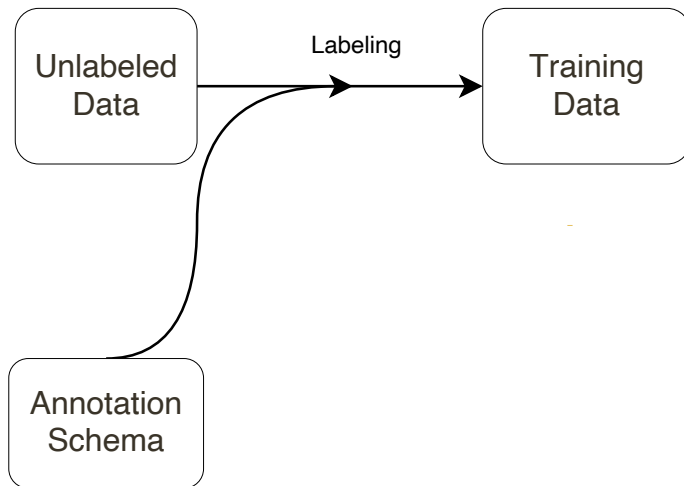
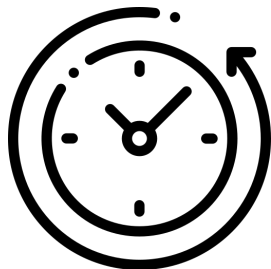


# MEGAnno: Exploratory Labeling for NLP in Computational Notebooks

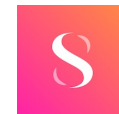
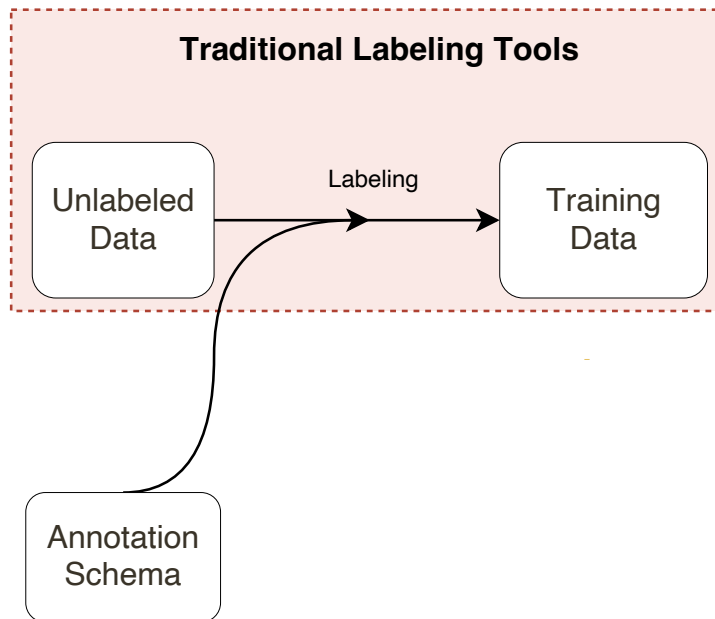
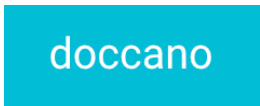
Dan Zhang, Hannah Kim, Rafael Li Chen, Eser Kandogan, Estevam Hruschka



