT Identity and Subjectivity in Relationships & Interactions

Welcome to today's Media Understanding Workshop hosted by Google and USC.

This workshop is being recorded.

Signup for CCMI's mailing list to hear about future events. (LINK)

Taxonomies for Relationships & Interactions

Hosted by Agata Lapedriza
Universitat Oberta de Catalunya

Presented by: Dylan Baker

Attendees:

- Digbalay Bose - USC SAIL
- Veena Vijai - USC SAIL
- Brendan Jou - Google
- Dylan Baker *(facilitator)* - Ethical AI, Google Research

Taxonomies for Relationships & Interactions

In what situations might we want to **compute and classify interactions between people**?

Taxonomies for Relationships & Interactions

In what situations might we want to **compute and classify interactions between people**?

What **factors** are important in understanding interactions between people?

What can be **perceived and measured**?

Taxonomies for Relationships & Interactions

- The choices we make to represent interactions **define the output space**

Taxonomies for Relationships & Interactions

- The choices we make to represent interactions **define the output space**
- **What applications** are these systems used for?
  - Are they surfaced explicitly? Are they used to inform future behavior?
  - How can they be used to uncover biases? To reinforce them? To change their impact?

Taxonomies for Relationships & Interactions

- The choices we make to represent interactions **define the output space**
- **What applications** are these systems used for?
  - Are they surfaced explicitly? Are they used to inform future behavior?
  - How can they be used to uncover biases?
- Consider **affect** and social dimensions
  - Emotions and affect aren’t exclusively social but inherently have a social component

Taxonomies for Relationships & Interactions

- The choices we make to represent interactions **define the output space**
- **What applications** are these systems used for?
  - Are they surfaced explicitly? Are they used to inform future behavior?
  - How can they be used to uncover biases?
- Consider **affect** and social dimensions
  - Emotions and affect aren't exclusively social but inherently have a social component
- **"Top-down" vs. "bottom-up"** approaches
  - Better to define a broad taxonomic space, then select a subset of labels for a particular context? Or define different taxonomies for each context?
  - Accepting biases in data ("the world as it is") vs. enforcing top-level objectives
  - Easy to enforce priors—for better or worse—with a "top-down" approach
  - How does one define an "interaction" to begin with?

Taxonomies for Relationships & Interactions

Towards what **applications** might we want to compute and classify interactions between people?

What **factors** are important in understanding interactions between people?

What can be **perceived and measured**?

When are **top-down vs bottom-up approaches** most useful or appropriate?

How does **affect** play a role in taxonomizing relationships and interactions?

Character Networks for Interaction Modelling

Hosted by Tanaya Guha
University of Warwick

Presented by: Amrutha Nadarajan
SAIL Lab, USC

Attendees:

- Joanne Garde-Hansen - University of Warwick
- Zeerak Wassem - University of Sheffield
- Sabyasache Baruah - SAIL Lab, USC
- Digbalay Bose - SAIL Lab, USC

Character-centric media understanding

- How characters interact with each other is crucial.

- Comprehensive dynamic model of character-character interaction
  - How to quantify “strength” of interaction?
  - How to detect interaction along dimensions like power, victim-perpetrator, emotional expressions?

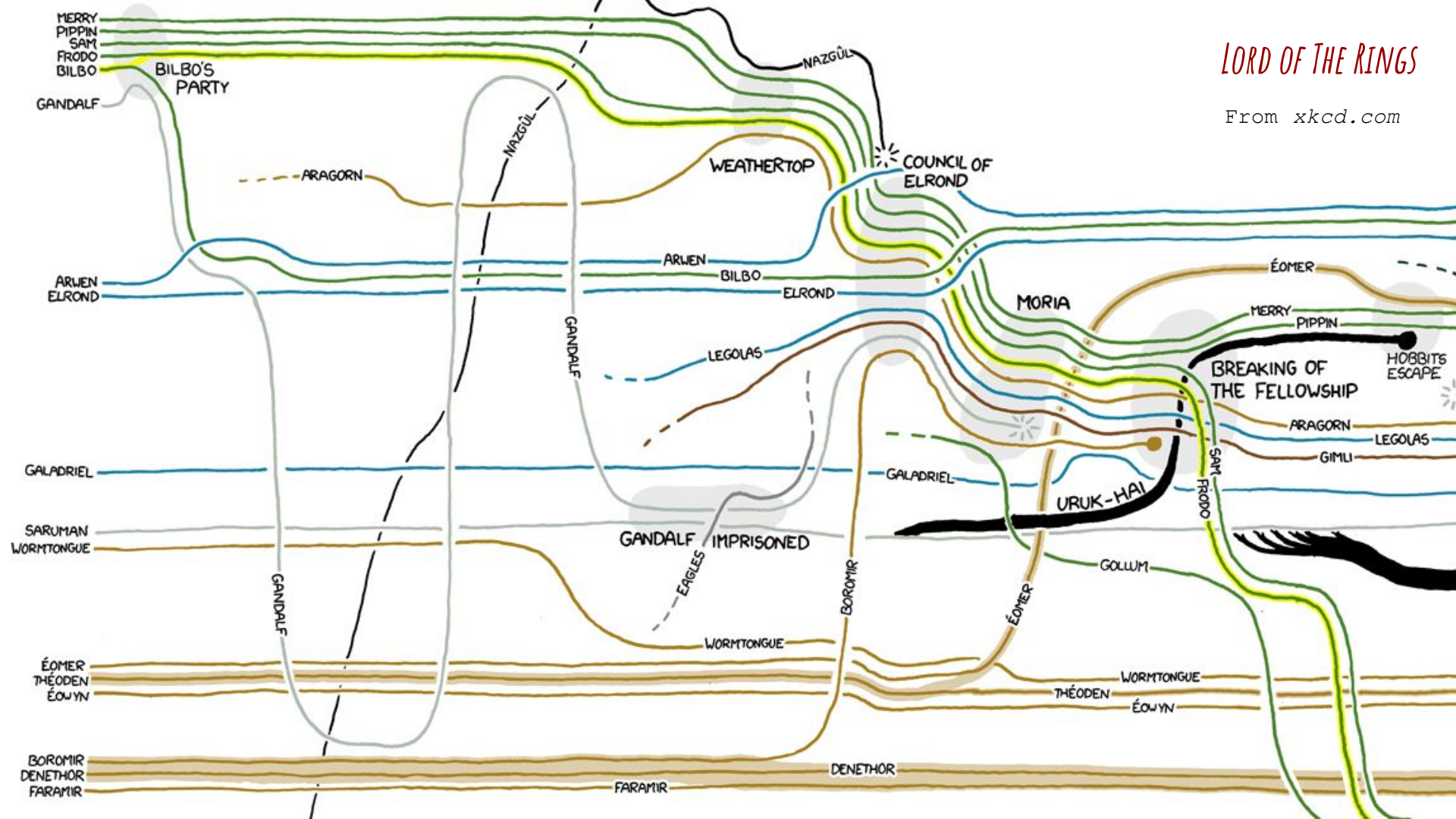
- Fuse information from different modalities
  - Visual stream, Audio steam, Subtitles
  - Screenplay/ movie scripts (not always available)

Character graphs are useful

- Answer summative questions:
  - *How important is a character?*
  - *What are the relationships between different characters?*
- Easily measure representation statistics (e.g. screen time)
- Discover social roles of characters
- Characters interactions indicate major events in storytelling
- Compare media stories/segments
- Retrieve similar stories using (sub)graph matching

