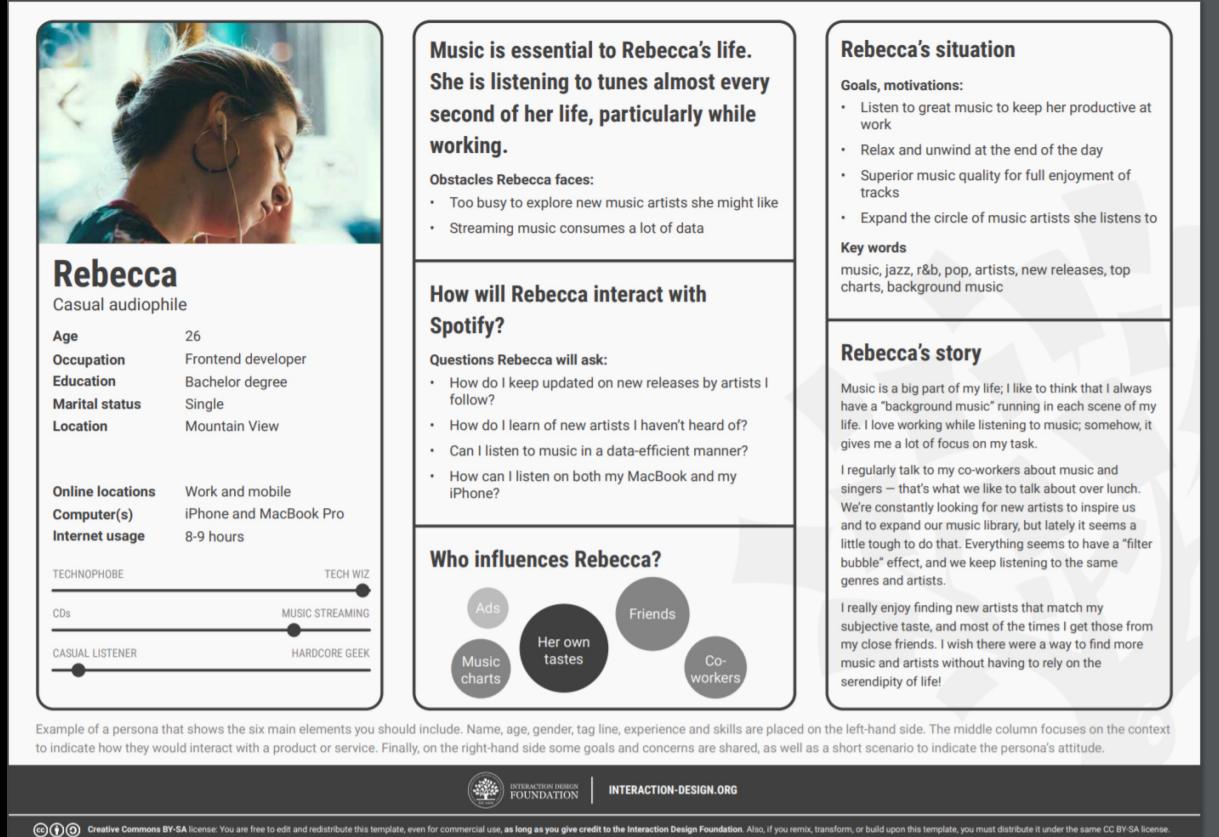


# Requirements and Intelligent Systems

Jin Guo  
SOCS McGill University

# User Personas



<https://www.interaction-design.org/literature/topics/personas>

# User Personas

- Discovered during the requirements discovery and user story development process
  - As a <type of user>, I want <some goal> so that <some reason>.***
- Represent *real* behavior patterns, attitudes, skillset, motivations and goals within the product's domain.
- Design for primary personas

What would  
Rebecca need  
to do?

Is this function going to make Rebecca's life easier?



## Rebecca

Casual audiophile

Age	26
Occupation	Frontend developer
Education	Bachelor degree
Marital status	Single
Location	Mountain View
Online locations	Work and mobile
Computer(s)	iPhone and MacBook Pro
Internet usage	8-9 hours
TECHNOPHOBIE	
CDs	MUSIC STREAMING
CASUAL LISTENER	
HARDCORE GEEK	

Example of a persona that shows the six main elements you should include. Name, age, gender, tag line, experience and skills are placed on the left-hand side. The middle column focuses on the context to indicate how they would interact with a product or service. Finally, on the right-hand side some goals and concerns are shared, as well as a short scenario to indicate the persona's attitude.

