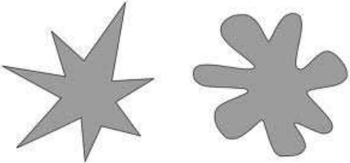
# From Graphs to Maps

Stephen Kobourov
University of Arizona

artist: Matt Cusick

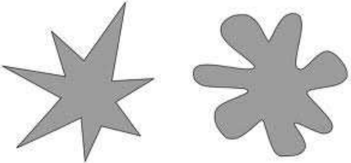
Karlsruhe Institute of Technology, June 23, 2015

# Motivation: why visualize?



- Images are powerful

# Motivation: why visualize?



- Images are powerful
- People are good at images

- Given metric data

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 92 | 04 | 06 | 13 | 03 | 14 | 10 | 13 | 46 | 05 | 22 | 03 | 25 | 34 | 06 | 06 | 09 | 35 | 23 | 06 | 37 | 13 | 17 | 12 | 07 | 03 | 02 | 07 | 05 | 05 | 08 | 06 | 05 | 06 | 02 | 03 | A |
| B | 05 | 84 | 37 | 31 | 05 | 28 | 17 | 21 | 05 | 19 | 34 | 40 | 06 | 10 | 12 | 22 | 25 | 16 | 18 | 02 | 18 | 34 | 08 | 84 | 30 | 42 | 12 | 17 | 14 | 40 | 32 | 74 | 43 | 17 | 04 | 04 | B |
| C | 04 | 38 | 87 | 17 | 04 | 29 | 13 | 07 | 11 | 19 | 24 | 35 | 14 | 03 | 09 | 51 | 34 | 24 | 14 | 06 | 06 | 11 | 14 | 32 | 82 | 38 | 13 | 15 | 31 | 14 | 10 | 30 | 28 | 24 | 18 | 12 | C |
| D | 08 | 62 | 17 | 88 | 07 | 23 | 40 | 36 | 09 | 13 | 81 | 56 | 08 | 07 | 09 | 27 | 09 | 45 | 29 | 06 | 17 | 20 | 27 | 40 | 15 | 33 | 03 | 09 | 06 | 11 | 09 | 19 | 08 | 10 | 05 | 06 | D |
| E | 06 | 13 | 14 | 06 | 97 | 02 | 04 | 04 | 17 | 01 | 05 | 06 | 04 | 04 | 05 | 01 | 05 | 10 | 07 | 67 | 03 | 03 | 02 | 05 | 06 | 05 | 04 | 03 | 05 | 03 | 05 | 02 | 04 | 02 | 03 | 03 | E |
| F | 04 | 51 | 33 | 19 | 02 | 90 | 10 | 29 | 05 | 33 | 16 | 50 | 07 | 06 | 10 | 42 | 12 | 35 | 14 | 02 | 21 | 27 | 25 | 19 | 27 | 13 | 08 | 16 | 47 | 25 | 26 | 24 | 21 | 05 | 05 | 05 | F |
| G | 09 | 18 | 27 | 38 | 01 | 14 | 90 | 06 | 05 | 22 | 33 | 16 | 14 | 13 | 82 | 52 | 23 | 21 | 05 | 03 | 15 | 14 | 32 | 21 | 23 | 39 | 15 | 14 | 05 | 10 | 04 | 10 | 17 | 23 | 20 | 11 | G |
| H | 03 | 45 | 23 | 25 | 09 | 32 | 08 | 87 | 10 | 10 | 09 | 29 | 05 | 08 | 08 | 14 | 08 | 17 | 37 | 04 | 36 | 59 | 09 | 33 | 14 | 11 | 03 | 09 | 15 | 43 | 70 | 35 | 17 | 04 | 03 | 03 | H |
| I | 64 | 07 | 07 | 13 | 10 | 08 | 06 | 12 | 93 | 03 | 05 | 16 | 13 | 30 | 07 | 03 | 05 | 19 | 35 | 16 | 10 | 05 | 08 | 02 | 05 | 07 | 02 | 05 | 08 | 09 | 06 | 08 | 05 | 02 | 04 | 05 | I |
| J | 07 | 09 | 38 | 09 | 02 | 24 | 18 | 05 | 04 | 85 | 22 | 31 | 08 | 03 | 21 | 63 | 47 | 11 | 02 | 07 | 09 | 09 | 09 | 22 | 32 | 28 | 67 | 66 | 33 | 15 | 07 | 11 | 28 | 29 | 26 | 23 | J |
| K | 05 | 24 | 38 | 73 | 01 | 17 | 25 | 11 | 05 | 27 | 91 | 33 | 10 | 12 | 31 | 14 | 31 | 22 | 02 | 02 | 23 | 17 | 33 | 63 | 16 | 18 | 05 | 09 | 17 | 08 | 08 | 18 | 14 | 13 | 05 | 06 | K |
| L | 02 | 69 | 43 | 45 | 10 | 24 | 12 | 26 | 09 | 30 | 27 | 86 | 06 | 02 | 09 | 37 | 36 | 28 | 12 | 05 | 16 | 19 | 20 | 31 | 25 | 59 | 12 | 13 | 17 | 15 | 26 | 29 | 36 | 16 | 07 | 03 | L |
| M | 24 | 12 | 05 | 14 | 07 | 17 | 29 | 08 | 08 | 11 | 23 | 08 | 96 | 62 | 11 | 19 | 15 | 20 | 07 | 09 | 13 | 04 | 21 | 09 | 18 | 08 | 05 | 07 | 06 | 06 | 05 | 07 | 11 | 07 | 10 | 04 | M |
| N | 31 | 04 | 13 | 30 | 08 | 12 | 10 | 16 | 13 | 03 | 16 | 08 | 59 | 93 | 05 | 09 | 05 | 28 | 12 | 10 | 16 | 04 | 12 | 04 | 06 | 11 | 05 | 02 | 03 | 04 | 04 | 06 | 02 | 02 | 10 | 02 | N |
| O | 07 | 07 | 20 | 06 | 05 | 09 | 76 | 07 | 02 | 39 | 26 | 10 | 04 | 08 | 86 | 37 | 35 | 10 | 03 | 04 | 11 | 14 | 25 | 35 | 27 | 27 | 19 | 17 | 07 | 07 | 06 | 18 | 14 | 11 | 20 | 12 | O |
| P | 05 | 22 | 33 | 12 | 05 | 36 | 22 | 12 | 03 | 78 | 14 | 46 | 05 | 06 | 21 | 83 | 43 | 23 | 09 | 04 | 12 | 19 | 19 | 19 | 41 | 30 | 34 | 44 | 24 | 11 | 15 | 17 | 24 | 23 | 25 | 13 | P |
| Q | 08 | 20 | 38 | 11 | 04 | 15 | 10 | 05 | 02 | 27 | 23 | 26 | 07 | 06 | 22 | 51 | 91 | 11 | 02 | 03 | 06 | 14 | 12 | 37 | 50 | 63 | 34 | 32 | 17 | 12 | 09 | 27 | 40 | 58 | 37 | 24 | Q |
| R | 13 | 14 | 16 | 23 | 05 | 34 | 26 | 15 | 07 | 12 | 21 | 37 | 14 | 12 | 12 | 29 | 08 | 87 | 16 | 02 | 23 | 23 | 62 | 14 | 12 | 13 | 07 | 10 | 13 | 04 | 07 | 12 | 07 | 09 | 01 | 02 | R |
| S | 17 | 24 | 05 | 30 | 11 | 26 | 05 | 59 | 16 | 03 | 13 | 10 | 05 | 17 | 06 | 06 | 03 | 18 | 96 | 09 | 56 | 24 | 12 | 10 | 06 | 07 | 08 | 02 | 02 | 15 | 28 | 09 | 05 | 05 | 05 | 02 | S |
| T | 13 | 10 | 01 | 05 | 46 | 03 | 06 | 06 | 14 | 06 | 14 | 07 | 06 | 05 | 06 | 11 | 04 | 04 | 07 | 96 | 08 | 05 | 04 | 02 | 02 | 06 | 05 | 05 | 03 | 03 | 03 | 08 | 07 | 06 | 14 | 06 | T |
| U | 14 | 29 | 12 | 32 | 04 | 32 | 11 | 34 | 21 | 07 | 44 | 32 | 11 | 13 | 06 | 20 | 12 | 40 | 51 | 06 | 93 | 57 | 34 | 17 | 09 | 11 | 06 | 06 | 16 | 34 | 10 | 09 | 09 | 07 | 04 | 03 | U |
| V | 05 | 17 | 24 | 16 | 09 | 29 | 06 | 39 | 05 | 11 | 26 | 43 | 04 | 01 | 09 | 17 | 10 | 17 | 11 | 06 | 32 | 92 | 17 | 57 | 35 | 10 | 10 | 14 | 28 | 79 | 44 | 36 | 25 | 10 | 01 | 05 | V |
| W | 09 | 21 | 30 | 22 | 09 | 36 | 25 | 15 | 04 | 25 | 29 | 18 | 15 | 06 | 26 | 20 | 25 | 61 | 12 | 04 | 19 | 20 | 86 | 22 | 25 | 22 | 10 | 22 | 19 | 16 | 05 | 09 | 11 | 06 | 03 | 07 | W |
| X | 07 | 64 | 45 | 19 | 03 | 28 | 11 | 06 | 01 | 35 | 50 | 42 | 10 | 08 | 24 | 32 | 61 | 10 | 12 | 03 | 12 | 17 | 21 | 91 | 48 | 26 | 12 | 20 | 24 | 27 | 16 | 57 | 29 | 16 | 17 | 06 | X |
| Y | 09 | 23 | 62 | 15 | 04 | 26 | 22 | 09 | 01 | 30 | 12 | 14 | 05 | 06 | 14 | 30 | 52 | 05 | 07 | 04 | 06 | 13 | 21 | 44 | 86 | 23 | 26 | 44 | 40 | 15 | 11 | 26 | 22 | 33 | 23 | 16 | Y |
| Z | 03 | 46 | 45 | 18 | 02 | 22 | 17 | 10 | 07 | 23 | 21 | 51 | 11 | 02 | 15 | 59 | 72 | 14 | 04 | 03 | 09 | 11 | 12 | 36 | 42 | 87 | 16 | 21 | 27 | 09 | 10 | 25 | 66 | 47 | 15 | 15 | Z |
| 1 | 02 | 05 | 10 | 03 | 03 | 05 | 13 | 04 | 02 | 29 | 05 | 14 | 09 | 07 | 14 | 30 | 28 | 09 | 04 | 02 | 03 | 12 | 14 | 17 | 19 | 22 | 84 | 63 | 13 | 08 | 10 | 08 | 19 | 32 | 57 | 55 | 1 |
| 2 | 07 | 14 | 22 | 05 | 04 | 20 | 13 | 03 | 25 | 26 | 09 | 14 | 02 | 03 | 17 | 37 | 28 | 06 | 05 | 03 | 06 | 10 | 11 | 17 | 30 | 13 | 62 | 89 | 54 | 20 | 05 | 14 | 20 | 21 | 16 | 11 | 2 |
| 3 | 03 | 08 | 21 | 05 | 04 | 32 | 06 | 12 | 02 | 23 | 06 | 13 | 05 | 02 | 05 | 37 | 19 | 09 | 07 | 06 | 04 | 16 | 06 | 22 | 25 | 12 | 18 | 64 | 86 | 31 | 23 | 41 | 16 | 17 | 08 | 10 | 3 |
| 4 | 06 | 19 | 19 | 12 | 06 | 25 | 14 | 16 | 07 | 21 | 13 | 19 | 03 | 03 | 02 | 17 | 29 | 11 | 09 | 03 | 17 | 55 | 08 | 37 | 24 | 03 | 05 | 26 | 44 | 89 | 42 | 44 | 32 | 10 | 03 | 03 | 4 |
| 5 | 08 | 45 | 15 | 14 | 02 | 45 | 04 | 67 | 07 | 14 | 04 | 41 | 02 | 00 | 04 | 13 | 07 | 09 | 27 | 02 | 14 | 45 | 07 | 45 | 10 | 10 | 14 | 10 | 30 | 69 | 90 | 42 | 24 | 10 | 06 | 05 | 5 |
| 6 | 07 | 80 | 30 | 17 | 04 | 23 | 04 | 14 | 02 | 11 | 11 | 27 | 06 | 02 | 07 | 16 | 30 | 11 | 14 | 03 | 12 | 30 | 09 | 58 | 38 | 39 | 15 | 14 | 26 | 24 | 17 | 86 | 69 | 14 | 05 | 14 | 6 |
| 7 | 06 | 33 | 22 | 14 | 05 | 25 | 06 | 04 | 06 | 24 | 13 | 32 | 07 | 06 | 07 | 36 | 39 | 12 | 06 | 02 | 03 | 13 | 09 | 30 | 30 | 50 | 22 | 29 | 18 | 15 | 12 | 61 | 85 | 70 | 20 | 13 | 7 |
| 8 | 03 | 23 | 40 | 06 | 03 | 15 | 15 | 06 | 02 | 33 | 10 | 14 | 03 | 06 | 14 | 12 | 45 | 02 | 06 | 04 | 06 | 07 | 05 | 24 | 35 | 50 | 42 | 29 | 16 | 16 | 09 | 30 | 60 | 89 | 61 | 26 | 8 |
| 9 | 03 | 14 | 23 | 03 | 01 | 06 | 14 | 05 | 02 | 30 | 06 | 07 | 16 | 11 | 10 | 31 | 32 | 05 | 06 | 07 | 06 | 03 | 08 | 11 | 21 | 24 | 57 | 39 | 09 | 12 | 04 | 11 | 42 | 56 | 91 | 78 | 9 |
| 0 | 09 | 03 | 11 | 02 | 05 | 07 | 14 | 04 | 05 | 30 | 08 | 03 | 02 | 03 | 25 | 21 | 29 | 02 | 03 | 04 | 05 | 03 | 02 | 12 | 15 | 20 | 50 | 26 | 09 | 11 | 05 | 22 | 17 | 52 | 81 | 94 | 0 |
| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | |

# Motivation: do something...

- Given metric data
- Visualize it "nicely"

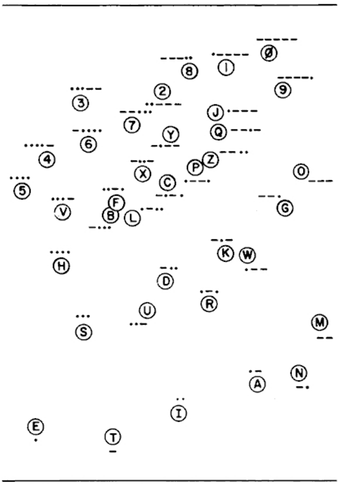
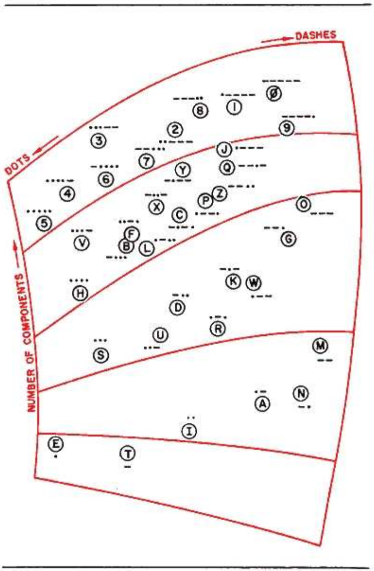



Figure 3A: Configuration Resulting from Morse Code Similarities

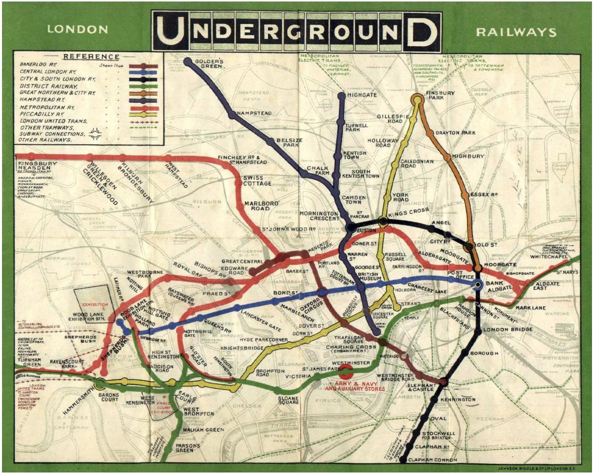
# Motivation: do something...

- Given metric data
- Visualize it "nicely"
- What is "nicely"?



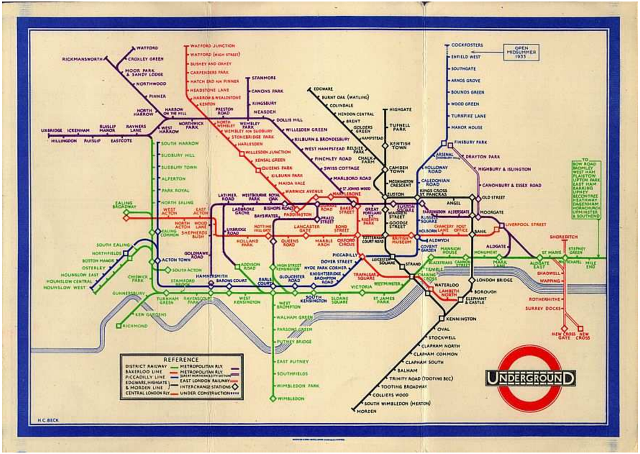
# Motivation: ... but not too much

A little help can go a long way!



