



THE UNIVERSITY
OF BRITISH COLUMBIA

Visual analytics for exploring online conversations



Soheil Alavi
salavis@cs.ubc.ca



Felipe González-Pizarro
felipegp@cs.ubc.ca

Topic modeling and InfoVis

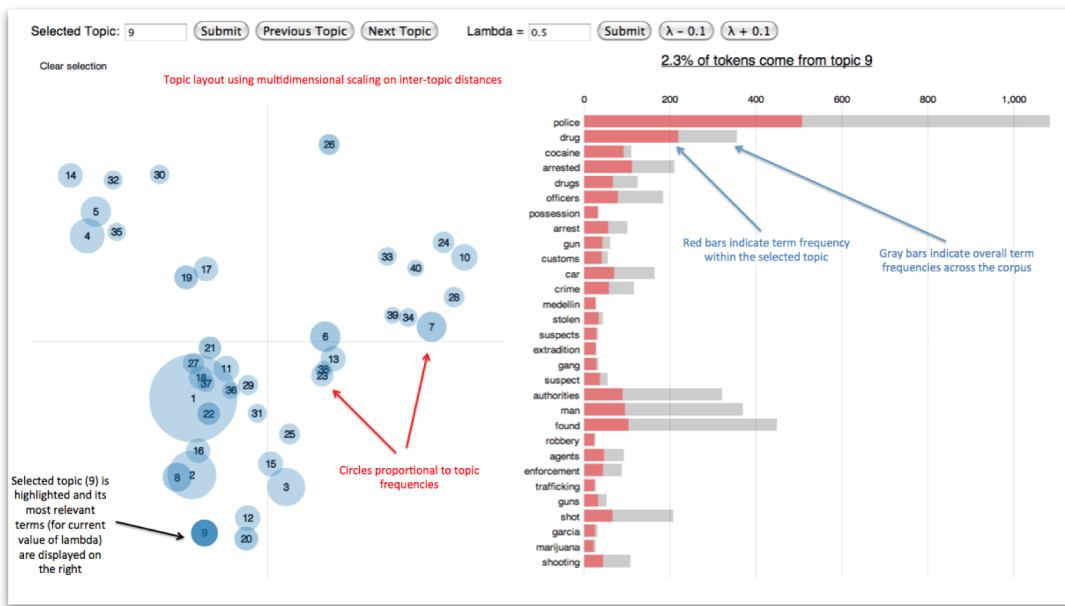
- ▶ It is needed to **summarize** and **understand** textual data
- ▶ Promising solution: Topic modeling
 - ▶ Statistical approach for extracting **topics** from large text corpora.
 - ▶ Topic models do not provide meanings and interpretation directly
 - ▶ **humans must be involved [1]**
- ▶ Humans who directly interact with and interpret the output of topic modeling **may rely on visualization tools to better interpret the results [2]**

[1] Dou, W., Wang, X., Chang, R., & Ribarsky, W. (2011, October). Paralleltopics: A probabilistic approach to exploring document collections. In *2011 IEEE conference on visual analytics science and technology (VAST)* (pp. 231-240). IEEE.

[2] Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems* (pp. 288-296).

Topic modeling interpretation

```
Símbolo del sistema
aniola sangre honra fuerza jamás honor hijos razón lucha libertad
0,5 nación viene pobres inglaterra spencer humanos ley llama naturalmente echar humana tr
abajadores gente busca pide cuya esclavitud remedio partes servidumbre
7,0,5 alma ideas tierra poeta luz verso tiempos ojos poema cielo frente pecho viene árbol c
osa estudio mente hijos madre mar
8,0,5 mundo poesía whitman aire canto libro amigos walt universo palabras natural muerte am
a religión siente libros composición crear mar vivos
9,0,5 libertad vida guerra cubano manos cuerpo cubanos hora corazón años sangre fuerzas his
toria útil mano gloria obra justicia universal deber
<950> LL/token: -9,84201
<960> LL/token: -9,84735
<970> LL/token: -9,85153
<980> LL/token: -9,84867
<990> LL/token: -9,8374
0,5 unidos moneda plata comisión comercio unión delegados washington política congreso ne
w internacional delegado york hispanoamérica argentina podía conferencia monetaria político
1,0,5 pueblos américa país pueblo mundo elementos gobierno común repúblicas natural interés
odio sistema países norte intereses cabeza grandeza peligro factores
2,0,5 hombres hombre tierra negro blanco raza naturaleza negros razón palabra educación úni
co tierras estudio camino pensamiento venir pie superior ciencia
3,0,5 hombre vida naturaleza hombres mente humana espíritu sol época ideas oro forma poetas
ven existencia grande alas humano fuego fuerza
4,0,5 méxico unidos derechos puesto riqueza tratado libre nación países york entra obra pro
ductos frutos cambio leyes ferrocarriles llegar proyecto paso
5,0,5 cuba república pueblo derecho españa patria voluntad independencia libertad gobierno
sangre española hijos honra libre derechos revolución honor universal lucha
6,0,5 paz pueblo esclavitud pobres idea nación viene inglaterra spencer humanos cuya llama
naturalmente trabajadores gente remedio provecho partes número esclavo
7,0,5 alma tierra poeta luz ojos verso tiempos viene pecho poema cielo árbol cosa frente ma
dre brazos ley casa pone acaso
8,0,5 mundo poesía whitman libro aire mar natural canto muerte vive amor amigos walt univer
so pies brazo ama palabras religión jóvenes
9,0,5 libertad guerra vida cubano manos hora años cubanos virtud carácter mano corazón just
icia historia cuerpo sangre útil alto fuerzas mérito
<1000> LL/token: -9,84364
Total time: 13 seconds
C:\mallet>
```

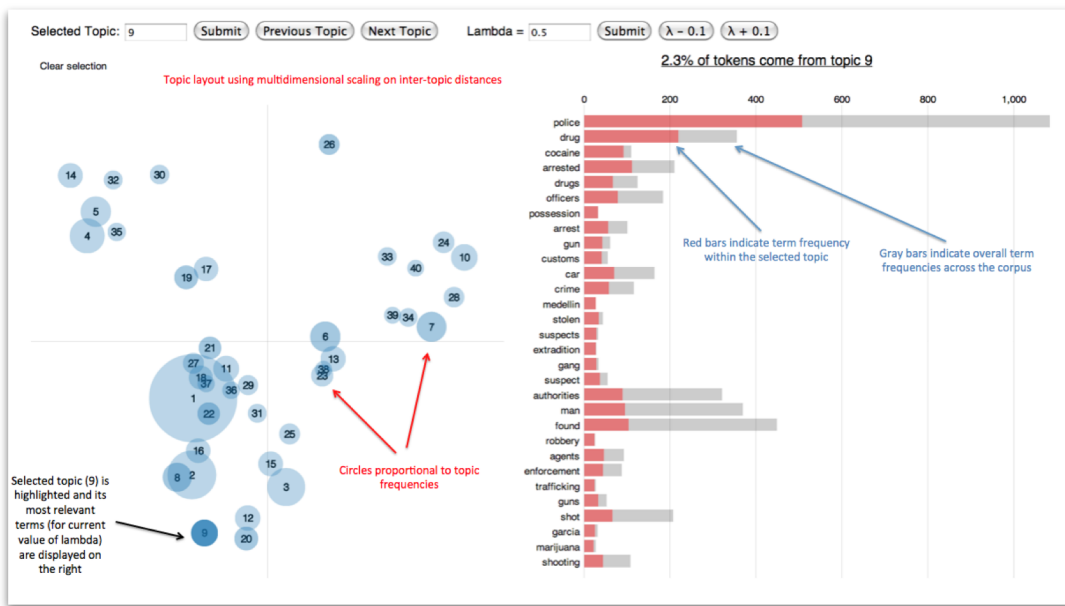


[1] Graham, S., Weingart, S., & Milligan, I. (2012). Getting started with topic modeling and MALLET. The Editorial Board of the Programming Historian.

[2] Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70).

