

Facilitating Conversational Interaction in Natural Language Interfaces for Visualization

Rishab Mitra

rmitra34@gatech.edu



Arpit Narechania

anarechania3@gatech.edu



Alex Endert

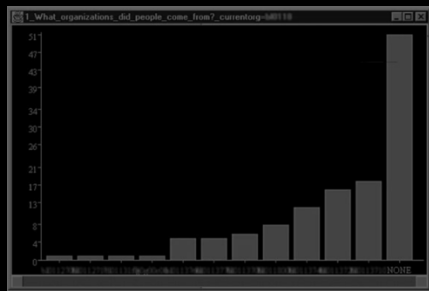
endert@gatech.edu



John Stasko

stasko@cc.gatech.edu





InfoStill 2001



Articulate 2011



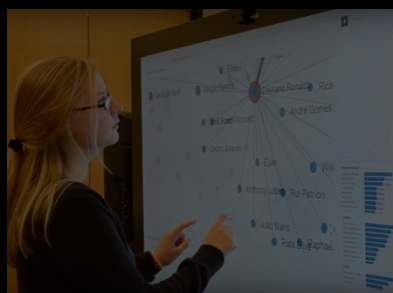
DataTone 2015



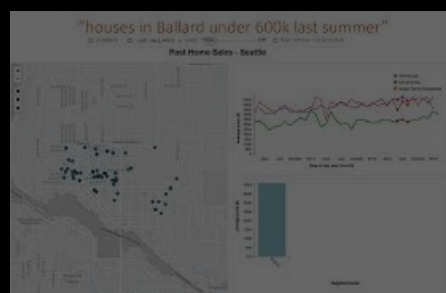
Eviza 2016



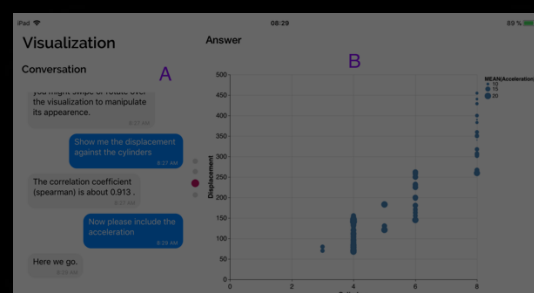
Articulate2 2016



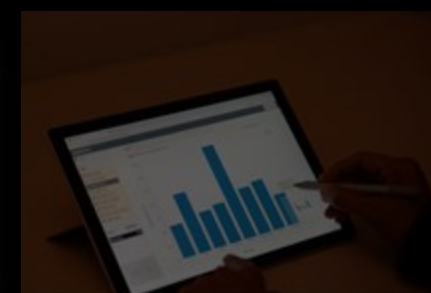
Orko 2018



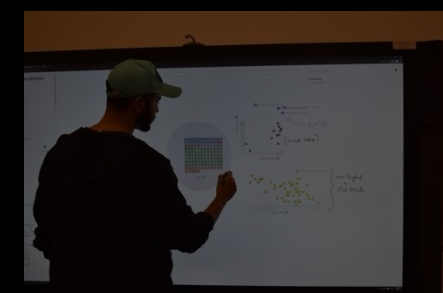
Evizeon 2018



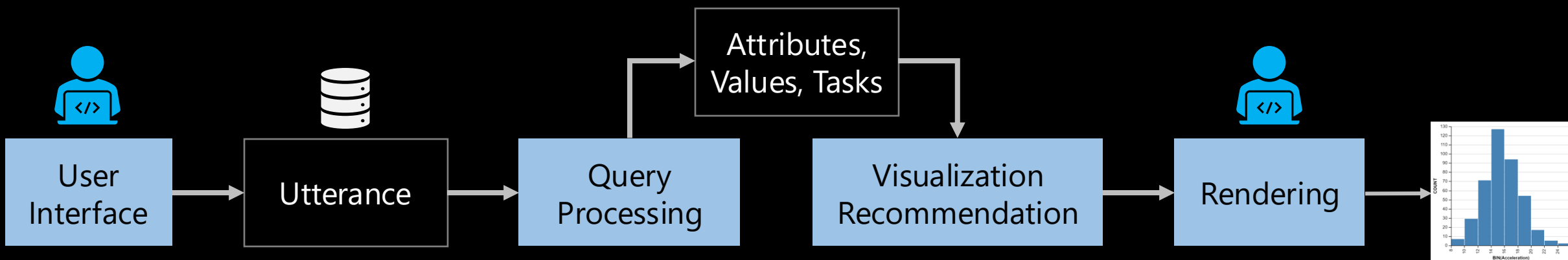
Valletto 2018

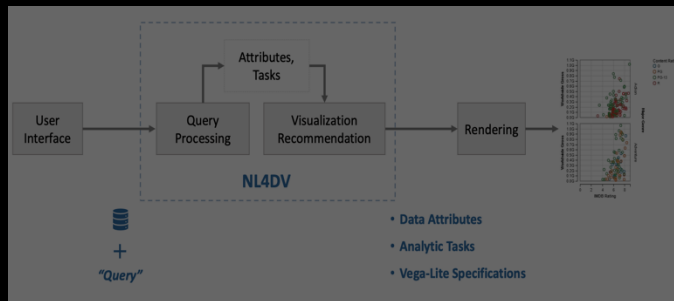


InChorus 2020

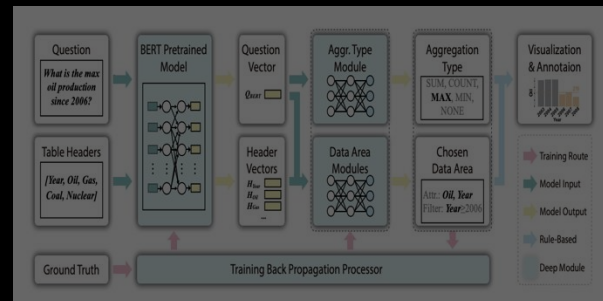