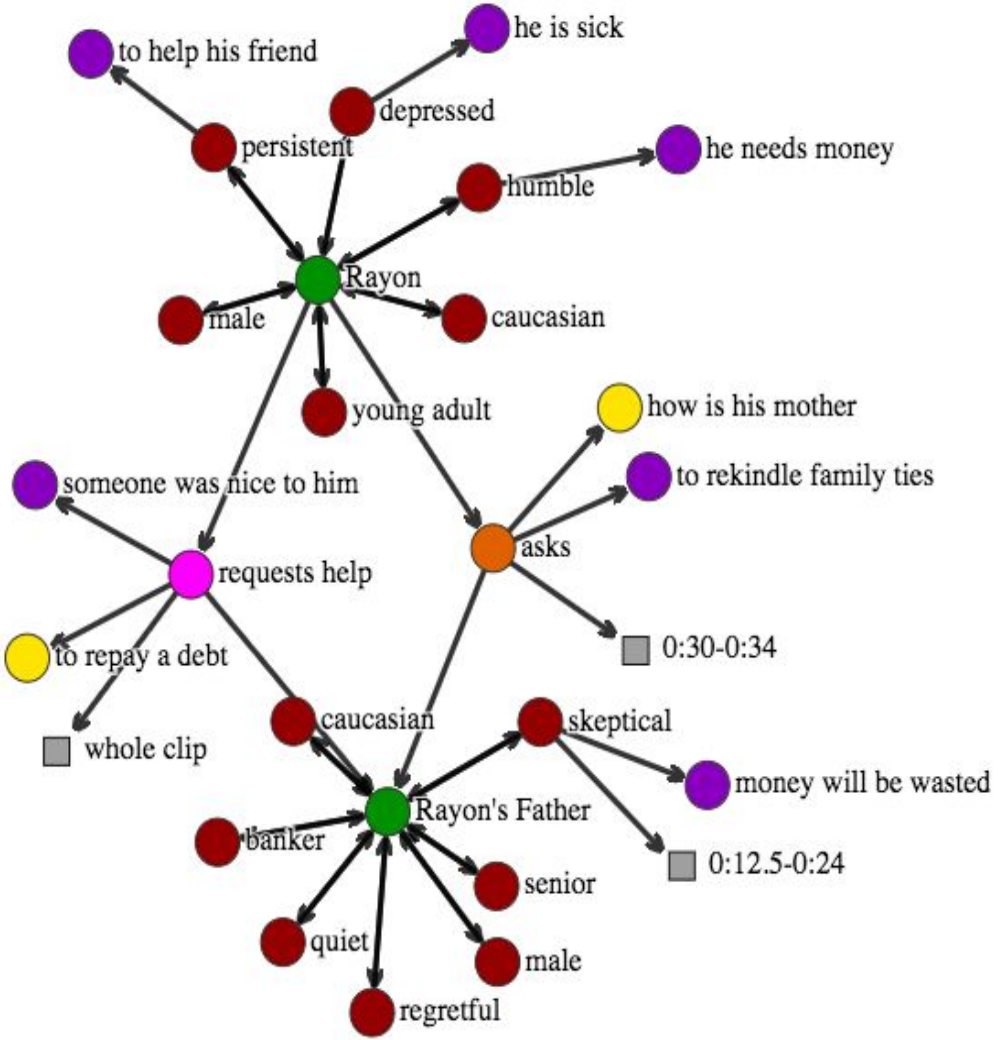
Lord of The Rings

From *xkcd.com*

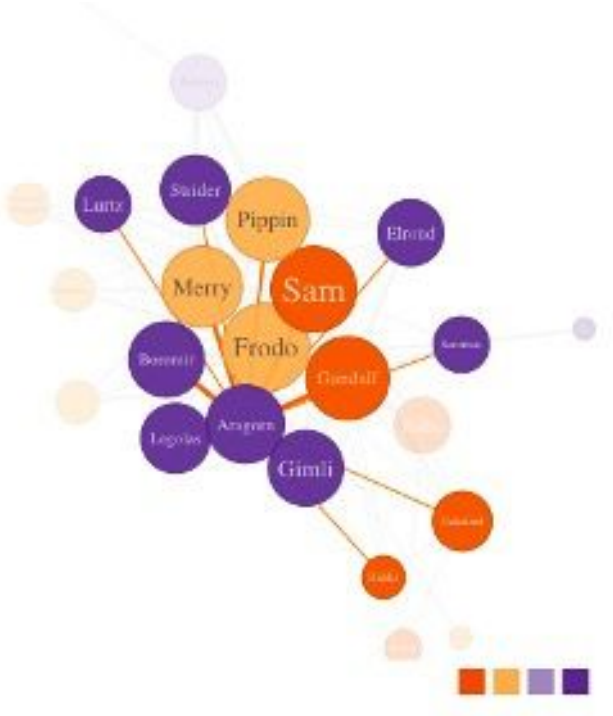


Character graphs

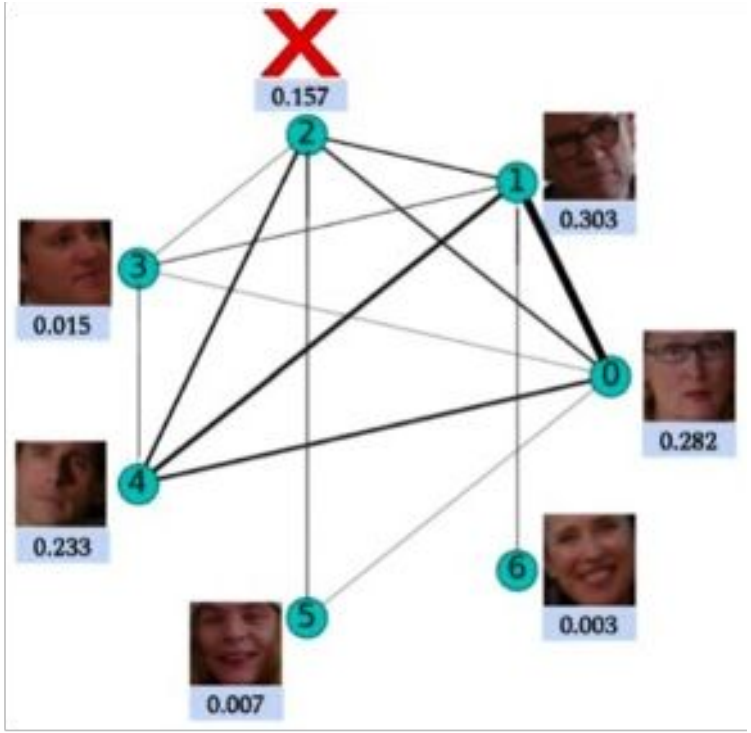
Social situations network from the MovieGraph database (hand annotated)



Character graphs



Movigalaxies.com ***LoTR*** *(2001) network using script parsing*



Hope Springs (2001) network using face clustering

Characterizing interactions beyond co-occurrences

- How to characterize strength of characters?
  - We are trying to get to a higher abstraction from simple co-occurrences
  - Want to characterize meaningful interactions more than just demographics/counts
  - Is ML there yet? we need to use a set of rules?
  - Not able to look at larger context than (some) temporal resolution
  - Can look at faces, body gestures

- Why ML and not statistics?
  - Because we don't always have access to script
  - Script is just one aspect - after post production, things change - work of art not always methodical

How to characterize dimensions like power?

(e.g., victim/perpetrator dynamics or affect)

Representation is one aspect

- Eg: How does representation change before content is produced and after content is produced?

Othering language (us vs them)

- in group and out group convo
- Not just identity - stranger and strangeness
- Groups of people misunderstanding each other (othering language)
- NLP work : looking at othering language computationally

Comprehensive understanding of interactions

- Incorporating information from different streams
  - How to merge, say a graph from video analysis vs a graph from text
  - Scene graphs are another tool - give low level interactions
- Could we add more attributes to the characters ?
  - Would prior knowledge of character importance help these graphs ?
  - For eg: if books → scripts, can we use this information to pre-inform relationships? If so, how?

Perception of content vs content as multiple views of the same problem

- Content is getting tailored to different audience
  - Segmentation of audience with personalization by platforms like netflix
- Perception experiment
  - Creating custom trailers
  - if ML could select out scenes and putting it together to create multiple trailers (and now have different sets of human annotations on this)
- Sampling: Selection of specific frames
  - We could do it as researchers
  - Audience does it in forms of memes etc (they communicate)
  - Industry selects (based on different parts of the world)
- How to get people to watch content that they might not want to?

Takeaway

- Need a taxonomy for relationships
  - There are short cut ways of getting to this, but those are not complete
- ML methods that can model complex relationships beyond “fancy counting”
  - Need methods to benchmark some of these ML efforts / annotations are sparse or hard to get
  - Fusing information from variable sources, modalities is a ongoing research area
- Perception of content just as important as content creation
  - needs to be factored in for effective analysis of complex relationships

The Ethics of Measuring Body Size

Hosted by Dr. Caroline Heldman
Geena Davis Institute for Gender in Media

Presented by:
Dr. Caroline Heldman

Attendees:

- Julio Vallejo, Pigmentocracia
- Krishna Somandepalli, SAIL, USC
- Meredith Conroy, GDIGM
- Nathan Cooper Jones, GDIGM
- Paxton Misra, Google
- Raghuveer Peri, SAIL, USC
- Dr. Rebecca Cooper, GDIGM

The Ethics of Measuring Body Size

WHY IS THIS IMPORTANT?

Because sizeism is a social justice issue.

SIZEISM STATS

- Medical fat shaming is common, causing patients to avoid seeking medical care and leading to sub-par care from physicians (NHS, 2015).
- Male jurors more likely to convict fat women for crimes (Yale Center for Food & Obesity, 2013)
- Fat people earn $1.50 less on average than others for wage workers (Council on Size & Weight Discrimination, 2017).
- 61% of fat adults face employment discrimination (Puhl, 2018).
- 45% of women and 28% of men face harassment/bullying (Puhl, 2018).
- 85% of fat kids and teens are bullied in school (Chalker, 2014).
- Parents less likely to purchase a car for a fat teen than others (CarParts, 2020).

MYTHS OF FATNESS

MYTH #1: Fatness is about willpower, not biology.
FACTS: Dozens of factors determine body size, including genetics, hyperthyroidism, depression, and an allergy to leptin, to name a few.

MYTH #2: People can lose weight by eating less.
FACTS: Different bodies process calories differently. Restricting calories actually causes some people to gain weight.

MYTH #3: Thin people are healthy. Fat people are unhealthy.
FACTS: A 2005 CDC study finds no connection between weight and death rates. 2013 meta-analysis finds “overweight” people have a 6% lower rate of dying than others, after controlling for smoking, age, and gender.

TERMS TO AVOID

- “Obese”
- “Morbidly Obese”
- “Heavy”
- “Plus-size”
- “Overweight”
- “Non-thin”

TERM TO USE

- **Fat**

- **“Large Body Type”**
  - Stigma-free
  - De-centers smaller bodies as the norm
  - Accurately suggests that body types aren’t very malleable

FAT TROPES IN MEDIA

1. **Comic Relief** (e.g., a character who exists for comic relief)
2. **Sidekick** (e.g., supportive buddy--often best friend to a pretty girl)
3. **Mamma Hen** (e.g., nurturing mother figure; great listener)
4. **Nympho** (e.g., hypersexual or sexually vulgar character; bordering on predatory)
5. **Loser** (e.g., has a bad job, frumpy clothing, bad hair)

1.