**Music is essential to Rebecca's life. She is listening to tunes almost every second of her life, particularly while working.**

**Obstacles Rebecca faces:**

- Too busy to explore new music artists she might like
- Streaming music consumes a lot of data

**Rebecca's situation**

**Goals, motivations:**

- Listen to great music to keep her productive at work
- Relax and unwind at the end of the day
- Superior music quality for full enjoyment of tracks
- Expand the circle of music artists she listens to

**Key words**

music, jazz, r&b, pop, artists, new releases, top charts, background music

**How will Rebecca interact with Spotify?**

**Questions Rebecca will ask:**

- How do I keep updated on new releases by artists I follow?
- How do I learn of new artists I haven't heard of?
- Can I listen to music in a data-efficient manner?
- How can I listen on both my MacBook and my iPhone?

**Rebecca's story**

Music is a big part of my life; I like to think that I always have a "background music" running in each scene of my life. I love working while listening to music; somehow, it gives me a lot of focus on my task.

I regularly talk to my co-workers about music and singers – that's what we like to talk about over lunch. We're constantly looking for new artists to inspire us and to expand our music library, but lately it seems a little tough to do that. Everything seems to have a "filter bubble" effect, and we keep listening to the same genres and artists.

I really enjoy finding new artists that match my subjective taste, and most of the times I get those from my close friends. I wish there were a way to find more music and artists without having to rely on the serendipity of life!

<https://www.interaction-design.org/literature/topics/personas>

# User Personas

What would  
Rebecca need  
to do?

Is this function going to  
make Rebecca's life  
easier?

**Rebecca**  
Casual audiophile  
Age 26  
Occupation Frontend developer  
Education Bachelor degree  
Marital status Single  
Location Mountain View

Online locations Work and mobile  
Computer(s) iPhone and MacBook Pro  
Internet usage 8-9 hours

TECHNOPHONE TECH WIZ  
CDE MUSIC STREAMING  
CASUAL LISTENER HARDCORE GEEK

**Music is essential to Rebecca's life. She is listening to tunes almost every second of her life, particularly while working.**

**Obstacles Rebecca faces:**

- Too busy to explore new music artists she might like
- Streaming music consumes a lot of data

**How will Rebecca interact with Spotify?**

**Questions Rebecca will ask:**

- How do I keep updated on new releases by artists I follow?
- How do I learn of new artists I haven't heard of?
- Can I listen to music in a data-efficient manner?
- How can I listen on both my MacBook and my iPhone?

**Who influences Rebecca?**

Ads, Her own tastes, Friends, Music charts, Co-workers

Example of a persona that shows the six main elements you should include. Name, age, gender, tag line, experience and skills are placed on the left-hand side. The middle column focuses on the context to indicate how they would interact with a product or service. Finally, on the right-hand side some goals and concerns are shared, as well as a short scenario to indicate the persona's attitude.

Creative Commons BY-SA license. You are free to edit and redistribute this template, even for commercial use, as long as you give credit to the Interaction Design Foundation. Also, if you remix, transform, or build upon this template, you must distribute it under the same CC BY-SA license.

<https://www.interaction-design.org/literature/topics/personas>

# User Personas

## Empathy

**Rebecca**  
Casual audiophile  
Age 26  
Occupation Frontend developer  
Education Bachelor degree  
Marital status Single  
Location Mountain View

Online locations Work and mobile  
Computer(s) iPhone and MacBook Pro  
Internet usage 8-9 hours

TECHNOPHONE TECH WIZ  
CDE MUSIC STREAMING  
CASUAL LISTENER HARDCORE GEEK

**Music is essential to Rebecca's life. She is listening to tunes almost every second of her life, particularly while working.**

**Obstacles Rebecca faces:**

- Too busy to explore new music artists she might like
- Streaming music consumes a lot of data

**How will Rebecca interact with Spotify?**

**Questions Rebecca will ask:**

- How do I keep updated on new releases by artists I follow?
- How do I learn of new artists I haven't heard of?
- Can I listen to music in a data-efficient manner?
- How can I listen on both my MacBook and my iPhone?