DataBreeze 2020

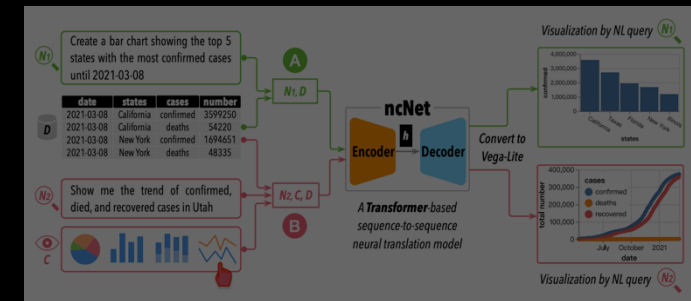




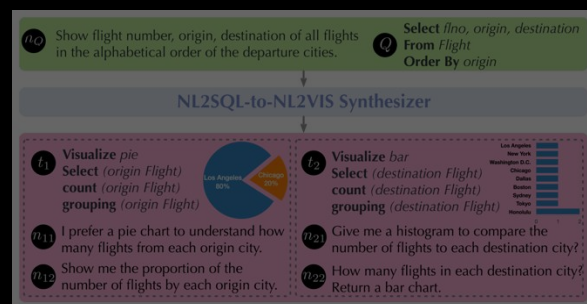
NL4DV 2020



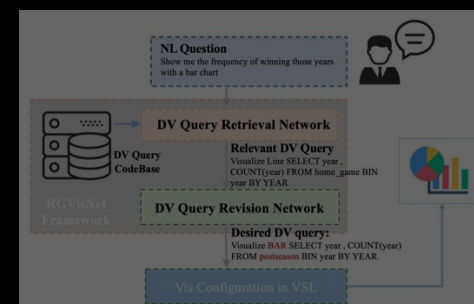
Advisor 2021



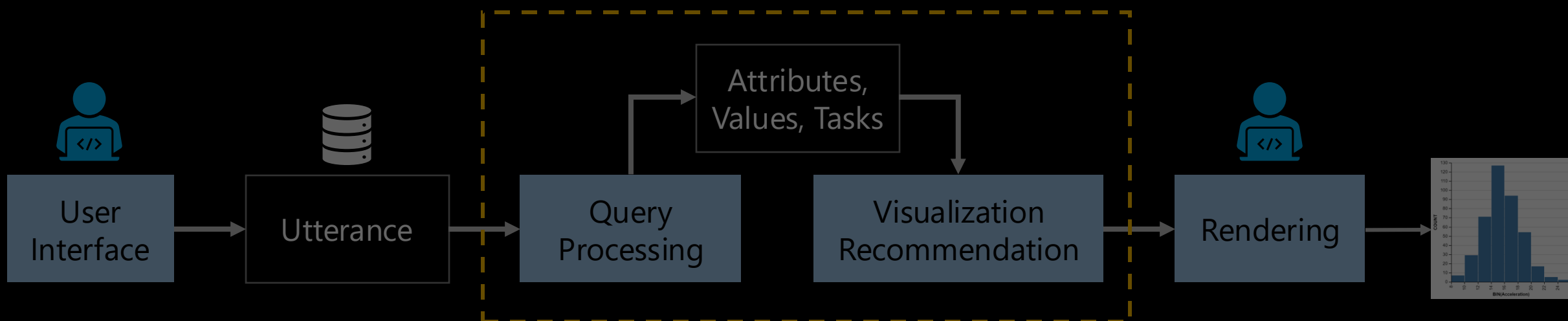
ncNet 2021



SEQ2VIS 2021



RGVisNet 2022





IMDB
Movies

Title	Worldwide Gross	Production Budget	Release Year	Content Rating	Running Time	Genre	Creative Type	Rotten Tomatoes Rating	IMDB Rating
Titanic	1.84G	200M	1997	PG-13	194	Thriller	Historical Fiction	82	7.4
Pirates of the Caribbean: Dead Man's Chest	1.07G	225M	2006	PG-13	151	Adventure	Fantasy	54	7.3
The Dark Knight	1.02G	185M	2008	PG-13	152	Action	Super Hero	93	8.9
Pirates of the Caribbean: At World's End	961M	300M	2007	PG-13	167	Adventure	Fantasy	45	7