# Data Labeling

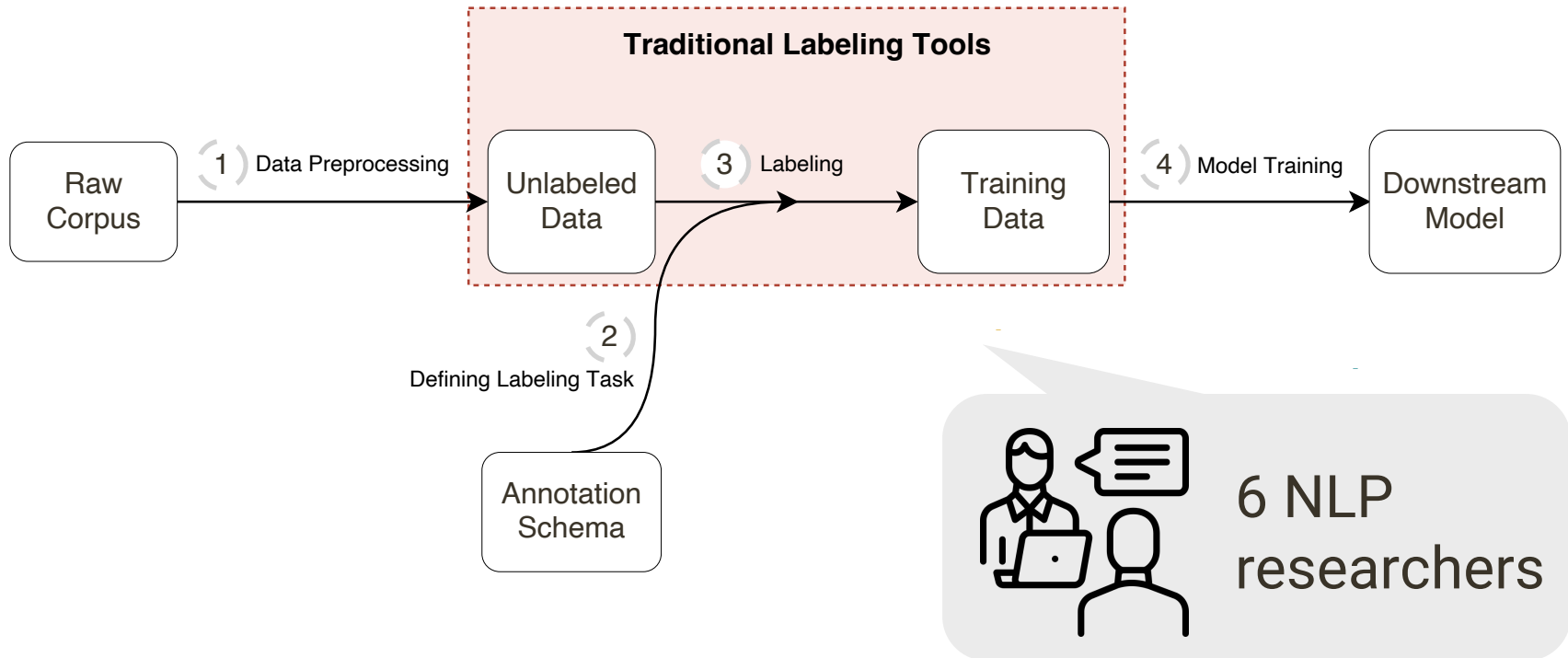


# [Standalone] Data