Chart Generation for Contextualized Responses to NL COVID-19 Queries

Hannah DeBalsi, Bianca Yu

CS 294W Final Presentation
Motivation

Alexa, how many COVID-19 cases are there currently worldwide?

There are 6.8 million confirmed cases of COVID-19 worldwide, as of June 7.

Ok.
Motivation

Alexa, how many COVID-19 cases are there currently worldwide?

There are 6.8 million confirmed cases of COVID-19 worldwide, as of June 7.
Motivation

Alexa, how many COVID-19 cases are there currently worldwide?

There are 6.8 million confirmed cases of COVID-19 worldwide, as of June 7.

Interesting ... looks like cases are increasing.
There are 6.8 million confirmed cases of COVID-19 worldwide, as of June 7.

Answered Question

Enhanced Understanding with Context

Alexa, how many COVID-19 cases are there currently worldwide?

Interesting ... looks like cases are increasing.
Outline

Motivation

Quick Recap
Final Results
Lessons Learned

User Research

Quick Recap
Final Results
Lessons Learned

Chart Generation

Quick Recap
Final Results
Lessons Learned

Future Work
User Research / Quick Recap

**Survey #1**

What do users want to know about COVID-19?
+ What are the most common quantitative questions?
User Research / Quick Recap

Survey #1
What do users want to know about COVID-19?
+ What are the most common quantitative questions?

Survey #2
What kinds of COVID-19 questions are best answered visually?
+ What is the most effective graph for these questions?
User Research / Quick Recap

Survey #1  What do users want to know about COVID-19?
+ What are the most common quantitative questions?

Survey #2  What kinds of COVID-19 questions are best answered visually?
+ What is the most effective graph for these questions?

Preliminary
Semantic Parser
User Research / Final Results

Survey #1

<table>
<thead>
<tr>
<th>Q Category</th>
<th>Count</th>
<th># Quantitative</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>COVID Factsheet</td>
<td>14</td>
<td>3</td>
<td>21%</td>
</tr>
<tr>
<td>Death Count</td>
<td>10</td>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>Case Count</td>
<td>9</td>
<td>8</td>
<td>89%</td>
</tr>
<tr>
<td>Symptoms</td>
<td>6</td>
<td>2</td>
<td>33%</td>
</tr>
<tr>
<td>Treatment</td>
<td>6</td>
<td>1</td>
<td>17%</td>
</tr>
<tr>
<td>Vaccine</td>
<td>5</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Infection</td>
<td>4</td>
<td>3</td>
<td>75%</td>
</tr>
<tr>
<td>Lockdown</td>
<td>4</td>
<td>1</td>
<td>25%</td>
</tr>
<tr>
<td>Testing</td>
<td>4</td>
<td>1</td>
<td>25%</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Personal Safety</td>
<td>3</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Recovery Count</td>
<td>1</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>Unemployment Count</td>
<td>1</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>71</strong></td>
<td><strong>31</strong></td>
<td></td>
</tr>
</tbody>
</table>
User Research / Final Results

Survey #2
User Research / Final Results

“What US city has the highest number of cases?”

New York City, New York has the highest number of confirmed COVID-19 cases.

Total COVID-19 Cases, Top 5 US Counties

Distribution of Total COVID-19 Cases for US Counties

Daily Change of COVID-19 Cases in New York
User Research / Final Results

**Significant agreement:**
the percentage of participants who assigned the response a particular rank is $\geq 50\%$

*Survey #2*
User Research / Final Results

**Survey #2**

*Significant agreement:* the percentage of participants who assigned the response a particular rank is \( \geq 50\% \)

**Qs w/ significant agreement on first choice**

BAR CHART

LINE CHART
User Research / Final Results

Survey #2

**Significant agreement:**
The percentage of participants who assigned the response a particular rank is \( \geq 50\% \)

- Qs w/ significant agreement on first choice
- Qs w/ significant agreement on last choice

**Graphs:**
- BAR CHART
- LINE CHART
- TEXT ONLY
User Research / Final Results

Preliminary
Semantic Parser
Q3

How many new cases have there been reported in Georgia today?
Q3

How many new cases have there been reported in Georgia today?
User Research / Final Results

Q3

How many new cases have there been reported in Georgia today?
Q3

How many new cases have there been reported in Georgia today?
Q3

How many new cases have there been reported in Georgia today?
Q3

How many new cases have there been reported in Georgia today?
Q3

How many new cases have there been reported in Georgia today?
User Research / Final Results

