Text Analytics: Finding Insights through words

30 May 2019
DASA-CE: Who are we?

- DASA-CE provides full scope cost and performance analytic capability development and execution across the data life-cycle.

- Employ four vital change enablers to facilitate analytic capability development:
  1. Costing Information Frameworks
  2. Policies, Processes, and Procedures
  3. People Development
  4. Technology Infrastructure

- In-house analytic capability development:
  - Data Scientists, BI Experts, ORSAs
  - Advanced Analytic Methods
  - Analytic Toolsets and Tool Creation
  - Enterprise Cost Modeling
  - SAP ERP Expertise (ECC, BIBW, HANA, S/4 HANA)

DASA-CE has IOC business intelligence and data analytics capabilities.
Text Analytics – The Need
Text Analytics – Industry

- **Social Media Monitoring**
  - Parse and monitor mentions of company online

- **Brand Monitoring**
  - In addition to social media parse through mentions through news etc

- **Customer Service/Feedback**
  - Find key complaints/praise for your products
  - Route service tickets to correct party

- **Sales and Marketing**
  - Determine why customers like or dislike your product and categorize
  - Find what is working in your sales and commercials
Text Analytics – Case Study

- As part of the Army’s ongoing Category Management efforts CE is developing a solution to parse through each contract and get a sense of the words contained.
- Object is to parse through thousands of contracts.
- Two main functionalities now:
  - Document/Word tagging:
    - Find specific wording input into system and see how many times it appears.
  - Topic modeling:
    - Use clustering to determine groupings most seen in each type of contract.
Text Analytics – The results (pt 2)

Terms search

<table>
<thead>
<tr>
<th>Term</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFEBS</td>
<td>System</td>
</tr>
<tr>
<td>LMP</td>
<td>System</td>
</tr>
<tr>
<td>GCSS A</td>
<td>System</td>
</tr>
<tr>
<td>OSMIS</td>
<td>System</td>
</tr>
<tr>
<td>FORCES</td>
<td>System</td>
</tr>
<tr>
<td>AMCOS</td>
<td>System</td>
</tr>
<tr>
<td>CPP</td>
<td>System</td>
</tr>
<tr>
<td>Cost Performance Portal</td>
<td>System</td>
</tr>
<tr>
<td>JIAT</td>
<td>System</td>
</tr>
<tr>
<td>TSS-E</td>
<td>System</td>
</tr>
<tr>
<td>GCSS Army</td>
<td>System</td>
</tr>
<tr>
<td>IT</td>
<td>Term</td>
</tr>
<tr>
<td>Information Technology</td>
<td>Term</td>
</tr>
</tbody>
</table>

Results found

<table>
<thead>
<tr>
<th>ngrams</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>it</td>
<td>13</td>
</tr>
</tbody>
</table>
Text Analytics – Predictive modeling

Our goal is to be able to have the system identify and/or predict once a contract comes into the solution what category it is most likely to be in:
- Allows future contracts to be allocated
- Allows the double checking of categories given separate data