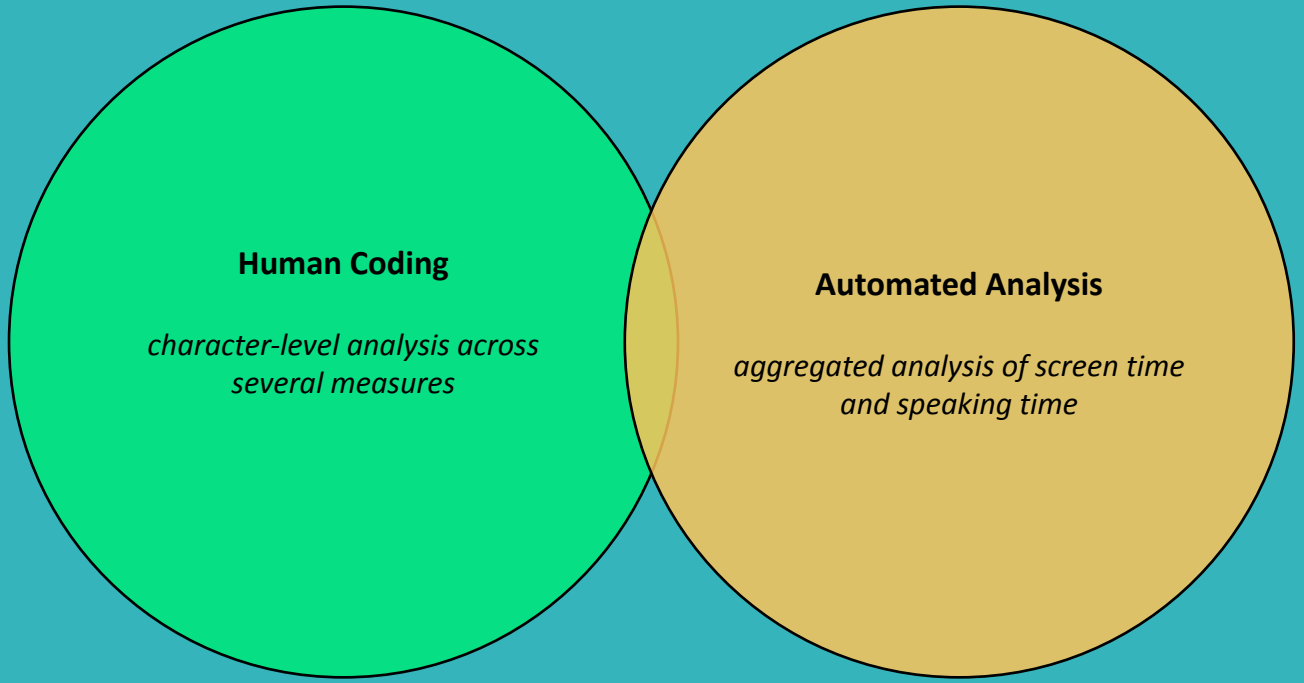

2.


3.


4.

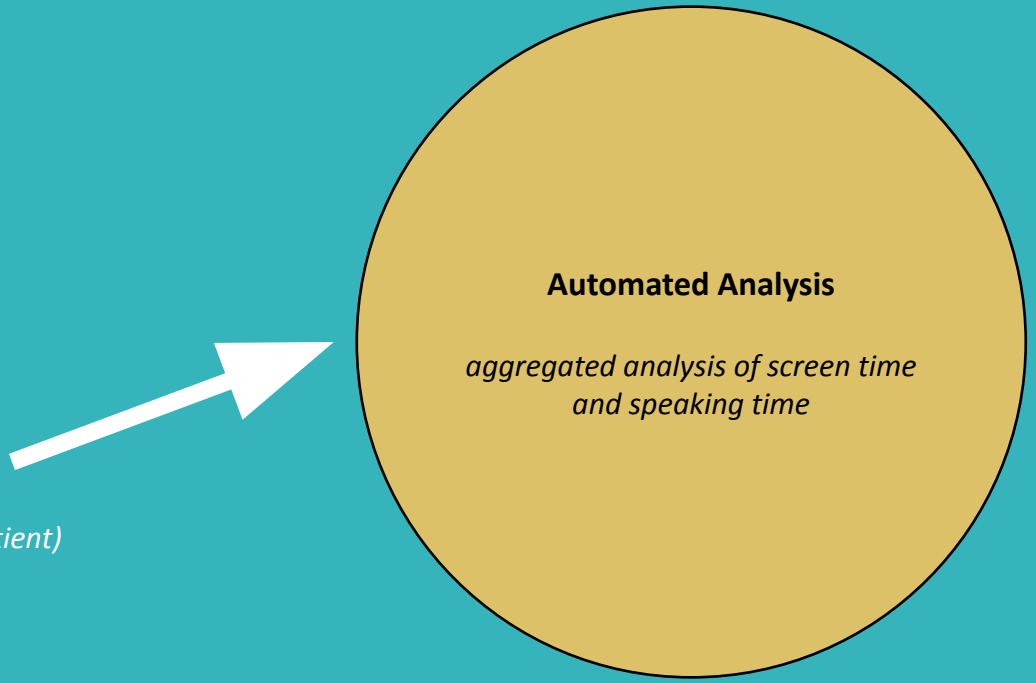

5.


AUTOMATING MEASURES OF BODY SIZE


Human Coding
character-level analysis across several measures
Automated Analysis
aggregated analysis of screen time and speaking time

GD-IQ

(Geena Davis Inclusion Quotient)



Can we automate body type classification for characters in film and TV?

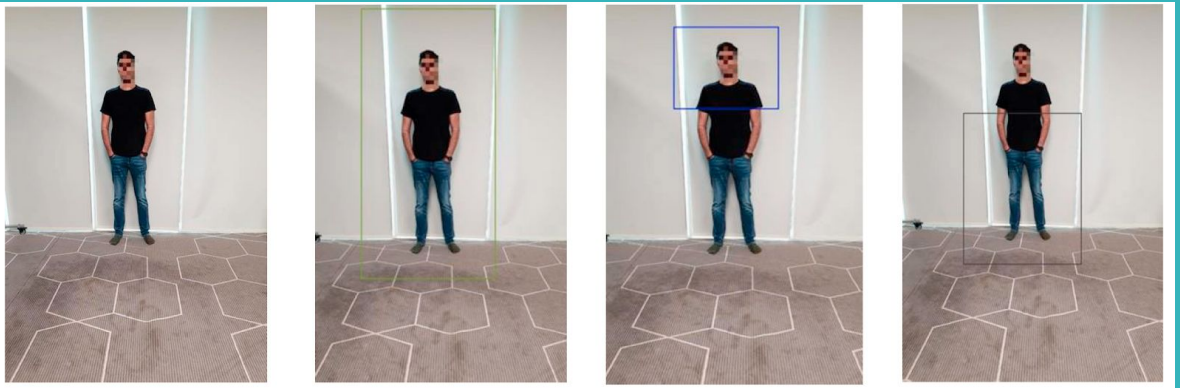


Fig. 1. Original image, full detector, upper detector and lower detector with front image.

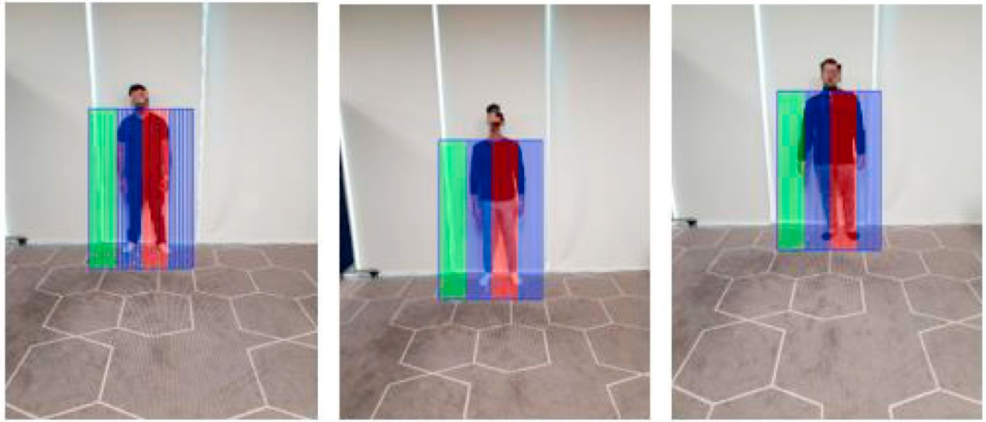
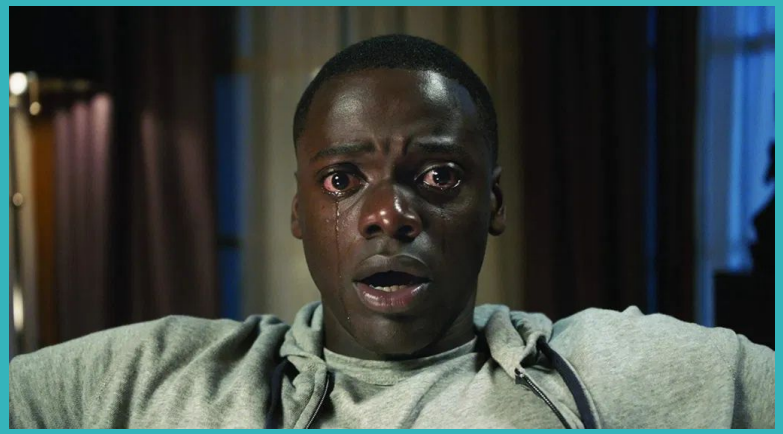


Fig. 2. Features extraction using segmented the detected image into parts.

Ashmawi et al. (2019)

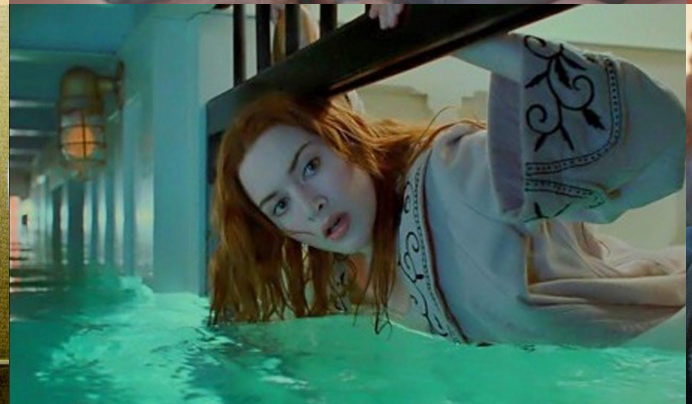
vs.