Topic modeling interpretation

```
Símbolo del sistema
aniola sangre honra fuerza jamás honor hijos razón lucha libertad
0,5 nación viene pobres inglaterra spencer humanos ley llama naturalmente echar humana tr
abajadores gente busca nide cuya esclavitud remedio partes servidumbre
7,0,5 alma ideas tierra poeta luz verso tiempos ojos poema cielo frente pecho viene árbol c
osa estudio mente hijos madre mar
8,0,5 mundo poesía whitman aire canto libro amigos walt universo palabras natural muerte am
a religión siente libros composición crear mar vivos
9,0,5 libertad vida guerra cubano manos cuerpo cubanos hora corazón años sangre fuerzas his
toria útil mano gloria obra justicia universal deber
<950> LL/token: -9,84201
<960> LL/token: -9,84735
<970> LL/token: -9,85153
<980> LL/token: -9,84867
<990> LL/token: -9,8374
>
0,5 unidos moneda plata comisión comercio unión delegados washington política congreso n
internacional delegado york hispanoamérica argentina podía conferencia monetaria político
0,5 pueblos américa país pueblo mundo elementos gobierno común repúblicas natural interé
odio sistema países norte intereses cabeza grandeza peligro factores
0,5 hombres hombre tierra negro blanco raza naturaleza negros razón palabra educación ún
o tierras estudio camino pensamiento venir pie superior ciencia
0,5 hombre vida naturaleza hombres mente humana espíritu sol época ideas oro forma poeta
ven existencia grande alas humano fuego fuerza
0,5 méxico unidos derechos puesto riqueza tratado libre nación países york entra obra pr
ductos frutos cambio leyes ferrocarriles llegar proyecto paso
0,5 cuba república pueblo derecho españa patria voluntad independencia libertad gobierno
s sangre española hijos honra libre derechos revolución honor universal lucha
0,5 paz pueblo esclavitud pobres idea nación viene inglaterra spencer humanos cuya llama
naturalmente trabajadores gente remedio provecho partes número esclavo
0,5 alma tierra poeta luz ojos verso tiempos viene pecho poema cielo árbol cosa frente m
re brazos ley casa pone acaso
0,5 mundo poesía whitman libro aire mar natural canto muerte vive amor amigos walt unive
o pies brazo ama palabras religión jóvenes
0,5 libertad guerra vida cubano manos hora años cubanos virtud carácter mano corazón jus
cia historia cuerpo sangre útil alto fuerzas mérito
Total time: 13 seconds
C:\mallet>
```

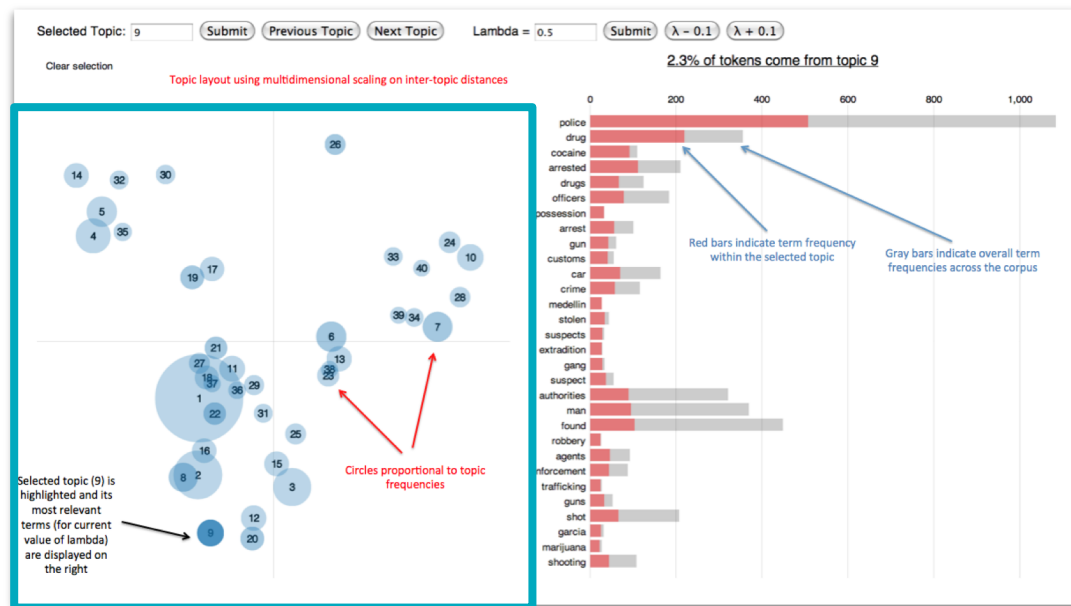


[1] Graham, S., Weingart, S., & Milligan, I. (2012). Getting started with topic modeling and MALLET. The Editorial Board of the Programming Historian.

[2] Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70).

Topic modeling interpretation

```
Símbolo del sistema
aniola sangre honra fuerza jamás honor hijos razón lucha libertad
0 0,5 nación viene pobres inglaterra spencer humanos ley llama naturalmente echar humana tr
abajadores gente busca pide cuya esclavitud remedio partes servidumbre
7 0,5 alma ideas tierra poeta luz verso tiempos ojos poema cielo frente pecho viene árbol c
osa estudio mente hijos madre mar
8 0,5 mundo poesía whitman aire canto libro amigos walt universo palabras natural muerte am
a religión siente libros composición crear mar vivos
9 0,5 libertad vida guerra cubano manos cuerpo cubanos hora corazón años sangre fuerzas his
toria útil mano gloria obra justicia universal deber
<950> LL/token: -9,84201
<960> LL/token: -9,84735
<970> LL/token: -9,85153
<980> LL/token: -9,84867
<990> LL/token: -9,8374
0 0,5 unidos moneda plata comisión comercio unión delegados washington política congreso ne
w internacional delegado york hispanoamérica argentina podía conferencia monetaria político
1 0,5 pueblos américa país pueblo mundo elementos gobierno común repúblicas natural interés
odio sistema países norte intereses cabeza grandeza peligro factores
2 0,5 hombres hombre tierra negro blanco raza naturaleza negros razón palabra educación úni
co tierras estudio camino pensamiento venir pie superior ciencia
3 0,5 hombre vida naturaleza hombres mente humana espíritu sol época ideas oro forma poetas
ven existencia grande alas humano fuego fuerza
4 0,5 méxico unidos derechos puesto riqueza tratado libre nación países york entra obra pro
ductos frutos cambio leyes ferrocarriles llegar proyecto paso
5 0,5 cuba república pueblo derecho españa patria voluntad independencia libertad gobierno
sangre española hijos honra libre derechos revolución honor universal lucha
6 0,5 paz pueblo esclavitud pobres idea nación viene inglaterra spencer humanos cuya llama
naturalmente trabajadores gente remedio provecho partes número esclavo
7 0,5 alma tierra poeta luz ojos verso tiempos viene pecho poema cielo árbol cosa frente ma
dre brazos ley casa pone acaso
8 0,5 mundo poesía whitman libro aire mar natural canto muerte vive amor amigos walt univer
so pies brazo ama palabras religión jóvenes
9 0,5 libertad guerra vida cubano manos hora años cubanos virtud carácter mano corazón just
icia historia cuerpo sangre útil alto fuerzas mérito
<1000> LL/token: -9,84364
Total time: 13 seconds
C:\mallet>
```

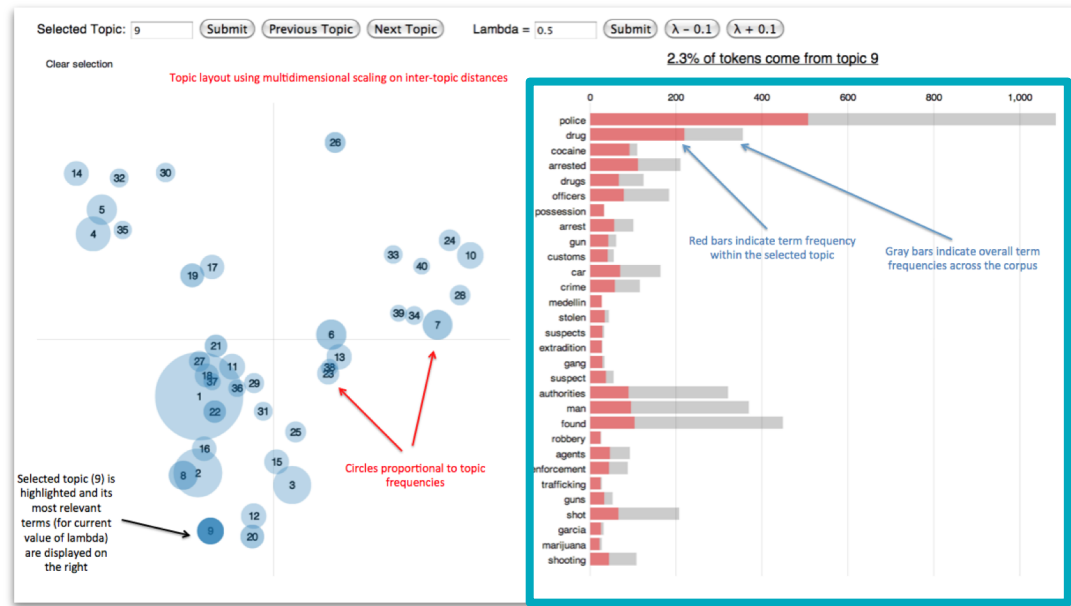


[1] Graham, S., Weingart, S., & Milligan, I. (2012). Getting started with topic modeling and MALLET. The Editorial Board of the Programming Historian.