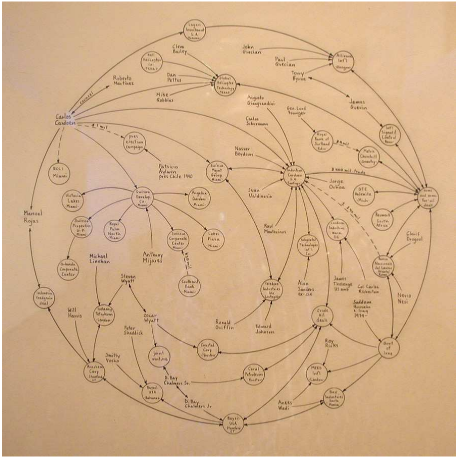
# Motivation: ... but not too much

A little help can go a long way!



# Graphs

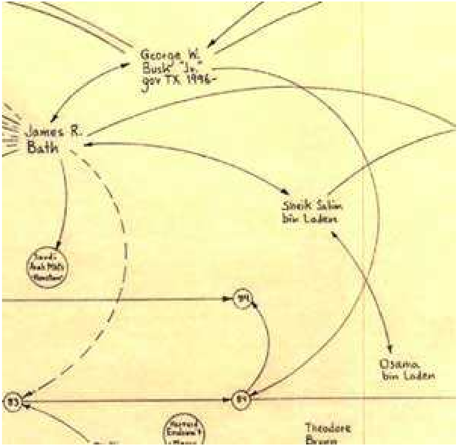
- Modeling with graphs



**Art by Mark Lombardi*

# Graphs

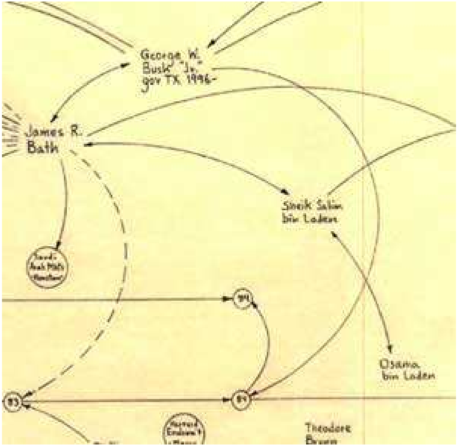
- Modeling with graphs
- Relational data as *graphs*
  - objects → vertices
  - relationships → edges




**Art by Mark Lombardi*

# Graphs

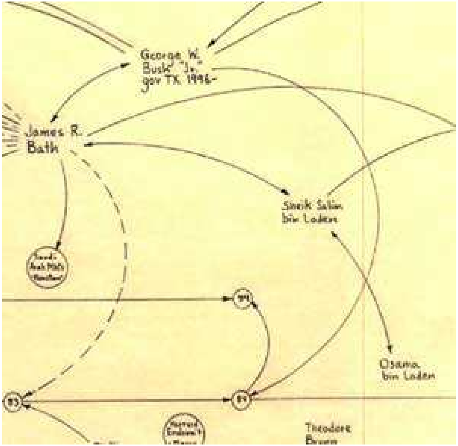
- Modeling with graphs
- Relational data as *graphs*
    - objects → vertices
    - relationships → edges
- Metric data as *graphs*
    - distance b/n points
    - points → vertices
    - distances → weighted edges




**Art by Mark Lombardi*

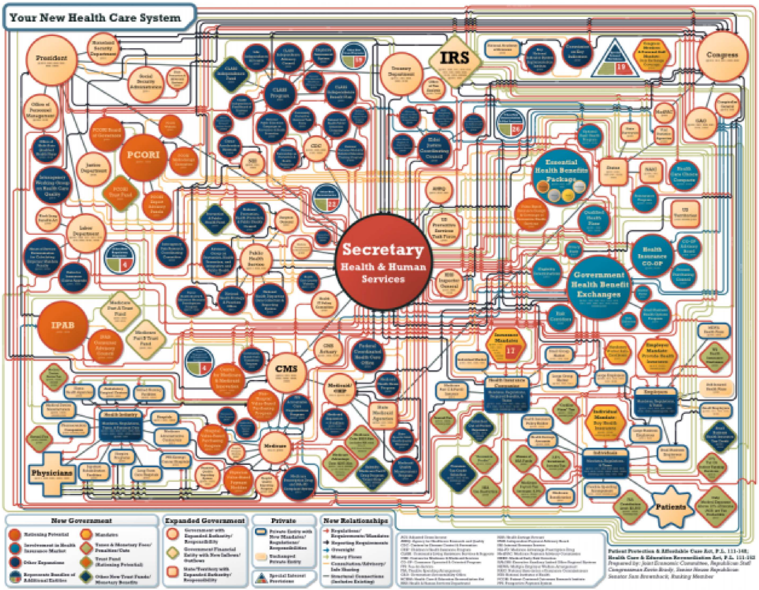
# Graphs

- Modeling with graphs
- Relational data as *graphs*
  - objects $\rightarrow$ vertices
  - relationships $\rightarrow$ edges
- Metric data as *graphs*
  - distance b/n points
  - points $\rightarrow$ vertices
  - distances $\rightarrow$ weighted edges
- Graph drawing
  - can produce great visualization
  - not so good at clusters, neighborhoods

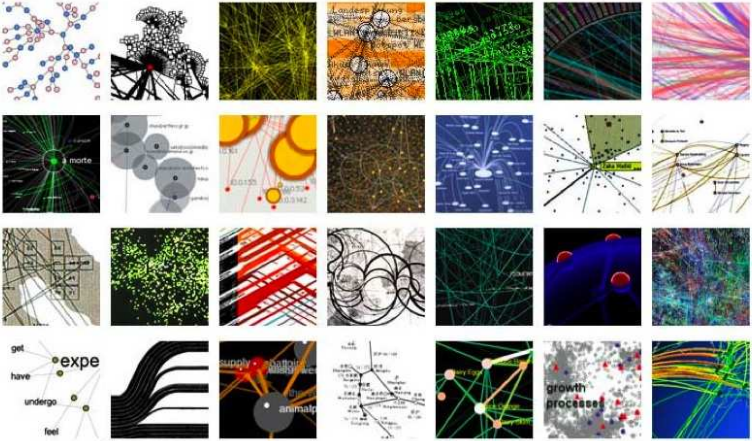



**Art by Mark Lombardi*

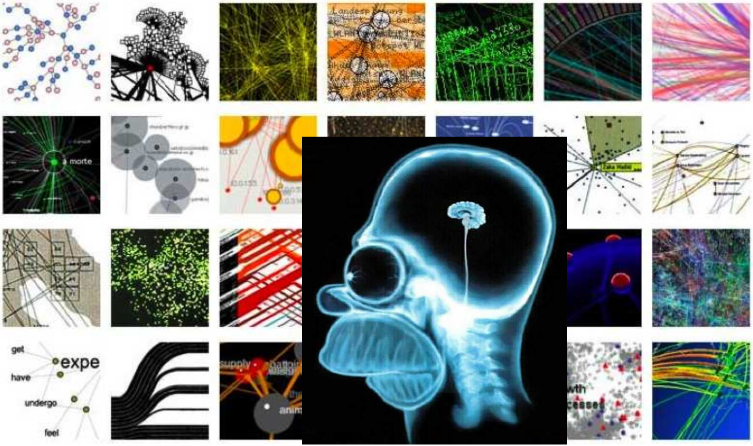
# Convince or Confuse?



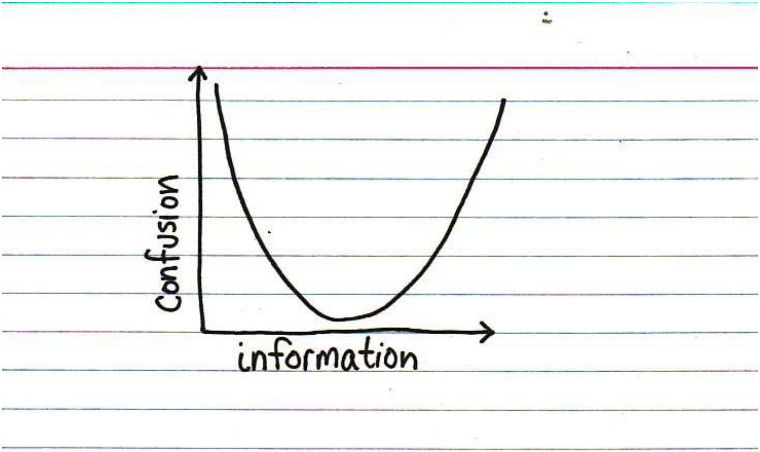
# All Vis, All the Time?



# All Vis, All the Time?

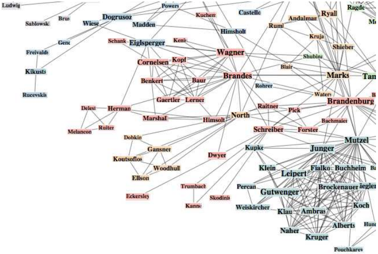


# Less is More



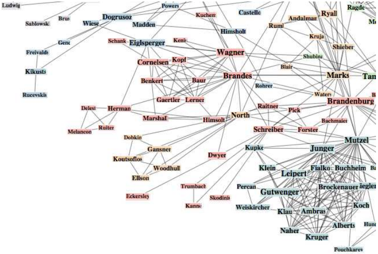
# Why Maps?

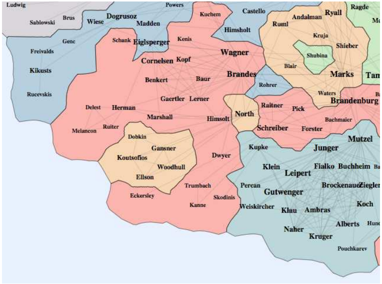
- Dimensionality reduction: $\mathcal{R}^n \rightarrow \mathcal{R}^2$



# Why Maps?

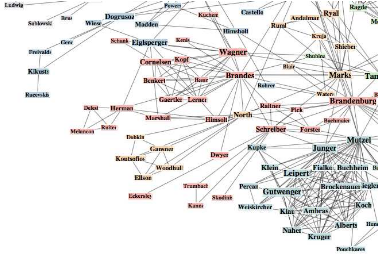
- Dimensionality reduction: $\mathcal{R}^n \rightarrow \mathcal{R}^2$
- Natural extension from graphs

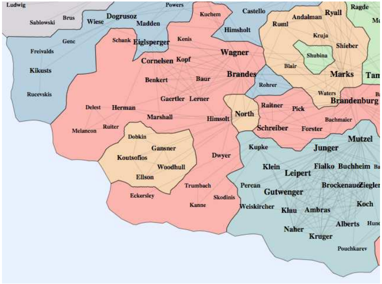


# Why Maps?

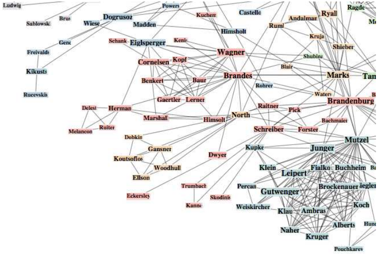
- Dimensionality reduction: $\mathcal{R}^n \to \mathcal{R}^2$
- Natural extension from graphs
- Explicit clustering: regions and colors

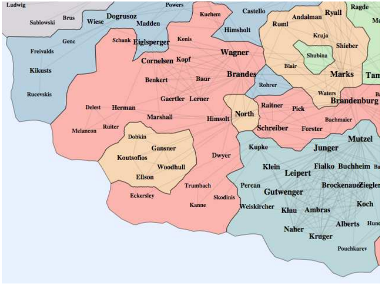


# Why Maps?

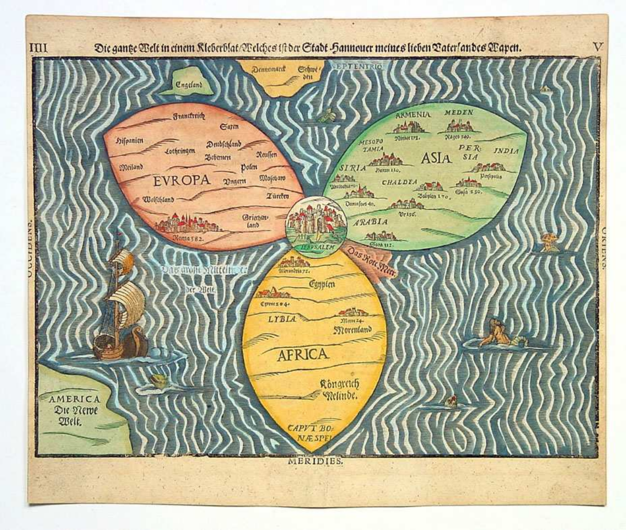
- Dimensionality reduction: $\mathcal{R}^n \to \mathcal{R}^2$
- Natural extension from graphs
- Explicit clustering: regions and colors
- Intuitive and familiar



# Maps, Maps, and More Maps

- The World circa 1500



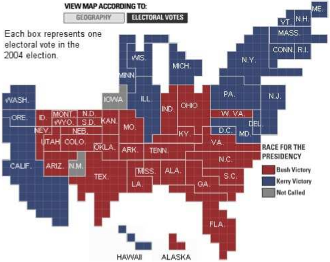
# Maps, Maps, and More Maps

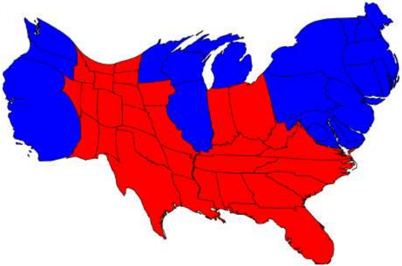
- The World circa 1500
- Europe circa 1500

# Maps, Maps, and More Maps

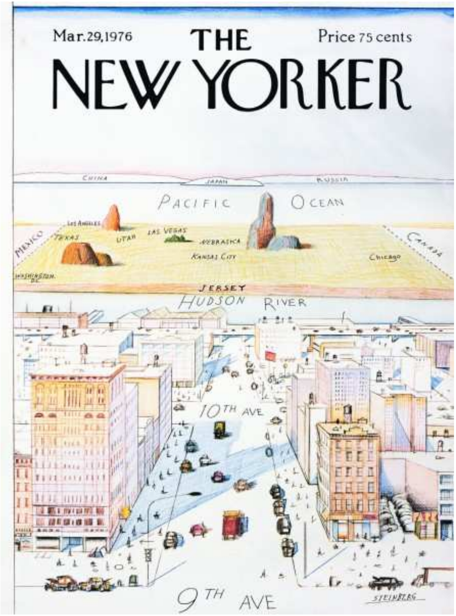
- Redrawing geographic maps subject to constraints

- Redrawing geographic maps subject to constraints
- Cartograms: 2004 US election results