**Who influences Rebecca?**

Ads, Her own tastes, Friends, Music charts, Co-workers

Example of a persona that shows the six main elements you should include. Name, age, gender, tag line, experience and skills are placed on the left-hand side. The middle column focuses on the context to indicate how they would interact with a product or service. Finally, on the right-hand side some goals and concerns are shared, as well as a short scenario to indicate the persona's attitude.

Creative Commons BY-SA license: You are free to edit and redistribute this template, even for commercial use, as long as you give credit to the Interaction Design Foundation. Also, if you remix, transform, or build upon this template, you must distribute it under the same CC BY-SA license.

INTERACTION DESIGN FOUNDATION | INTERACTION-DESIGN.ORG

<https://www.interaction-design.org/literature/topics/personas>

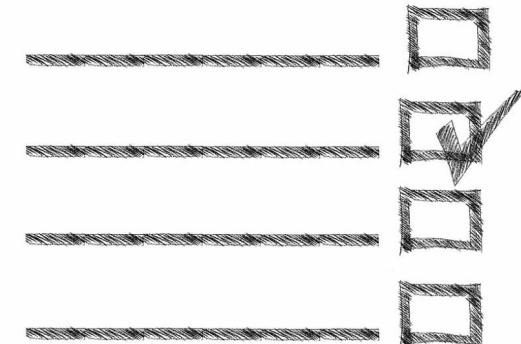
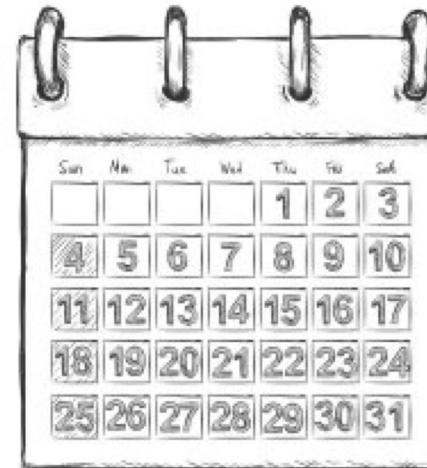
*"On a project for servicing coin-operated vending machines, I introduced Dolly the Serviceperson and Ralph the Warehouse Supervisor. We wrote scenarios for them and they became part of the project team—virtually."*

Software Requirements by Karl Eugene Wiegers, Joy Beatty, Microsoft Press, 2013

# Activity

Design a schedule management experience for

*The person next to you*



## **Step1: Interview**

8 mins (4 mins each)

"Who are involved with the scheduling task? Why was it important? How actually managed the scheduling tasks"  
"What was difficult about performing that task?"

Notes from the first interview

## Step2: Dig Deeper

6 mins (3 mins each)

follow up on things that intrigued you during the first interview. Try to dig for **stories, feelings, and emotion**. Ask 'WHY?' often

Notes from the second interview

## Step3: Capture Findings

3 mins

### Goals and Wishes

What is your partner trying to achieve through interaction with scheduling tool?

