LG 467 Computers in Linguistics

[1-2021] Wrap-up

Sakol Suethanapornkul

We've reached the end...

We want computers to be able to process & understand natural languages..

..so that they can complete useful tasks for us:

- from reminding us about upcoming events
- through booking a flight
- to answering medical questions

We've reached the end...

- *Theoretical computational linguistics*:
  - some theories (finite-state automata, context-free grammar, etc)

- *Applied computational linguistics*:
  - text processing
  - NLP applications as discussed in *Language and Computers* (tokenization, morphological analyzer, POS tagging, etc.)

We've reached the end...

mostly solved

Spam detection
OK, let's meet by the big ... ✓
D1ck too small? Buy V1AGRA ... ✗

Text categorization
Phillies shut down Rangers 2-0 — SPORTS
Jobless rate hits two-year low — BUSINESS

Part-of-speech (POS) tagging
ADJ ADJ NOUN VERB ADV
Colorless green ideas sleep furiously.

Named entity recognition (NER)
PERSON ORG LOC
Obama met with UAW leaders in Detroit ...

Information extraction (IE)
You're invited to our bunga bunga party, Friday May 27 at 8:30pm in Cordura Hall
Party
May 27
add

making good progress

Sentiment analysis
The pho was authentic and yummy.
Waiter ignored us for 20 minutes.

Coreference resolution
Obama told Mubarak he shouldn't run again.

Word sense disambiguation (WSD)
I need new batteries for my *mouse*.

Syntactic parsing
I can see Russia from my house!

Machine translation (MT)
Our specialty is panda fried rice.
我们的专长是熊猫炒饭

still really hard

Semantic search
people protesting globalization — Search
...demonstrators stormed IMF offices...

Question answering (QA)
Q. What currency is used in China?
A. The yuan

Textual inference & paraphrase
T. Thirteen soldiers lost their lives ...
H. Several troops were killed in the ... — YES

Summarization
Sheen continues rant against ...
Sheen is nuts

Discourse & dialog
Where is Thor playing in SF?
Metreon at 4:30 and 7:30

Adapted from: Bill MacCartney's Sym Sys 100 lecture at Stanford

A few things we didn't cover

Sentiment analysis

Is this movie awesome or awful? Is this restaurant a good place for dinner?

...zany characters and richly applied satire, and some great plot twists

It was pathetic. The worst part about it was the boxing scenes...

...awesome caramel sauce and sweet toasty almonds. I love this place!

...awful pizza and ridiculously overpriced...

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis

Is this movie awesome or awful? Is this restaurant a good place for dinner?

 *...zany characters and richly applied satire, and some great plot twists*

 *It was pathetic. The worst part about it was the boxing scenes...*

 *...awesome caramel sauce and sweet toasty almonds. I love this place!*

 *...awful pizza and ridiculously overpriced...*

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis

Is this movie awesome or awful? Is this restaurant a good place for dinner?

...zany characters and richly applied satire, and some great plot twists

It was pathetic. The worst part about it was the boxing scenes...

...awesome caramel sauce and sweet toasty almonds. I love this place!

...awful pizza and ridiculously overpriced...

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis

Sentiment analysis is one common **text categorization** task. Other examples include:

- spam detection (spam: YES or NO)
- authorship identification (Jefferson, Hamilton, etc.)
- language Identification (English, German, Swedish)
- assigning subject categories, topics, or genres (History, Geography, ...)

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis: Classification

More formally, any classification task (**classifier**) requires:

Input:

- a document d
- a fixed set of classes $C = \{c_1, c_2, c_3, \ldots, c_J\}$
- a training set of m hand-labeled documents $(d_1,c_1),\ldots,(d_m,c_m)$