"Show the relationship between budget and IMDB rating for action and adventure movies that grossed over 100M"

NL4DV
(VIS'20)

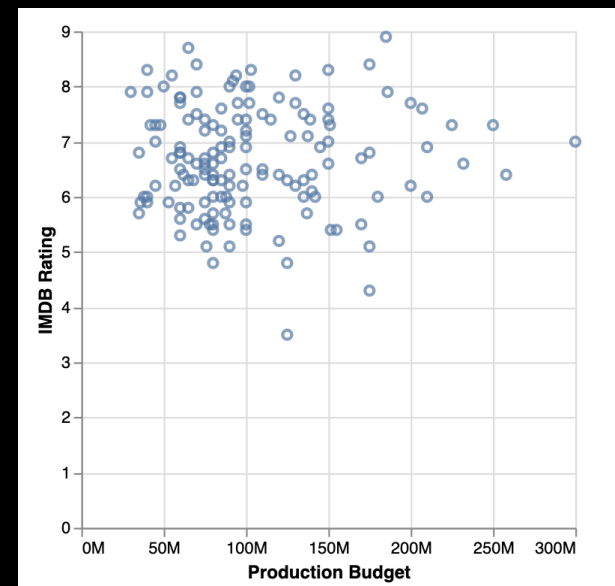


Attributes

```
{
  "Production Budget": {...},
  "IMDB Rating": {...},
  "Genre": {...},
  "Worldwide Gross": {...},
}
```

Tasks

```
{
  "Correlation": ["Production Budget",
                  "IMDB Rating"],
  "Filter": ["Genre" IN ["Action", "Adventure"],
             "Worldwide Gross" > 100000000]
}
```





IMDB
Movies

Title	Worldwide Gross	Production Budget	Release Year	Content Rating	Running Time	Genre	Creative Type	Rotten Tomatoes Rating	IMDB Rating
Titanic	1.84G	200M	1997	PG-13	194	Thriller	Historical Fiction	82	7.4
Pirates of the Caribbean: Dead Man's Chest	1.07G	225M	2006	PG-13	151	Adventure	Fantasy	54	7.3
The Dark Knight	1.02G	185M	2008	PG-13	152	Action	Super Hero	93	8.9
Pirates of the Caribbean: At World's End	961M	300M	2007	PG-13	167	Adventure	Fantasy	45	7.1

Eviza: A Natural Language Interface for Visual Analysis

Vidya Setlur, Sarah E. Battersby, Melanie Tory, Rich Gossweiler, Angel X. Chang in ACM UIST'16

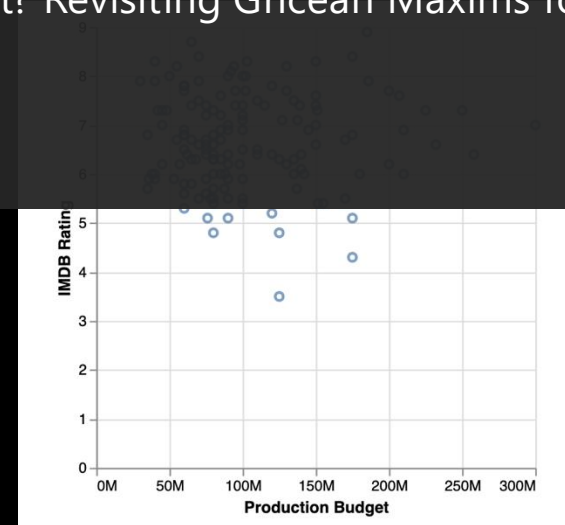
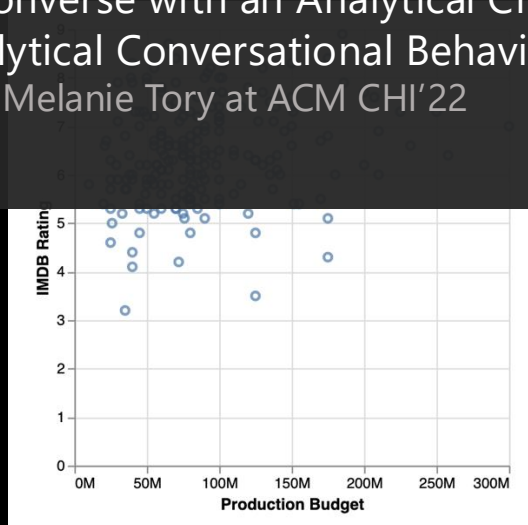
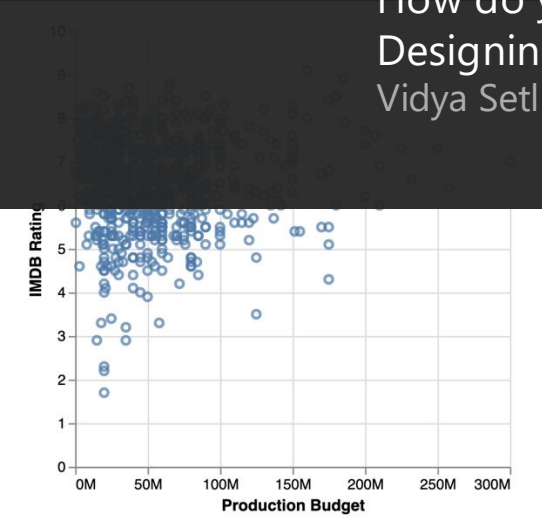
Applying Pragmatics Principles for Interaction with Visual Analytics