*User verbs*

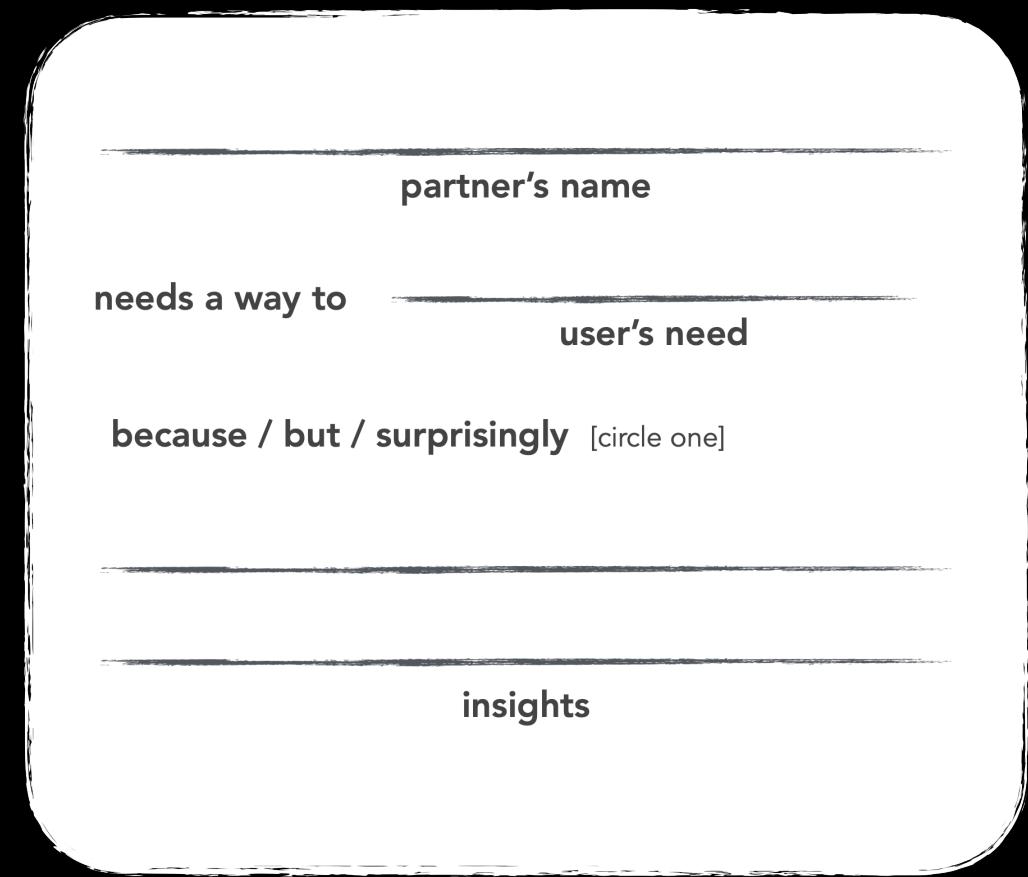
### Insights

New learnings about your partner's feelings and motivations. what's something you see about your partner's experience that maybe s/he doesn't see?\*