Labeling

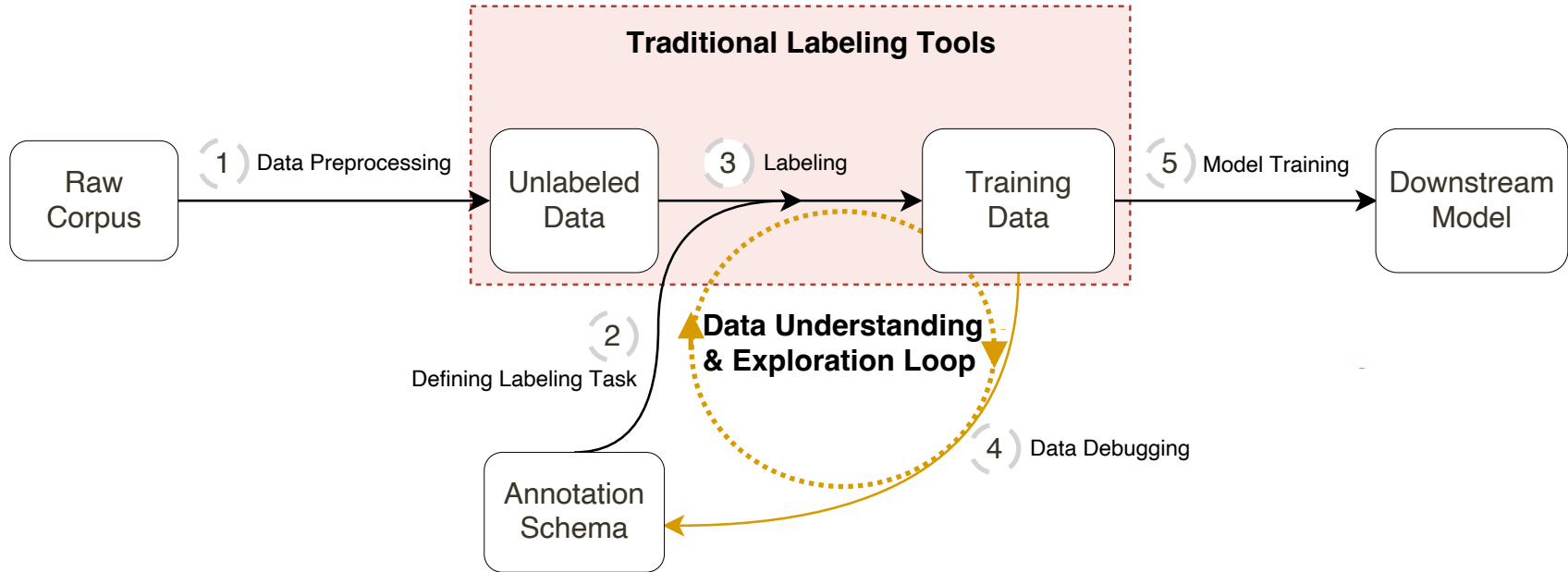


And many more...