[2] Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70).

Topic Modeling Interpretation

```
Símbolo del sistema
ánola sangre honra fuerza jamás honor hijos razón lucha libertad
0 0,5 nación viene pobres inglaterra spencer humanos ley llama naturalmente echar humana tr
abajadores gente busca pide cuya esclavitud remedio partes servidumbre
7 0,5 alma ideas tierra poeta luz verso tiempos ojos poema cielo frente pecho viene árbol c
osa estudio mente hijos madre mar
8 0,5 mundo poesía whitman aire canto libro amigos walt universo palabras natural muerte am
a religión siente libros composición crear mar vivos
9 0,5 libertad vida guerra cubano manos cuerpo cubanos hora corazón años sangre fuerzas his
toria útil mano gloria obra justicia universal deber
<950> LL/token: -9,84201
<960> LL/token: -9,84735
<970> LL/token: -9,85153
<980> LL/token: -9,84867
<990> LL/token: -9,8374
0 0,5 unidos moneda plata comisión comercio unión delegados washington política congreso ne
w internacional delegado york hispanoamérica argentina podía conferencia monetaria político
1 0,5 pueblos américa país pueblo mundo elementos gobierno común repúblicas natural interés
odio sistema países norte intereses cabeza grandeza peligro factores
2 0,5 hombres hombre tierra negro blanco raza naturaleza negros razón palabra educación úni
co tierras estudio camino pensamiento venir pie superior ciencia
3 0,5 hombre vida naturaleza hombres mente humana espíritu sol época ideas oro forma poetas
ven existencia grande alas humano fuego fuerza
4 0,5 méxico unidos derechos puesto riqueza tratado libre nación países york entra obra pro
ductos frutos cambio leyes ferrocarriles llegar proyecto paso
5 0,5 cuba república pueblo derecho españa patria voluntad independencia libertad gobierno
sangre española hijos honra libre derechos revolución honor universal lucha
6 0,5 paz pueblo esclavitud pobres idea nación viene inglaterra spencer humanos cuya llama
naturalmente trabajadores gente remedio provecho partes número esclavo
7 0,5 alma tierra poeta luz ojos verso tiempos viene pecho poema cielo árbol cosa frente ma
dre brazos ley casa pone acaso
8 0,5 mundo poesía whitman libro aire mar natural canto muerte vive amor amigos walt univer
so pies brazo ama palabras religión jóvenes
9 0,5 libertad guerra vida cubano manos hora años cubanos virtud carácter mano corazón just
icia historia cuerpo sangre útil alto fuerzas mérito
<1000> LL/token: -9,84364
Total time: 13 seconds
C:\mallet>
```



[1] Graham, S., Weingart, S., & Milligan, I. (2012). Getting started with topic modeling and MALLET. The Editorial Board of the Programming Historian.

[2] Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70).

Finding mechanisms to improve these visual representations is still an open challenge [1]

[1] Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.



Name this search as and

Save

Item	Remove
topic 92. (union, workers, labor) == 2%	X
topic 81. (police, officers, officer) == 3%	X
Locations == NEW YORK CITY	X

Aggregate Statistics (16 documents) Show

Document 106: Mayor Rudolph W. Giuliani threatened yesterday to dismiss any officer who staged a work slowdown in protest of Gov. George E. Pataki's veto of a police arbitration bill that would probably have led to higher pay. Police officials said there was no sign of any major labor action, but that 100 officers may have written fewer parking tickets over the weekend. The action was seen by precinct commanders as individual acts encouraged by a few police union delegates if the Police Commissioner catches ...

Tag:

Search builds on previous searches. It does not reset each time.



TWEETS (10469)

MORGANBLACKXXX Fri, 10/26/2012 at 11:16AM
@ThaMMarilyn @tsalm56...ha ha...ifs Friday! Headed out for a haircut & s...

RONGEORGE Fri, 10/26/2012 at 11:16AM
Tropical Storm warning? I call it a jetski opportunity. #storm

THE UNIVERSITY OF TEXAS AT AUSTIN
NLP

Datasets Favorites

Words Topics

Sort by: Document

Filters

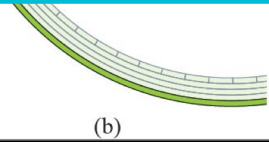
Metric: Coherence Greater than

Add a filter by

119: ...
47: ...
101: ...
120: ...
8: con...
104: ...
142: g...
1: ...
38: ec...
2: ...
41: ...
60: b...
17: ...
32: ...
136: history ...
66: dollars reform
47: mccain senator
70: health care
97: iraq troops
16: young service

work	Intelligent design	[law, state, case]	related topics	[theory, work, human]	[film, series, show]	[god, call, give]
human	Immanuel Kant	[son, year, death]				
idea	Philosophy of mathematics	[woman, child, man]				
term	History of science	[god, call, give]				
study	Free will	[black, white, people]				
view	Truth	[film, series, show]				
science	Psychoanalysis	[law, force, army]				
concept	Charles Peirce	[language, word, form]				
form	Existentialism	[@car@, make, design]				

Charles Peirce
Social sciences
Idealism
Franz Boas
Karl Marx
Mysticism
Truth
Free will
Scientism



36/2012 at 11:20AM
this hurricane is figuring
id in who's...

0/26/2012 at 11:19AM
ane Sandy; Region
ntially devastating s...

CTR Fri, 10/26/2012
ed Cross Hurricane
a great tool...

0/26/2012 at 11:22AM
: @MissZee_umm i am
onda thank u ver...

26/2012 at 11:22AM
ast yr we did not have
ck because not i had p...

Fri, 10/26/2012 at
work washington reports
naphy agency online legal
tage document
tap nsa post at web
sho collect
terish content link
snowden muscular
filter cameron server
parent david grom



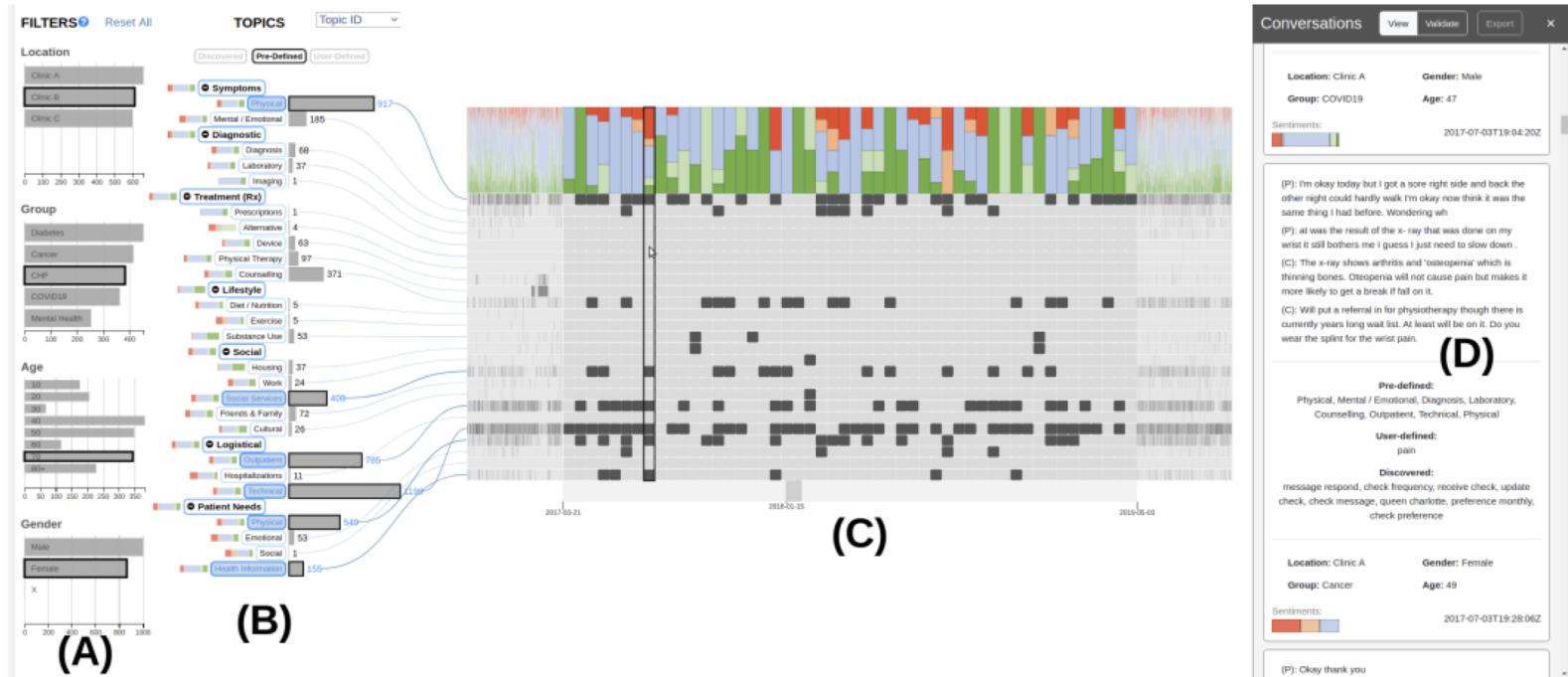
(c)