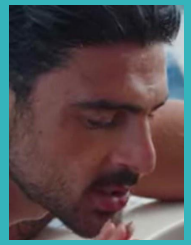


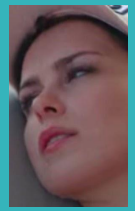


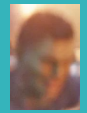








Automate body type classification for characters in film and TV using a character's face.

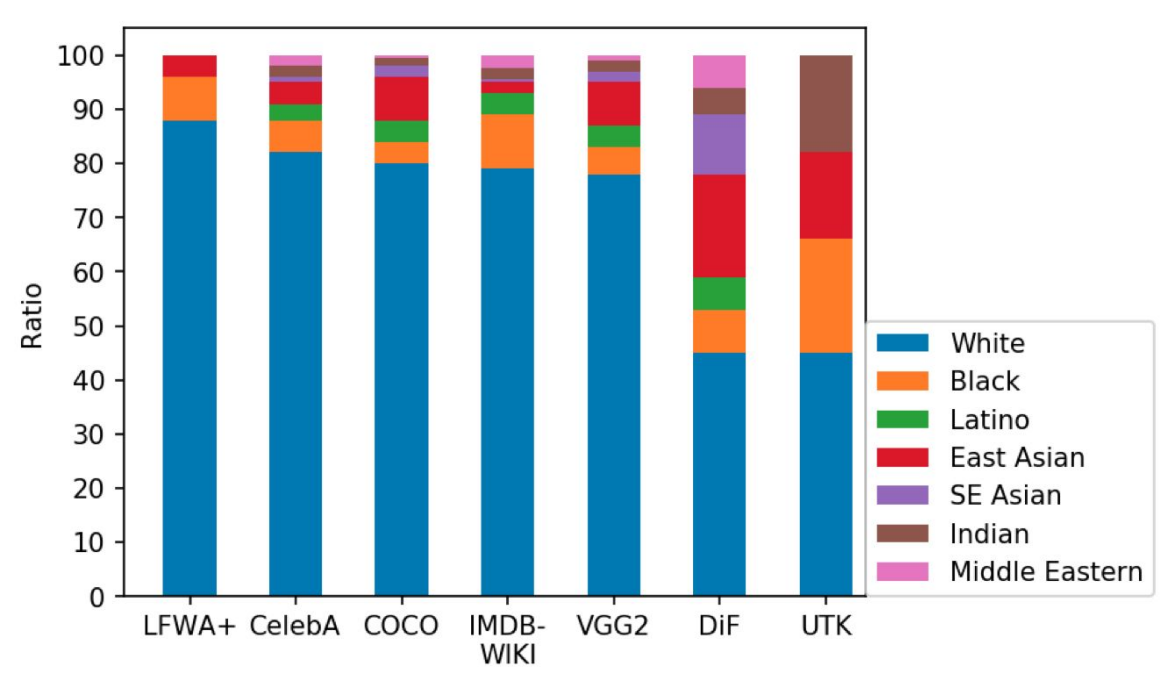
Body Type

Gender

Age

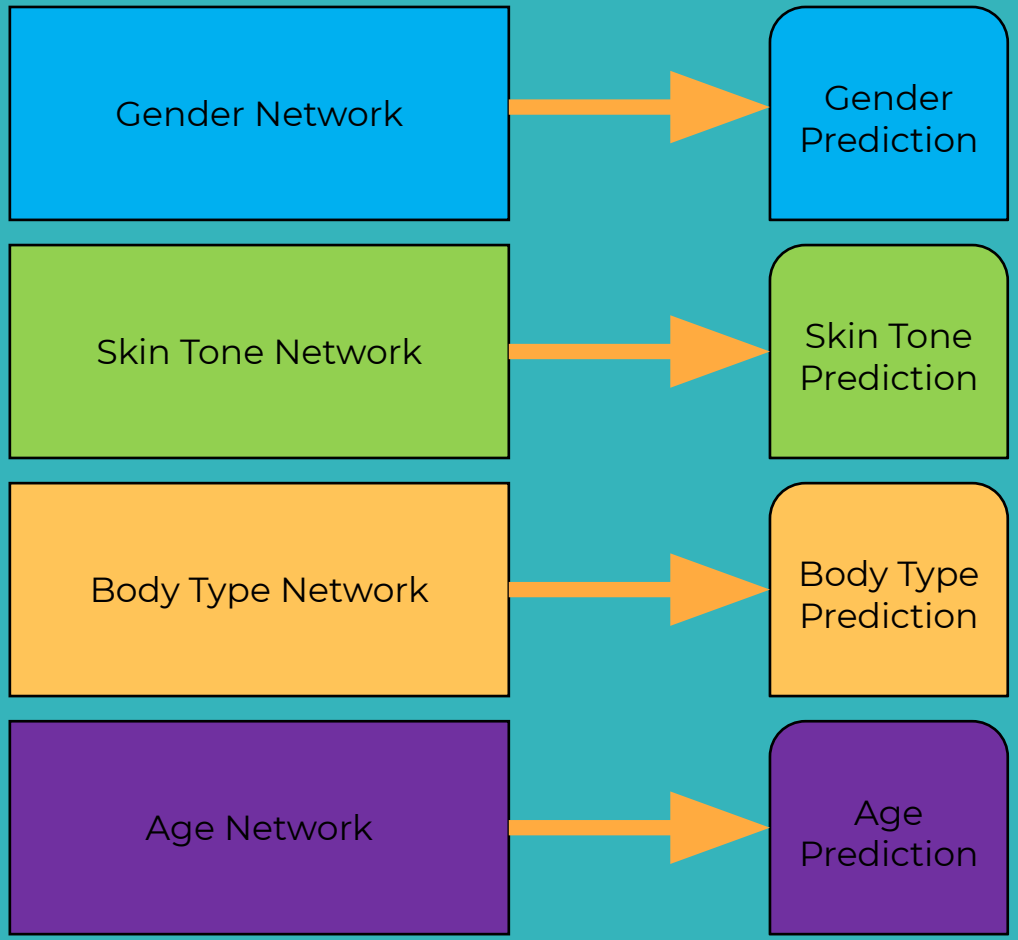
Skin Tone

Automate body type, gender, age, and skin tone classification for characters in film and TV using a character's face.

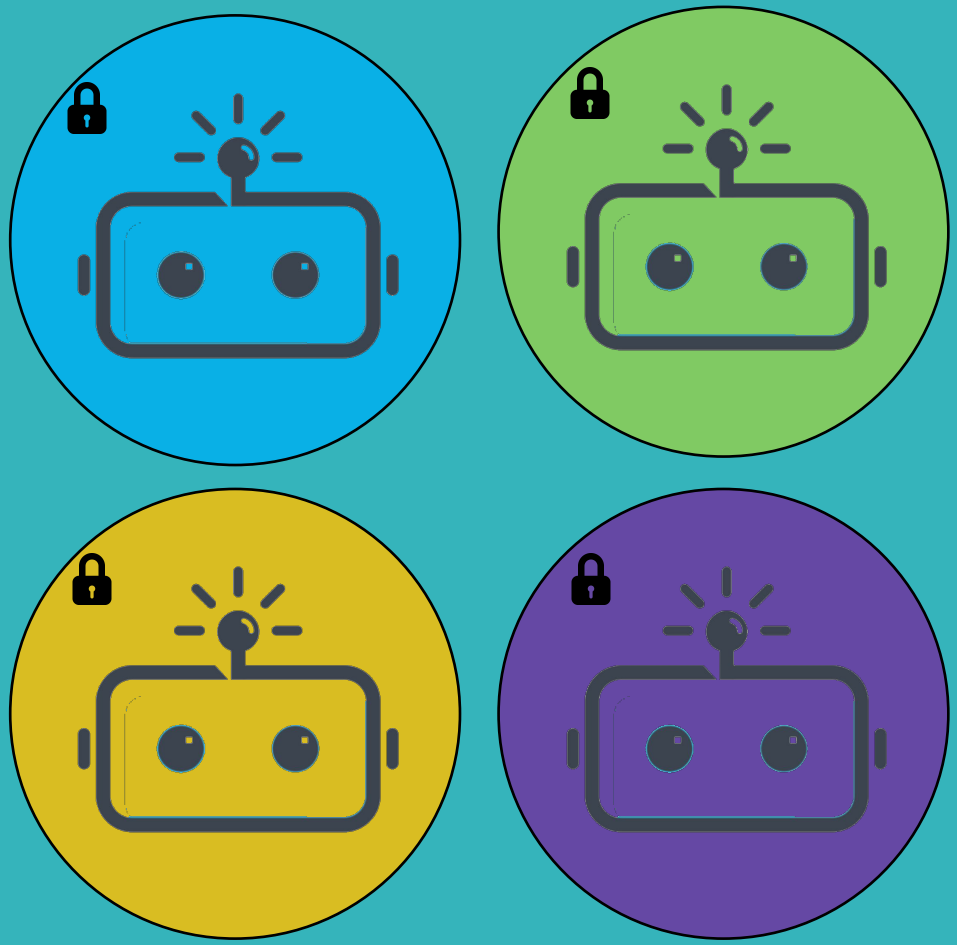

Ratio
100
90
80
70
60
50
40
30
20
10
0
LFWA+
CelebA
COCO
IMDB-WIKI
VGG2
DiF
UTK
White
Black
Latino
East Asian
SE Asian
Indian
Middle Eastern

