This notebook displays visualizations for subsequent runs of a LDA training on the same corpus of data.

In [1]:
```
from IPython.core.display import display, HTML
import ktm_prepviz, pyLDAvis
vis = ktm_prepviz.prepviz("tokenization")
```

In [9]:
```
display(HTML("<h1>Tokenized with large dictionary</h1>"))
pyLDAvis.display(vis[0])
```

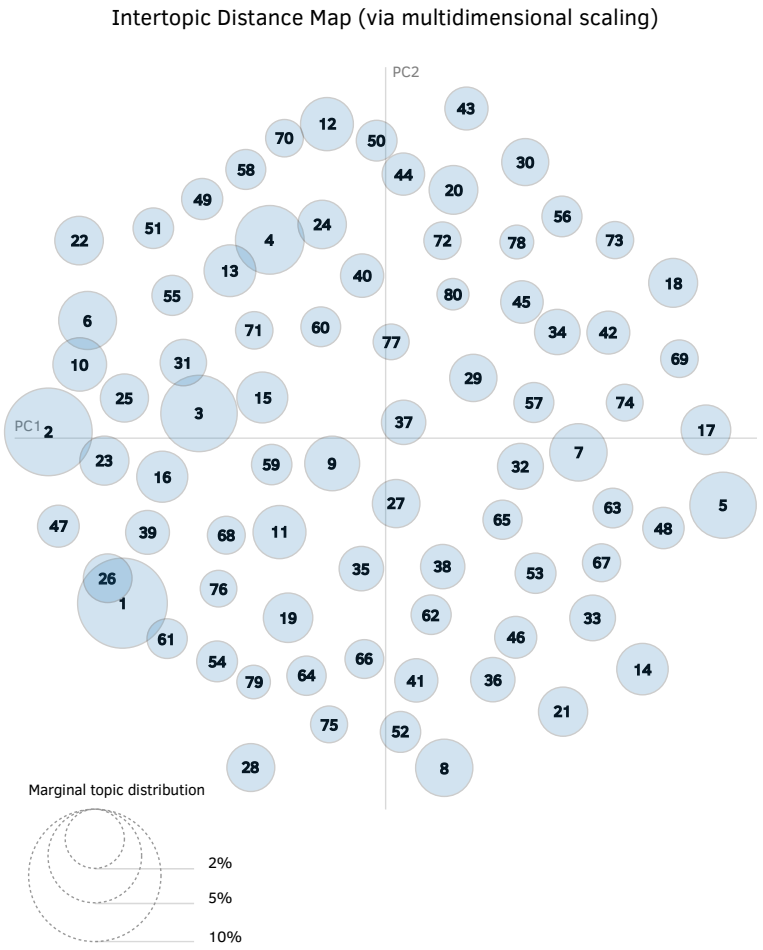
Out[9]:

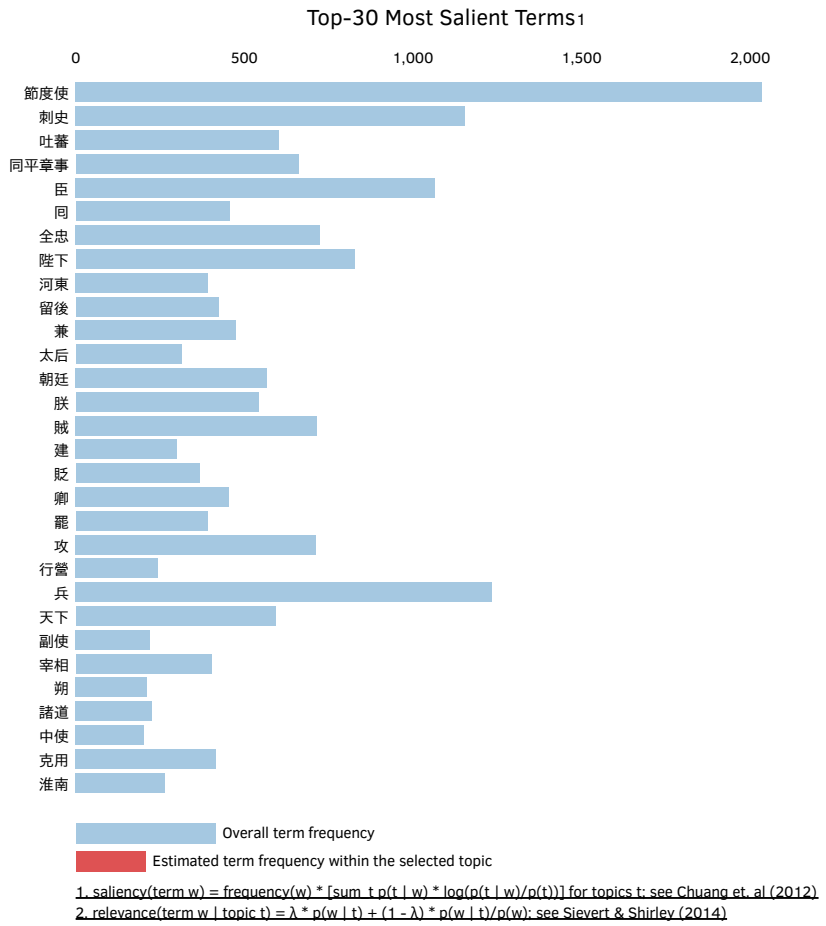
Tokenized with large dictionary

Selected Topic: 0 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:(2) λ = 1