Enamul Hoque, Vidya Setlur, Melanie Tory, Isaac Dykeman in IEEE TVCG'17

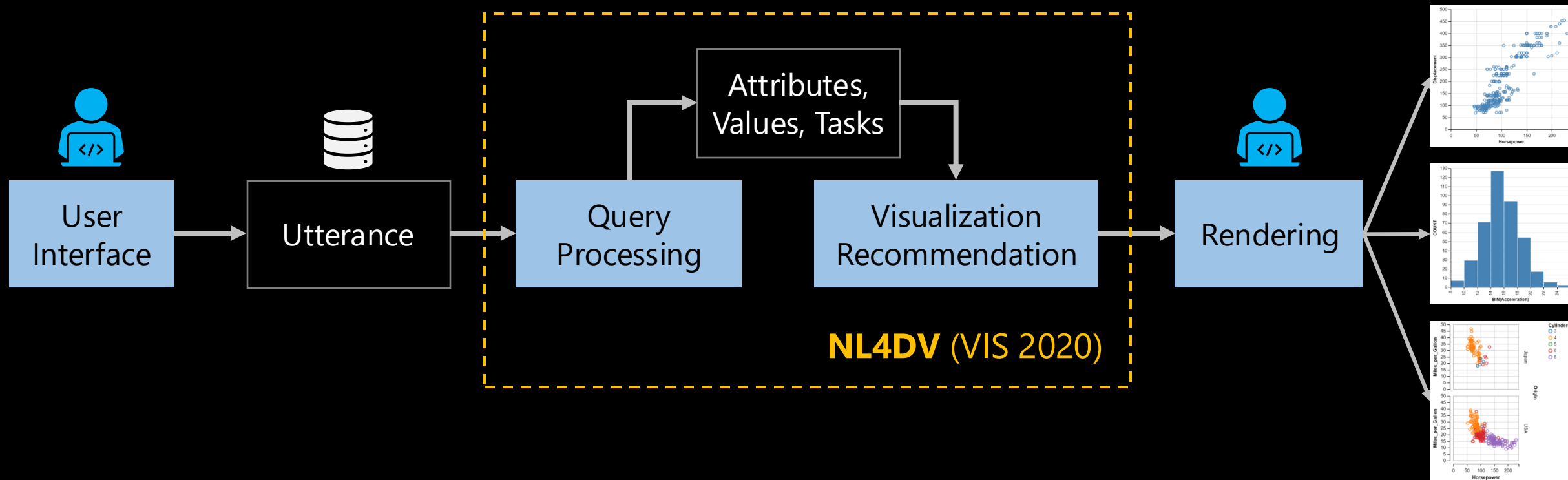
How do you Converse with an Analytical Chatbot? Revisiting Gricean Maxims for Designing Analytical Conversational Behavior