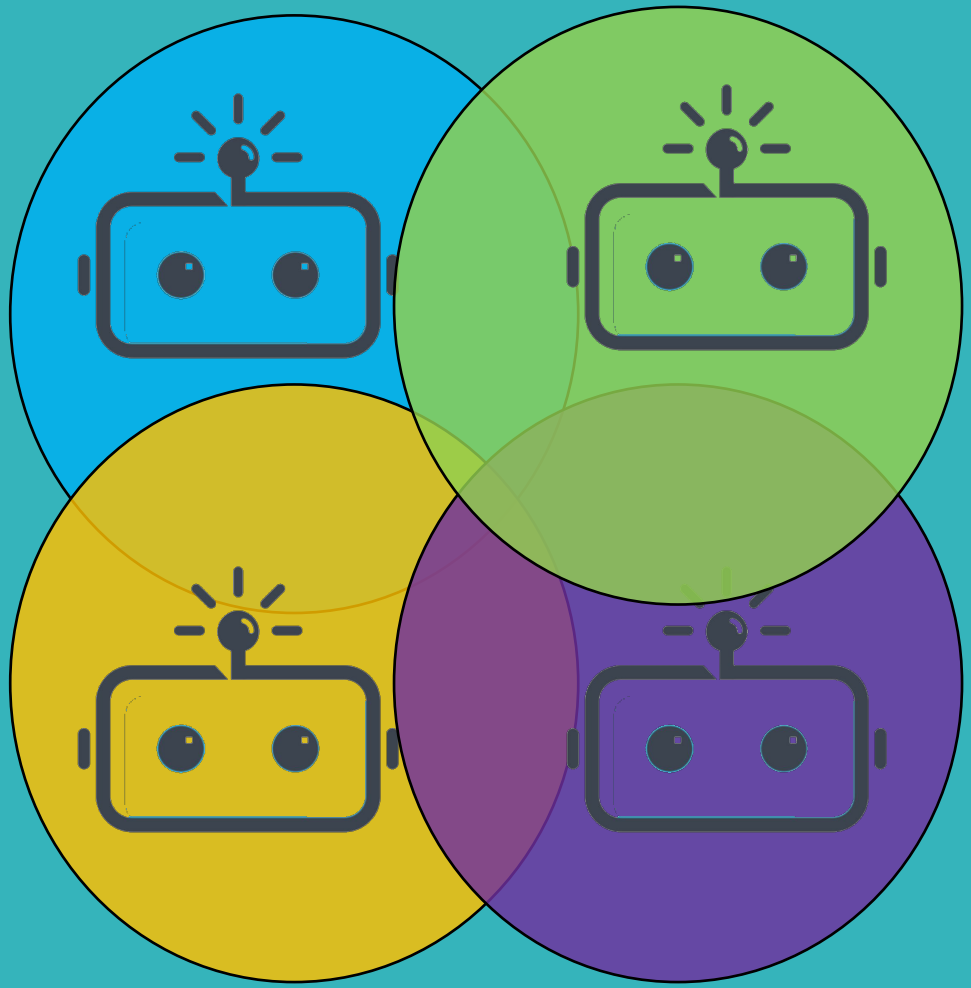
All Fire Saga data comes *exclusively* from movies and TV shows.

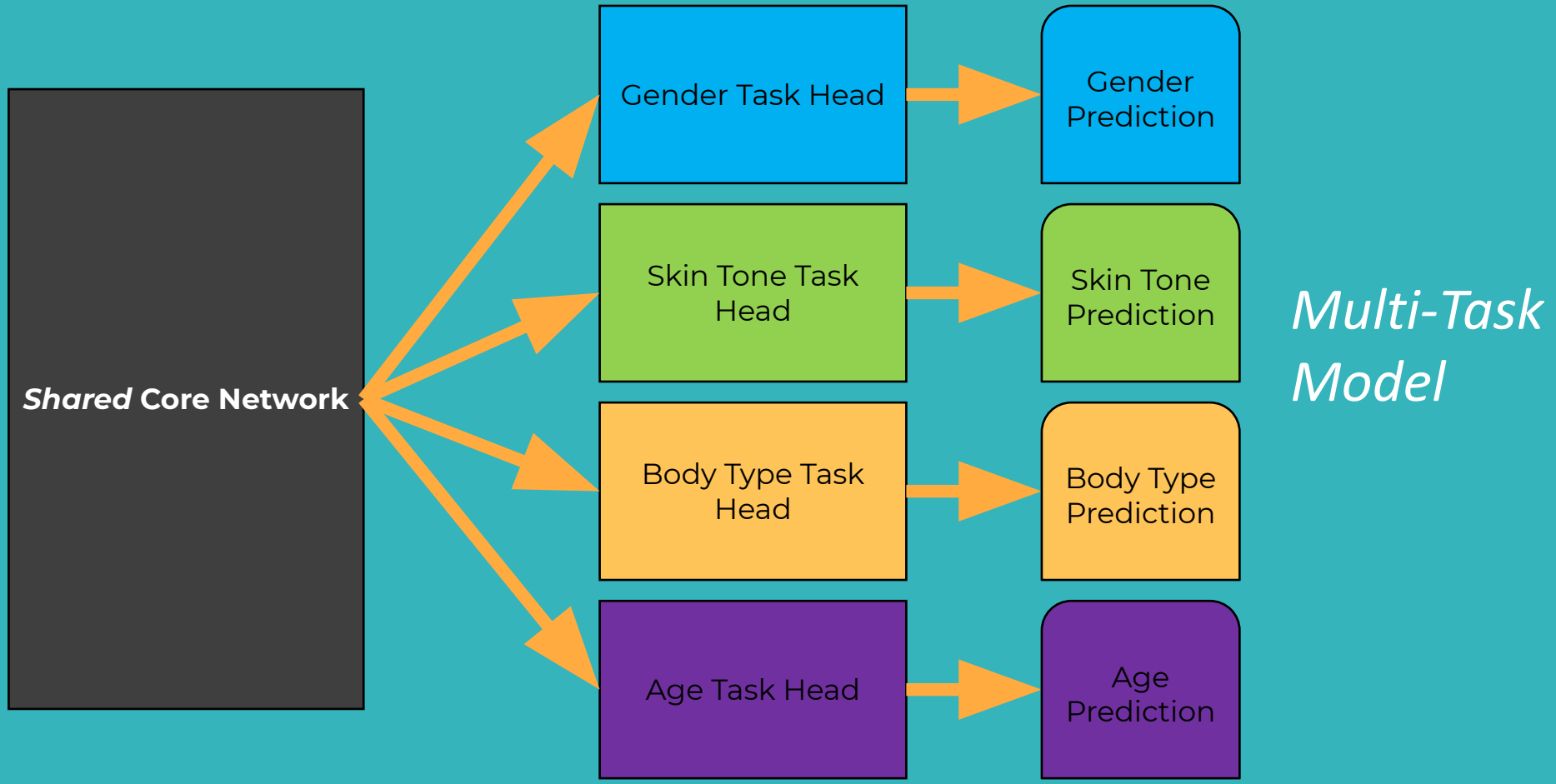
Baseline	Fire Saga
Women comprise **51%** of the U.S. population	Women comprise **50%** of the Fire Saga dataset
People of color comprise **38%** of the U.S. population	People of color comprise **49%** of the Fire Saga dataset
People ages 60+ comprise **19%** of the U.S. population	People ages 60+ comprise **22%** of the Fire Saga dataset
People with large body types comprise **39%** of the U.S. population	People with large body types comprise **26%** of the Fire Saga dataset



Single Task Models






Shared Core Network
Gender Task Head
Skin Tone Task Head
Body Type Task Head
Age Task Head
Gender Prediction
Skin Tone Prediction
Body Type Prediction
Age Prediction
Multi-Task Model

Now, each task can **share information** and **learn together** without overfitting or cheating the task.

RESULTS

Metric	Score
Gender Accuracy	**97.1%**
Skin Tone Accuracy	**92.4%**
Body Type Accuracy	**97.3%**
Age Accuracy	**98.2%**

** via 95/5 stratified train/test split.*


Gender
Skin Tone
Age
Body Type

Key questions and takeaways

1. Is measuring body size unique compared to race, and gender because it is highly stigmatized?
    a. Why do we think of it differently than other observable human characteristics?
2. Can we reclaim the word “fat”?

3. Action: Create an unconscious bias test to measure fatphobia and sizeism.

Break

Be Back @ 11:20am PT

Understanding Conflict and Abusive Language

Hosted by Wendy Chun
Simon Fraser University

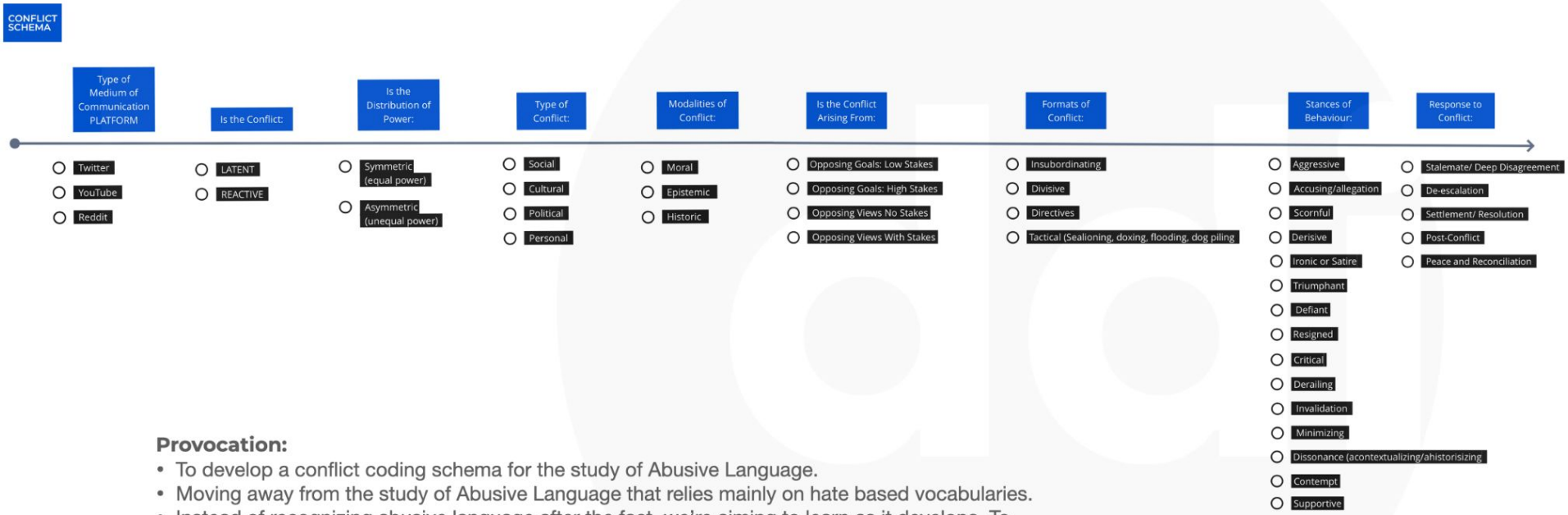
Presented by:

Attendees:

- Gillian Russell (facilitator) -DDI, SFU
- Christine Tomlinson -DDI, SFU
- Kree Cole-McLaughlin -Google
- Sabyasachee Baruah -USC
- Sarah Ciston -USC
- Shrikanath Narayanan -USC
- Victor Martinez -USC

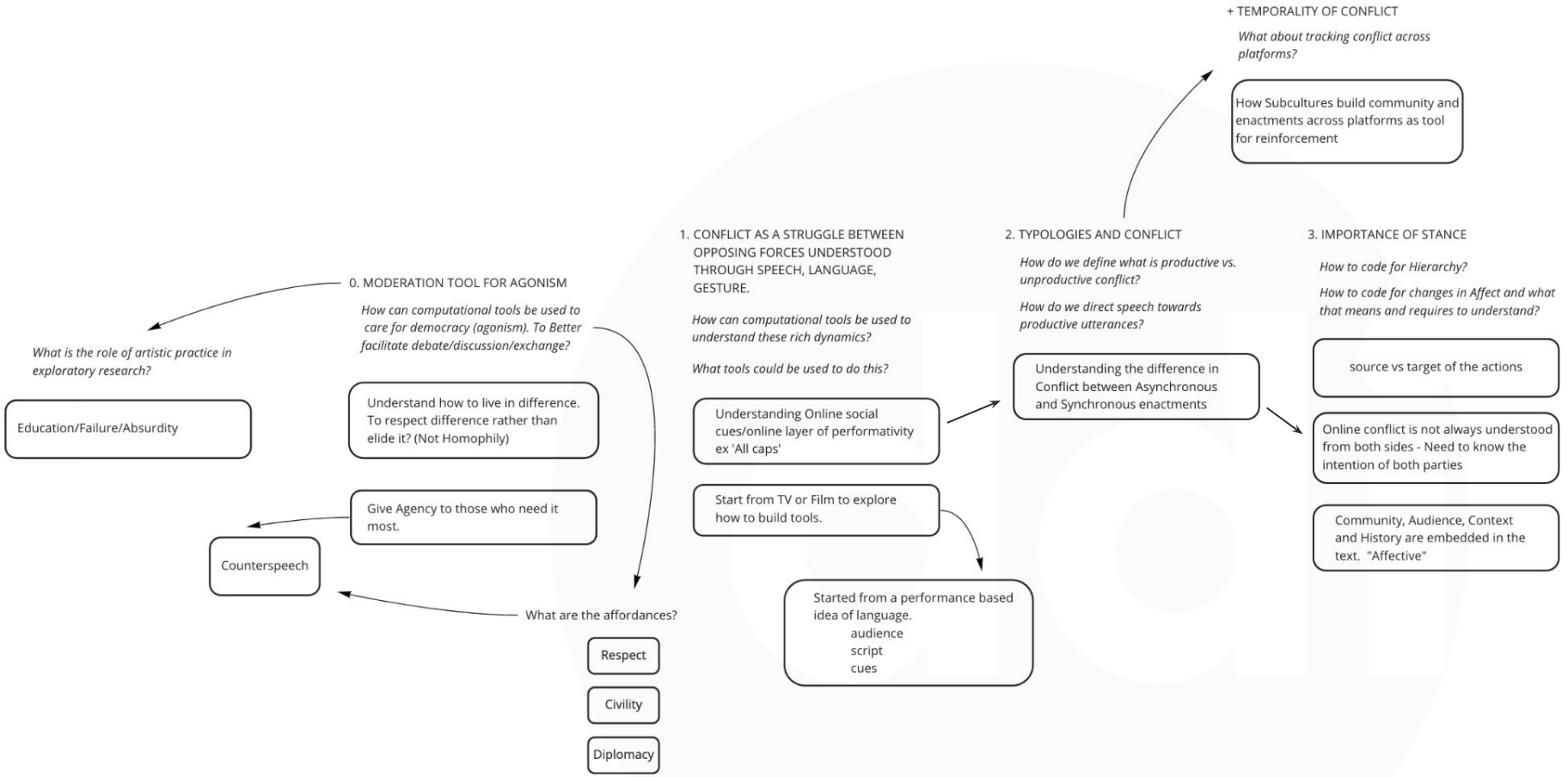
UNDERSTANDING CONFLICT AND ABUSIVE LANGUAGE

What knowledges and imaginaries are necessary to employ machine learning to create more responsive (rather than retroactive or reactionary) interventions to abuse? How might we direct AI to engage in conflict to foster user- and community-based empowerment - to understand the capacity of what emerges in conflict as a productive space to equitably intervene in and redirect abusive conversations?



Provocation:

- To develop a conflict coding schema for the study of Abusive Language.
- Moving away from the study of Abusive Language that relies mainly on hate based vocabularies.
- Instead of recognizing abusive language after the fact, we're aiming to learn as it develops. To gather contextual signals and speech patterns.


+ TEMPORALITY OF CONFLICT
What about tracking conflict across platforms?
How Subcultures build community and enactments across platforms as tool for reinforcement
1. CONFLICT AS A STRUGGLE BETWEEN OPPOSING FORCES UNDERSTOOD THROUGH SPEECH, LANGUAGE, GESTURE.
How can computational tools be used to understand these rich dynamics?
What tools could be used to do this?
Understanding Online social cues/online layer of performativity ex 'All caps'
Start from TV or Film to explore how to build tools.
Started from a performance based idea of language.
audience
script
cues
2. TYPOLOGIES AND CONFLICT
How do we define what is productive vs. unproductive conflict?
How do we direct speech towards productive utterances?
Understanding the difference in Conflict between Asynchronous and Synchronous enactments
3. IMPORTANCE OF STANCE
How to code for Hierarchy?
How to code for changes in Affect and what that means and requires to understand?
source vs target of the actions
Online conflict is not always understood from both sides - Need to know the intention of both parties
Community, Audience, Context and History are embedded in the text. "Affective"
0. MODERATION TOOL FOR AGONISM
How can computational tools be used to care for democracy (agonism). To Better facilitate debate/discussion/exchange?
What is the role of artistic practice in exploratory research?
Education/Failure/Absurdity
Understand how to live in difference. To respect difference rather than elide it? (Not Homophily)
Give Agency to those who need it most.
Counterspeech
What are the affordances?
Respect
Civility
Diplomacy

Discussion

How might we direct AI to engage in conflict to foster user- and community-based empowerment - to understand the capacity of what emerges in conflict as a productive space to equitably intervene in and redirect abusive conversations?

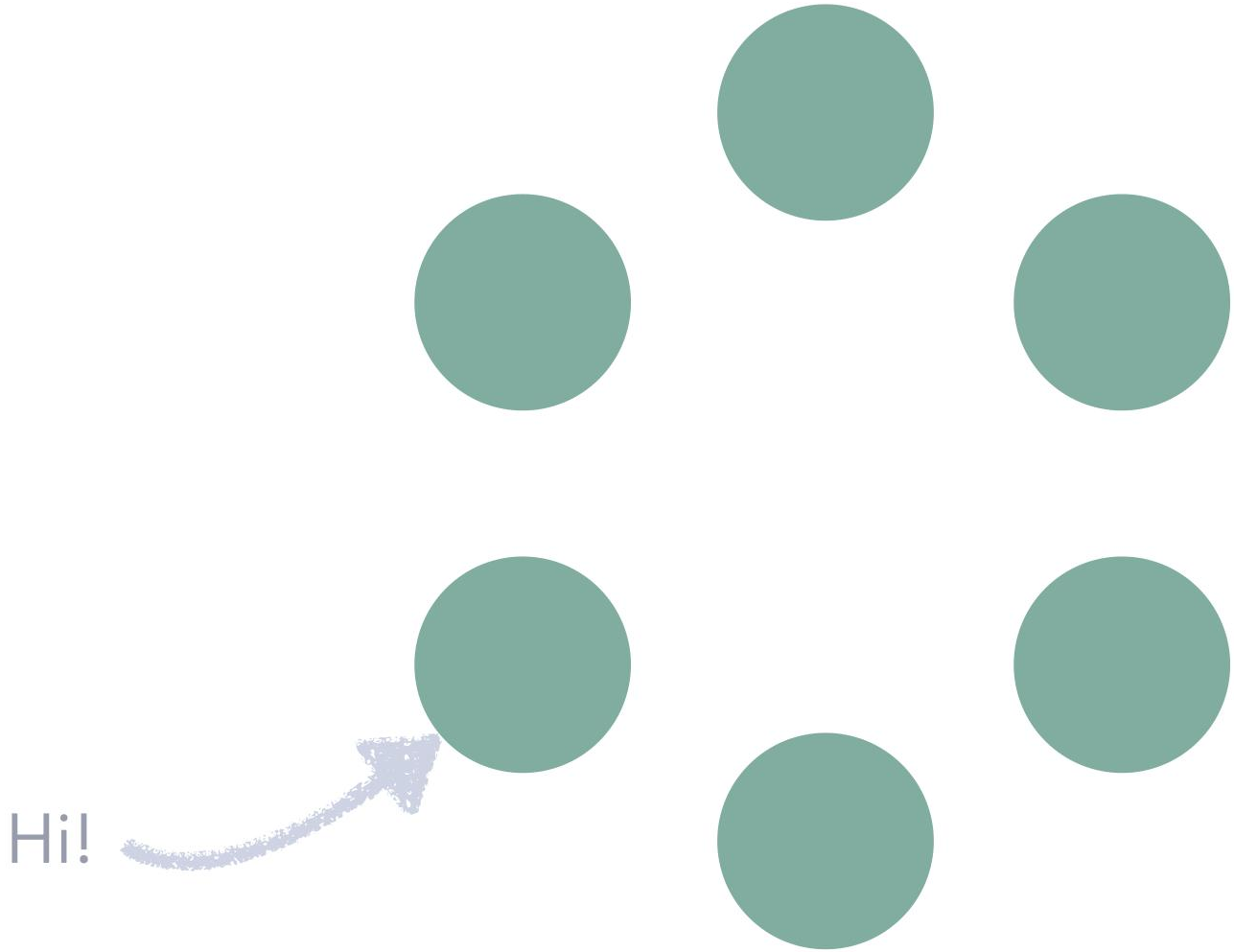
Identity and Subjectivity in Relationships & Interactions