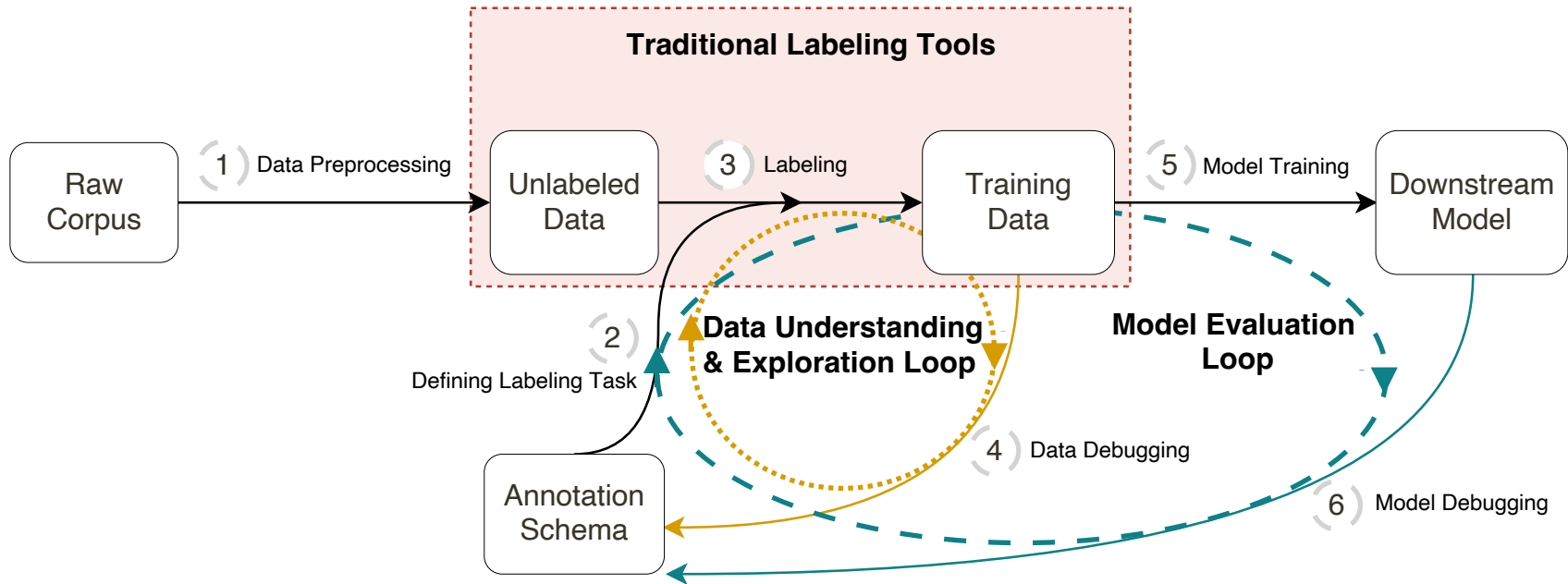
# Data Labeling *within* ML Workflow



# Dual-Loop Model of Data Labeling



# Dual-Loop Model of Data Labeling

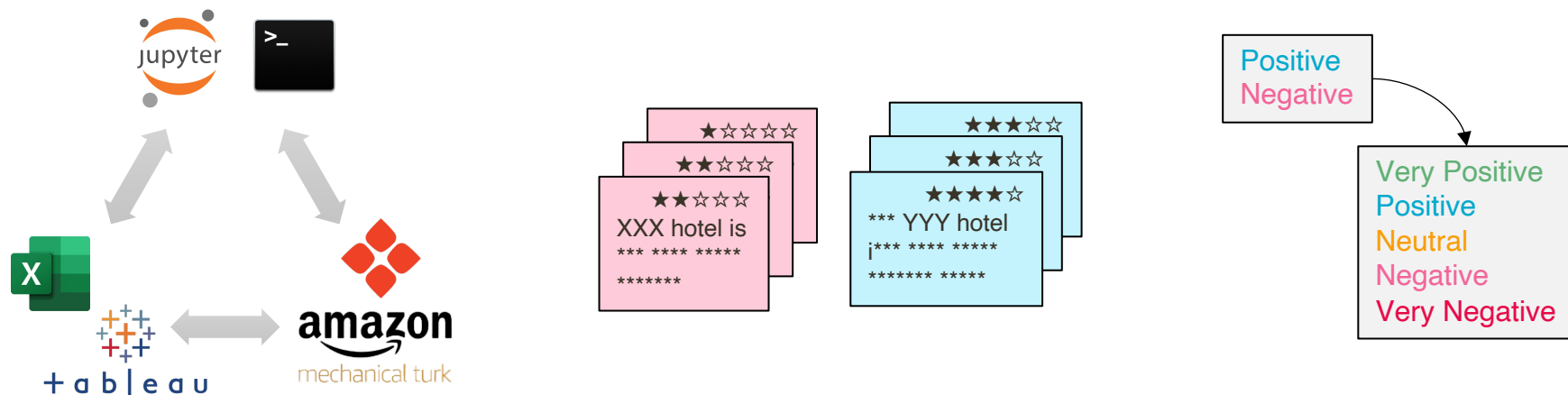


# Challenges from Dual-Loop Model

Gaps between ML tools

Lack of customizable and fine-grained control

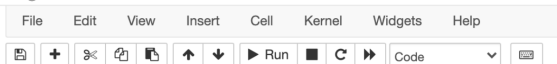
Lack of support for project evolution



# We present MEGAnno, a flexible, exploratory, efficient, and seamless labeling framework for NLP researchers and practitioners.