(d)

(b)

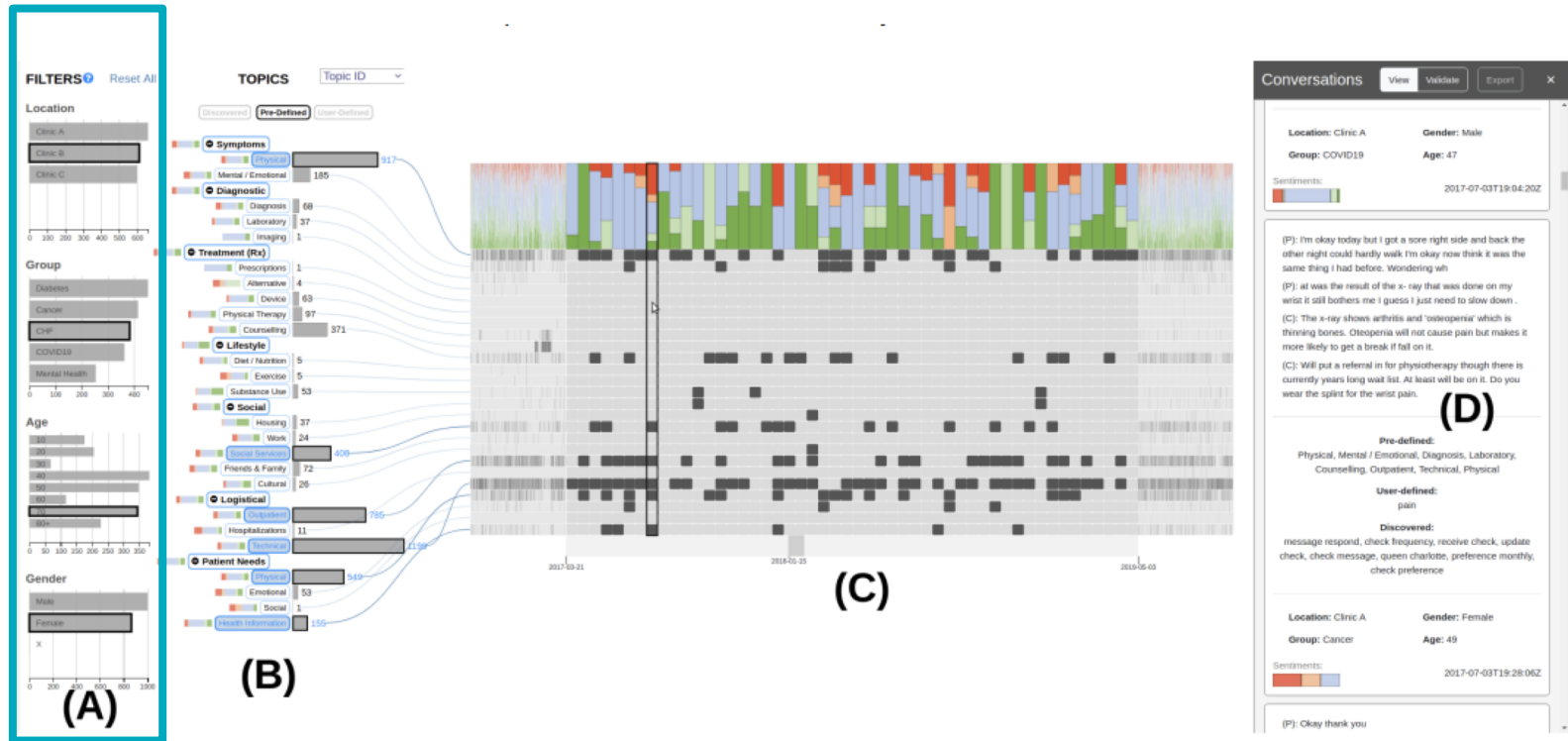
Recent research efforts



ConVIscope Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D).

Li, R., Hoque, E., Carenini, G., Lester, R., & Chau, R. (2021). ConVIscope: Visual Analytics for Exploring Patient Conversations. *arXiv preprint arXiv:2108.13514*.

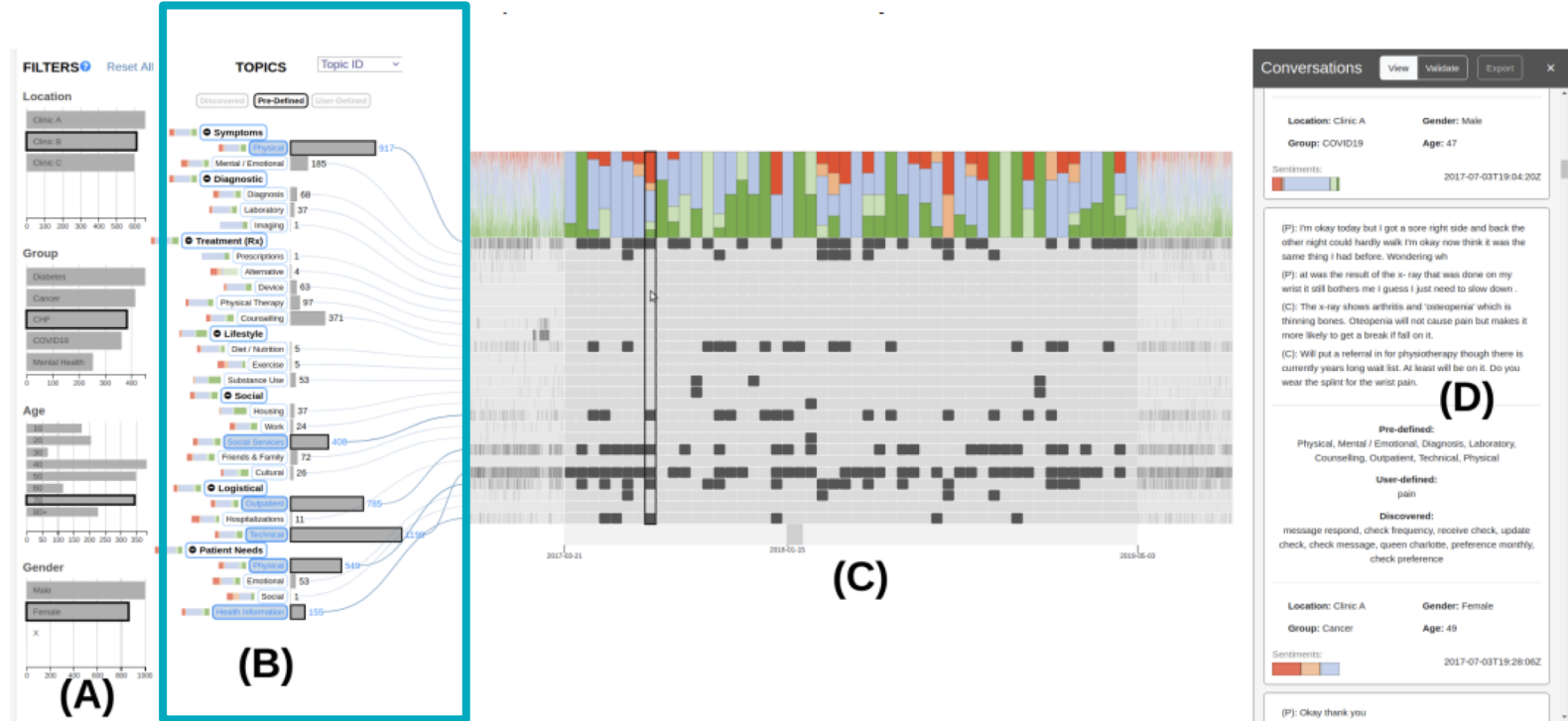
Recent research efforts



ConVIscope Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D).