Supervised machine learning

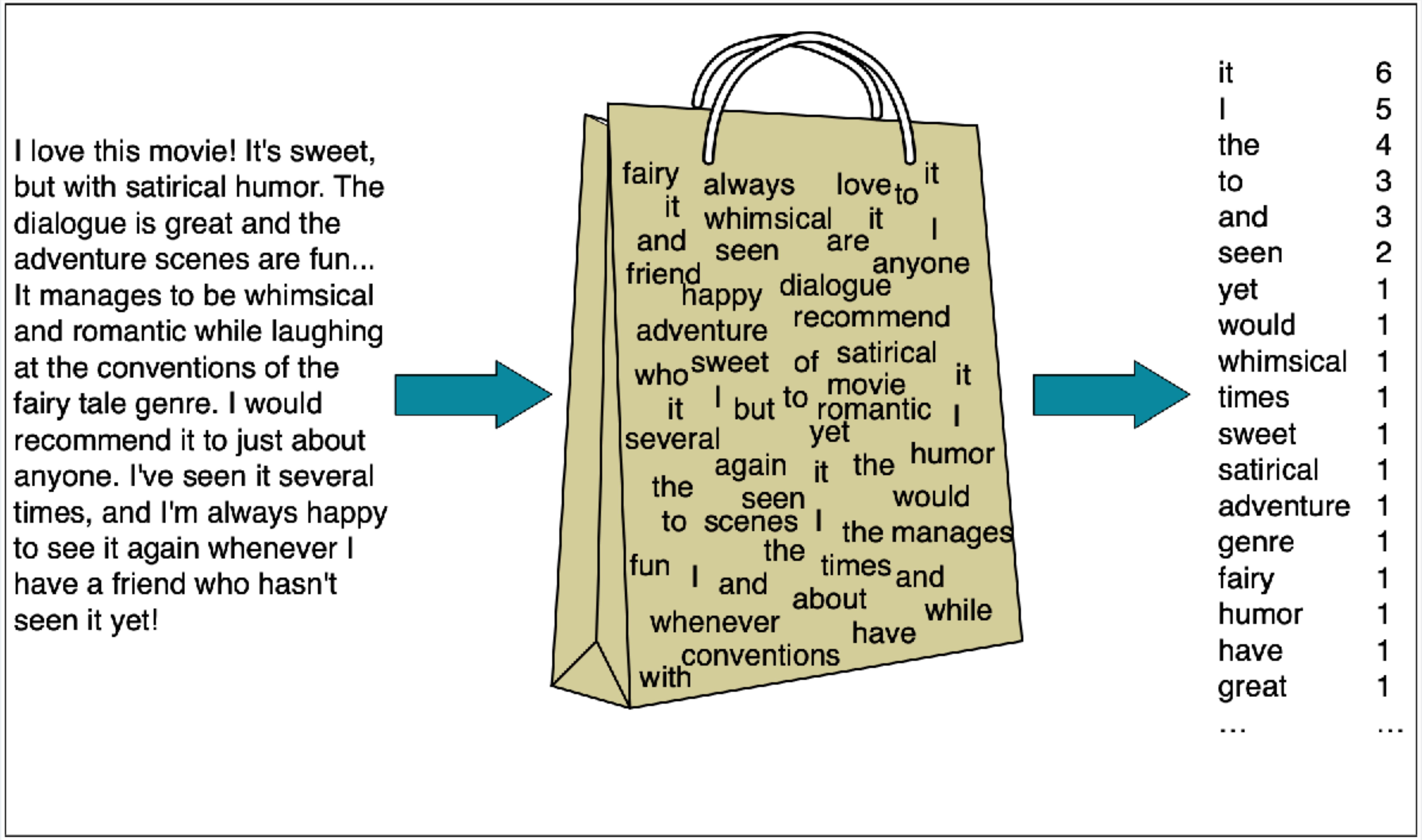
Output:

- a learned classifier γ:d →c

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis: Naive Bayes Classifier

Intuition: A text is represented as a **bag of words** (unordered set of words)



Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis: Naive Bayes Classifier

Let's take the following miniature training and test documents:

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis: Naive Bayes Classifier

Let's take the following miniature training and test documents:

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

Prior: $$p(-) = \frac{3}{5}; p(+) = \frac{2}{5}$$

Adapted from: Jurafsky & Martin [chapter 4 PPT]

Sentiment analysis: Naive Bayes Classifier

Let's take the following miniature training and test documents:

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

Prior: $p(-) = \frac{3}{5}; p(+) = \frac{2}{5}$

Likelihood:

$$p(''predictable''|-) = \frac{1+1}{14+20} = \frac{2}{34}$$

$$p(''no''|-) = \frac{1+1}{14+20} = \frac{2}{34}$$

$$p(''fun''|-) = \frac{0+1}{14+20} = \frac{1}{34}$$

Sentiment analysis: Naive Bayes Classifier

Let's take the following miniature training and test documents:

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

Prior: $p(-) = \frac{3}{5}; p(+) = \frac{2}{5}$

Likelihood:

Posterior:

$$p(-)p(S|-) = \frac{3}{5} \times \frac{2 \times 2 \times 1}{34^3} ✔$$

$$p(+)p(S|+) = \frac{2}{5} \times \frac{1 \times 1 \times 2}{29^3}$$

Conversational Agents

Task-oriented dialogue agents designed to help users solve tasks ✓

- Setting a timer
- Making a travel reservation
- Playing a song
- Buying a product

chatbots designed for extended conversations ✗

- mimicking unstructured conversations or chats

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Conversation

Conversation between humans is an intricate and complex joint activity

C_1: . . . I need to travel in May.
A_2: And, what day in May did you want to travel?
C_3: OK uh I need to be there for a meeting that's from the 12th to the 15th.
A_4: And you're flying into what city?
C_5: Seattle.
A_6: And what time would you like to leave Pittsburgh?
C_7: Uh hmm I don't think there's many options for non-stop.
A_8: Right. There's three non-stops today.
C_9: What are they?
A_{10}: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C_{11}: OK I'll take the 5ish flight on the night before on the 11th.
A_{12}: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C_{13}: OK.
A_{14}: And you said returning on May 15th?
C_{15}: Uh, yeah, at the end of the day.
A_{16}: OK. There's #two non-stops . . . #
C_{17}: #Act. . . actually #, what day of the week is the 15th?
A_{18}: It's a Friday.
C_{19}: Uh hmm. I would consider staying there an extra day til Sunday.
A_{20}: OK. . . OK. On Sunday I have . . .

Turn 10: Multiple sentences

Turn 13: One word

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Conversation

Conversation between humans is an intricate and complex joint activity

C_1: ...I need to travel in May.
A_2: And, what day in May did you want to travel?
C_3: OK uh I need to be there for a meeting that's from the 12th to the 15th.
A_4: And you're flying into what city?
C_5: Seattle.
A_6: And what time would you like to leave Pittsburgh?
C_7: Uh hmm I don't think there's many options for non-stop.
A_8: Right. There's three non-stops today.
C_9: What are they?
A_{10}: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C_{11}: OK I'll take the 5ish flight on the night before on the 11th.
A_{12}: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C_{13}: OK.
A_{14}: And you said returning on May 15th?
C_{15}: Uh, yeah, at the end of the day.
A_{16}: OK. There's #two non-stops ...#
C_{17}: #Act...actually #, what day of the week is the 15th?
A_{18}: It's a Friday.
C_{19}: Uh hmm. I would consider staying there an extra day til Sunday.
A_{20}: OK...OK. On Sunday I have ...

Turn-taking issues

- When to take the floor?
- When to yield the floor?

Interruptions

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Conversation

Participants in conversation need to establish **common ground**

- Speech is action! Speakers need to ground each other's utterances
- **Grounding**: acknowledging that the hearer has understood

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Conversation

Grounding: acknowledging that the hearer has understood

C_1: ...I need to travel in May.
A_2: And, what day in May did you want to travel?
C_3: OK uh I need to be there for a meeting that's from the 12th to the 15th.
A_4: And you're flying into what city?
C_5: Seattle.
A_6: And what time would you like to leave Pittsburgh?
C_7: Uh hmm I don't think there's many options for non-stop.
A_8: Right. There's three non-stops today.
C_9: What are they?
A_{10}: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C_{11}: OK I'll take the 5ish flight on the night before on the 11th.
A_{12}: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C_{13}: OK.
A_{14}: And you said returning on May 15th?
C_{15}: Uh, yeah, at the end of the day.
A_{16}: OK. There's #two non-stops ...#
C_{17}: #Act...actually #, what day of the week is the 15th?
A_{18}: It's a Friday.
C_{19}: Uh hmm. I would consider staying there an extra day til Sunday.
A_{20}: OK...OK. On Sunday I have ...