Jupyter Notebook



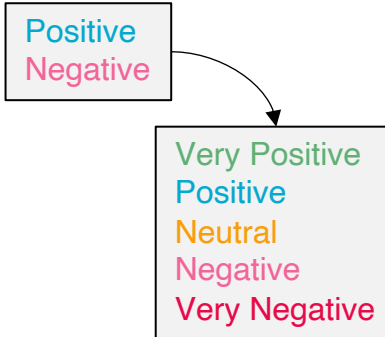
In [1]: Python API

**Widget**

MEGAnno

	data	sentiment_label
0	@JetBlue I've had no shampoo sad, no winter coat sad, no deodorant sad, no flatiron sad, and the rest if a	neg
1	@JetBlue right now completely pissed off sad	neg
2	@SouthwestAir no self help way to put in tsa pre check number for existing reservation. Very annoying sad. 800 is not reachable.	neg
3	@SouthwestAir no self help way to put in tsa pre check number for existing reservation. Very annoying sad. 800 is not reachable.	neg

300 showing



Q /#\*\*fai[^\ ]\*/ regex

Matching on failures, faith, fail, failing!!, failure., #fail, #fail..., fair, fails, fail., faithful, failed, failscustomers, fairs, failure, failures., fairs., #fail., fair."

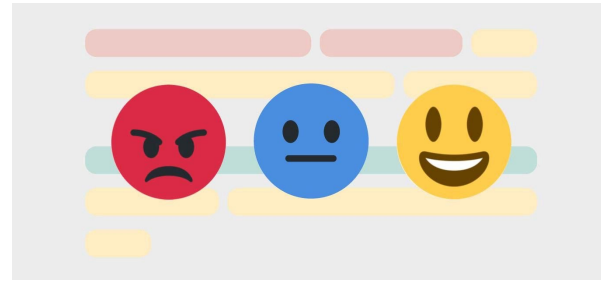
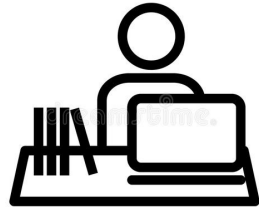
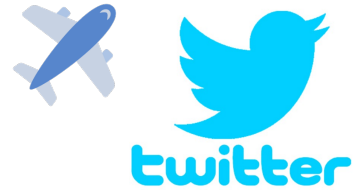
Suggestions

^ ^ey ^jh ^jp ^kn ^kp

Search Mode Tips



# Use Case: Sentiment Analysis



# Import Data and Set Up Schema

```
1 from meganno import Project
2 L = Project(<auth>, "Tweet Sentiment")
3 L.import_data(<doc_url>, format="csv")
```

```
1 label_schema = [{
2     "label_name": "sentiment_label",
3     "level": "record",
4     "options": ["Positive", "Negative"]
5 }]
6 L.set_config(config1)
```

# Explore and Annotate with Our Widget

```
# search results => subset s1  
s1 = demo.search(keyword='', limit=10, start=0)  
# bring up a widget  
s1.show()
```

Switching between  
**Single View** and **Table View**

**Batch** labeling

**Sort and filter** based on text,  
label, and metadata (e.g., #  
of RT, timestamp, author, ...)

# Exploratory Labeling - Search

## API

### By keyword:

```
L.search(keyword='amazing')
```

### By regex:

```
L.search(regex='.*#fai[^\ ]* .*')
```

### By user-defined metadata:

```
L.search(pattern="terrible <ADJ> <NOUN>",  
tagger=spacy.load("en_core_w  
eb_sm"))
```

## Widget

```
# iterate by id  
L.search_by_id(limit=3000, skip = 0).show(widget_config)
```

showing demo's annotations

### Twitter Sentiment

Submit | Filter

		data
1	<input type="checkbox"/>	@USAirways tried twice today on hold for 30 minutes
2	<input type="checkbox"/>	@SouthwestAir thanks, got put on the am flight for
3	<input type="checkbox"/>	@united It is super frustrating that the folks at the United Ticket Counter in Pittsburg aren't honoring their ov
4	<input type="checkbox"/>	@JetBlue #fail My wife on the phone asking to switch flight times. In mid switch gets disconnected. Now Cancelled Flight! No new time! No call back
5	<input type="checkbox"/>	
6	<input type="checkbox"/>	@JetBlue #fail My wife on the phone asking to switch flight times. In mid switch gets disconnected. Now Car
7	<input type="checkbox"/>	@JetBlue Guys, really bad @JetBlue #fail . Someone better call my wife back to get this handled. 203-382-3

demo 36 of 2242 showing

Search: /##fai[^\ ]\*/

Matching on failures, faith, fail, failing!, failure., #fail, #fail..., fair, fails, fail., faithful, failed, failscustomers, fairs, failure, failures., fairs., #fail., fair."