Li, R., Hoque, E., Carenini, G., Lester, R., & Chau, R. (2021). ConVIscope: Visual Analytics for Exploring Patient Conversations. *arXiv preprint arXiv:2108.13514*.

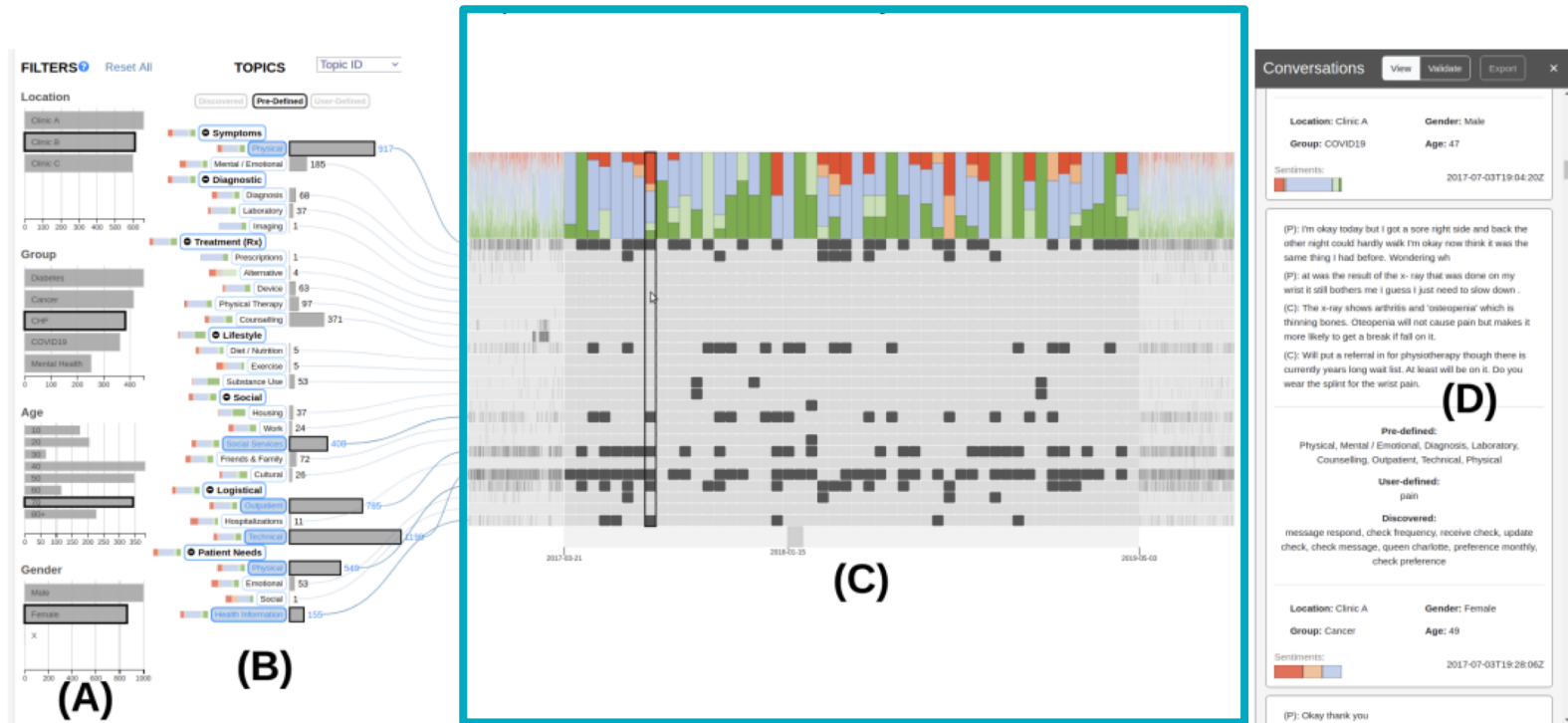
Recent research efforts



ConVIscope Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D).

Li, R., Hoque, E., Carenini, G., Lester, R., & Chau, R. (2021). ConVIscope: Visual Analytics for Exploring Patient Conversations. *arXiv preprint arXiv:2108.13514*.

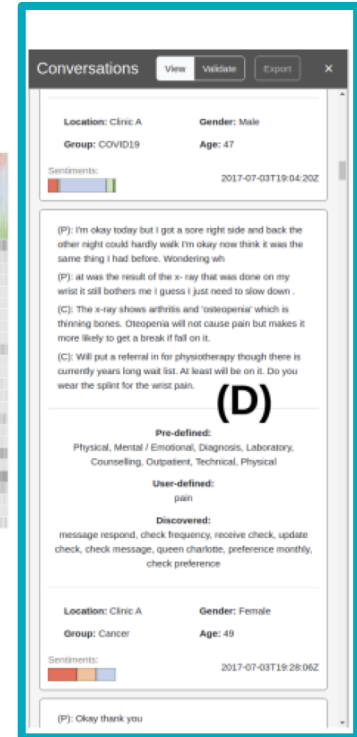
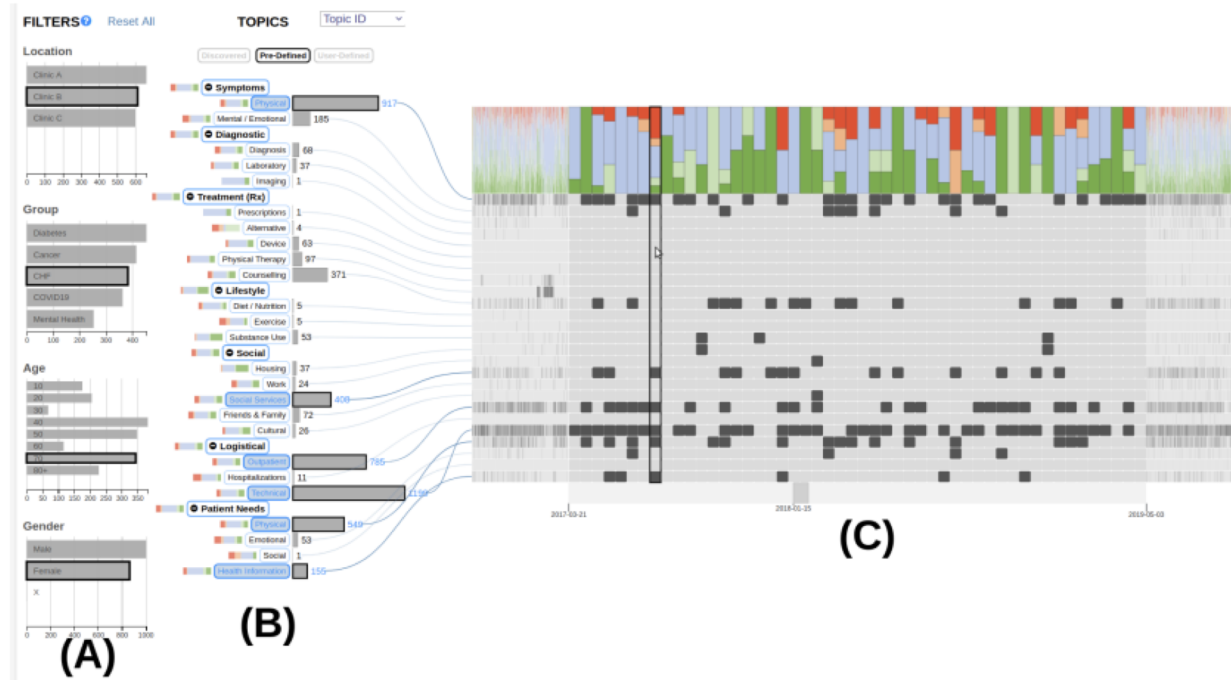
Recent research efforts



ConVIscope Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D).

Li, R., Hoque, E., Carenini, G., Lester, R., & Chau, R. (2021). ConVIscope: Visual Analytics for Exploring Patient Conversations. *arXiv preprint arXiv:2108.13514*.