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Conversation

Grounding is very important for computers too! Consider the following hypothetical conversation:

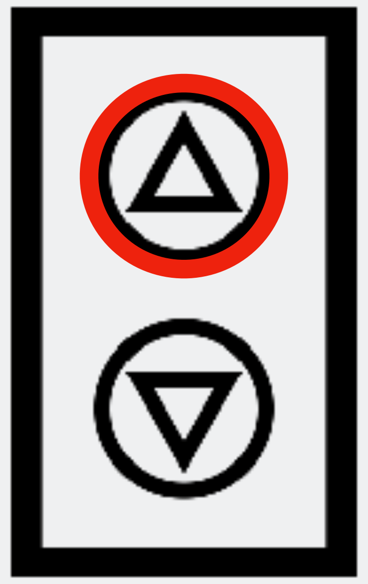
System: Did you want to review some more of your profile?

User: No.

System: What's next?

Awkward!

A button light in an elevator!



Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Frames

How can we make human-computer conservations natural sounding? To do that, we need to talk about **frame-based architecture**

Frame: a knowledge structure consisting of a collection of slots

- Each slot can be filled with information of a given type
- Each slot is associated with a question to the user

Slot	Type	Question Template
ORIGIN CITY	city	"From what city are you leaving?"
DESTINATION CITY	city	"Where are you going?"
DEPARTURE TIME	time	"When would you like to leave?"
DEPARTURE DATE	date	"What day would you like to leave?"
ARRIVAL TIME	time	"When do you want to arrive?"
ARRIVAL DATE	date	"What day would you like to arrive?"

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Frames

- System asks questions of user, filling any slots that user specifies
- User might fill many slots at a time:
    - *I want a flight from San Francisco to Denver one way leaving after five p.m. on Tuesday*
- When frame is filled, do database query

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: Frames

System can have multiple frames like:

- car or hotel reservations
- general route information
  - Which airlines fly from Boston to San Francisco?
- information about airfare practices
  - Do I have to stay a specific number of days to get a decent airfare?)

Frame detection:

- System must detect which slot of which frame user is filling...
- and switch dialogue control to that frame

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: NLU filling slots

1. Domain classification
    - Asking weather? Booking a flight? Programming alarm clock?
2. Intent Determination
    - Find a Movie, Show Flight, Remove Calendar Appt
3. Slot Filling
    - Extract the actual slots and fillers

Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: NLU filling slots

```
Wake me up tomorrow at six.

DOMAIN: ALARM-CLOCK
INTENT: SET-ALARM
TIME: 2021-11-25 0600
```


Adapted from: Jurafsky & Martin [chapter 24 PPT]

Conversational Agents: NLU filling slots

```
Show me morning flights from Boston to SF on Tuesday.

DOMAIN: AIR-TRAVEL
INTENT: SHOW-FLIGHTS
ORIGIN-CITY: Boston
ORIGIN-DATE: Tuesday
ORIGIN-TIME: morning
DEST-CITY: San Francisco
```


Adapted from: Jurafsky & Martin [chapter 24 PPT]

Before we wrap...

Two sides to NLP

Two sides to every coin: intended use and unintended consequences

The same is also true of NLP...

Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

Two sides to NLP

Positive use case: Mental health assessment (Benton et al., 2017)

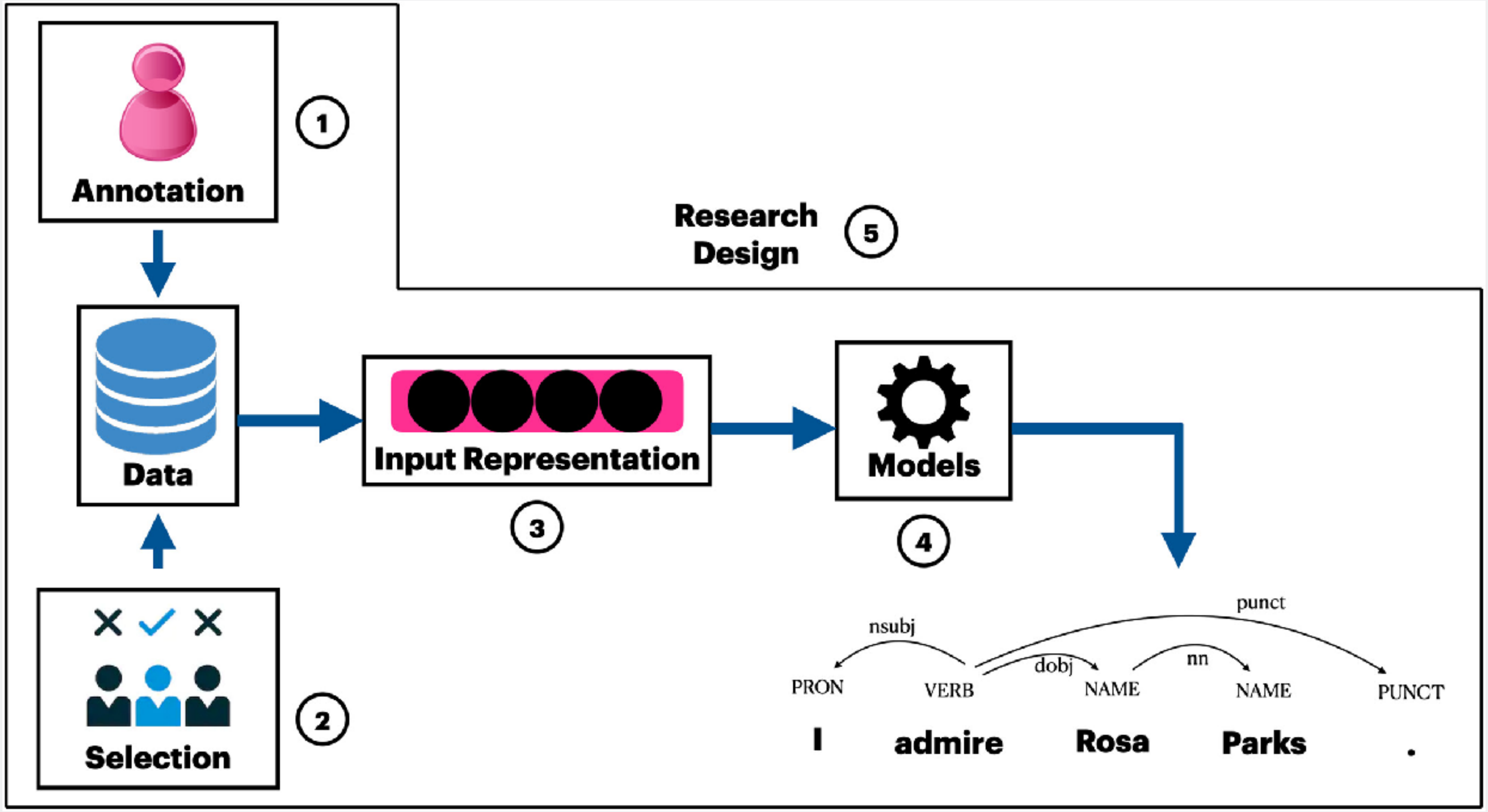
Potential misuse: Profiling users across sites (Wang et al., 2018)

Less overt but still biased: In speech recognition

- strong bias toward native speakers of any language
- Lower level of accuracy/understanding with dialect speakers or children