<table>
<thead>
<tr>
<th>LABELS</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrogative phrase</td>
<td>“How many”</td>
<td></td>
</tr>
<tr>
<td>Answer type (quantity)</td>
<td>“cases”</td>
<td></td>
</tr>
<tr>
<td>Answer type (noun)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>“Georgia”</td>
<td></td>
</tr>
<tr>
<td>Filter (quantitative)</td>
<td>“new”</td>
<td></td>
</tr>
<tr>
<td>Filter (date/time)</td>
<td>“today”</td>
<td></td>
</tr>
<tr>
<td>LABELS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interrogative phrase</strong></td>
<td>“How many”</td>
<td>→</td>
</tr>
<tr>
<td><strong>Answer type (quantity)</strong></td>
<td>“cases”</td>
<td>→</td>
</tr>
<tr>
<td><strong>Answer type (noun)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td>“Georgia”</td>
<td>→</td>
</tr>
<tr>
<td><strong>Filter (quantitative)</strong></td>
<td>“new”</td>
<td>→</td>
</tr>
<tr>
<td><strong>Filter (date/time)</strong></td>
<td>“today”</td>
<td>→</td>
</tr>
</tbody>
</table>
A text + graph response is more helpful than just text.

For quantitative data, text + graph response seems to always be more helpful.
A text + graph response is more helpful than just text.

Users are mostly interested in change over time.
A text + graph response is more helpful than just text.

Users are mostly interested in change over time.

The graph content itself is critical.

For quantitative data, text + graph response seems to always be more helpful.
Chart Generation / Quick Recap

- **Goal:** Generate graphical representations to questions about COVID-19

- **Started with:**
  - Generating plots for questions already supported by ThingTalk (ie weather)
  - Generating plots in Python and then displaying an image of the plot to the user

- **Moved to:**
  - Generating plots for questions about coronavirus
  - Generating plots in Javascript using Chart.js
Outline

Motivation

User Research

Chart Generation

Future Work

Quick Recap

Final Results

Lessons Learned


<table>
<thead>
<tr>
<th>Date</th>
<th>New Confirmed Cases</th>
<th>New Deaths</th>
<th>Total Confirmed Cases</th>
<th>Recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thursday, March 19</td>
<td>139</td>
<td>2</td>
<td>1,063</td>
<td>0</td>
</tr>
<tr>
<td>Wednesday, March 18</td>
<td>313</td>
<td>5</td>
<td>924</td>
<td>0</td>
</tr>
<tr>
<td>Tuesday, March 17</td>
<td>128</td>
<td>2</td>
<td>611</td>
<td>0</td>
</tr>
<tr>
<td>Monday, March 16</td>
<td>148</td>
<td>5</td>
<td>483</td>
<td>0</td>
</tr>
<tr>
<td>Sunday, March 15</td>
<td>42</td>
<td>1</td>
<td>335</td>
<td>0</td>
</tr>
<tr>
<td>Saturday, March 14</td>
<td>41</td>
<td>0</td>
<td>293</td>
<td>0</td>
</tr>
<tr>
<td>Friday, March 13</td>
<td>50</td>
<td>1</td>
<td>252</td>
<td>0</td>
</tr>
<tr>
<td>Thursday, March 12</td>
<td>0</td>
<td>0</td>
<td>202</td>
<td>0</td>
</tr>
</tbody>
</table>
Tuesday, March 3, 2020: 0 new confirmed cases, 0 new death in ca. In total, 53 confirmed cases, 0 death, 0 recovered.
Friday, March 6, 2020: 0 new confirmed cases, 0 new death in ct. In total, 0 confirmed cases, 0 death, 0 recovered.

t now => @com.covidtracking.us() => notify;

Tuesday, January 21, 2020: 0 new confirmed cases, 0 new death in US. In total, 1 confirmed cases, 0 death, 0 recovered.
Chart Generation / Final Results

State COVID cases

- **ny**: 50,000 cases
- **ct**: 25,000 cases
- **ca**: 20,000 cases
Outline

Motivation

User Research

Chart Generation

Future Work

Quick Recap

Final Results

Lessons Learned
Chart Generation / Lessons Learned

JavaScript

High level ideas → actual implementation

People
Future Work

Connect the pipeline

More robust mappings

More types of charts