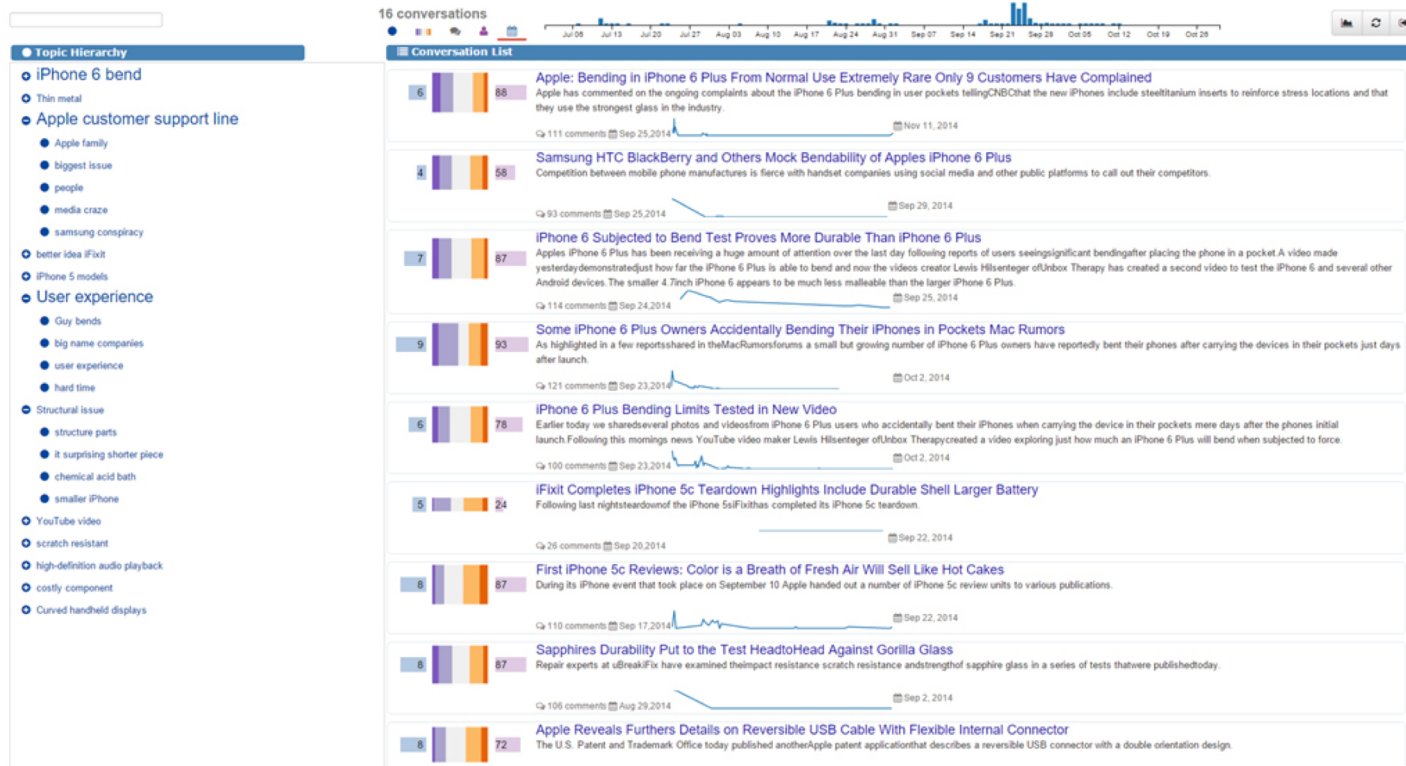
Recent research efforts



ConVIScope Views (see §4): Metadata (A), Topics (B), Analysis (C) and Conversations (D).

Li, R., Hoque, E., Carenini, G., Lester, R., & Chau, R. (2021). ConVIScope: Visual Analytics for Exploring Patient Conversations. *arXiv preprint arXiv:2108.13514*.

Recent research efforts



Hoque, E., & Carenini, G. (2019). Interactive topic hierarchy revision for exploring a collection of online conversations. *Information Visualization*, 18(3), 318–338. <https://doi.org/10.1177/1473871618757228>



Recent research efforts



Hoque, E., & Carenini, G. (2019). Interactive topic hierarchy revision for exploring a collection of online conversations. *Information Visualization*, 18(3), 318–338. <https://doi.org/10.1177/1473871618757228>



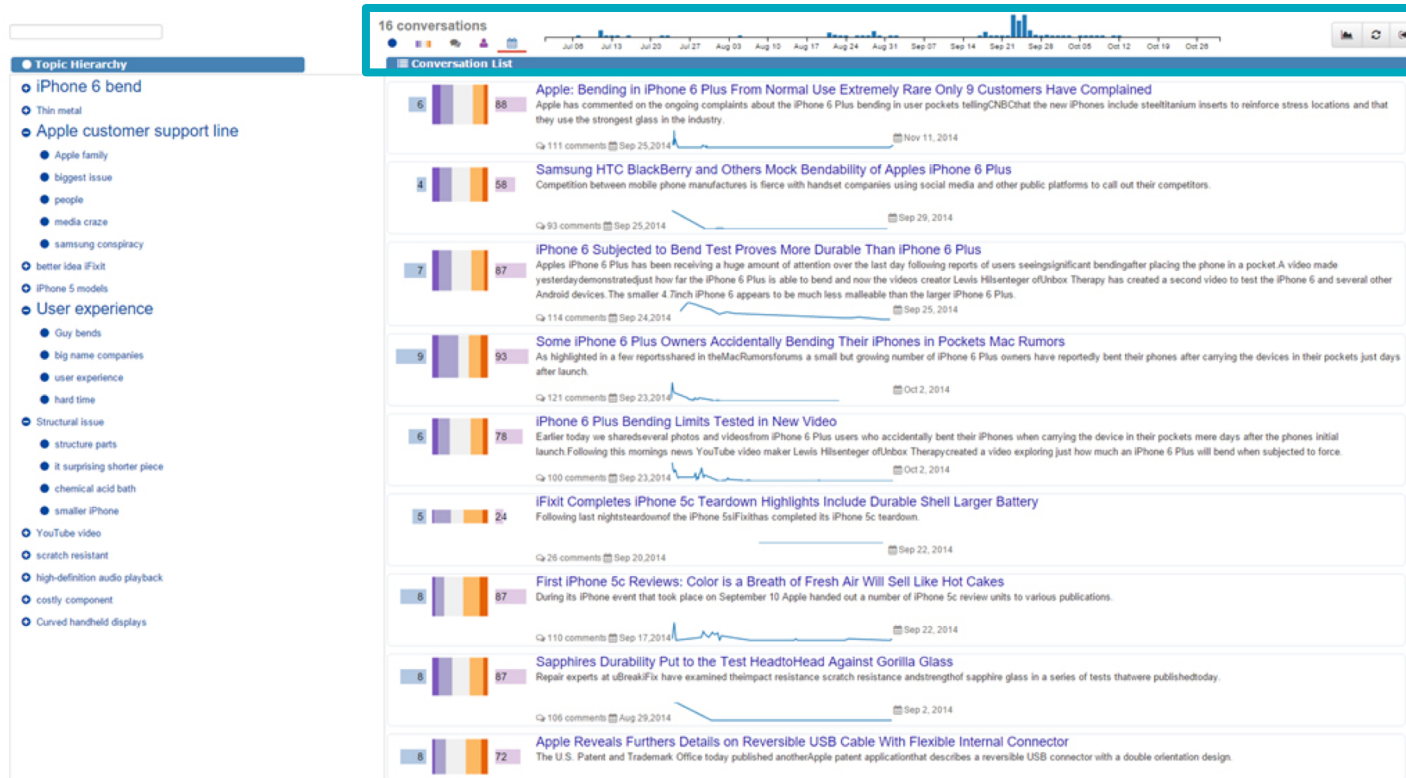
Recent research efforts



Hoque, E., & Carenini, G. (2019). Interactive topic hierarchy revision for exploring a collection of online conversations. *Information Visualization*, 18(3), 318–338. <https://doi.org/10.1177/1473871618757228>