Suggestions: ^, ^ey, ^jh, ^jp, ^kn, ^kp

Search Mode Tips

# Exploratory Labeling - Suggestion

```
# Example 1, adding sentence-bert embedding.  
# metadata searches.  
from sentence_transformers import SentenceTransformer  
model = SentenceTransformer('all-MiniLM-L6-v2')  
# set metadata generation function  
sv.set_metadata("bert-embedding", lambda x: list(model.encode(x).astype(float)), 500)
```

Find similar or diverse samples to label

```
s3 = sv.search(keyword='delay', limit=3, start=0)  
s4 = s3.suggest_similar('bert-embedding', limit=4) # needs to provide a valid meta_name  
s4.show()
```

```
show(7fd1aff79a60, self=<labeler_ui.widgets.Annotation.Annotation object at 0x7fd1b1033d00>)
```

### Annotation

 Single Table

Submit Filter Grid Annotating Settings

		data	sentiment
	<input type="checkbox"/>		<span>Bulk label</span>
1	<input type="checkbox"/>	@USAirways UR service is so shitty. Pilot never showed up so we waited hours because another pi	-
2	<input type="checkbox"/>	@SouthwestAir Being old I will miss my connection and can sleep at OAK or get a hotel on my own	-
3	<input type="checkbox"/>	@united my flight was at 1 pm. Still at the airport gee what service. Do they care, nope. Delayed ur	-

# Update Annotation Schema

```
sv.get_schemas().set_schemas({
  'label_schema': [
    {
      "name": "sentiment",
      "level": "record",
      "options": [
        { "value": "pos", "text": "positive" },
        { "value": "neg", "text": "negative" },
        { "value": "neu", "text": "neutral" },
      ]
    }
  ]
})
```

## sentiment

- positive (pos)
- negative (neg)
- neutral (neu)



```
sv.get_schemas().set_schemas({
  'label_schema': [
    {
      "name": "sentiment",
      "level": "record",
      "options": [
        { "value": "pos", "text": "positive" },
        { "value": "neg", "text": "negative" },
        { "value": "neu", "text": "neutral" },
      ]
    },
    # adding a span-level label
    {
      "name": "sp",
      "level": "span_ch",
      "options": [
        { "value": "pos", "text": "positive" },
        { "value": "neg", "text": "negative" },
      ]
    }
  ]
})
```

@united You allow shady 3rd party services to sell your tickets and then refuse to help the customer, who's left holding the bag. #flyunited

# Manage Annotation Status

**Dashboard** Aggregated

[Overview](#) [Annotator](#) [Projection](#)

**Overall Progress**

Annotated  data points (with at least 1 label) out of  total data points

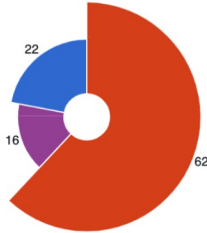
**Class Label - Distributions**

Label class distribution for  subtask: aggregate

When there is a tie in the voting, they are categorized under "t"