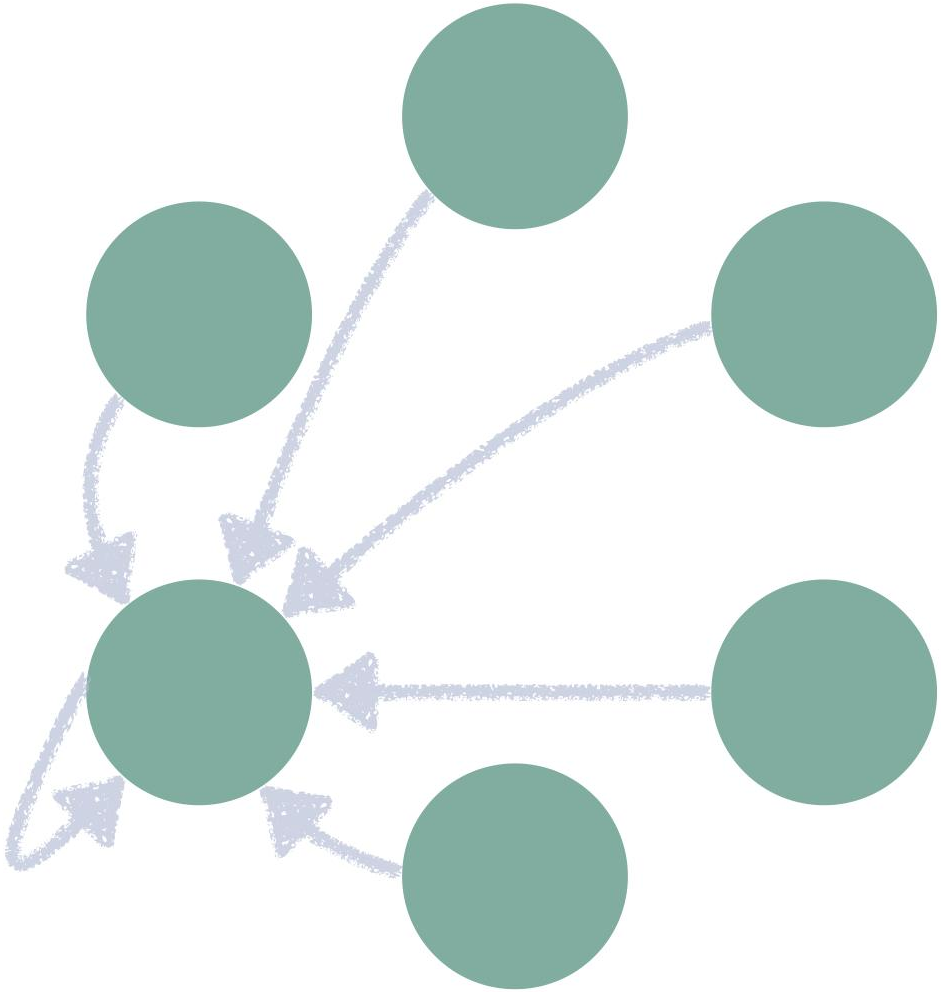
Hosted by Jed Brubaker
University of Colorado Boulder

Presented by: Jed & Kylee

Attendees:

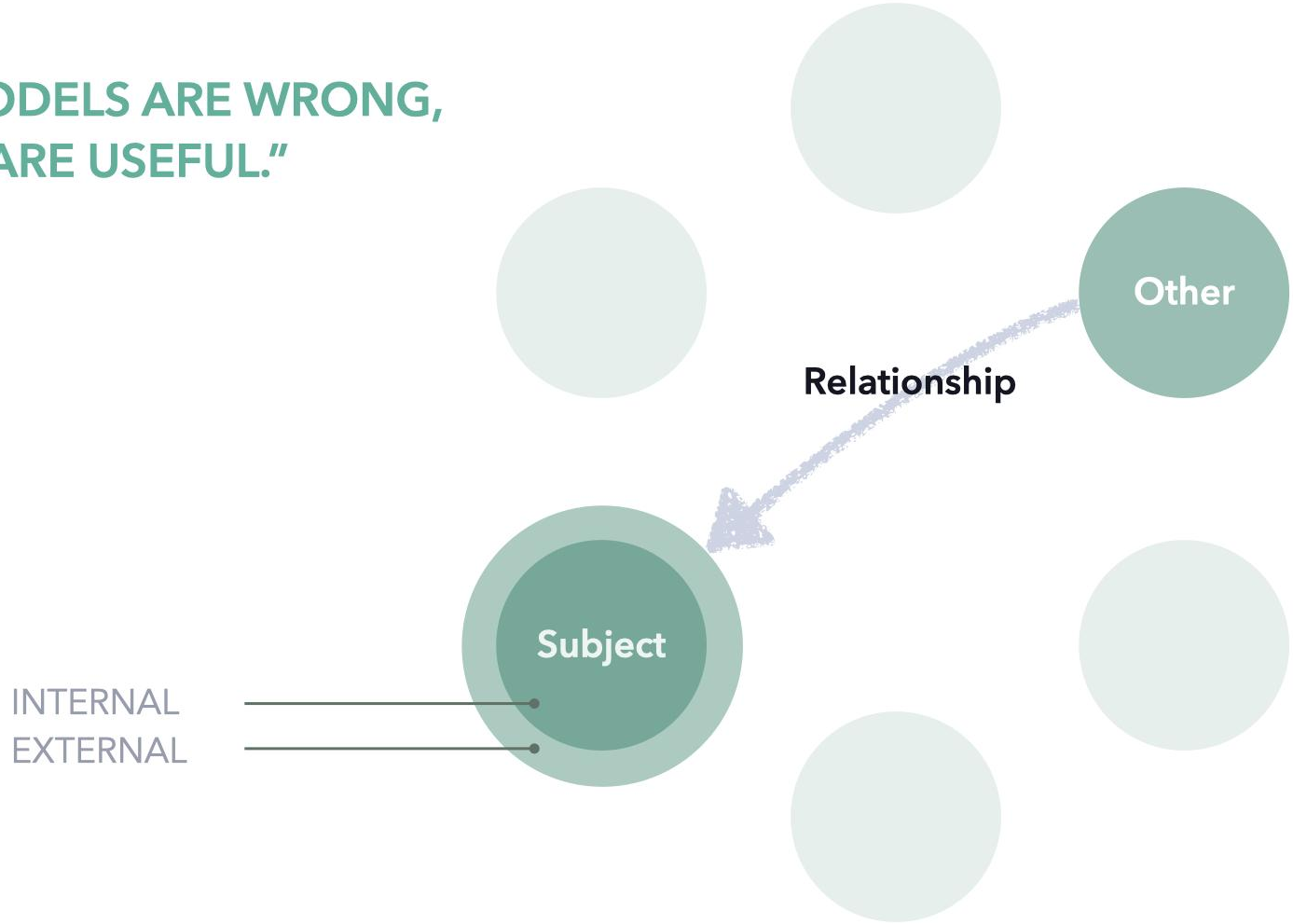
- Kylee Jaye - MUSE Team, Google Research
- Hannah Holtzclaw - Simon Fraser University
- Susanna Ricco - Soapbox Team, Google Research


Hi!



SO WHERE
IS IDENTITY?

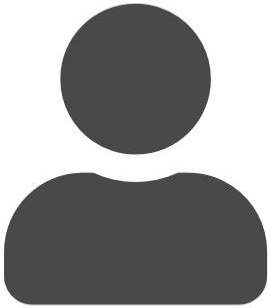
"ALL MODELS ARE WRONG,
SOME ARE USEFUL."