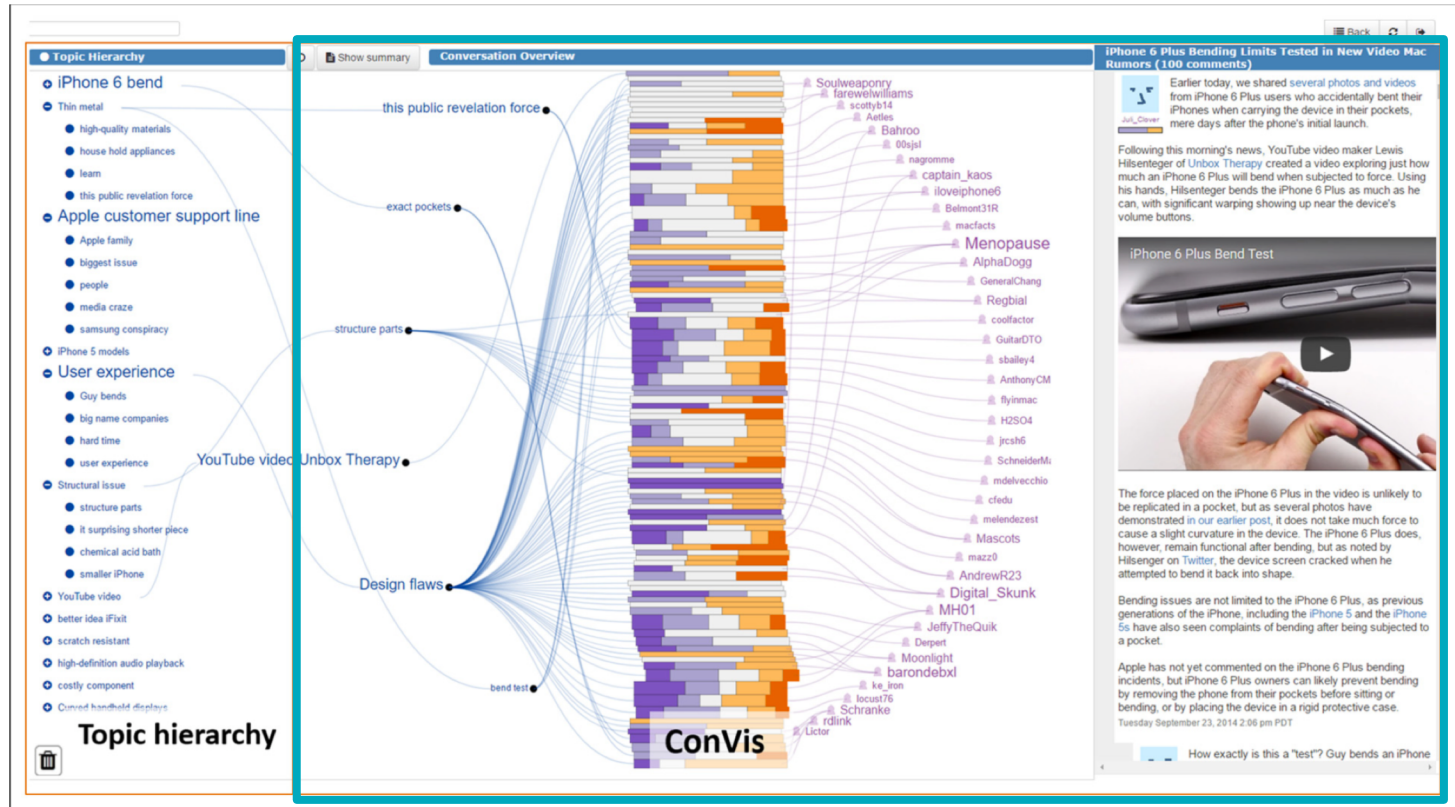
Recent research efforts



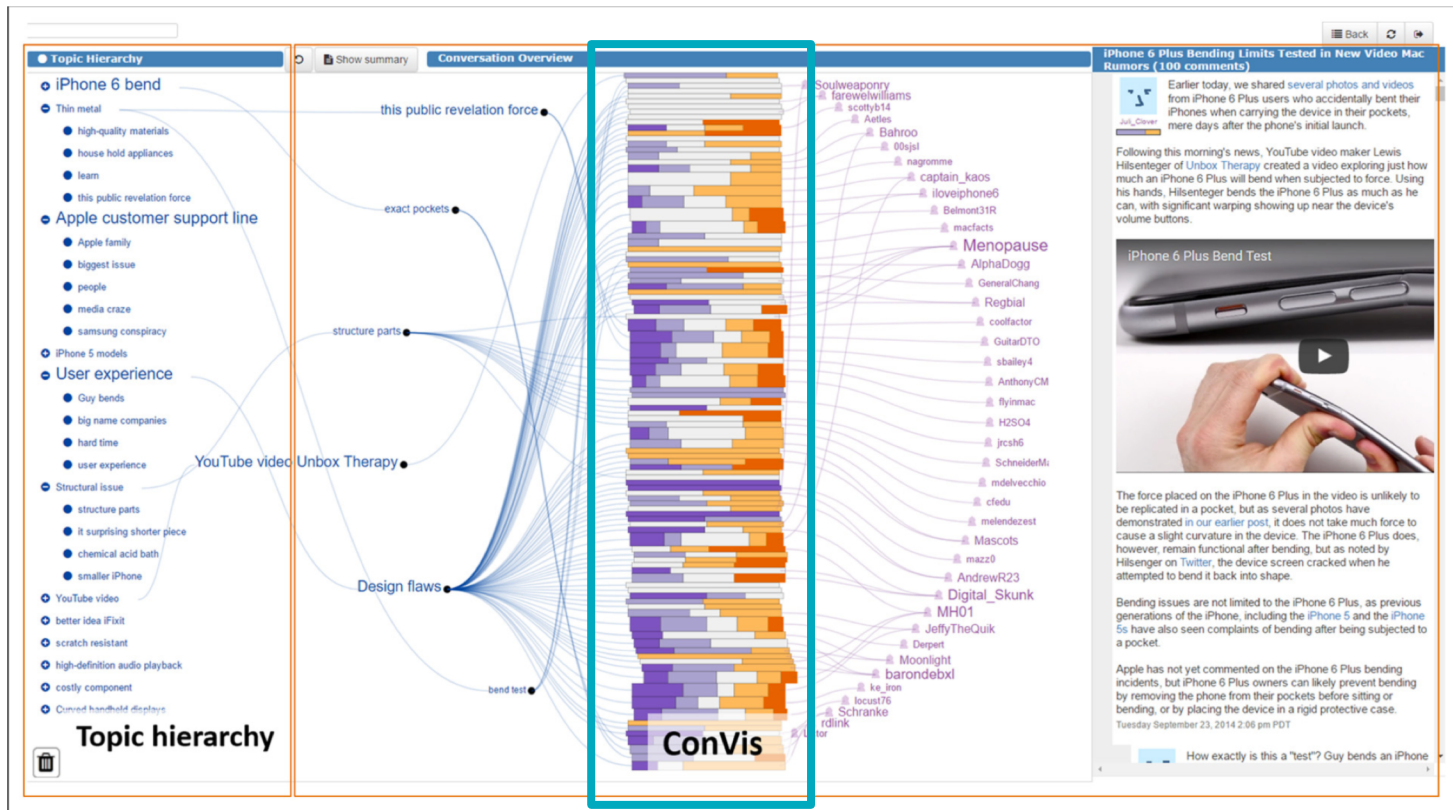
Hoque, E., & Carenini, G. (2019). Interactive topic hierarchy revision for exploring a collection of online conversations. *Information Visualization*, 18(3), 318–338. <https://doi.org/10.1177/1473871618757228>