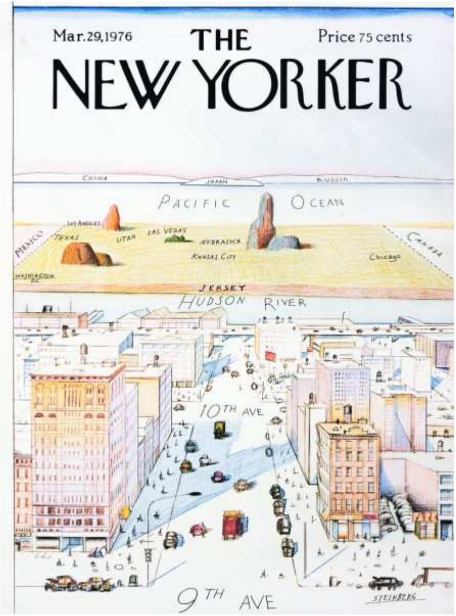



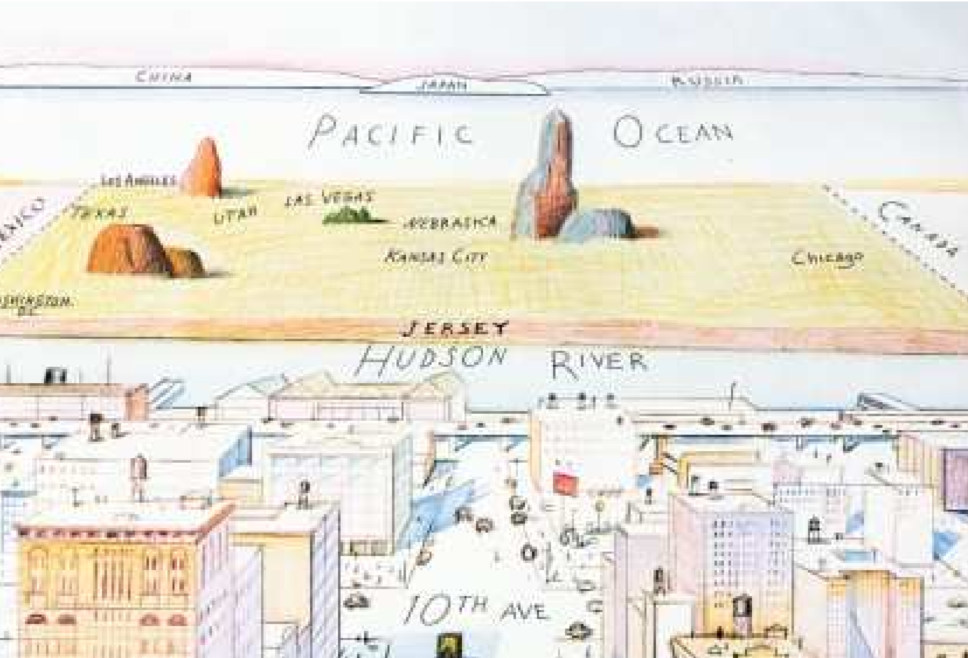
# Maps: New Yorker '76



# Maps: New Yorker '76



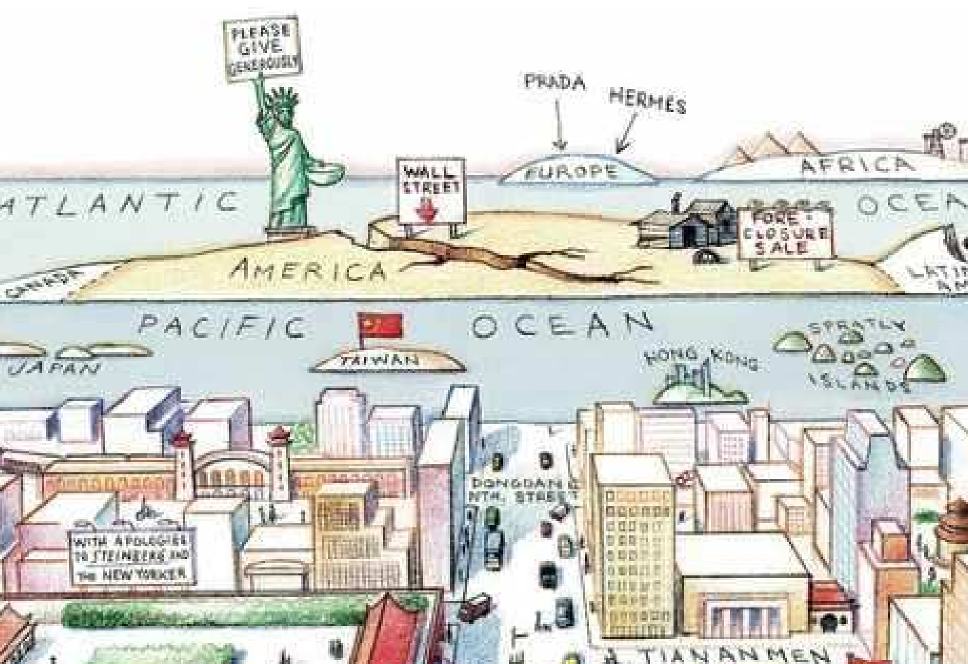
# Maps: New Yorker '76



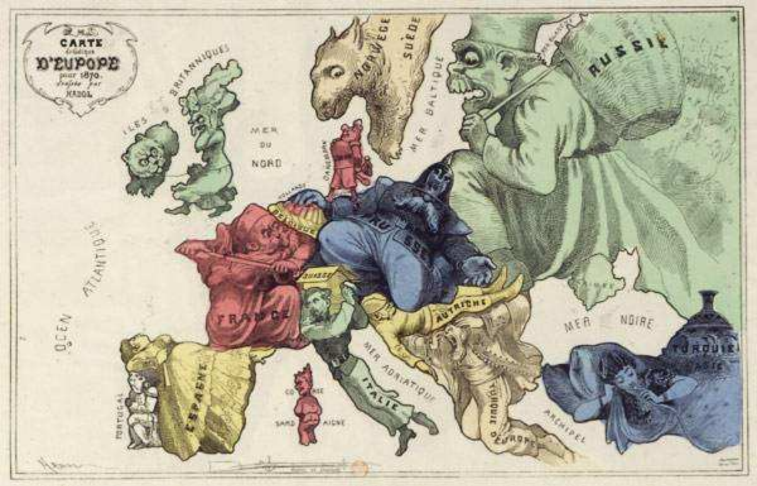
# Maps: Economist '09


The Economist
MARCH 21ST–27TH 2009
Economist.com
Shameless greed at AIG
One more try on Iran
What's wrong with General Electric
Let Michelle be Michelle
Remembering New Labour
How China sees the world
PLEASE GIVE GENEROUSLY
PRADA
HERMES
WALL STREET
EUROPE
AFRICA
ATLANTIC
OCEAN
AMERICA
FORE-CLOSURE SALE
LATIN AMERICA
CANADA
PACIFIC
OCEAN
JAPAN
TAIWAN
HONG KONG
SPRATLY ISLANDS
DONGDAN NTH. STREET
WITH APOLOGIES TO STEINBERG AND THE NEW YORKER
TIANANMEN SQUARE
IMPERIAL PALACE
CHANG'AN STREET
£4.00

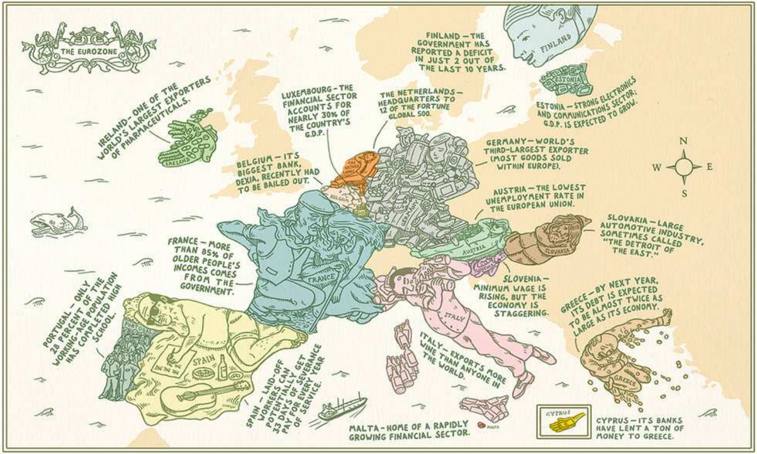
# Maps: Economist '09



- Whimsical maps: Europe in 1870

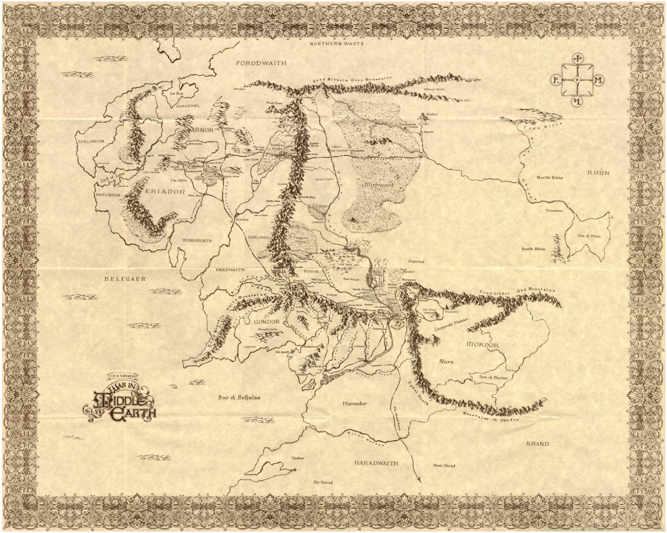


# Maps: Europe in 2011

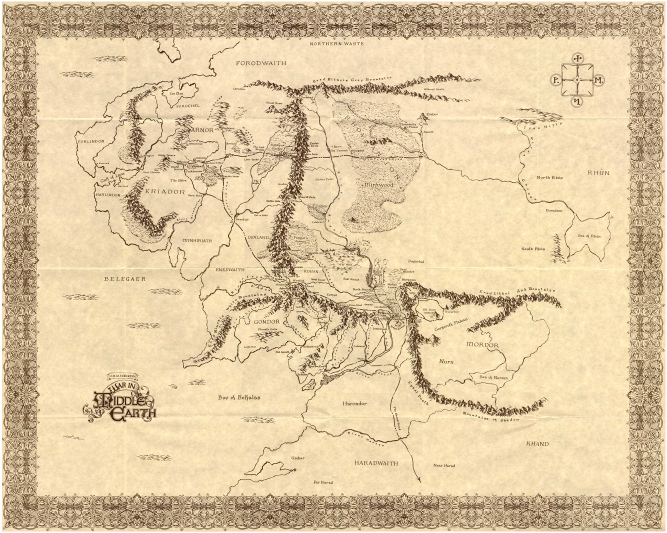



BANK
Í ALVÖRU.
ÉG ER BARA AÐ
GANGA Í GEGNUM
TÍMABIL.
NOR
FINNL
UK
DAN
ÞÝSK
PÓL
FRAKKL
SPÁNN
UNG
RÚM
ÍTALÍA
GRIKKL

# Maps: Tolkien's 1930's Middle Earth

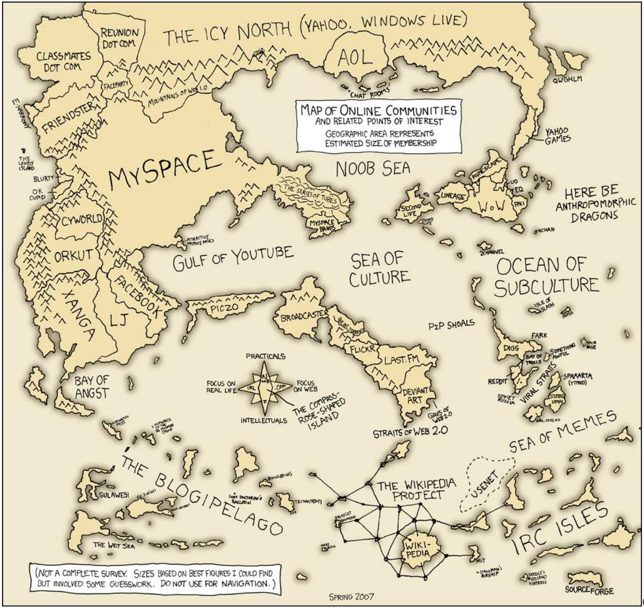


# Maps: Tolkien's 1930's Middle Earth



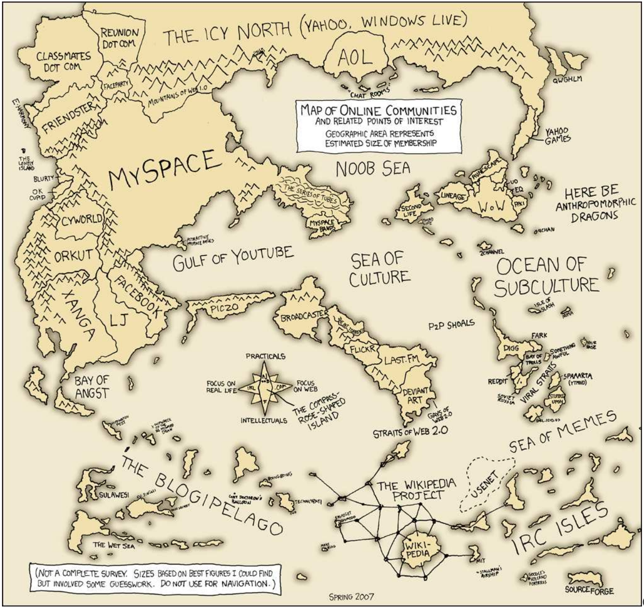
# Maps: Tolkien's 1930's Middle Earth

# Maps: xkcd online communities 2007

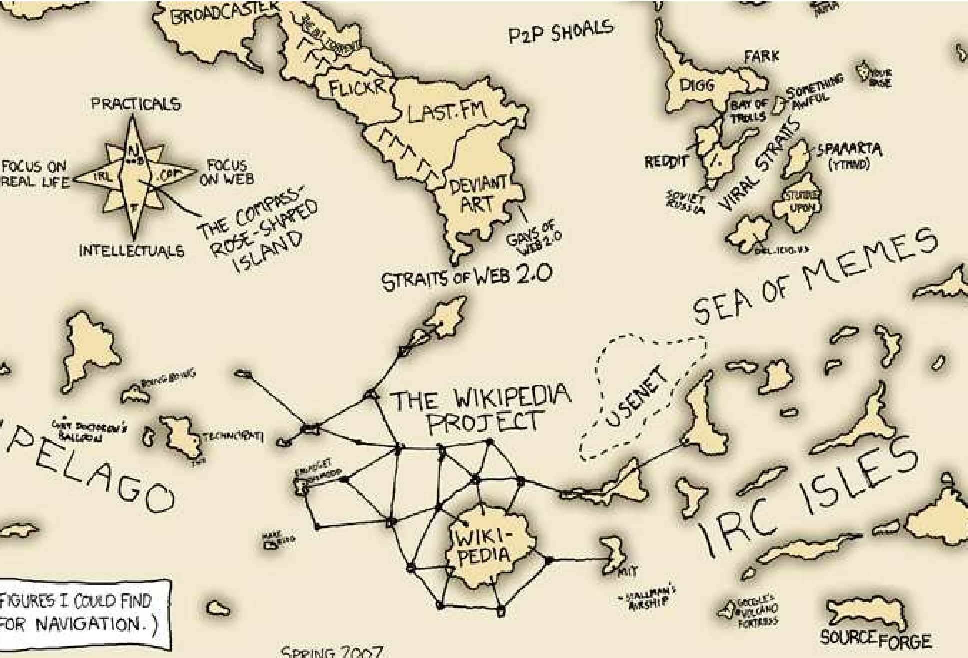



©xkcd

# Maps: xkcd online communities 2007



# Maps: xkcd online communities 2007



# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?

# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?

# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?
  - What type of a conference is SODA?

# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?
  - What type of a conference is SODA?


# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?
  - What type of a conference is SODA?
  - Which departments have strengths in Computer Vision?
  - How does a journal evolve over time?


# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?
  - What type of a conference is SODA?
  - Which departments have strengths in Computer Vision?
  - How does a journal evolve over time?
- With a map, of course!
  - Where do we get the data?
  - What exactly is on the map?
  - How do we make the map?


# Maps of Computer Science: Motivation

- How do I answer such questions?
  - What exactly does Prof. Michael Jordan do at Berkeley?
  - What type of a conference is SODA?
  - Which departments have strengths in Computer Vision?
  - How does a journal evolve over time?
- With a map, of course!
  - Where do we get the data?
  - What exactly is on the map?
  - How do we make the map?
- Map of Computer Science (MoCS)


# MoCS: Dataset

The DBLP bibliography server

# MoCS: Dataset

The DBLP bibliography server (DataBase systems and Logic Programming)

# MoCS: Dataset

The DBLP bibliography server (DataBase systems and Logic Programming)

- covers most CS journals/conf. (about 6,000 different ones)
- over 2.1 million indexed publications
- includes titles and bibliographic information

# MoCS: Dataset

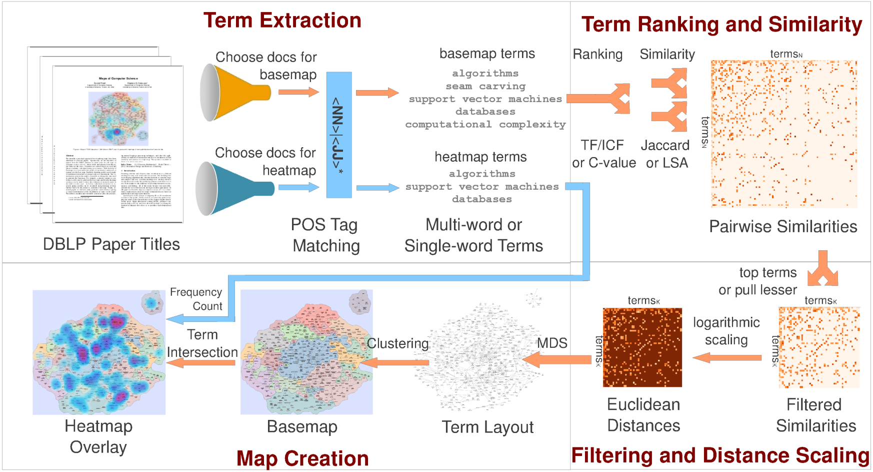
The DBLP bibliography server (DataBase systems and Logic Programming)

- covers most CS journals/conf. (about 6,000 different ones)
- over 2.1 million indexed publications
- includes titles and bibliographic information

Main challenges:

- large dataset (448,374 different words; 2,089,736 phrases)
- short text (titles with 10 words on average)
- graph vertices (terms representing research topics)
- graph edge selection (relations between terms)

# MoCS: Overview



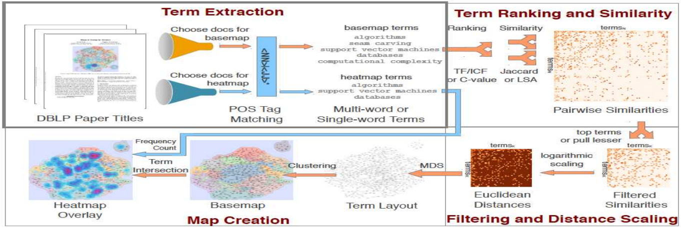
# Term Extraction

Multi-word phrases ("collocations" in NLP)

- Specificity: "wireless sensor networks" are a type of "network"
- Context: "Travelling Salesman Problem", not "Salesman"
- POS tagging and filtering

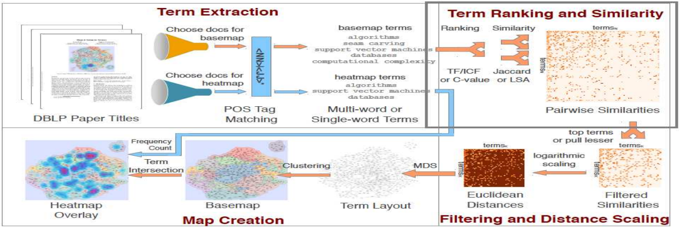
```
    ADJ         NN         NN
travelling  salesman  problem
```