Historical Tech Approaches

AUTHENTICATION

DATABASE
RECORDS



THE "USER"

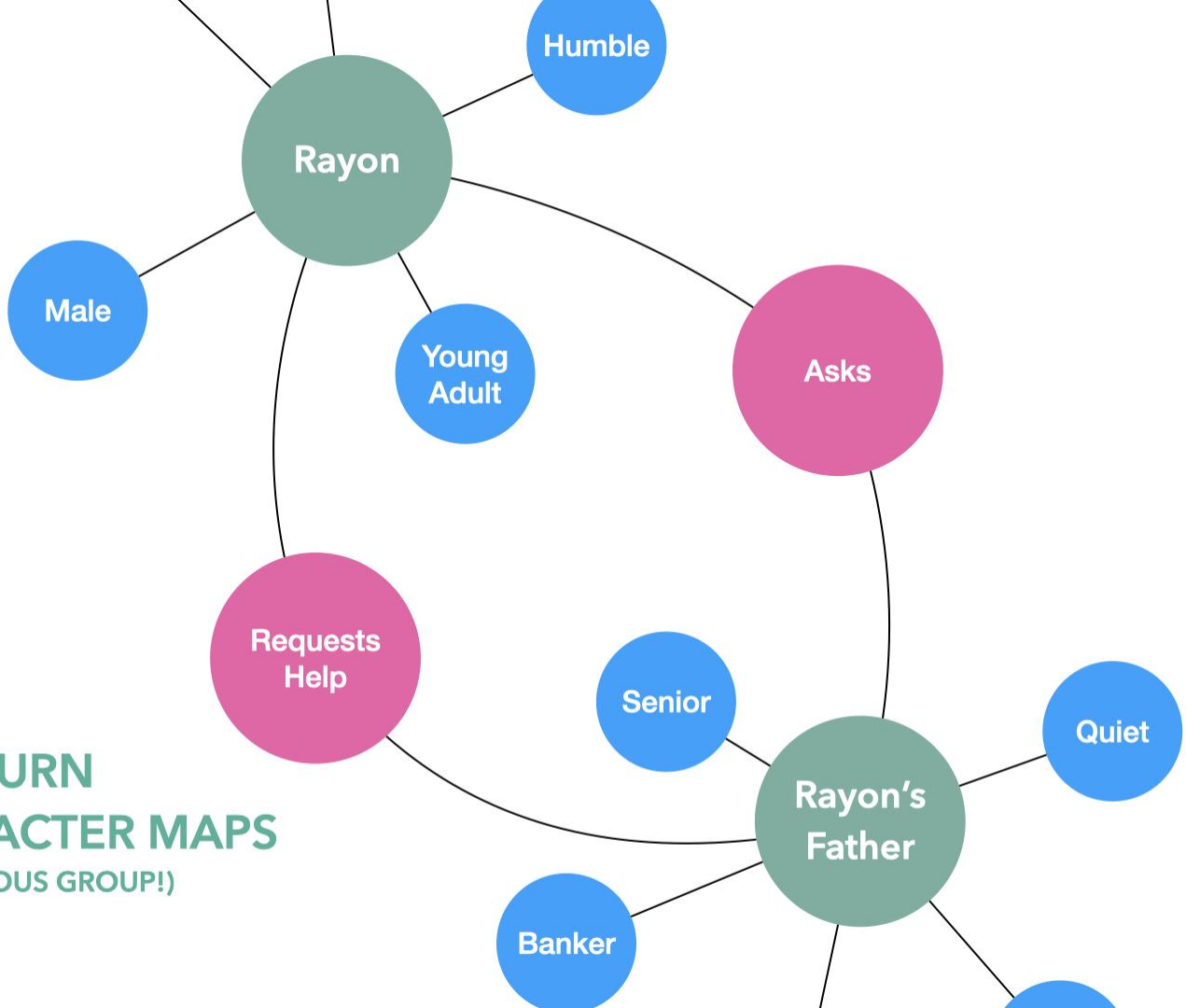
Is Kylee an artist or a scientist? Identity is situational, and the byproduct of relationships with others.

We need to reconsider our relationship with labels: Labels define particular world views that are stable in ways that our identities are not. Continuums, correlations, and affinities all invite us to think about the relationships between people and attributes in more tenuous, tentative, dynamic, and playful ways.

Labels act as proxies that occlude more interesting patterns and correlations in data. Proxies are just the symptoms of underlying problems that actually deserve our attention.

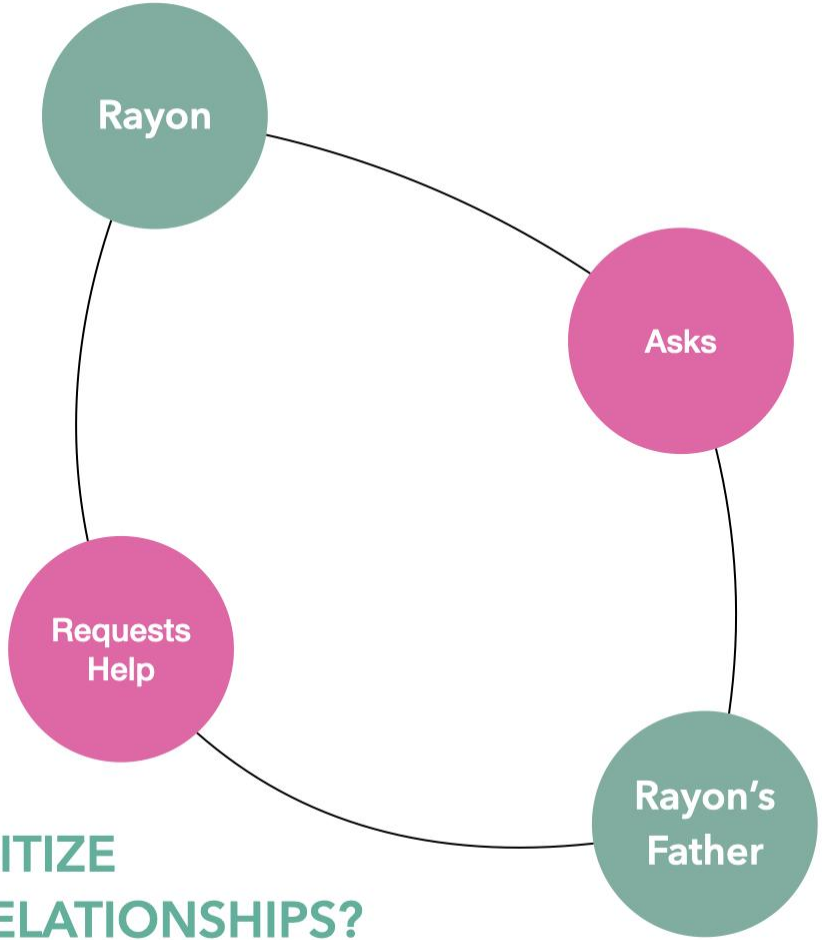
Change presents a challenge for CMI:

- Concepts evolve: How the definition of “woman” changes over time.
- People evolve: How one’s relationship with “woman” as a term changes over time.

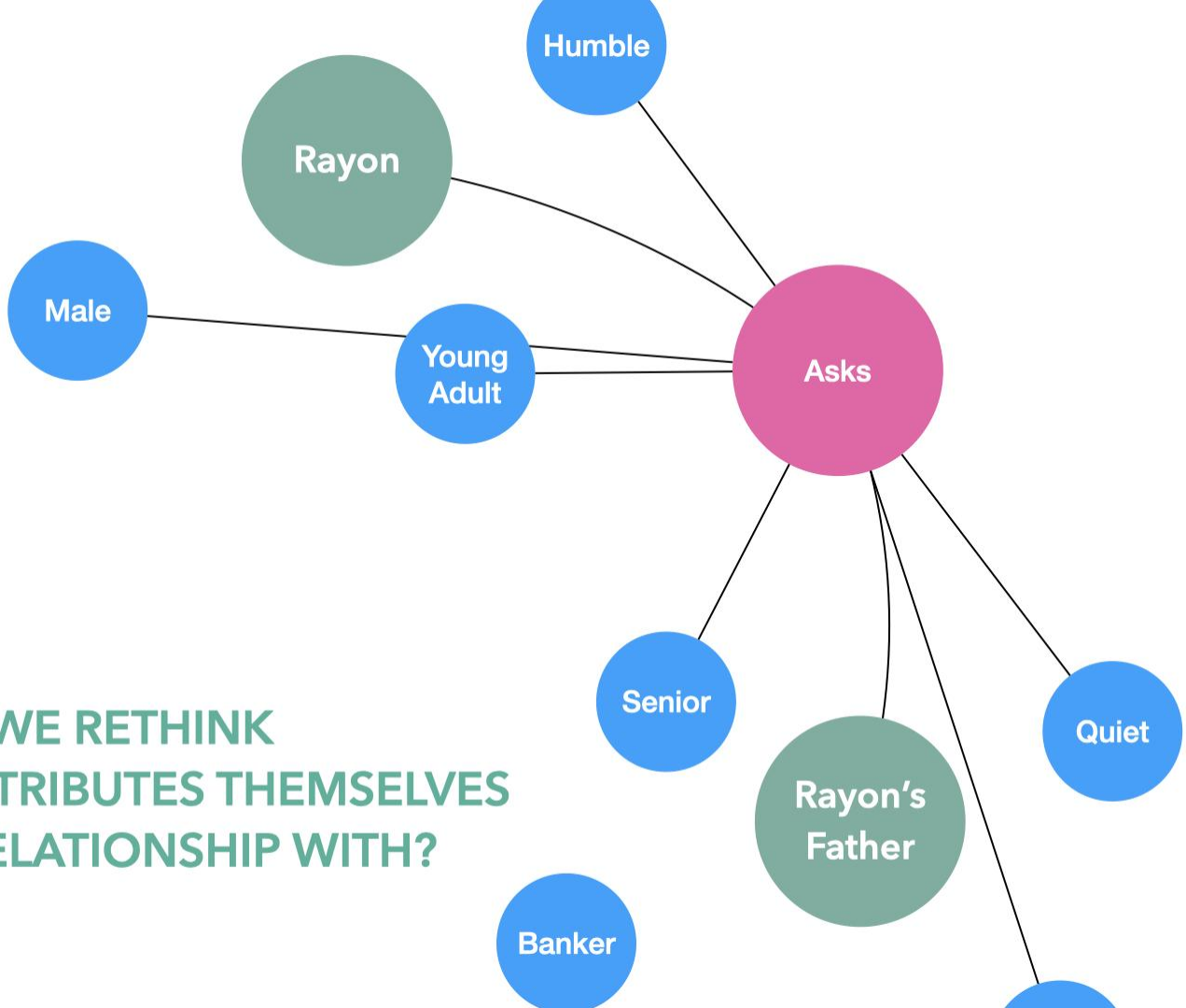


LET'S RETURN TO CHARACTER MAPS

(THANKS PREVIOUS GROUP!)



SHOULD WE PRIORITIZE INTERPERSONAL RELATIONSHIPS?



WHAT IF WE RETHINK WHAT ATTRIBUTES THEMSELVES ARE IN RELATIONSHIP WITH?

Thank You!