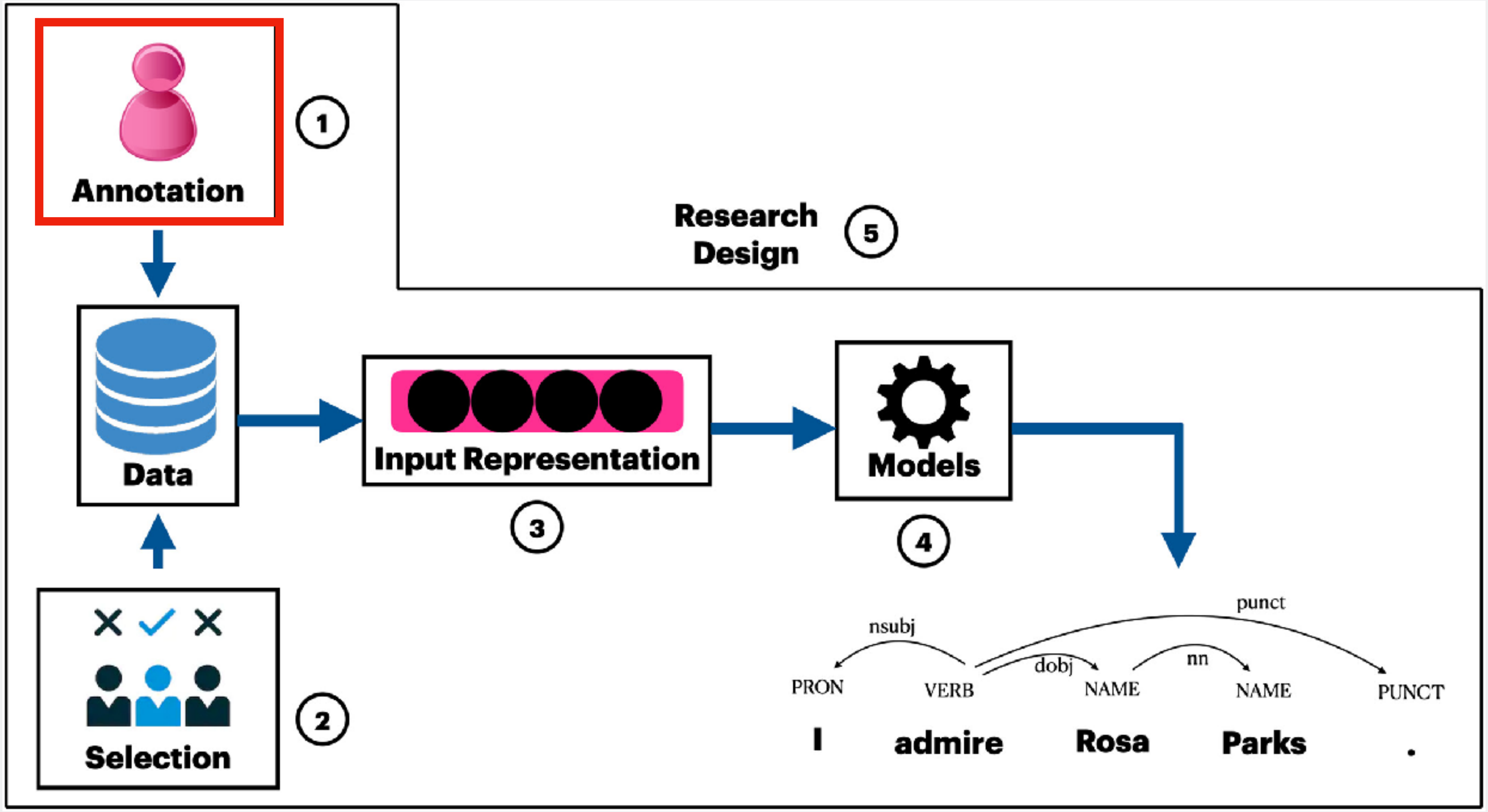
Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