- Extract noun and adjective subsequences
- Extra difficult due to short titles



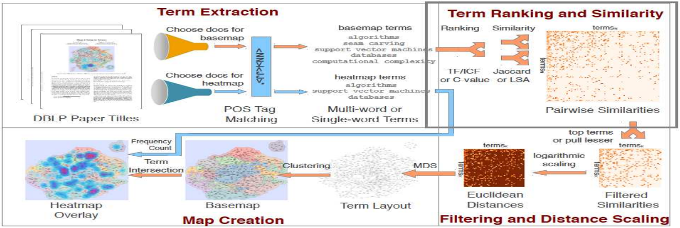
# Term Ranking

- Simplest possible ranking: by frequency
- TF/IDF: *term frequency / inverse document frequency*
- Extra difficult due to short titles (IDF is meaningless)
- TF/ICF: *term frequency / inverse corpus frequency*



# Term Similarity

- Idea: terms are similar if they are used together in titles
- Treat as set similarity: $S_i$ is the set of documents with term $i$
- Jaccard coefficient: $Jacc(S_i, S_j) = \frac{|S_i \cap S_j|}{|S_i \cup S_j|}$
- Extra difficult due to short titles
- Partial match Jaccard: count co-occurrence if terms overlap



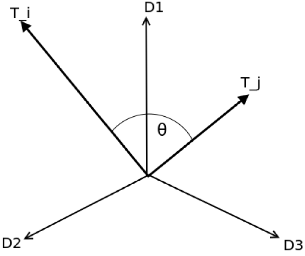
# Term Similarity: LSA

Latent Semantic Analysis (LSA)

- Term-document matrix $A$
- Compare terms as $n$-D vectors
- Cosine distance: compare angles

$$A = \begin{array}{c|cccc} & \mathrm{D}_1 & \mathrm{D}_2 & \cdots & \mathrm{D}_n \\ \hline \mathrm{T}_1 & \mathrm{tf}_{1,1} & \mathrm{tf}_{1,2} & \cdots & \mathrm{tf}_{1,t} \\ \mathrm{T}_2 & \mathrm{tf}_{2,1} & \mathrm{tf}_{2,2} & \cdots & \mathrm{tf}_{2,t} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathrm{T}_t & \mathrm{tf}_{n,1} & \mathrm{tf}_{n,2} & \cdots & \mathrm{tf}_{n,t} \end{array}$$

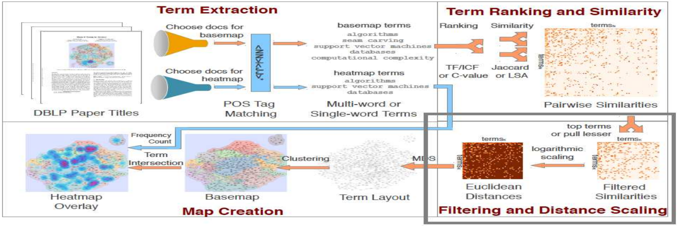
$$Dist(T_i, T_j) = \frac{T_i \cdot T_j}{||T_i||\,||T_j||}$$




- Small angle $\rightarrow$ large cosine: similar
- Large angle $\rightarrow$ small cosine: dissimilar
- Term co-occurrence matrix $AA^\top$

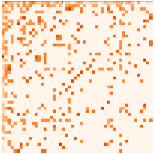
# Filtering and Distance Scaling

- LSA and Jaccard return similarity values between 0 and 1
- GMap takes a distance matrix
- Inverse logarithmic scaling
- Top Terms: only plot $N$ highest-ranked terms
- Pull Lesser Terms: plot $K$ most similar terms for each term $t$

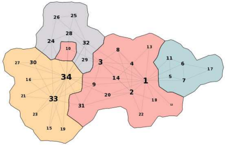


# Making a Map with GMap

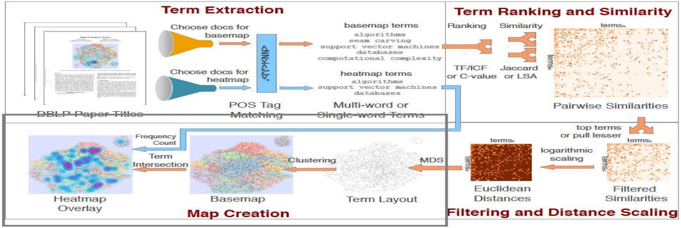
- Input: vertex-weighted, edge-weighted graph $G = (V, E)$
- Output: map, with clusters as countries and vertices as cities



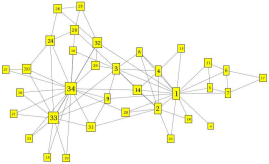
$\Longrightarrow$



- GMap: a framework for embedding + clustering + mapping
  - different algorithms: embedding, clustering, mapping
  - different overlays: journal profile, author profile, paper profile



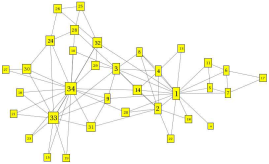
# GMap Overview

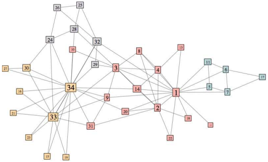


- Embedding
  - PCA, MDS, LLE, IsoMap, ...
  - GMap: *scalable force-directed method*

# GMap Overview

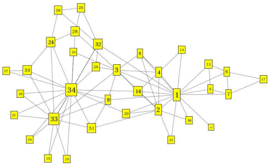
- Embedding
  - PCA, MDS, LLE, IsoMap, ...
  - GMap: *scalable force-directed method*
- Clustering
  - agglomerative, *k*-means, spectral, ...
  - GMap: *modularity clustering*

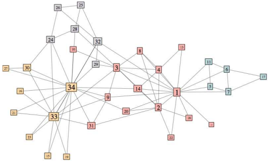


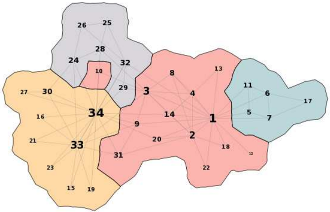


# GMap Overview

- Embedding
  - PCA, MDS, LLE, IsoMap, ...
  - GMap: *scalable force-directed method*
- Clustering
  - agglomerative, $k$-means, spectral, ...
  - GMap: *modularity clustering*
- Mapping
  - not much work on graph$\rightarrow$ map
  - GMap: *modified Voronoi Diagram*

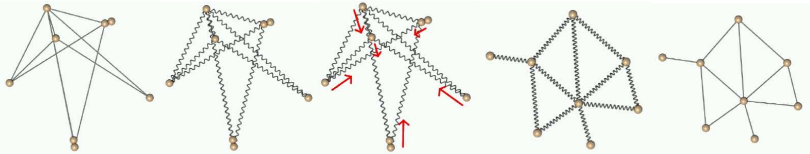




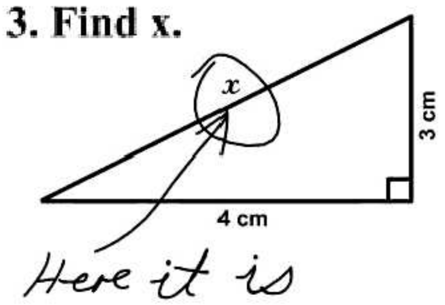


# GMap: Embedding

- Given a graph $G = (V, E)$
  - place vertices as points in $\mathcal{R}^d$
  - route edges in $\mathcal{R}^d$
- Force directed methods define an energy function on layouts
  - based on attractive/repulsive forces (Fruchterman-Reingold)
  - based on graph distances (Kamada-Kawai)
- Energy model
  - iterative improvement
  - minimal energy $\Rightarrow$ good layout



# ... and now for some math...



# GMap: Embedding, cont.

- Fruchterman-Reingold: balance of attraction and repulsion

$$F(v) = F_r(v) + F_a(v)$$

$$F_r(v) = \sum_{\forall u \in V} \frac{\kappa^2}{\|pos[u] - pos[v]\|^2}(pos[u] - pos[v])$$

$$F_a(v) = \sum_{u \in Adj(v)} \frac{\|pos[u] - pos[v]\|^2}{\kappa^2}(pos[u] - pos[v])$$

- $\kappa = \sqrt{A_{frame}/|E|}$, ideal edge length
- Kamada-Kawai: match Euclidean distance to graph distance

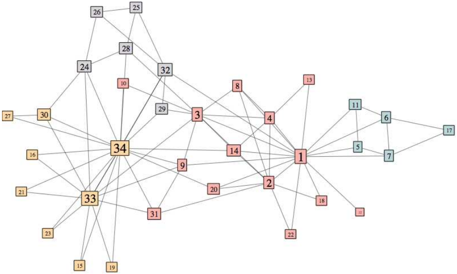
$$F(v) = \sum_{u \in V} \left( \frac{\|pos[u] - pos[v]\|^2}{(\kappa \times dist_G(u, v))^2} - 1 \right) (pos[u] - pos[v])$$

# GMap: Clustering

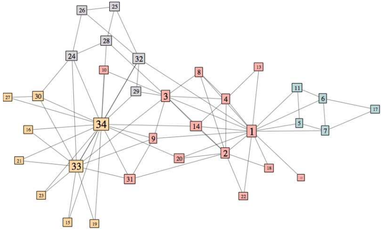
- Modularity
  - measure of the quality of vertex grouping
  - high edge density *within* groups
  - low edge density *between* groups

$$\frac{1}{2|E|}\sum_{\forall(u,v)}\left\{e(u,v)-\frac{deg(u)deg(v)}{2|E|}\right\}\delta(c_u,c_v)$$

- Computing modularity
  - value: $[-1, 1]$
  - opt. modularity is NP-hard
  - spectral heuristics are fast

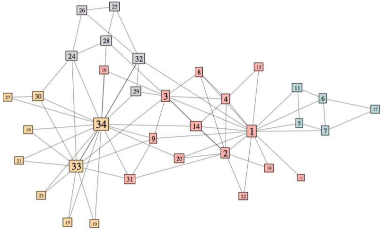


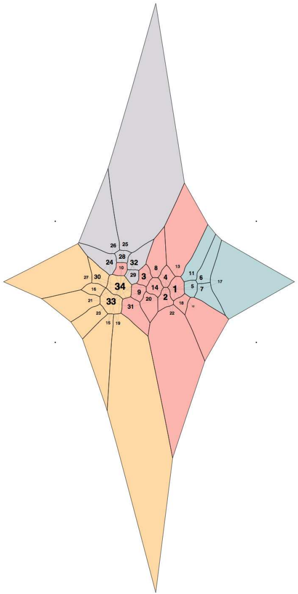
# GMap: Mapping




- Use Voronoi Diagram for borders
- Add bounding box for a finite VD
- Still not map-like:
  - outer boundary is not “form-fitting”
  - inner boundaries are too “jagged”

# GMap: Mapping

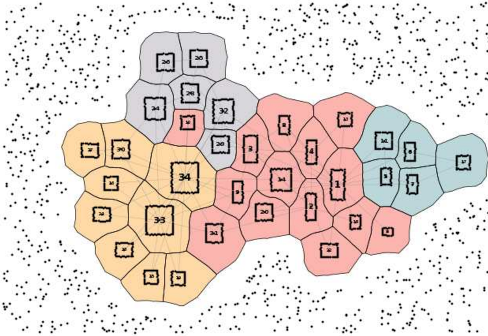


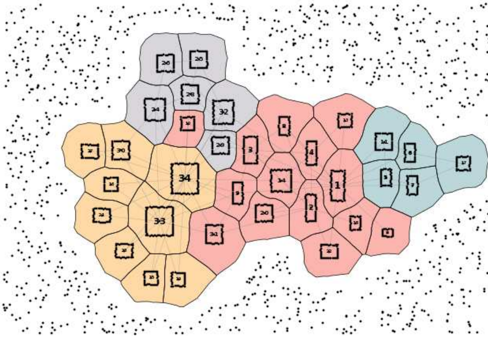

- Use Voronoi Diagram for borders
- Add bounding box for a finite VD
- Still not map-like:
  - outer boundary is not "form-fitting"
  - inner boundaries are too "jagged"

# GMap: Mapping, cont.

- Real and dummy points
  - real: real labeled data points
  - box: proportional regions
  - noise: form-fitting coasts

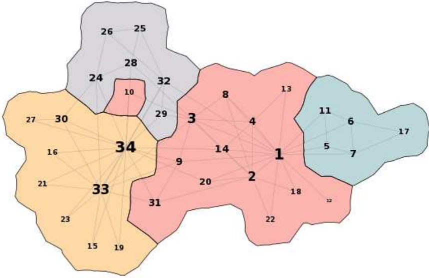


# GMap: Mapping, cont.



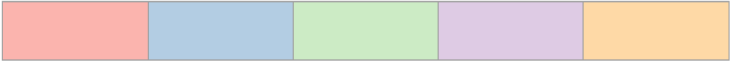
- Real and dummy points
  - real: real labeled data points
  - box: proportional regions
  - noise: form-fitting coasts

- Merge adjacent cells
  - European-style borders
  - natural-looking coastlines

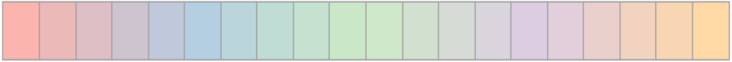


# GMap: Coloring

- Four Color Theorem: any map can be colored with **4** colors
- Our countries may not be contiguous: # countries = # colors
- Color selection: ColorBrewer "map-like" pastel colors



- blend them to get as many colors as needed

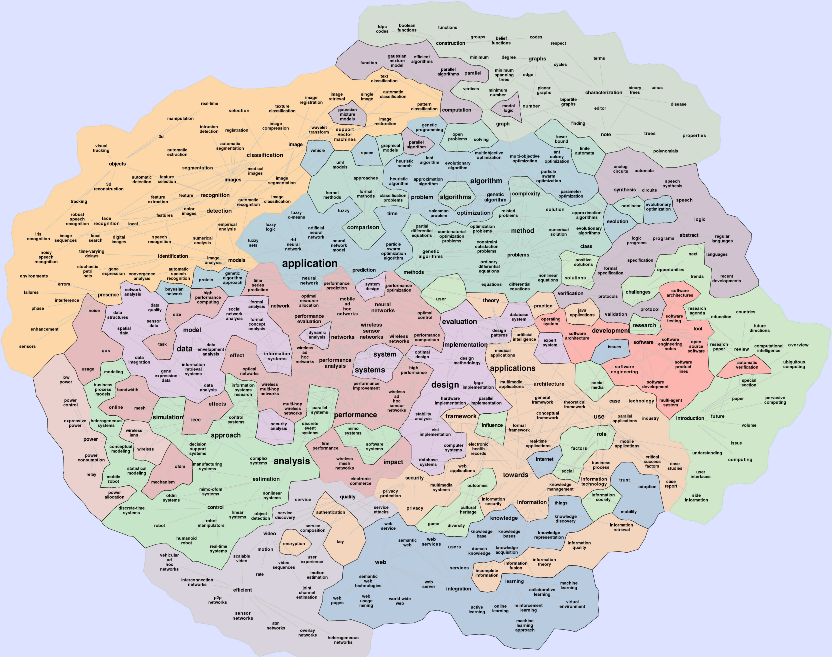


- Country color assignment via "country graph" $G_c = (V_c, E_c)$
  - vertices in $V_c$ are countries
  - edge $(i,j) \in E_c \iff$ countries $i$ and $j$ share a border
  - vertex labeling $c : V \rightarrow [1, |V_c|]$ for max diff. b/n neighbors

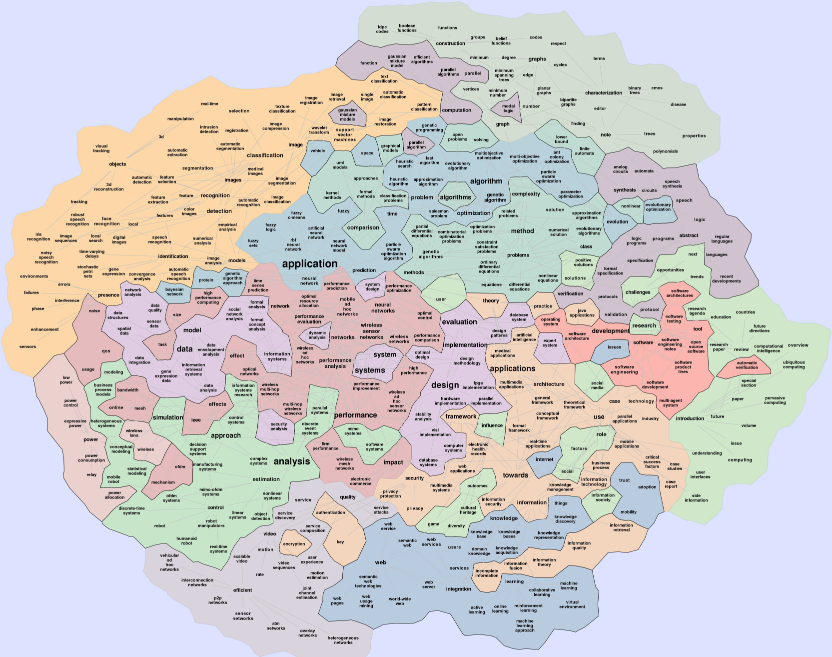
$$\max \min_{(i,j)\in E_c} w_{i,j} \left| c_i - c_j \right|, \quad \{c_1, c_2, \ldots, c_k\} \text{ is a permutation}$$