*Make inferences from what you heard*

## **Step4: Taking a stand with a point-of-view**

3 mins



## **Step5: Make a case for or against AI feature**

3 mins

I think AI can / cannot [circle one] **help solve**

**user's need**

**because** \_\_\_\_\_

## **Step6: Designing the reward function for the new feature**

5 mins

Our AI model will be optimized for

related to user's need

because \_\_\_\_\_

# Requirement Elicitation

- A collaborative and analytical process

Interviews

Workshops

Focus Groups

Observations

Questionnaires

# Workshop

- Facilitated and formal
- Different stakeholders
- Deliverables



# Focus Group

- Less formal
- Representative users
- Exploration



# Requirement Elicitation

- A collaborative and analytical process

Interviews

System Interface Analysis

Workshops

User Interface Analysis

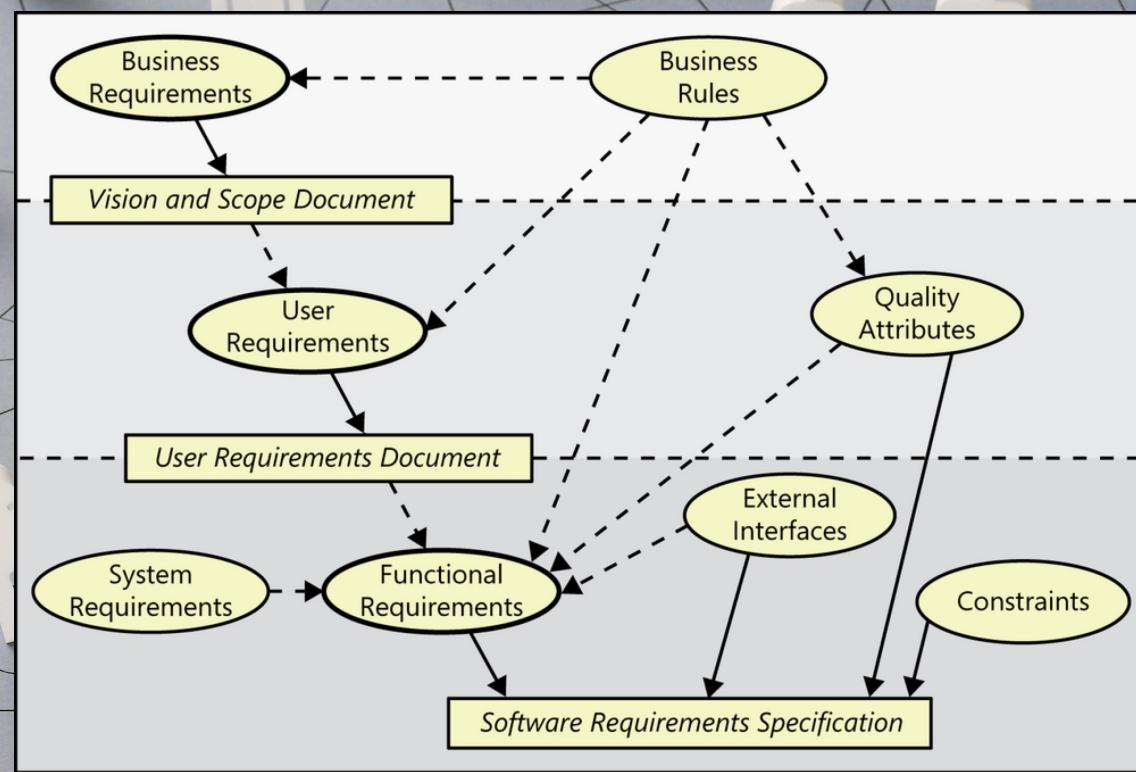
Focus Groups

Document Analysis

Observations

Questionnaires

# What is Requirement again?



# Requirements Traps

- Not requirements
- Assumed, Implied Requirements
- Missing requirements

# What is Requirement again?



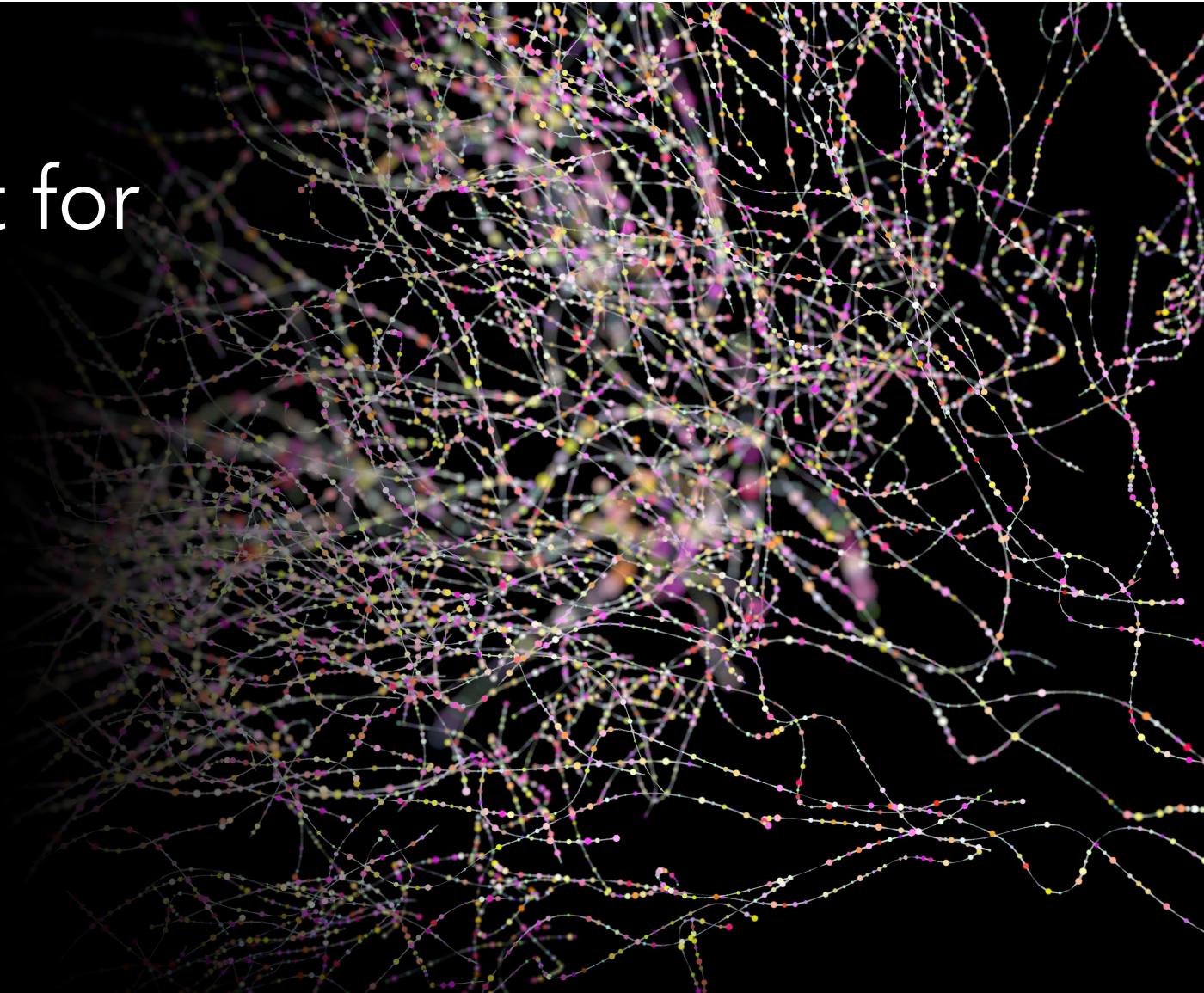
Is what enough?