Vidya Setlur and Melanie Tory at ACM CHI'22

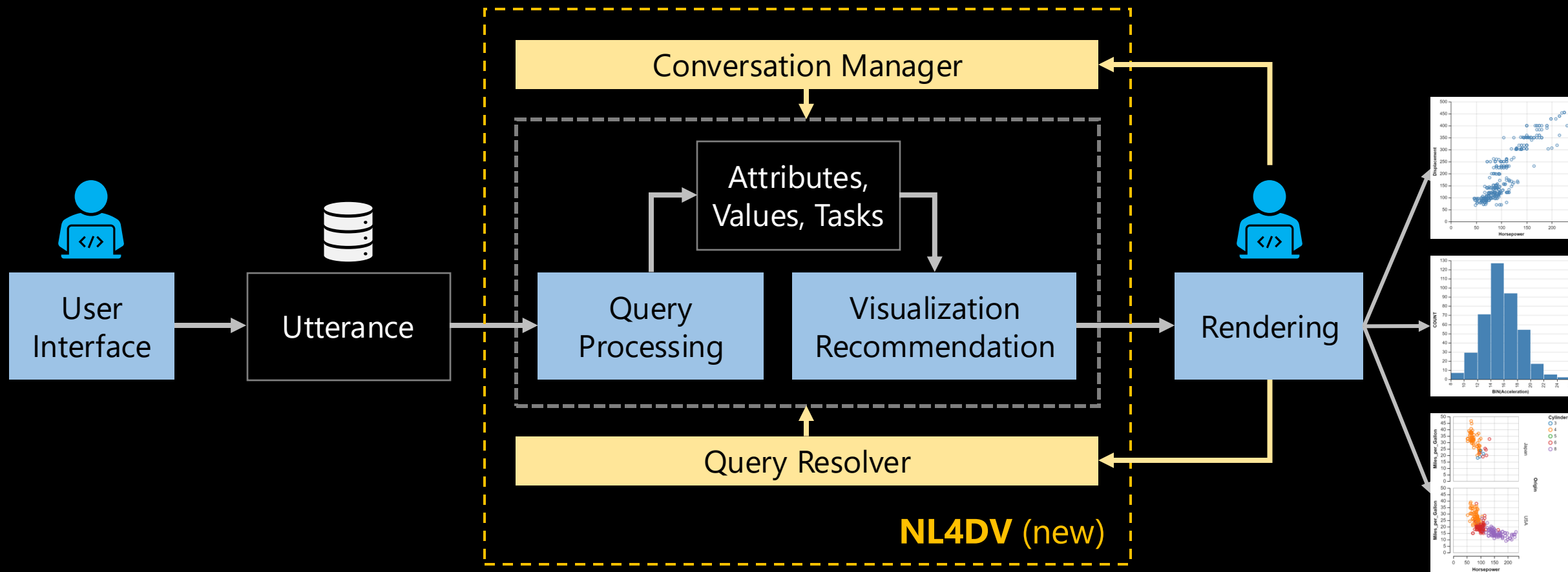
Conversational Interaction

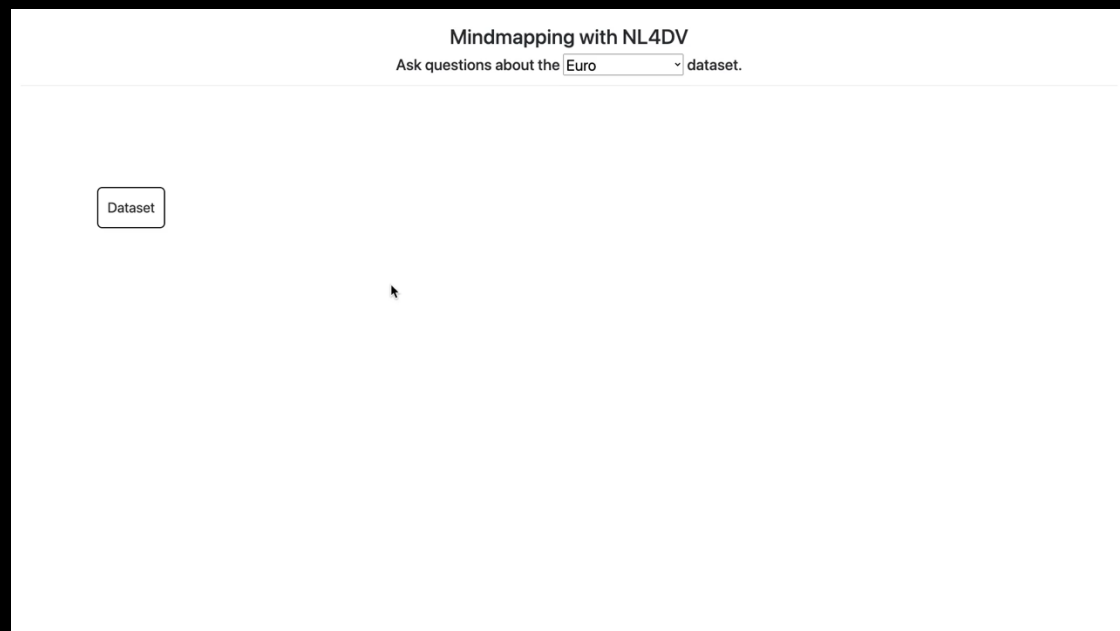


NL4DV: A Toolkit for Generating Analytic Specifications for Data Visualization from Natural Language Queries