- NP-hard problem; GMap uses spectral and greedy heuristics

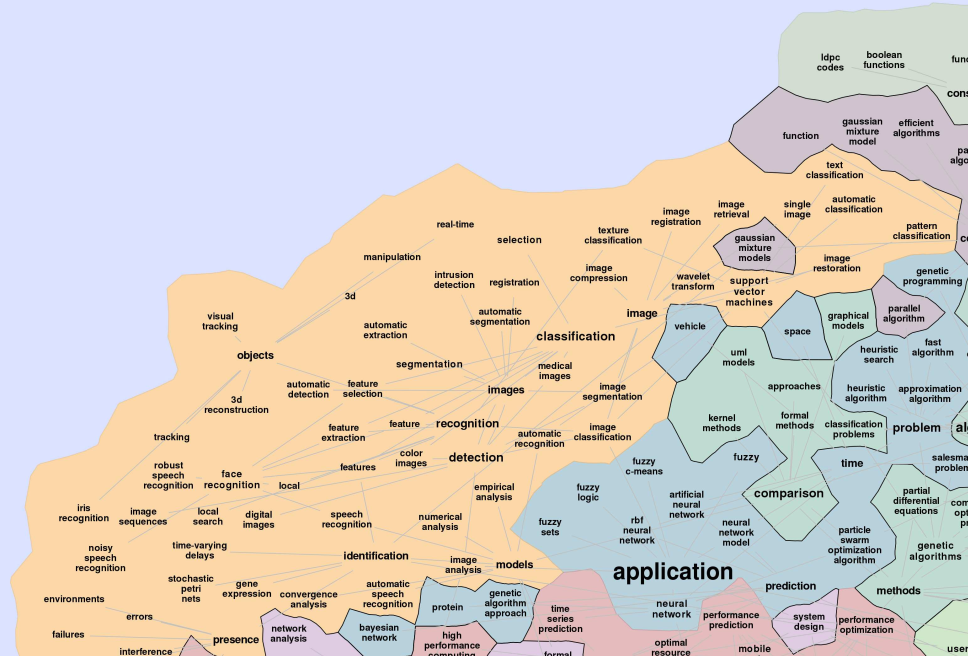
# Base Map of CS



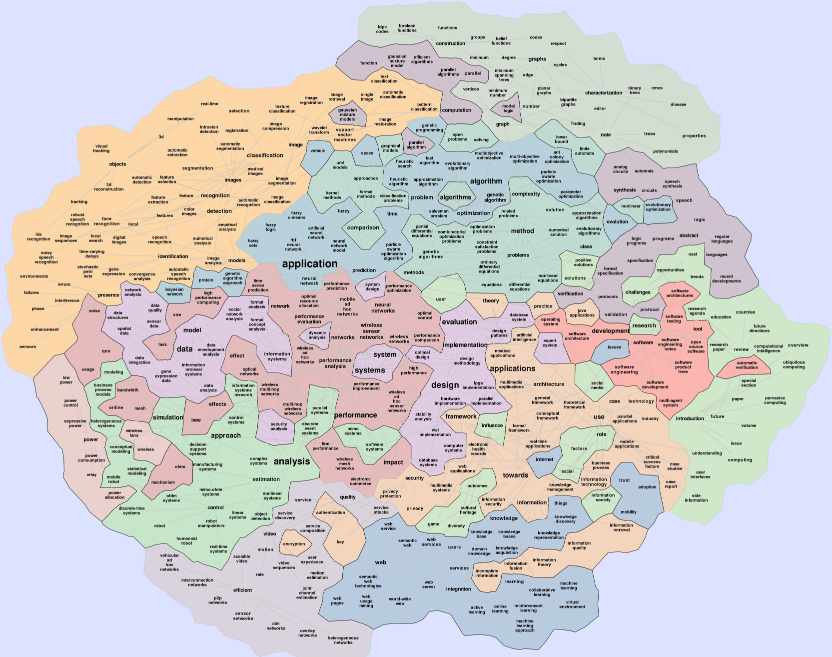
# Base Map of CS



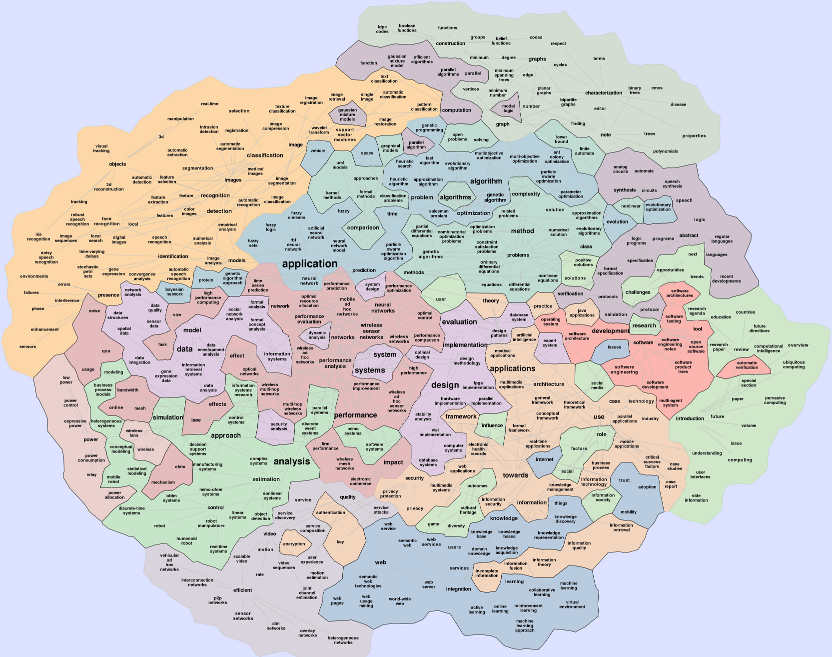
# Base Map of CS



# Base Map of CS

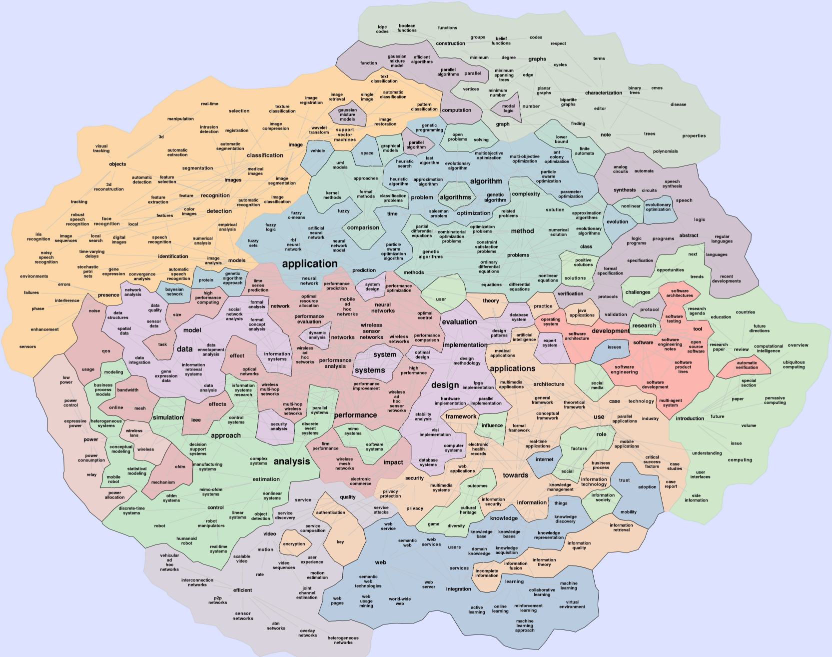


# Base Map of CS

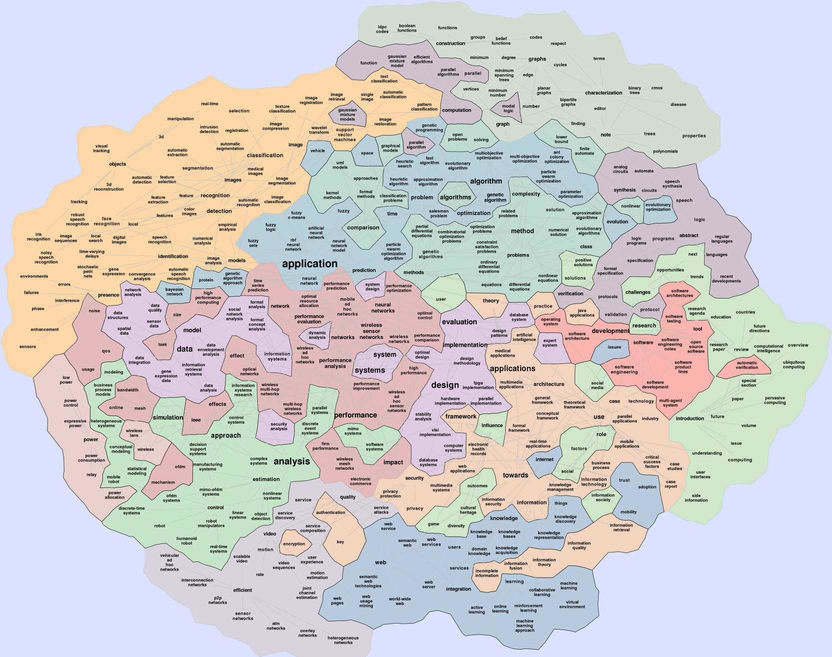


# Base Map of CS

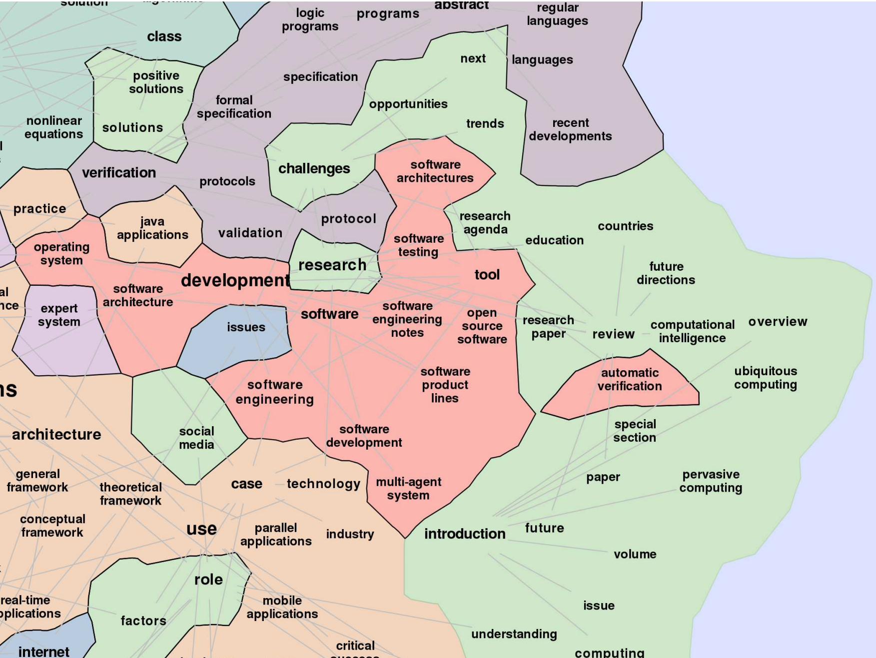
# Base Map of CS



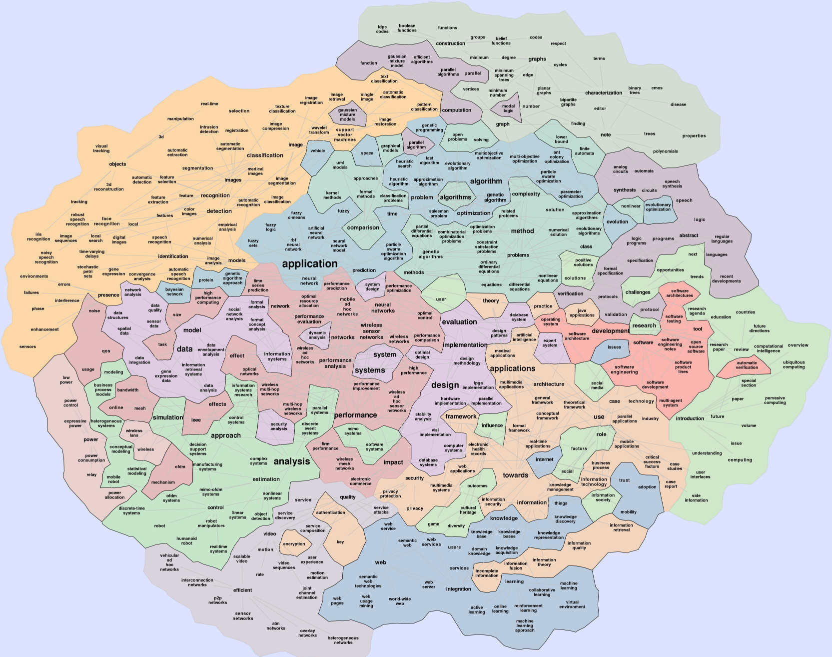
# Base Map of CS



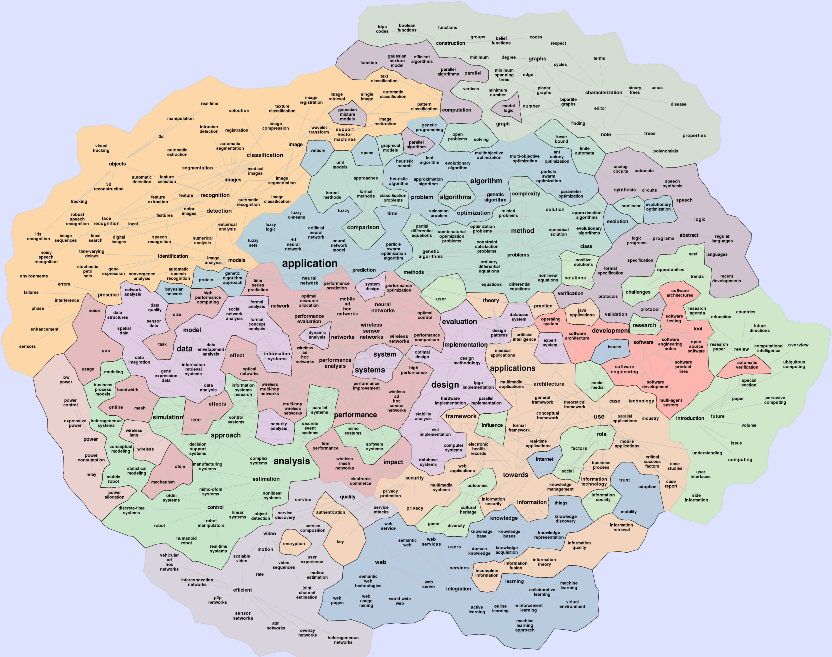
# Base Map of CS



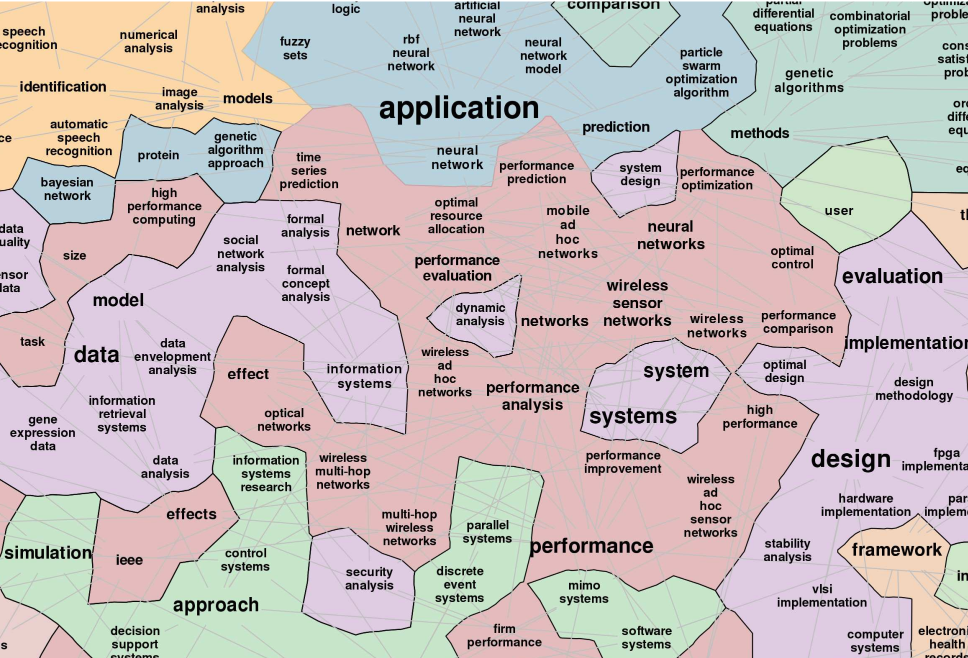
# Base Map of CS



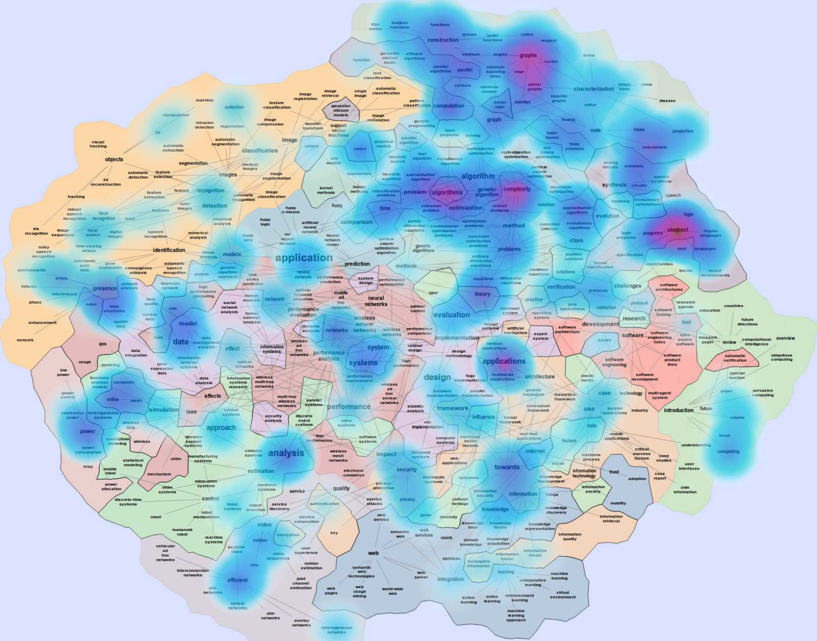
# Base Map of CS



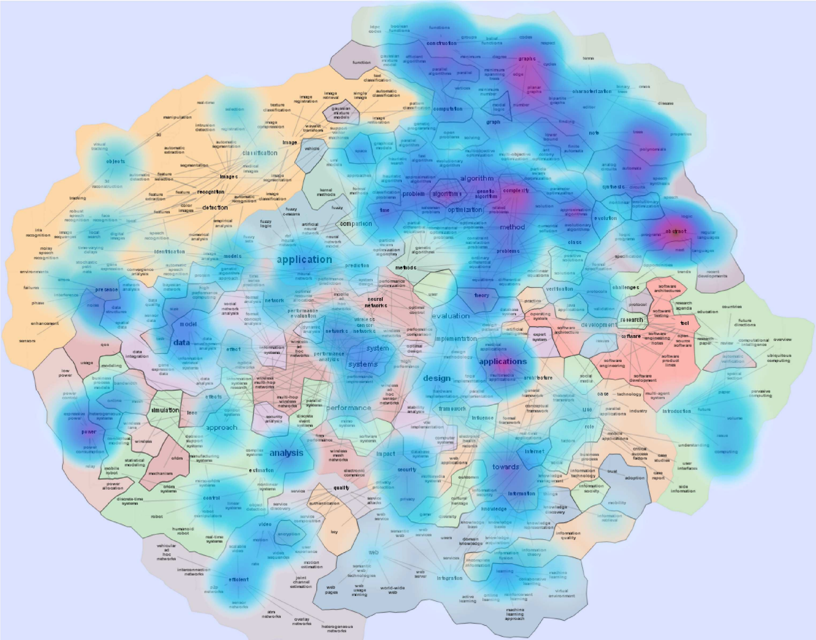