The problem of bias in NLP



Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

The problem of bias in NLP

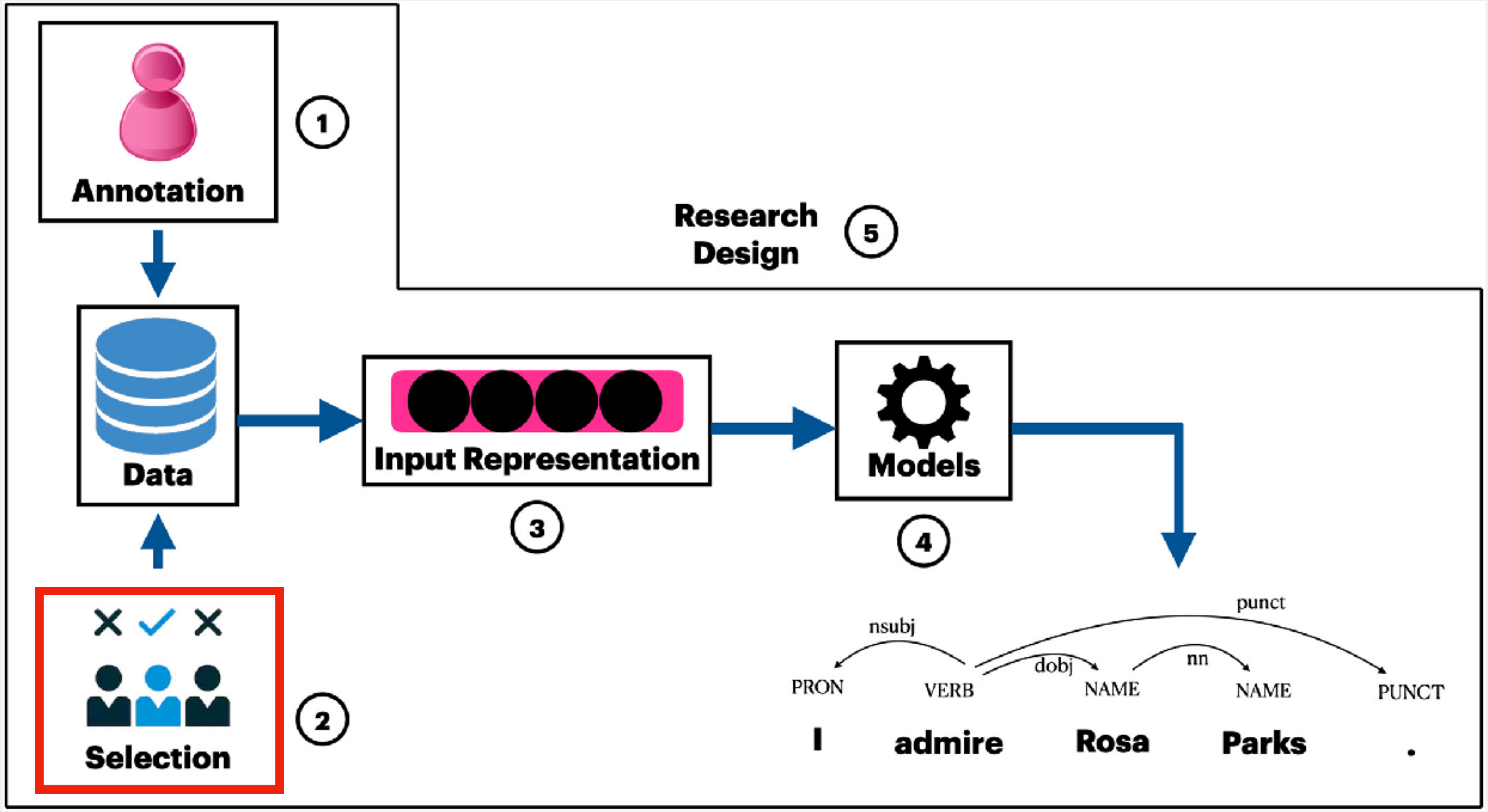


Label bias:

- Assigning "wrong" labels
  1. "social media"
  2. "This is so sick"

Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

The problem of bias in NLP

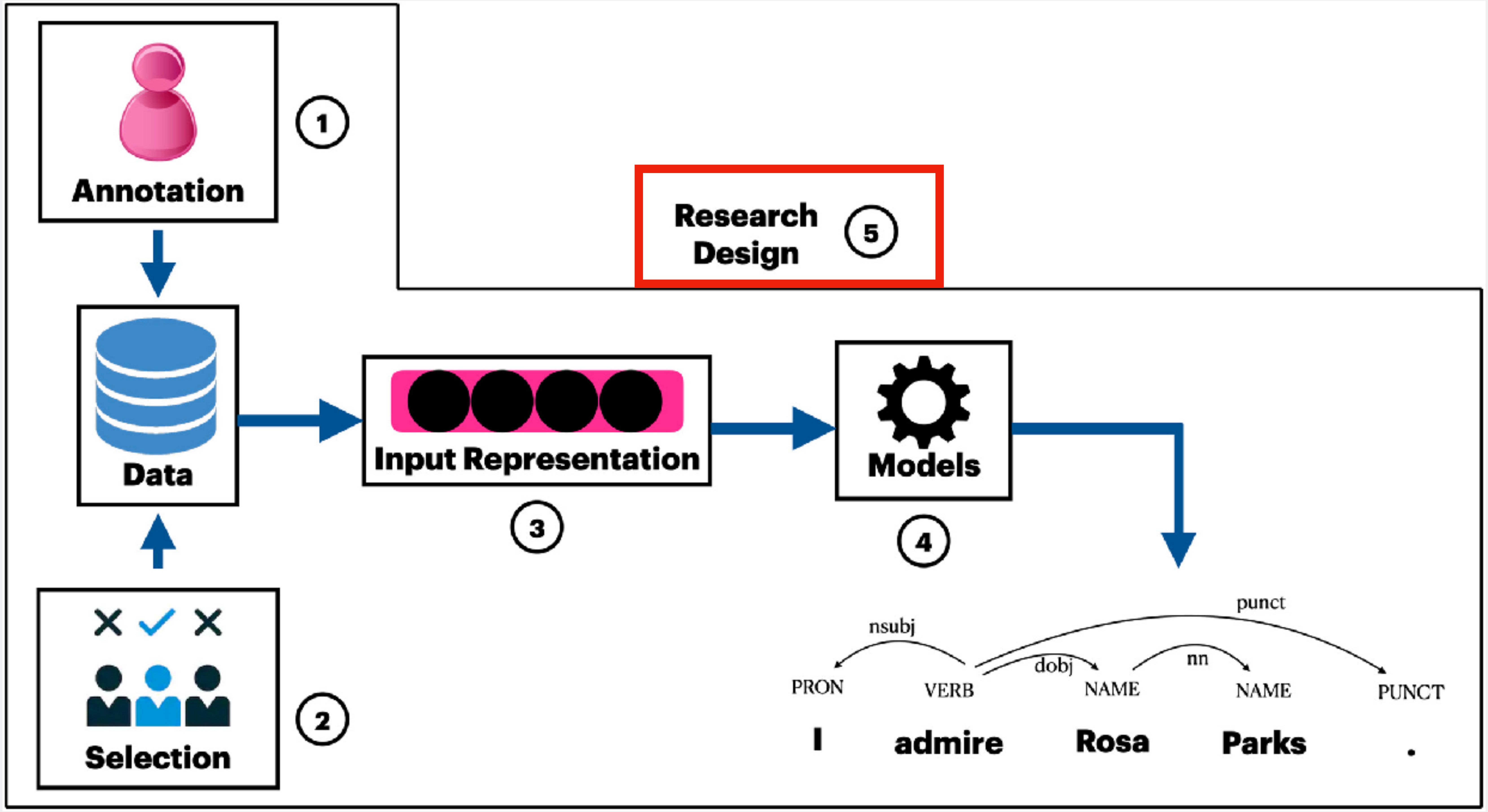


Selection bias:

- Models trained on "old" data
- Performance declined (e.g., POS tagging)
- Ageist, racist, and sexist models

Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

The problem of bias in NLP



Design bias:

- Most research on English (= the "default" language)
- N-gram models do not work as well in other Lx

Source: Hovy & Prabhumoye (2021). *Lang Linguist Compass*

The end

There are a lot more we did not have time to cover. Do not quit. Persevere!

Course evaluation

Off-the-record feedback or comment

And finally, thank you!