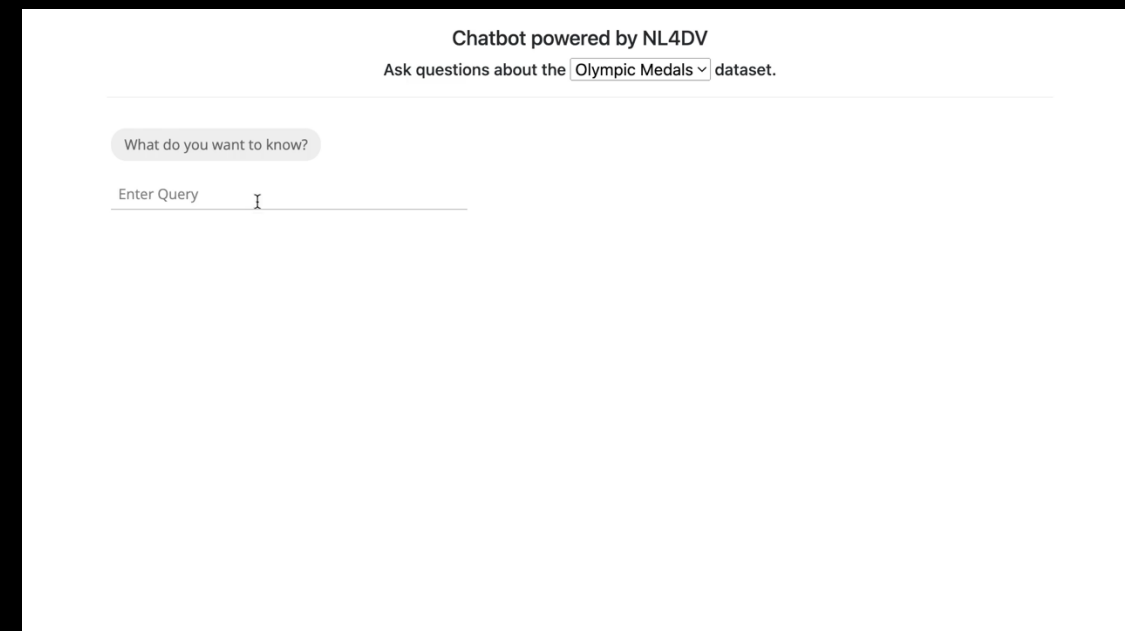


Facilitating Conversational Interaction in Natural Language Interfaces for Visualization





Example 1: **Mind Map**



Example 2: **Chatbot**

Example 1:

Mind Map

Mindmapping with NL4DV

Ask questions about the dataset.

Dataset



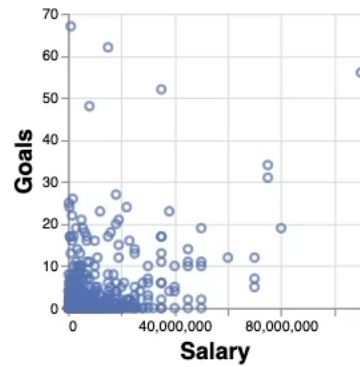
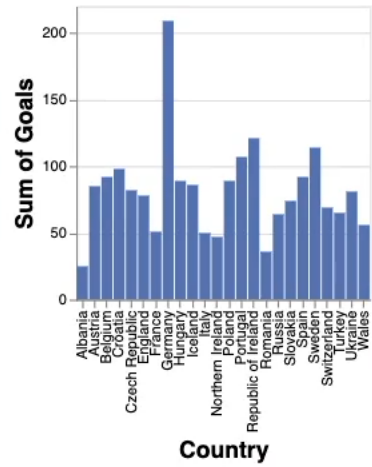
Mindmapping with NL4DV

Ask questions about the dataset.