# Base Map of CS



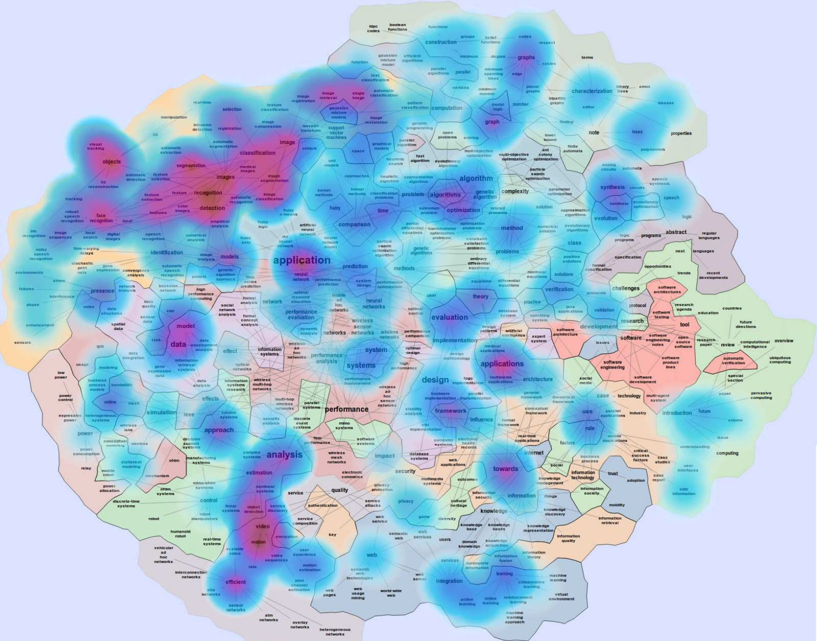
# Conference Heatmaps: STOC



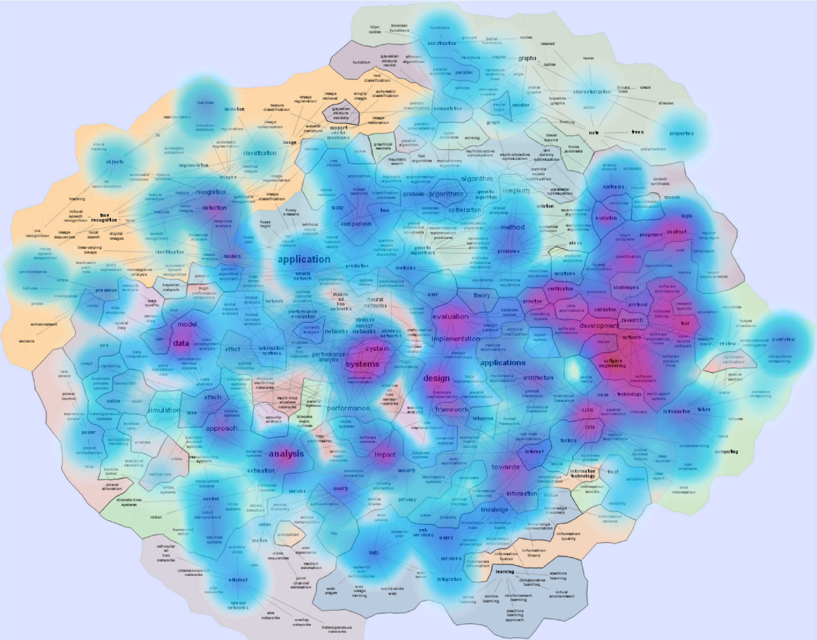
# Conference Heatmaps: FOCS



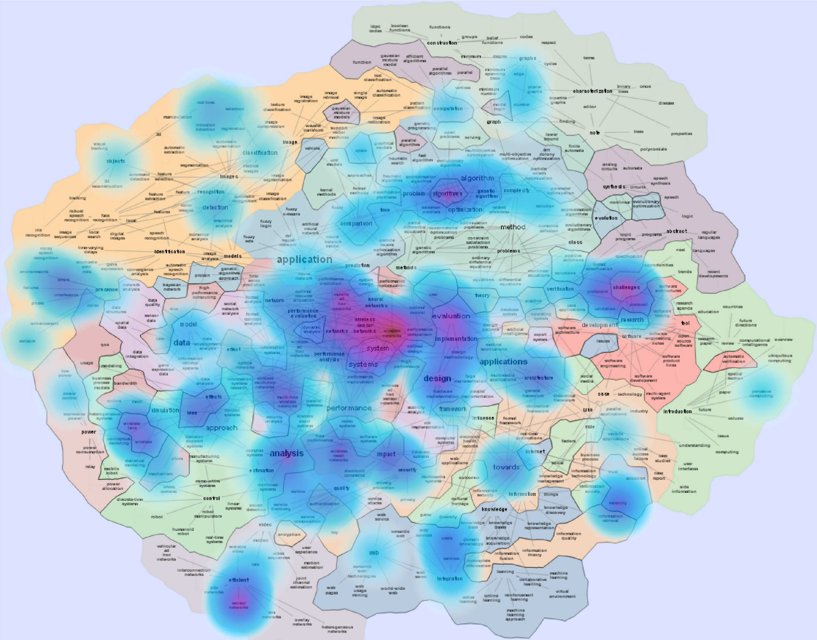
# Conference Heatmaps: CVPR



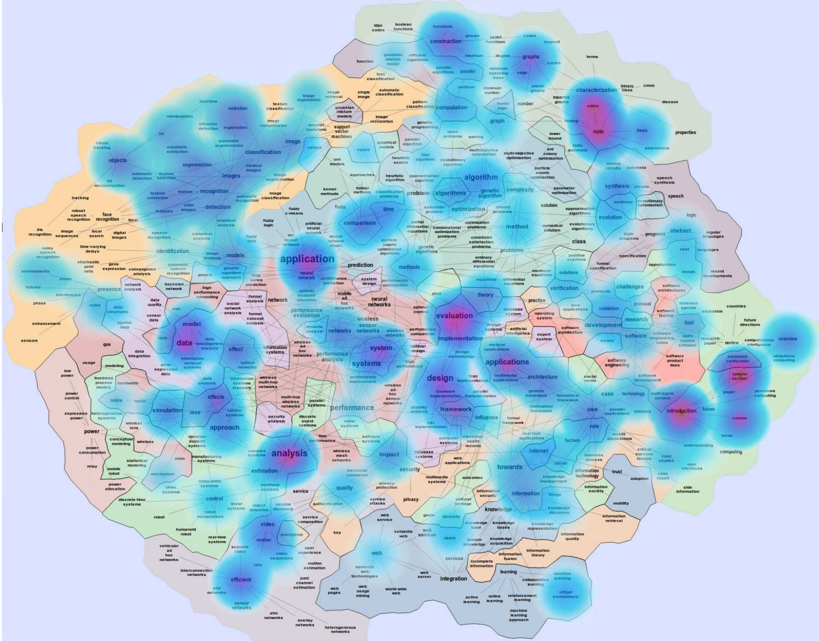
# Conference Heatmaps: ICSE



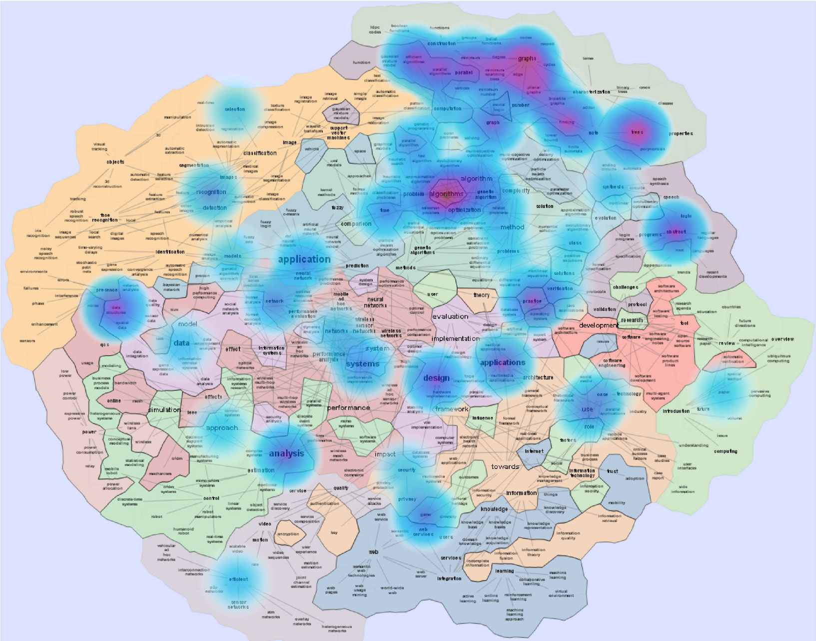
# Conference Heatmaps: MOBICOM



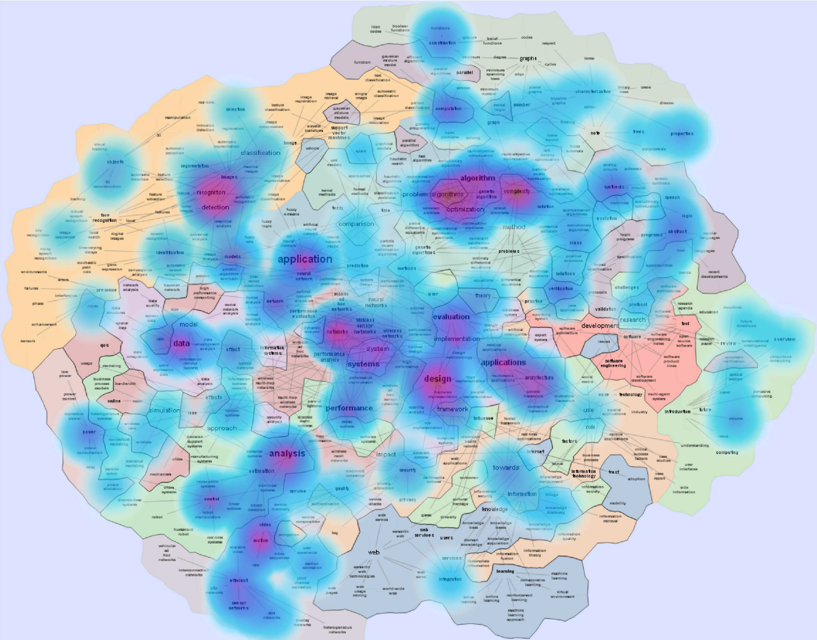
# Conference Heatmaps: TVCG



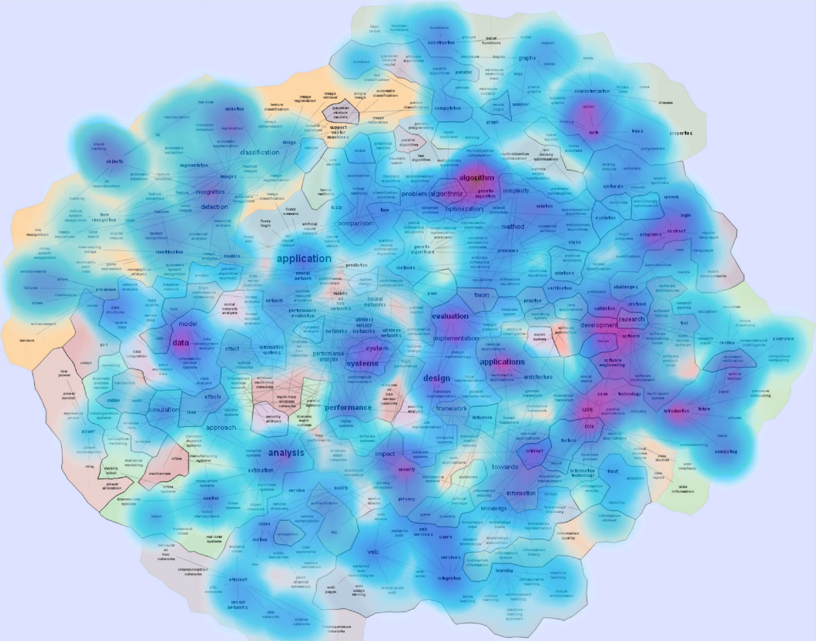
# Author Heatmaps: Tarjan over CS



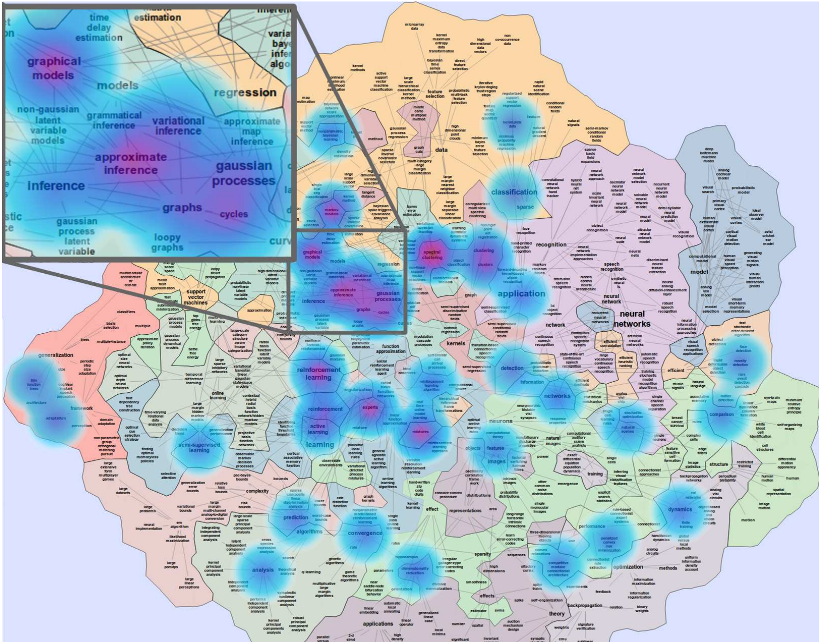
# Department Heatmaps: Arizona



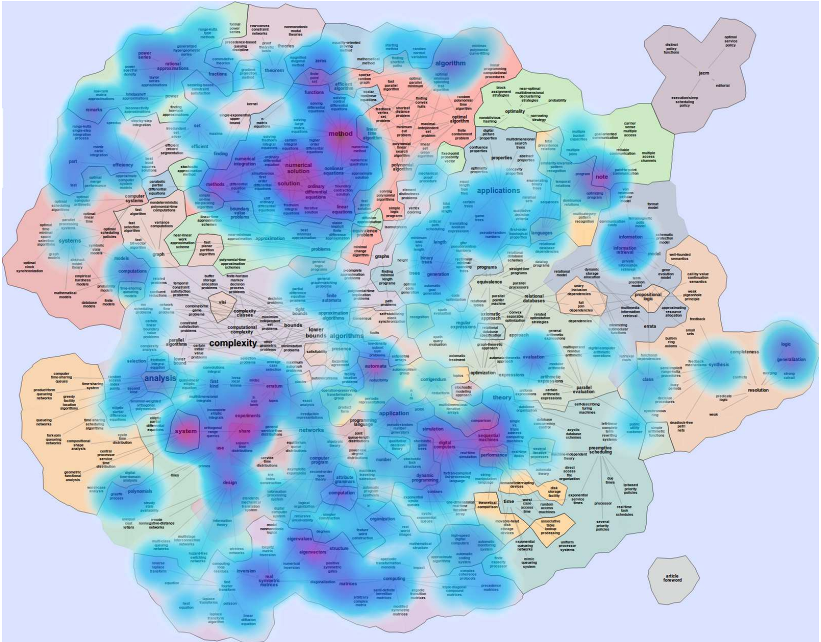
# Department Heatmaps: Georgia Tech



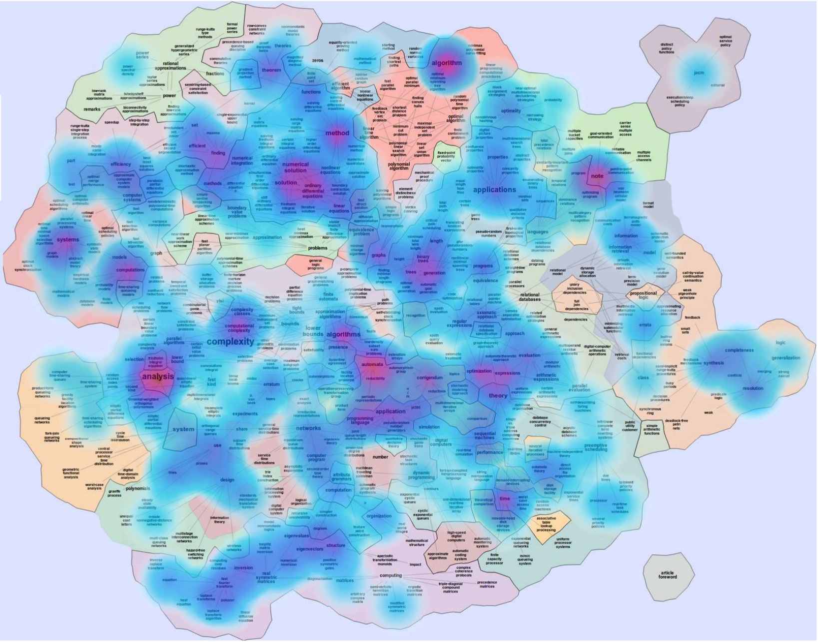
# Author Heatmaps: Michael Jordan over NIPS