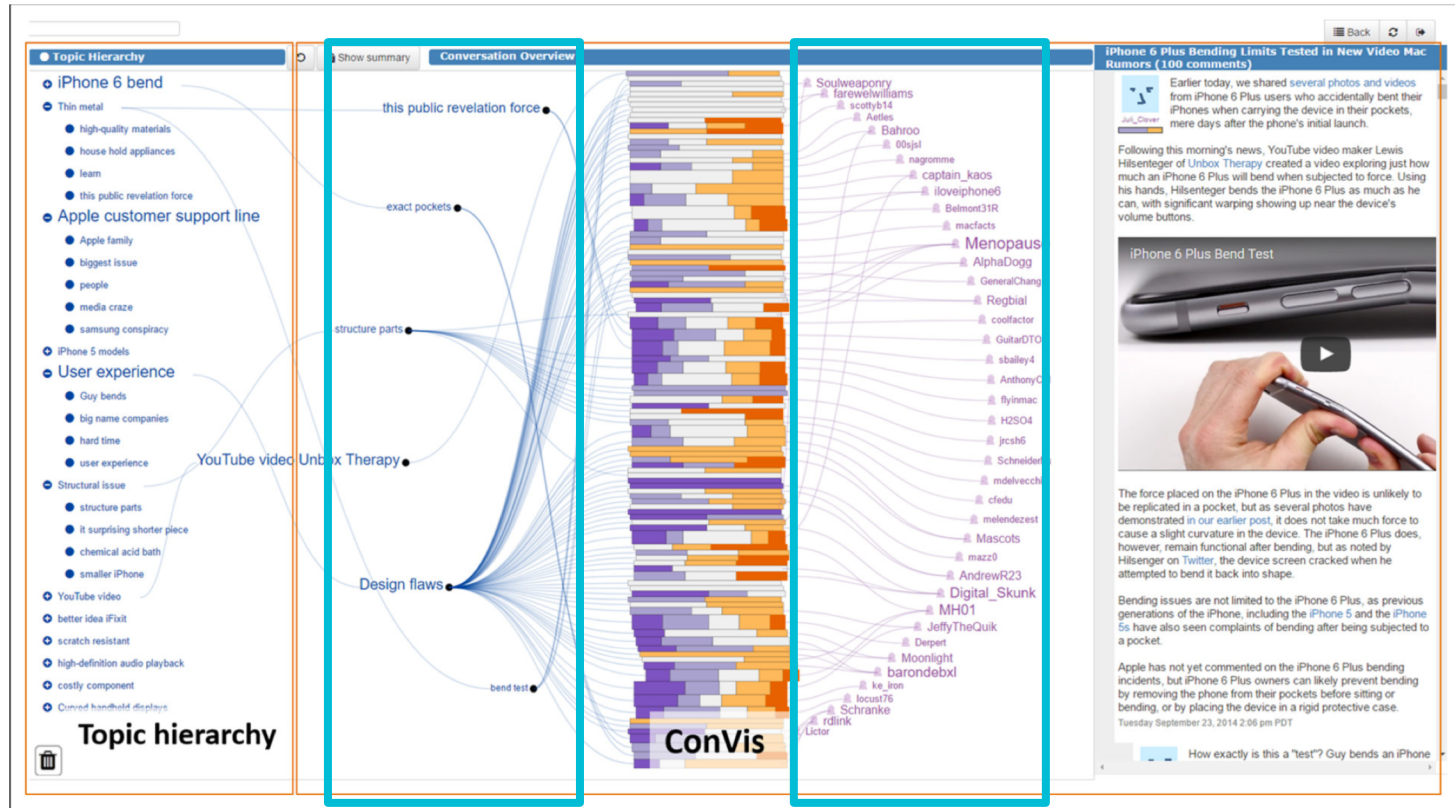
Recent research efforts



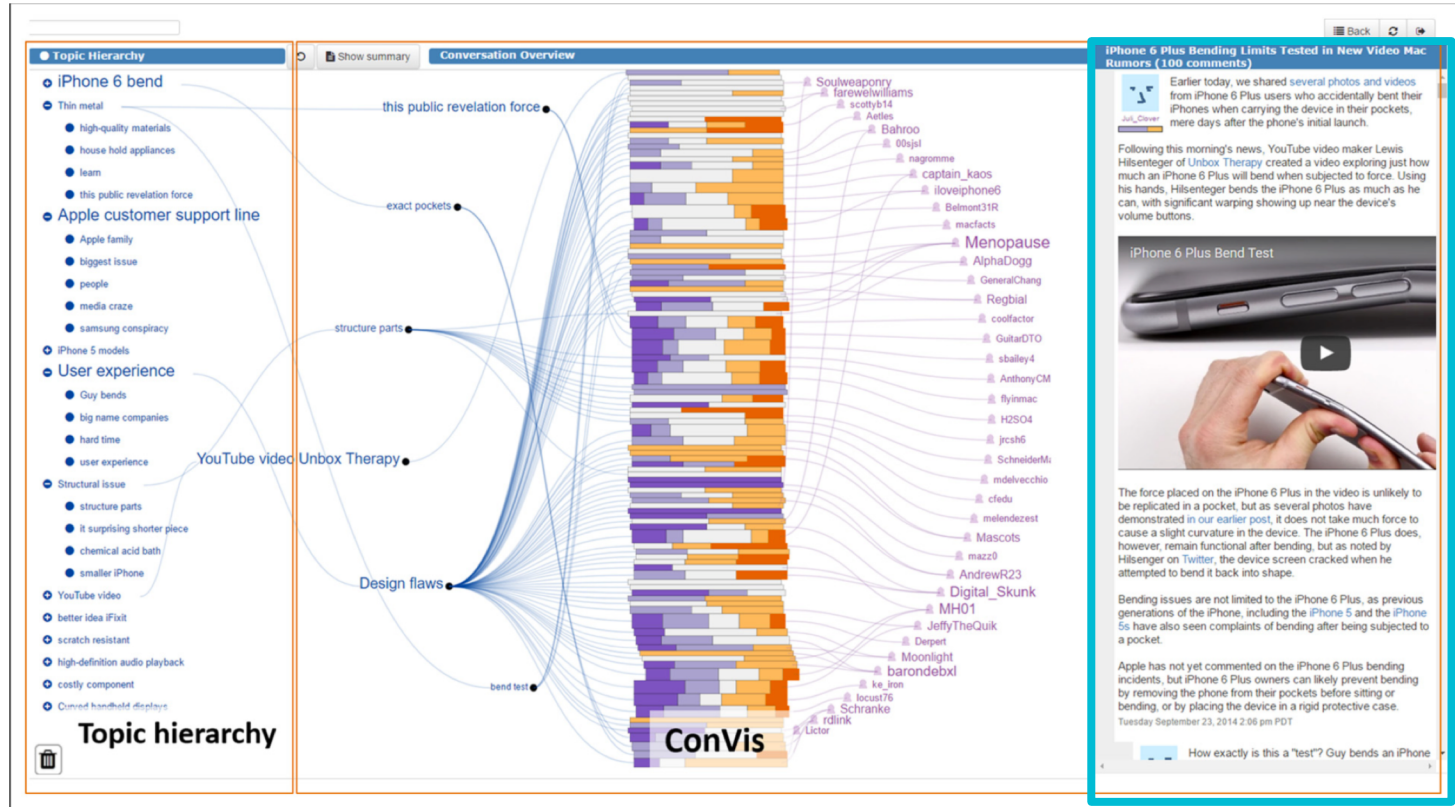
Recent research efforts



Recent research efforts



Recent research efforts



iPhone 6 Plus Bending Limits Tested In New Video Mac Rumors (100 comments)

Earlier today, we shared several photos and videos from iPhone 6 Plus users who accidentally bent their iPhones when carrying the device in their pockets, mere days after the phone's initial launch.

Following this morning's news, YouTube video maker Lewis Hilsentger of Unbox Therapy created a video exploring just how much an iPhone 6 Plus will bend when subjected to force. Using his hands, Hilsentger bends the iPhone 6 Plus as much as he can, with significant warping showing up near the device's volume buttons.

iPhone 6 Plus Bend Test

The force placed on the iPhone 6 Plus in the video is unlikely to be replicated in a pocket, but as several photos have demonstrated in our earlier post, it does not take much force to cause a slight curvature in the device. The iPhone 6 Plus does, however, remain functional after bending, but as noted by Hilsentger on Twitter, the device screen cracked when he attempted to bend it back into shape.

Bending issues are not limited to the iPhone 6 Plus, as previous generations of the iPhone, including the iPhone 5 and the iPhone 5s have also seen complaints of bending after being subjected to a pocket.

Apple has not yet commented on the iPhone 6 Plus bending incidents, but iPhone 6 Plus owners can likely prevent bending by removing the phone from their pockets before sitting or bending, or by placing the device in a rigid protective case.

Tuesday, September 23, 2014 2:06 pm PDT

How exactly is this a "test"? Guy bends an iPhone



Datasets

Text:

- ▶ Ubuntu Dialogue Dataset
- ▶ Amazon QA dataset: <https://jmcauley.ucsd.edu/data/amazon/qa/>
- ▶ Alexa Topical Dataset (Alexa prize competition)
- ▶ Movie Dialogue Corpus
- ▶ More?
 - ▶ ParlAI Framework has incorporated a variety of text datasets for downstream NLP Tasks: <https://github.com/facebookresearch/ParlAI>

Multimodal (Image + Text) :

- ▶ 4chan <https://zenodo.org/record/3606810#.YU-wSLhKiUk>
- ▶ MMHS150K Dataset. Hate speech in Twitter. <https://gomburu.github.io/2019/10/09/MMHS/>



Potential future work

- ▶ Provide support for Multimodal (Image - text) datasets
 - ▶ Adapt previous tools to support image content

- ▶ Seek options to:
 - ▶ Visualize emotions in addition to the sentiment analysis
 - ▶ Scalability: Visualize larger number of topics (e.g, 100 topics)



Questions?



Thank You! :)