Dataset

correlate goals and salary

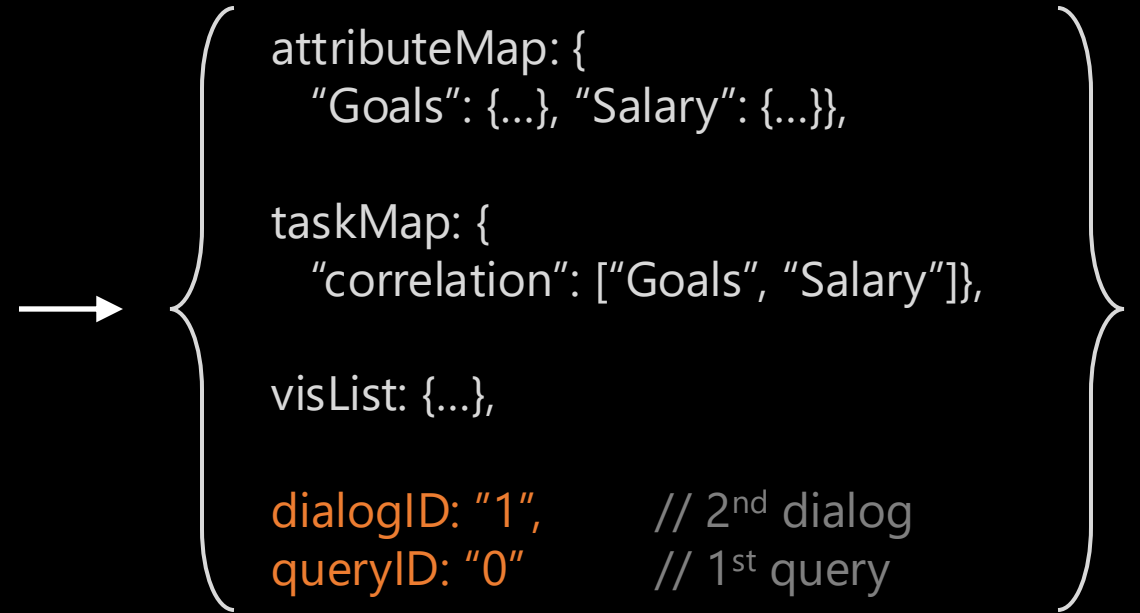
show total goals by country





query = "Correlate goals and salary"

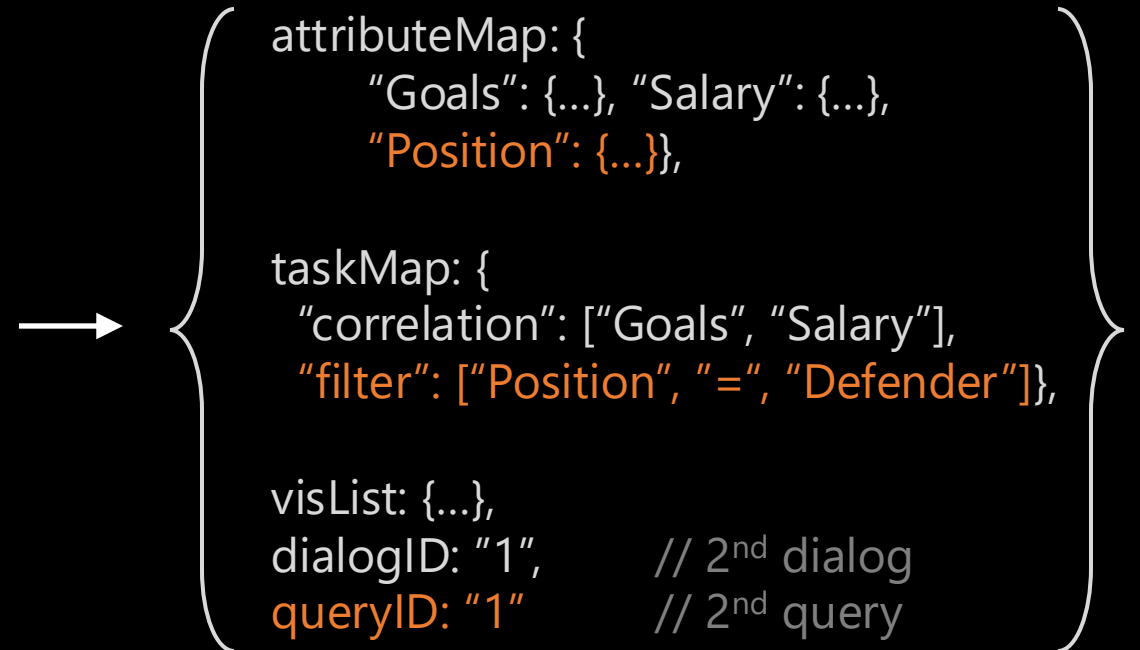
NL4DV.analyze_query(query)



query = "Now show only defenders"

NL4DV.analyze_query(query, dialog=True)

NL4DV.analyze_query(query, dialog=True,
 dialog_id="1", query_id="0")



Example 2:

Chatbot

Chatbot powered by NL4DV

Ask questions about the Olympic Medals ▾ dataset.

What do you want to know?

Enter Query



Chatbot powered by NL4DV

Ask questions about the Olympic Medals ▾ dataset.

What do you want to know?

Show medals for hockey and skating by country

What did you mean by "medal"?

Total Medal



Chatbot powered by NL4DV

Ask questions about the Olympic Medals ▾ dataset.

What do you want to know?

Show medals for hockey and skating by country

What did you mean by "medal"?

Total Medal

What did you mean by "skating"?

Speed Skating

What did you mean by "hockey"?