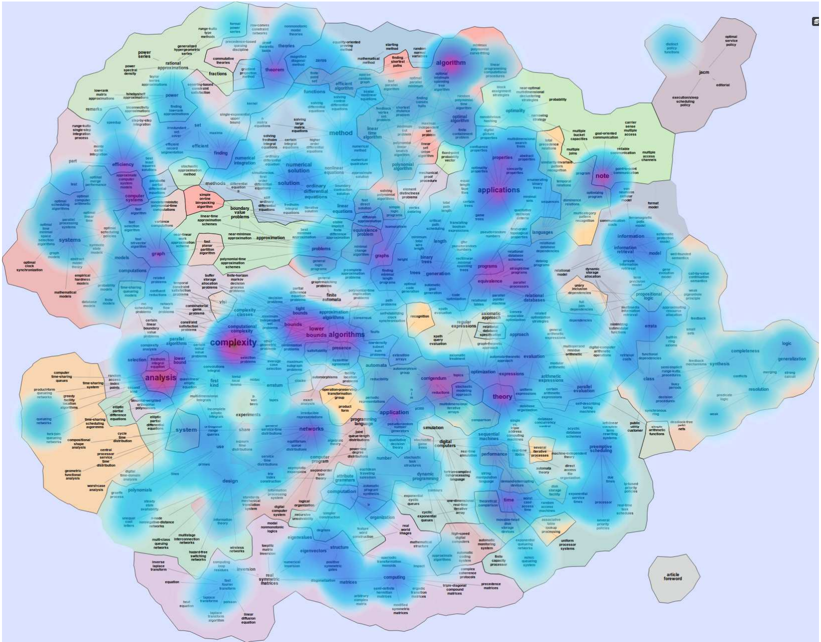
# Temporal Heatmaps: JACM 1954-1963



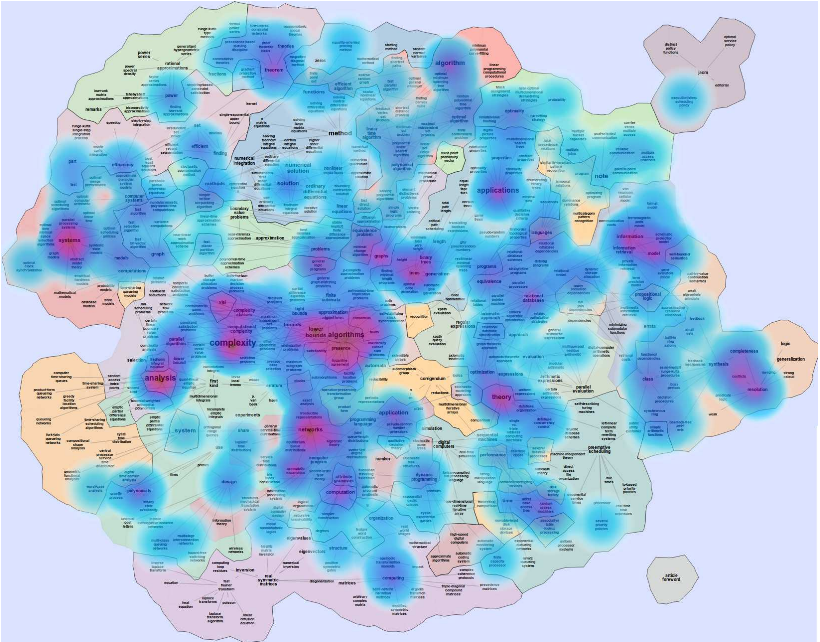
# Temporal Heatmaps: JACM 1964-1973



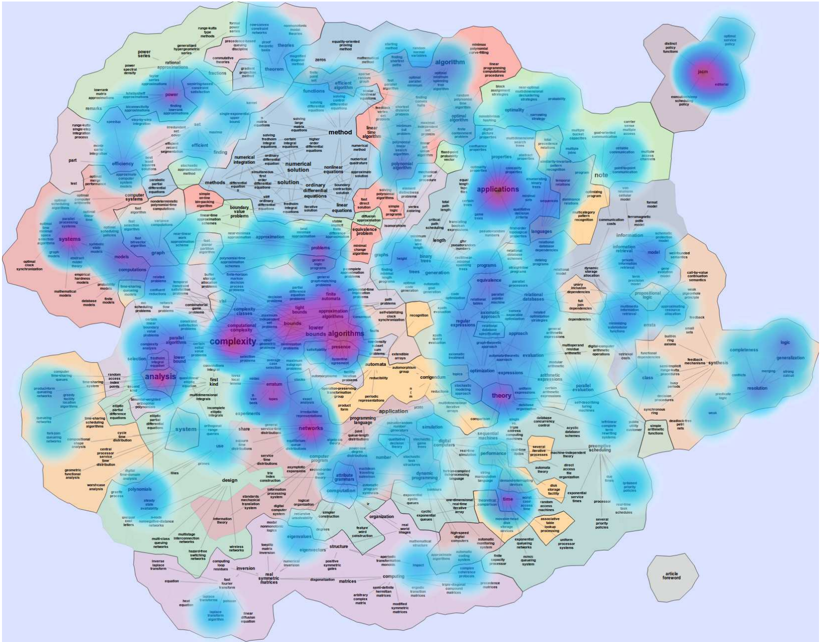
# Temporal Heatmaps: JACM 1974-1983



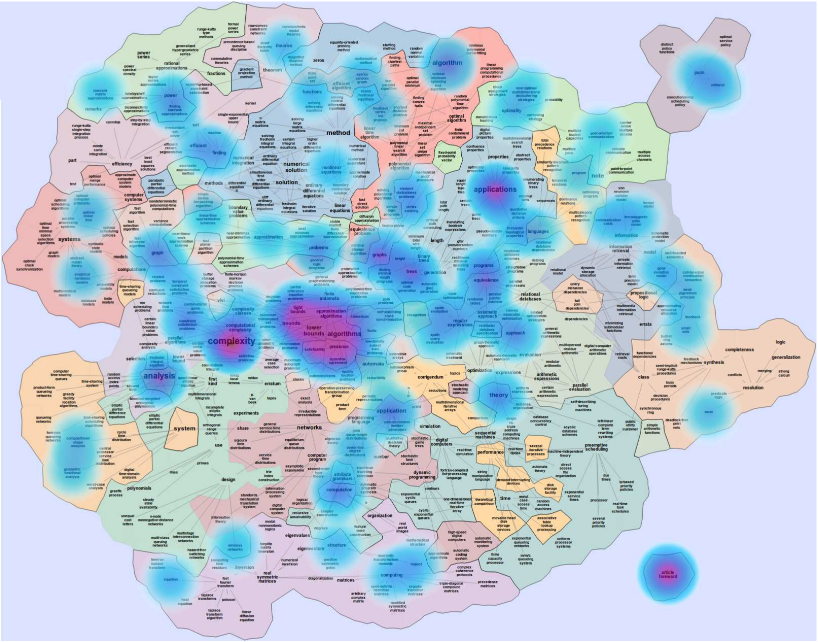
# Temporal Heatmaps: JACM 1984-1993



# Temporal Heatmaps: JACM 1994-2003



# Temporal Heatmaps: JACM 2004-2013



# Future Work

- MoCS++
  - do MoCS match what we expect to see?
  - what makes a canonical map good?
  - are titles enough? abstracts? papers?
  - MoX: Maps of X (physics, biology, ...)
  - all sources online (better NLP, layout, ...)
- GMap++
  - semantic zooming, levels of detail
  - terrain, fjords, rivers, ...
  - all source-code online (terrain, rivers, ...)
- Experiments
  - are maps "better" than graphs?
  - what types of maps are better?

`http://mocs.cs.arizona.edu`

`http://gmap.cs.arizona.edu`

# Acknowledgments

- Colleagues
  - Yifan Hu, Emden Gansner, AT&T
  - Daisuke Mashima, Georgia Tech
  - D. Fried, J. Fowler, S. Pupyrev, ... , Arizona
  - T. Johnson, L. Lazos, P. Simonetto ... , Arizona
- Workshops
  - Dagstuhl
  - Bertinoro
  - Barbados
- Funding
  - NSF
  - ONR
  - Humboldt

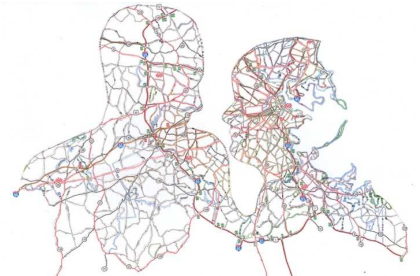
Alexander von Humboldt
Stiftung/Foundation

# Thanks!

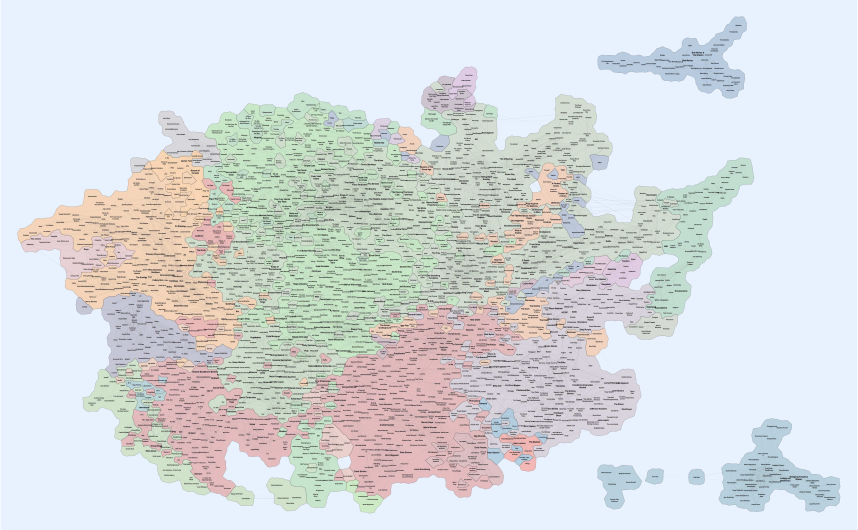
**T. Wallace and D. Huffman, "Atlas of Design"**

*Design and aesthetics matter, because form is not secondary to function; form is integral to function. A map cannot function if it remains unread. To truly engage map users requires that we present them with something worth looking at. Something that they will want to spend time studying. Something that acknowledges the human need for beauty. Something that causes them to think about the map in terms beyond whether or not it simply "works."*