**Labels**

- neg
- neu
- pos



Label	Count
neg	62
pos	22
neu	16

Monitor progress, distribution

Reconcile conflicts from multiple annotators

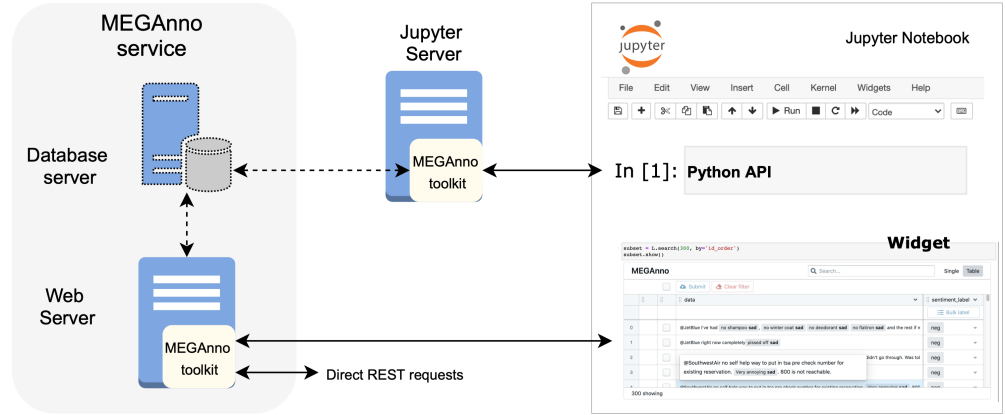
**Annotation**  Single Table

Reconcile Reconciling

		data	sent...	
	<input type="checkbox"/>			<input type="text" value="Bulk label"/>
1	<input type="checkbox"/>	@USAirways Your whole handling of this was a joke.Boston	-	neg: 100.00%
2	<input type="checkbox"/>	@SouthwestAir Why won't you let me leave Newark!?	-	neg: 100.00%
3	<input type="checkbox"/>	@united Hours on the phone...hung up on...all I want is to b	neg	neg: 50.00% pos: 50.00%
4	<input type="checkbox"/>	@united You allow shady 3rd party services to sell your tick	-	neg: 100.00%
5	<input type="checkbox"/>	@JetBlue they couldn't do it sun then it was supposed to be	-	neg: 100.00%
6	<input type="checkbox"/>	@JetBlue now when I call it's still at the warehouse and mic	-	neu: 100.00%

# MEGAnno Features

- Data Model
- Task Schema
- API
- Interactive Widgets
- ...



Please refer to our paper!



```

# Iterate by id
L.search_by_id(limit=1000, skip=0).show(widget_config)

```

showing demo's annotations

Twitter Sentiment

id	data	sent_label
1	@US Airways tried twice today on hold for 30 min...	-
2	@SouthwestAir thanks, got put on the am flight to...	-
3	@united It is super frustrating that the folks at the United Ticket Counter in Pittsburgh aren't honoring their...	-
4	@JetBlue My wife on the phone asking to switch flight times. In mid switch gets disconnected. Now Cancelled Flighted! No new time! No call back	-
5	@JetBlue My wife on the phone asking to switch flight times. In mid switch gets disconnected. Now Car...	-
7	@JetBlue Guys, really bad @JetBlue. Someone better call my wife back to get this handled. 203-382-3...	-

demo: 36 of 2242 showing

Sorting & filtering

sent\_label

sentiment\_label

- Sort asc.
- Sort desc.

Filter by

- Positive pos
- Negative neg
- Neutral neu

# MEGAnno:

## Exploratory Labeling for NLP in Computational Notebooks

```

subset=L.search_by_label(label_name='sentiment_label', filter_value='neg', limit=10)
subset.show(widget_config)

```

showing demo's annotations

Twitter Sentiment

Happy (hap) Sad (sad)

@JetBlue is REALLY getting on my nerves!! 😡😡 #nothappy sad

Label as

- Happy hap

Remove labels

sentiment\_label

- Positive (pos)
- Negative (neg)
- Neutral (neu)

demo: 7 showing 1 selected

Flexible annotation schema to support evolving projects

Efficient, exploratory labeling powered by our API and widget

Seamlessly integrated into existing ML environments

Reconcile Reconciling

id	data	sen...
20	@VirginAmerica now it's just t-minus 32 minutes t...	pos: 100.00%
21	@SouthwestAir Black History Commercial is really	neg: 50.00% pos: 50.00%
22	@US Airways sits on a throne of lies	neg: 100.00%

Demo at our booth (#7) & [meganno.github.io](https://meganno.github.io)



Megagon Labs