Hockey

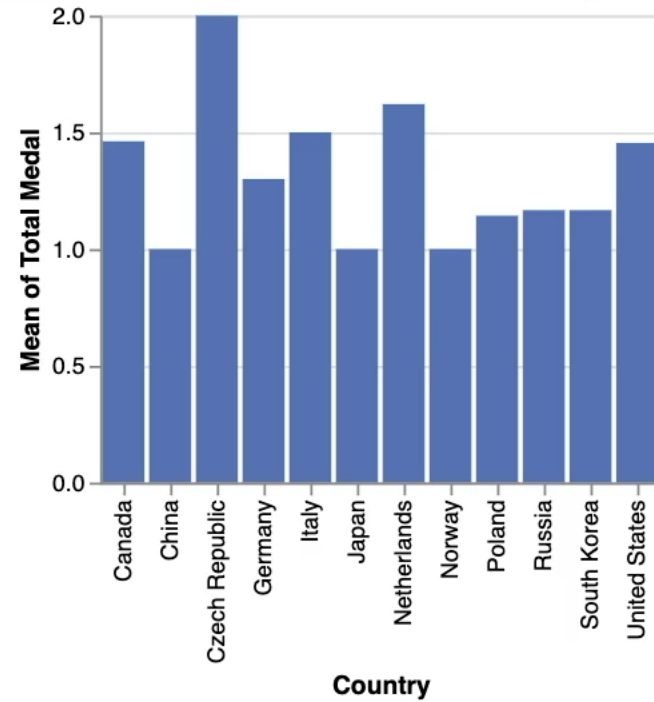


Chatbot powered by NL4DV

Ask questions about the Olympic Medals ▾ dataset.

Hockey

Show medals for hockey and skating by country



What do you want to know?

Enter Query

"Show medals for hockey and skating by country"

```

"ambiguities": {
  "attribute": {
    "medals": {
      "options": ["Gold Medal", ...],
      "selected": None
    },
  },
  "value": {
    "hockey": {
      "options": ["Hockey", ...],
      "selected": None
    },
    "skating": {
      "options": ["Speed Skating", ...],
      "selected": None
    }
  }
}

→ object = {
  "attribute": {
    "medals": "Total Medal",
  },
  "value": {
    "hockey": "Ice Hockey",
    "skating": "Speed Skating"
  }
}

→ NL4DV.update_query(object) →

"ambiguities": {
  "attribute": {
    "medals": {
      "options": ["Gold Medal", ...],
      "selected": "Total Medal"
    },
  },
  "value": {
    "hockey": {
      "options": ["Hockey", ...],
      "selected": "Ice Hockey"
    },
    "skating": {
      "options": ["Speed Skating", ...],
      "selected": "Speed Skating"
    }
  }
}

```

Going Forward

Improving Follow-up Inference and Interpretation

- Detecting and resolving multi turn conversational ambiguities.

"Show only Action movies." → "What about R-rated movies?"

- Automatically predicting compatible "dialog_id" and "query_id".

1: "Show total gross across genres"

2: "Correlate budget and gross"

→ 3. "Now replace genre with content rating"

- Going Beyond Heuristics towards Learning-based approaches.

NLVCorpus (CHI'21) and nvBench (SIGMOD'21) datasets mostly contain one-off queries.

Get started with
pip install nl4dv

For documentation and examples, check out
nl4dv.github.io

20

Search or jump to... Pull requests Issues Marketplace Explore

nl4dv / nl4dv Unwatch

<> Code ! Issues 🔗 Pull requests 🎬 Actions 📁 Projects 📖 Wiki ⚠ Security 📈 Insights

🔗 master 2 branches 2 tags Go to file Add file Code

Arpit Narechania bump up version 65a8ccf 12 days ago 56 commits

docs	add jupyter-notebook teaser	13 days ago
examples	remove readme for examples (it has been moved to the w...	12 days ago
nl4dv	also compute total execution time	2 months ago
.gitignore	move sample.py to ignore	4 months ago
LICENSE.txt	initial commit	4 months ago
MANIFEST.in	initial commit	4 months ago
README.md	update readme with link to the publication	12 days ago
README.txt	initial commit	4 months ago
requirements.txt	bump up nltk version	4 months ago
setup.py	bump up version	12 days ago

README.md

NL4DV

NL4DV stands for **Natural Language toolkit for Data Visualization**. It takes a **language query** about a given **dataset** as input and outputs a **structured JSON** containing:

- Data attributes,
- Analytic tasks, and

Thanks!

NL4DV.analyze_query(query, **dialog** = **True**)

	Attribute	Task	Visualization
Add	Show production budget as well	Sort in the descending order	As a bar chart
Remove	Remove budget	Remove the sum	-
Replace	Replace budget with gross	Show the max instead of min	Replace line chart with a bar chart

NL4DV.update_query({ **object** })

Attribute-level	Value-level
<i>"medals"</i> = ["Gold Medal", "Silver Medal", "Bronze Medal" , "Total Medal"]	<i>"skating"</i> = ["Speed Skating" , "Figure Skating", "Short Speed Skating"]

```
NL4DV.analyze_query( query, dialog = "auto")
```

- *query = "Correlate goals and salary"*

- *query = "Now show only defenders"*

```
NL4DV.analyze_query(query, dialog="auto")
```



```
{  
  attributeMap: {  
    "Goals": {...}, "Salary": {...},  
    "Position": {...}},  
  taskMap: {  
    "correlation": ["Goals", "Salary"],  
    "filter": ["Position", "=", "Defender"]},  
  visList: {...},  
  dialogID: "1",  
  queryID: "1",  
  followUpConfidence: "high"  
}
```