0.0 0.2 0.4 0.6 0.8 1





1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

In [8]:
```
display(HTML("<h1>Tokenized with small dictionary</h1>"))
pyLDAvis.display(vis[1])
```

Out[8]:

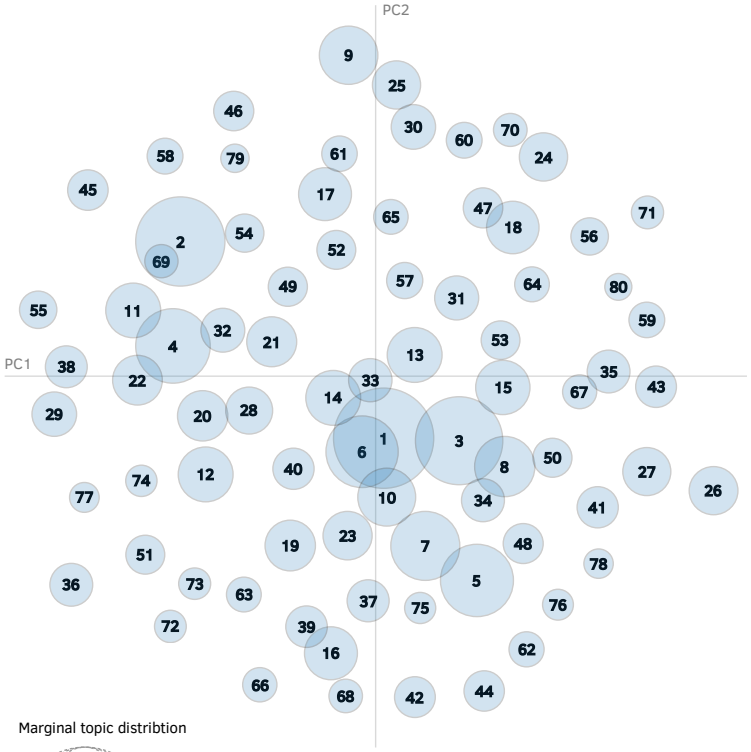
Tokenized with small dictionary

Selected Topic: 0 | Previous Topic | Next Topic | Clear Topic

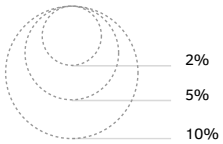
Slide to adjust relevance metric:(2) λ = 1

0.0 0.2 0.4 0.6 0.8 1

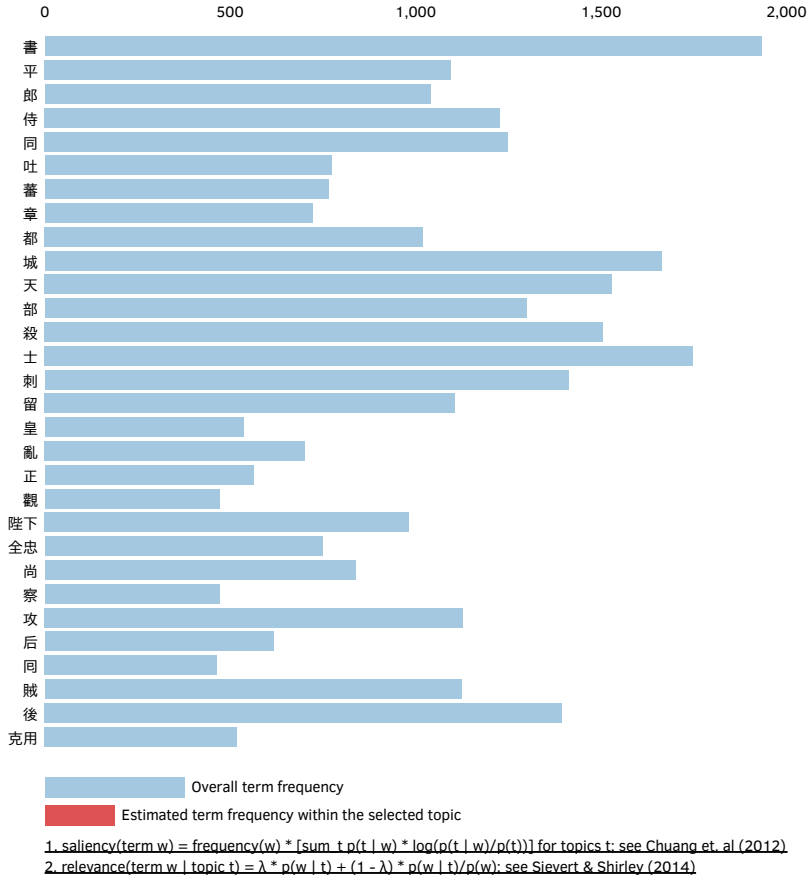
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribtion



Top-30 Most Salient Terms(1)



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

In [7]:
```
display(HTML("<h1>Tokenized by single character</h1>"))
pyLDAvis.display(vis[2])
```

Out[7]:

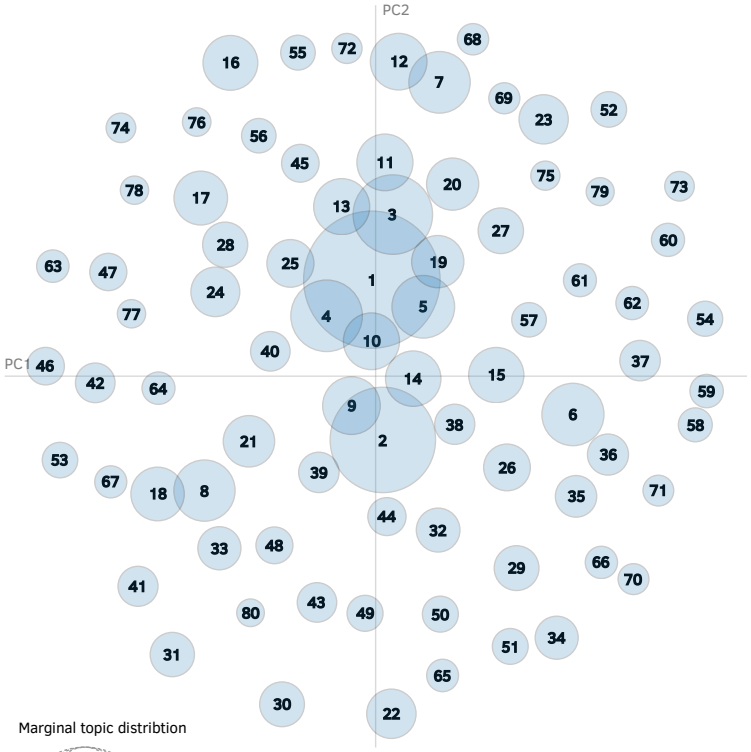
Tokenized by single character

Selected Topic: 0 | Previous Topic | Next Topic | Clear Topic

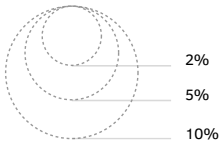
Slide to adjust relevance metric:(2) λ = 1

0.0 0.2 0.4 0.6 0.8 1

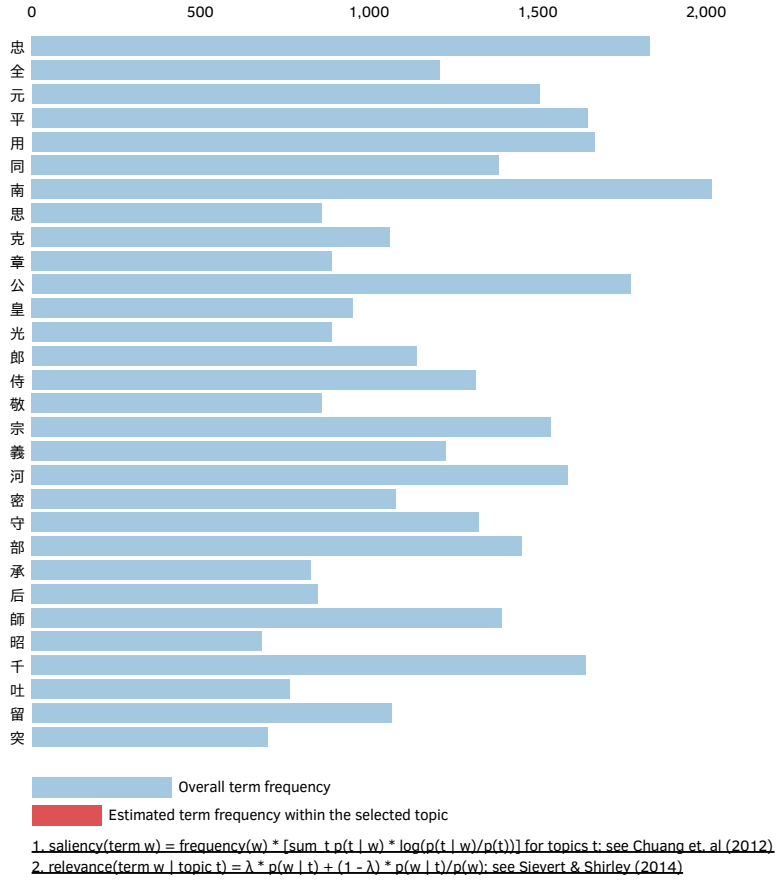
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribtion



Top-30 Most Salient Terms(1)



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

In [5]:
```
#pyLDAvis.display(vis[3])
```

In [6]:
```
pyLDAvis.display(vis[4])
```

```
---------------------------------------------------------------------------
IndexError                                Traceback (most recent call last)
<ipython-input-6-d1350acb6302> in <module>()
----> 1 pyLDAvis.display(vis[4])

IndexError: list index out of range
```

In [ ]:
```

```