Ask the why and  
who question

[Keynote talk by Amy  
Ko at RE 2021](#)

# Requirement for AI systems



# Acceptance Test

***As a consumer, I am always aware of my current***

***As a consumer, I always see current energy pricing reflected on my portal and on-premise devices so that I know that my energy usage costs are accurate and reflect any utility pricing***

- ***Verify the current pricing is always used and the calculated numbers are displayed correctly on the portal and each on-premise device (see attachment for formats).***
- ***Verify the pricing and the calculated numbers are updated correctly when the price changes.***
- ***Verify the “current price” field itself is updated according to the scheduled time.***
- ***Verify the info/error messages when there is a fault in the pricing (see approved error messages attached).***

# Intersection of User needs and AI strength

As a *<type of user>*, I want *<some goal>* so that *<some reason>*.

I think AI can / cannot [circle one] help solve

user's need

because

Our AI model will be optimized for

related to user's need

because

# Intersection of User needs and AI strength

*As a <type of user>, I want <some goal> so that <some reason>.*

Recommending different content to different users?

Prediction of future events?

Recognize patterns in images?

Understand natural language?

.....

Maintaining predictability?

Provide static information?

Complete transparency?

.....

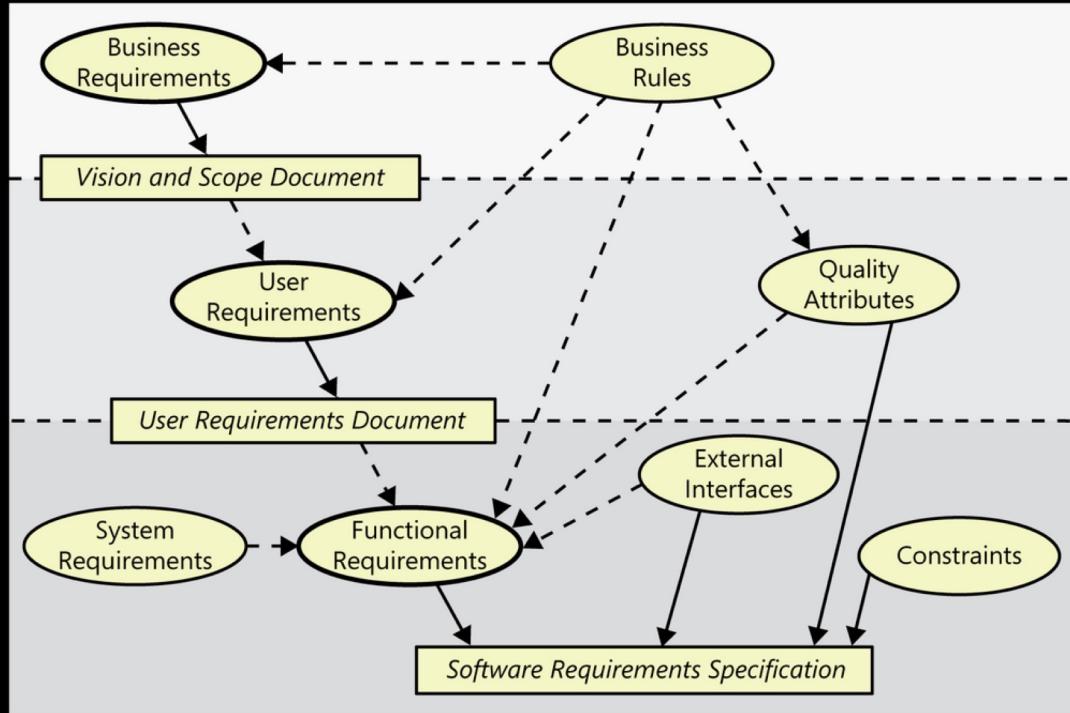
# Specification for ML Components

- **Formal methods perspectives**
  - System level specification - standard specification languages
  - Input-output robustness - Local and Global resilient
  - Input-output relations - pre-condition, post condition
  - Semantic Invariance - Domain-specific, e.g., scale invariant
  - Distribution Assumptions - statistical modelling, e.g., Probabilistic Programming
  - Fairness - e.g., individual fairness, demographic parity, etc. still evolving area
  - .....

Seshia, Sanjit A., et al. "Formal specification for deep neural networks." International Symposium on Automated Technology for Verification and Analysis. Springer, Cham, 2018.

# Challenges

- Understand and define non-functional requirements
- Reasoning the relations between different level of requirements



Rahimi, Mona, Jin LC Guo, Sahar Kokaly, and Marsha Chechik.  
"Toward Requirements Specification for Machine-Learned Components."  
In 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW), pp. 241-244. IEEE, 2019.

Dreossi, Tommaso, Alexandre Donzé, and Sanjit A. Seshia.  
"Compositional falsification of cyber-physical systems with machine learning components."  
Journal of Automated Reasoning 63.4 (2019): 1031-1053.

# Dataset Specification and Documentation

- Why is the dataset created? (e.g., is there a specific intended task gap that needed to be filled? Who funded the creation of the dataset?)
- What preprocessing/cleaning is done? (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances)
- If it relates to people, are they told what the dataset would be used for and did they consent? If so, how? Were they provided with any mechanism to revoke their consent in the future or for certain uses?
- Will the dataset be updated? How often, by whom?

# Datasheet

## Motivation for Dataset Creation

**Why was the dataset created?** (e.g., were specific tasks in mind, or a specific gap that needed to be filled?)

**What (other) tasks could the dataset be used for?** (e.g., are there obvious tasks for which it should not be used?)

**Has the dataset been used for any tasks where are the results so others can compare them?** (e.g., published papers?)

**Who funded the creation of the dataset?** (e.g., provide the name of the grant, provide the grant number)

**Any other comments?**

Gebru, Timnit, Jamie Morgenstern, Hanna Wallach, Hal Daumé II

## Data Collection Process

**How was the data collected?** (e.g., hardware apparatus/sensor, manual human curation, software program, software interface/API; how were these constructs/measures/methods validated?)

**Who was involved in the data collection process?** (e.g., students, crowdworkers) How were they compensated? (e.g., how much were crowdworkers paid?)

**Over what time-frame was the data collected?** Does the collection time-frame match the creation time-frame?

**How was the data associated with each instance acquired?** Was the data directly observable (e.g., raw text, movie ratings), reported by subjects (e.g., survey responses), or indirectly inferred/derived from other data (e.g., part of speech tags; model-based guesses for age or language)? If the latter two, were they validated/verified and if so how?

**Does the dataset contain all possible instances?** Or is it, for instance, a sample (not necessarily random) from a larger set of instances?

**If the dataset is a sample, then what is the population?** What was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)? Is the sample representative of the larger set (e.g., geographic coverage)? If not, why not (e.g., to cover a more diverse range of instances)? How does this affect possible uses?

**Is there information missing from the dataset and why?** (this does not include intentionally dropped instances; it might include, e.g., redacted text, withheld documents) Is this data missing because it was unavailable?

**Are there any known errors, sources of noise, or redundancies in the data?**

**Any other comments?**

## Dataset Composition

**What are the instances?** (that is, examples; e.g., documents, images, people, countries) Are there multiple types of instances? (e.g., movies, users, ratings; people, interactions between them; nodes, edges)

**Are relationships between instances made explicit in the data?** (e.g., social network links, user/movie ratings, etc.)

**How many instances of each type are there?**

**What data does each instance consist of?** “Raw” data (e.g., unprocessed text or images)? Features/attributes? Is there a label/target associated with instances? If the instances are related to people, are subpopulations identified (e.g., by age, gender, etc.) and what is their distribution?

**Is everything included or does the data rely on external resources?** (e.g., websites, tweets, datasets) If external resources, a) are there guarantees that they will exist, and remain constant, over time; b) is there an official archival version. Are there licenses, fees or rights associated with any of the data?

**Are there recommended data splits or evaluation measures?** (e.g., training, development, testing; accuracy/AUC)

**What experiments were initially run on this dataset?** Have a summary of those results and, if available, provide the link to a paper with more information here.

**Any other comments?**

TOE Datasets

<b>Data Preprocessing</b>
<b>What preprocessing/cleaning was done?</b> (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values, etc.)
<b>Was the “raw” data saved in addition to the preprocessed/cleaned data?</b> (e.g., to support unanticipated future uses)
<b>Is the preprocessing software available?</b>
<b>Does this dataset collection/processing procedure achieve the motivation for creating the dataset stated in the first section of this datasheet?</b>
<b>Any other comments?</b>

<b>Dataset Distribution</b>
<b>How is the dataset distributed?</b> (e.g., website, API, etc.; does the data have a DOI; is it archived redundantly?)
<b>When will the dataset be released/first distributed?</b> (Is there a canonical paper/reference for this dataset?)
<b>What license (if any) is it distributed under?</b> Are there any copyrights on the data?
<b>Are there any fees or access/export restrictions?</b>
<b>Any other comments?</b>
<b>Dataset Maintenance</b>
<b>Who is supporting/hosting/maintaining the dataset?</b> How does one contact the owner/curator/manager of the dataset (e.g. email address, or other contact info)?
<b>Will the dataset be updated?</b> How often and by whom? How will updates/revisions be documented and communicated (e.g., mailing list, GitHub)? Is there an erratum?
<b>If the dataset becomes obsolete how will this be communicated?</b>
<b>Is there a repository to link to any/all papers/systems that use this dataset?</b>
<b>If others want to extend/augment/build on this dataset, is there a mechanism for them to do so?</b> If so, is there a process for tracking/assessing the quality of those contributions. What is the process for communicating/distributing these contributions to users?
<b>Any other comments?</b>

<b>Legal &amp; Ethical Considerations</b>
<b>If the dataset relates to people (e.g., their attributes) or was generated by people, were they informed about the data collection?</b> (e.g., datasets that collect writing, photos, interactions, transactions, etc.)
<b>If it relates to other ethically protected subjects, have appropriate obligations been met?</b> (e.g., medical data might include information collected from animals)
<b>If it relates to people, were there any ethical review applications/reviews/approvals?</b> (e.g. Institutional Review Board applications)
<b>If it relates to people, were they told what the dataset would be used for and did they consent? What community norms exist for data collected from human communications?</b> If consent was obtained, how? Were the people provided with any mechanism to revoke their consent in the future or for certain uses?
<b>If it relates to people, could this dataset expose people to harm or legal action?</b> (e.g., financial social or otherwise) What was done to mitigate or reduce the potential for harm?
<b>If it relates to people, does it unfairly advantage or disadvantage a particular social group?</b> In what ways? How was this mitigated?
<b>If it relates to people, were they provided with privacy guarantees?</b> If so, what guarantees and how are these ensured?
<b>Does the dataset comply with the EU General Data Protection Regulation (GDPR)?</b> Does it comply with any other standards, such as the US Equal Employment Opportunity Act?
<b>Does the dataset contain information that might be considered sensitive or confidential?</b> (e.g., personally identifying information)
<b>Does the dataset contain information that might be considered inappropriate or offensive?</b>
<b>Any other comments?</b>

On Thursday:

Team and Collaboration