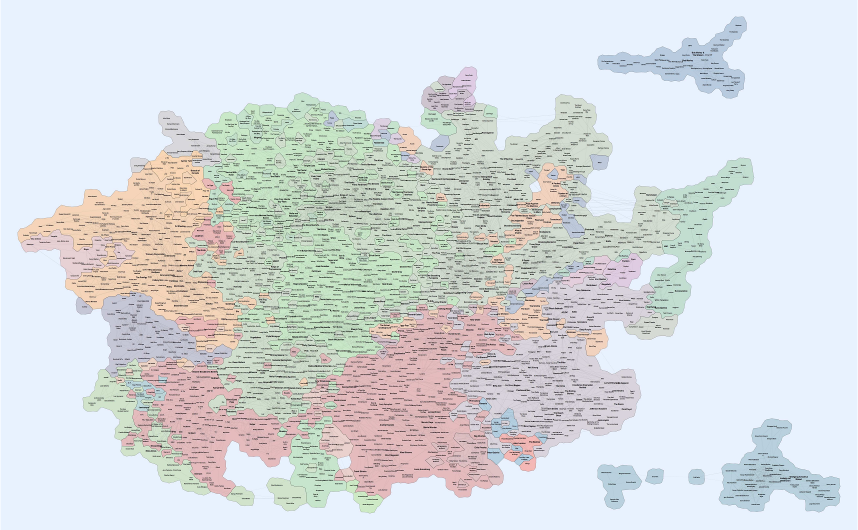
artist: Nikky Rosatto



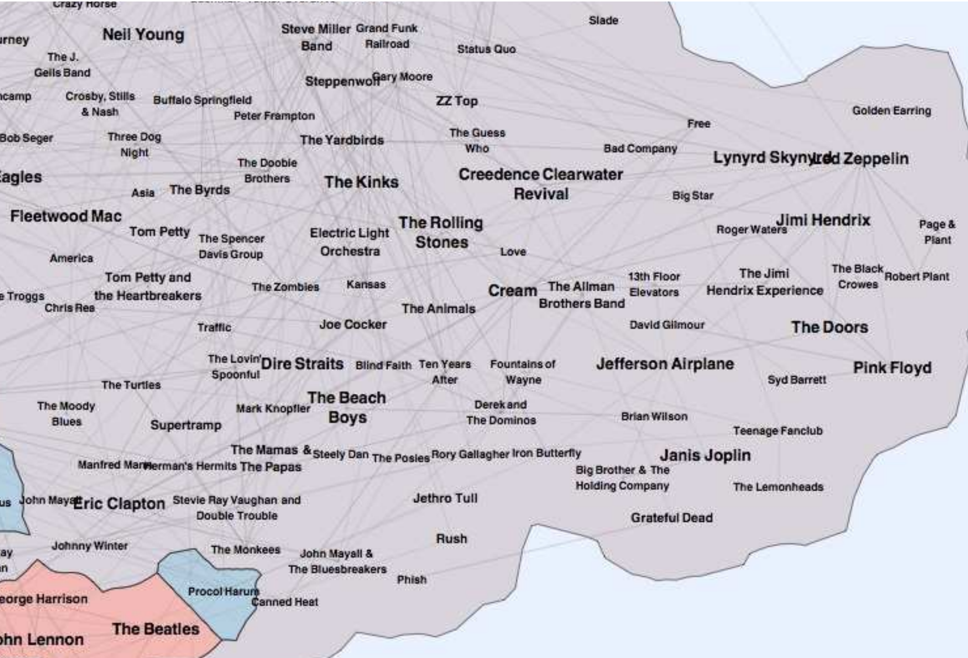
# Map of Music



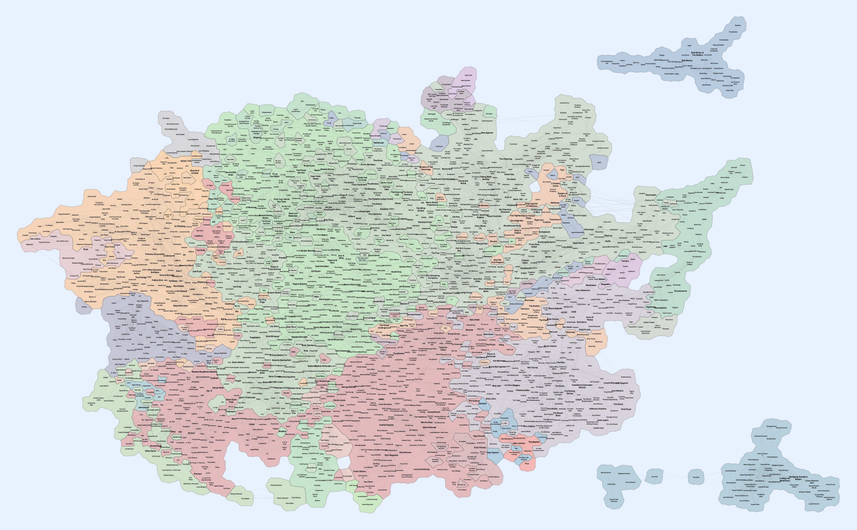
# Map of Music



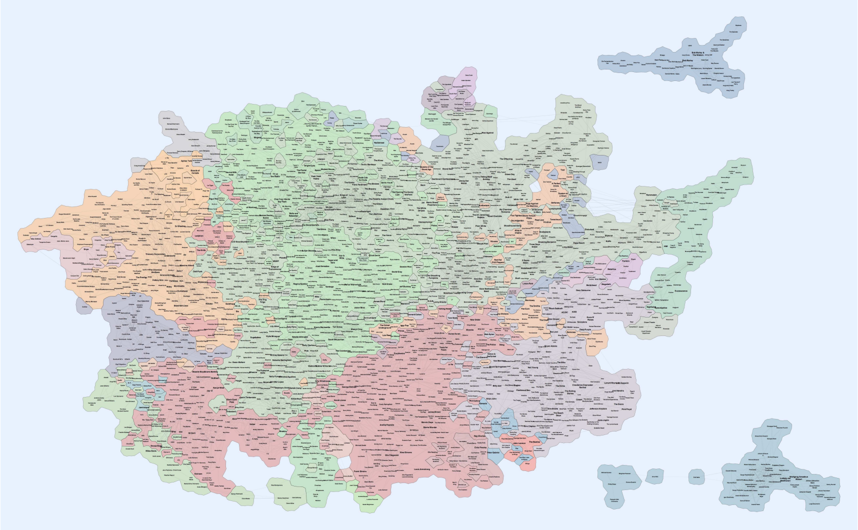
# Map of Music


Crazy Horse
Neil Young
Steve Miller Band
Grand Funk Railroad
Slade
The J. Geils Band
Status Quo
Steppenwolf
Gary Moore
Crosby, Stills & Nash
Buffalo Springfield
ZZ Top
Peter Frampton
Golden Earring
Free
Bob Seger
Three Dog Night
The Yardbirds
The Guess Who
Bad Company
Lynyrd Skynyrd
Led Zeppelin
The Doobie Brothers
Creedence Clearwater Revival
The Kinks
Asia
The Byrds
Big Star
Fleetwood Mac
Jimi Hendrix
Roger Waters
Page & Plant
Tom Petty
The Spencer Davis Group
Electric Light Orchestra
The Rolling Stones
Love
America
Tom Petty and the Heartbreakers
The Zombies
Kansas
Cream
The Allman Brothers Band
13th Floor Elevators
The Jimi Hendrix Experience
The Black Crowes
Robert Plant
Chris Rea
The Animals
Traffic
Joe Cocker
David Gilmour
The Doors
The Lovin' Spoonful
Dire Straits
Blind Faith
Ten Years After
Fountains of Wayne
Jefferson Airplane
Pink Floyd
Syd Barrett
The Turtles
The Moody Blues
Mark Knopfler
The Beach Boys
Derek and The Dominos
Brian Wilson
Supertramp
Teenage Fanclub
The Mamas & The Papas
Steely Dan
The Posies
Rory Gallagher
Iron Butterfly
Janis Joplin
Herman's Hermits
Big Brother & The Holding Company
The Lemonheads
Eric Clapton
Stevie Ray Vaughan and Double Trouble
Jethro Tull
Grateful Dead
Johnny Winter
Rush
The Monkees
John Mayall & The Bluesbreakers
Phish
Procol Harum
Canned Heat
The Beatles

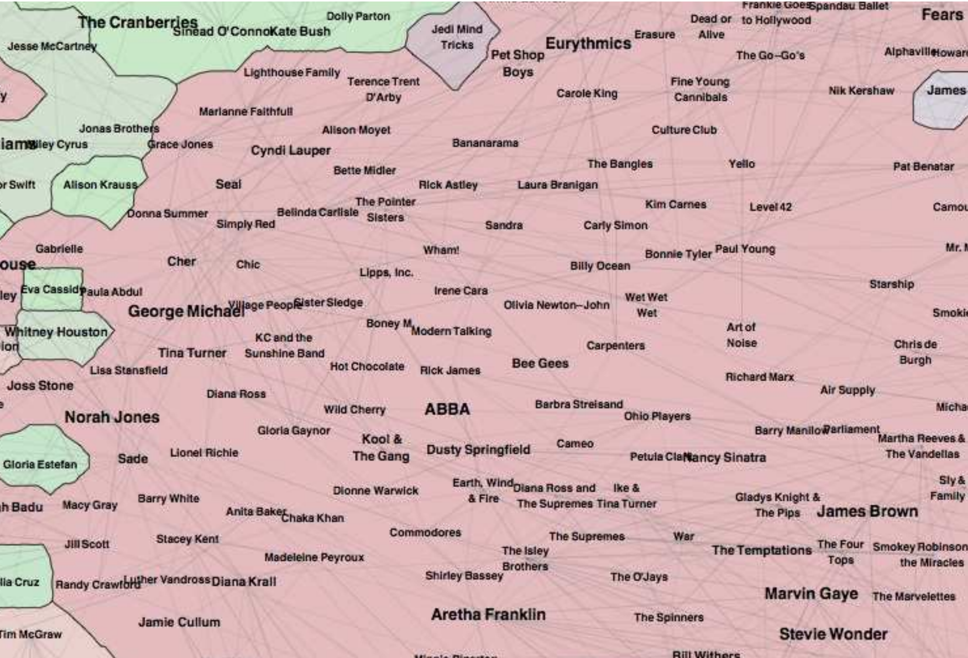
# Map of Music



# Map of Music

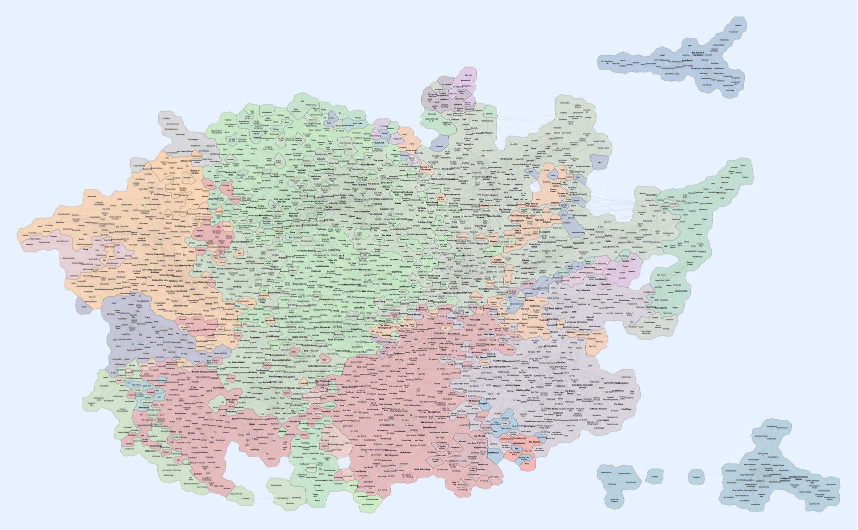


# Map of Music

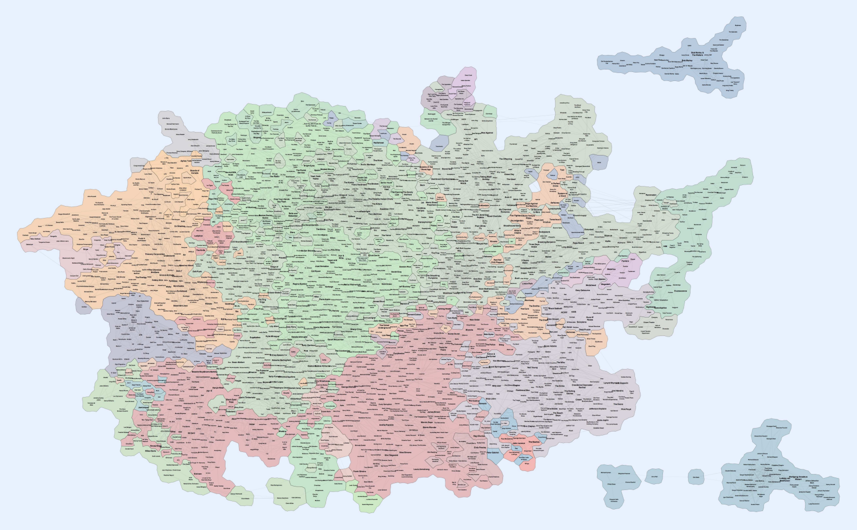


The Cranberries
Dolly Parton
Sinéad O'Connor
Kate Bush
Jedi Mind Tricks
Pet Shop Boys
Eurythmics
Erasure
Dead or Alive
Frankie Goes to Hollywood
Spandau Ballet
Jesse McCartney
Lighthouse Family
Terence Trent D'Arby
Carole King
Fine Young Cannibals
The Go-Go's
Alphaville
Nik Kershaw
Marianne Faithfull
Jonas Brothers
Miley Cyrus
Grace Jones
Alison Moyet
Cyndi Lauper
Bananarama
Culture Club
Alison Krauss
Seal
Bette Midler
Rick Astley
The Bangles
Yello
Pat Benatar
Laura Branigan
Donna Summer
Simply Red
Belinda Carlisle
The Pointer Sisters
Sandra
Kim Carnes
Carly Simon
Level 42
Gabrielle
Cher
Chic
Wham!
Lipps, Inc.
Billy Ocean
Bonnie Tyler
Paul Young
Eva Cassidy
Paula Abdul
Irene Cara
Starship
George Michael
Village People
Sister Sledge
Olivia Newton-John
Wet Wet Wet
Whitney Houston
Boney M.
Modern Talking
KC and the Sunshine Band
Tina Turner
Carpenters
Art of Noise
Chris de Burgh
Lisa Stansfield
Hot Chocolate
Rick James
Bee Gees
Richard Marx
Joss Stone
Diana Ross
Air Supply
Wild Cherry
ABBA
Barbra Streisand
Ohio Players
Norah Jones
Gloria Gaynor
Kool & The Gang
Barry Manilow
Parliament
Martha Reeves & The Vandellas
Dusty Springfield
Cameo
Petula Clark
Nancy Sinatra
Gloria Estefan
Sade
Lionel Richie
Sly & Family
Earth, Wind & Fire
Dionne Warwick
Diana Ross and The Supremes
Ike & Tina Turner
Gladys Knight & The Pips
James Brown
Macy Gray
Barry White
Anita Baker
Chaka Khan
Commodores
Jill Scott
Stacey Kent
The Supremes
War
The Temptations
The Four Tops
The Isley Brothers
Madeleine Peyroux
Randy Crawford
Luther Vandross
Diana Krall
Shirley Bassey
The O'Jays
Marvin Gaye
The Marvelettes
Jamie Cullum
Aretha Franklin
The Spinners
Stevie Wonder
Tim McGraw
Bill Withers

# Map of Music



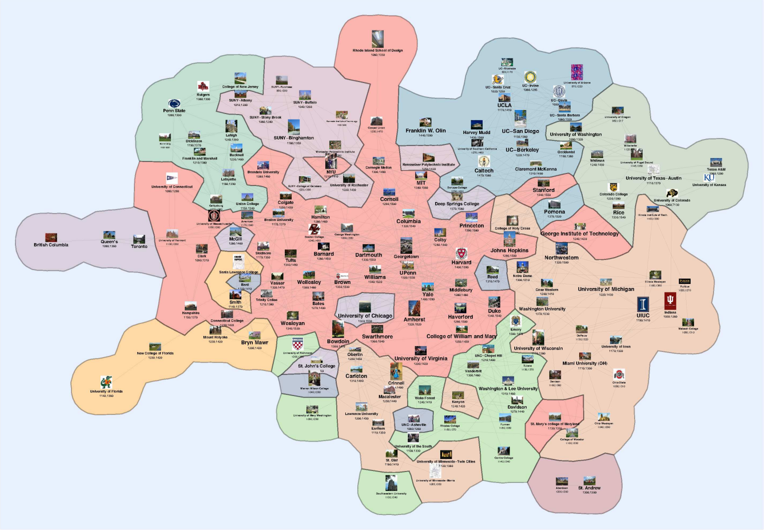
# Map of Music



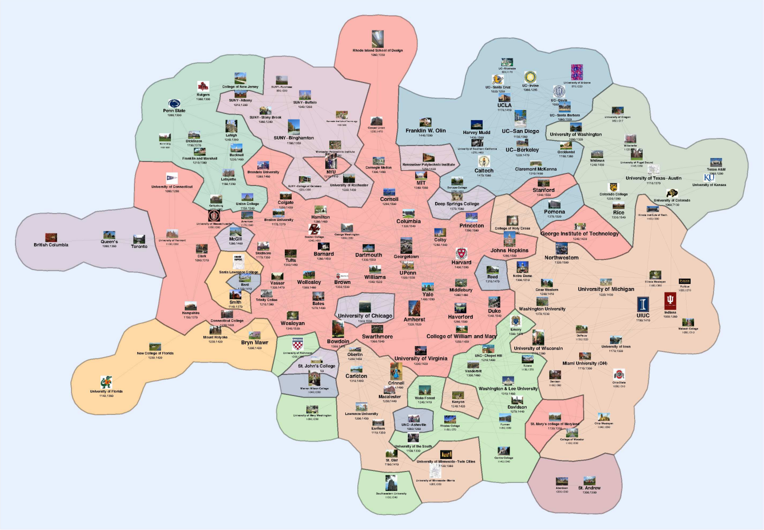
# Map of Music

Giuseppe Verdi
Giacomo Puccini
Gioacchino Rossini
Georges Bizet
Johann Strauss II
Richard Wagner
Richard Strauss
Robert Schumann
Hector Berlioz
Frédéric Chopin
Pyotr Ilyich Tchaikovsky
Franz Liszt
Claude Debussy
Franz Schubert
Sergei Rachmaninoff
Johannes Brahms
Edvard Grieg
Wolfgang Amadeus Mozart
Ludwig van Beethoven
Gabriel Fauré
Gustav Mahler
Jean Sibelius
Felix Mendelssohn
Johann Sebastian Bach
Maurice Ravel
Modest Petrovich Mussorgsky
Franz Joseph Haydn
Georg Friedrich Händel
Antonín Dvořák
Sergei Prokofiev
Camille Saint-Saëns
Antonio Vivaldi
Johann Pachelbel
Sir Edward Elgar
Georg Philipp Telemann
Dmitri Shostakovich
Igor Stravinsky
Samuel Barber
Luigi Boccherini
Gustav Holst

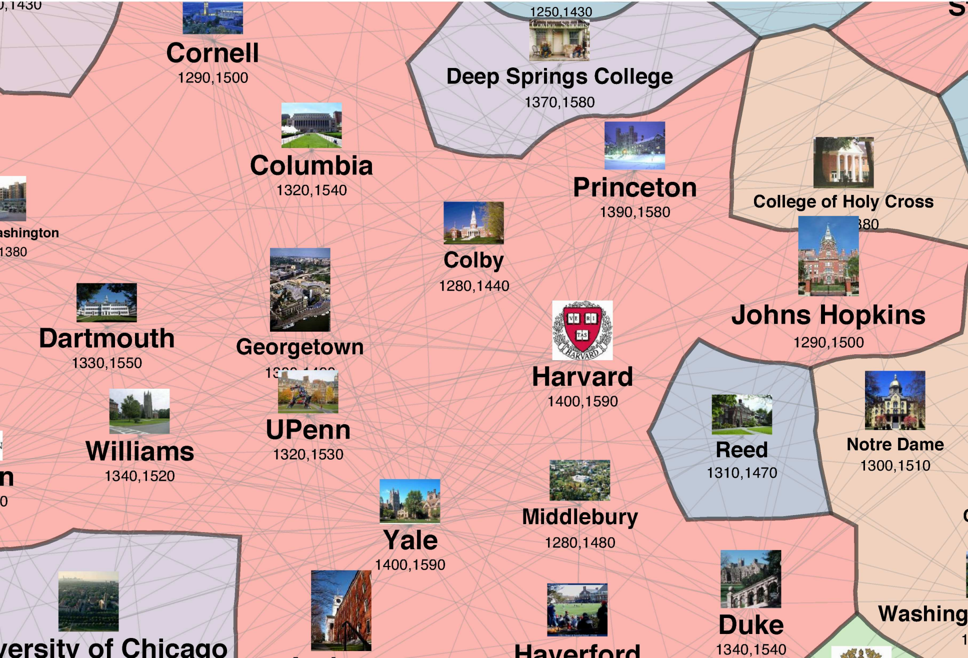
# Map of Academia



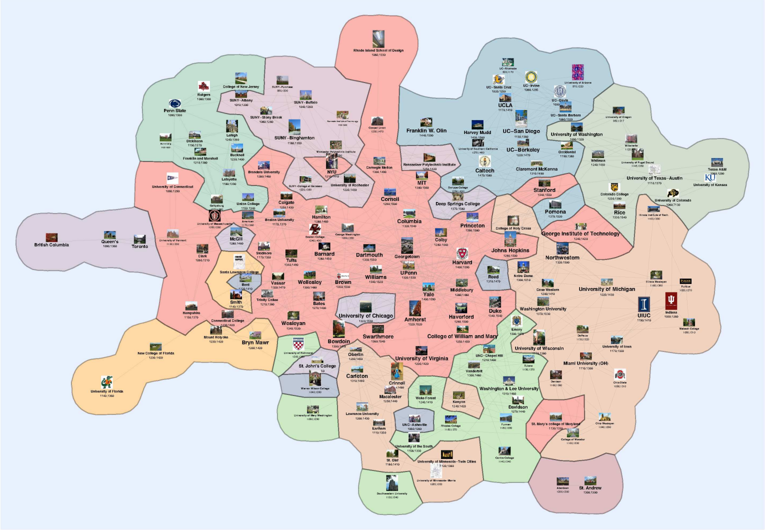
# Map of Academia



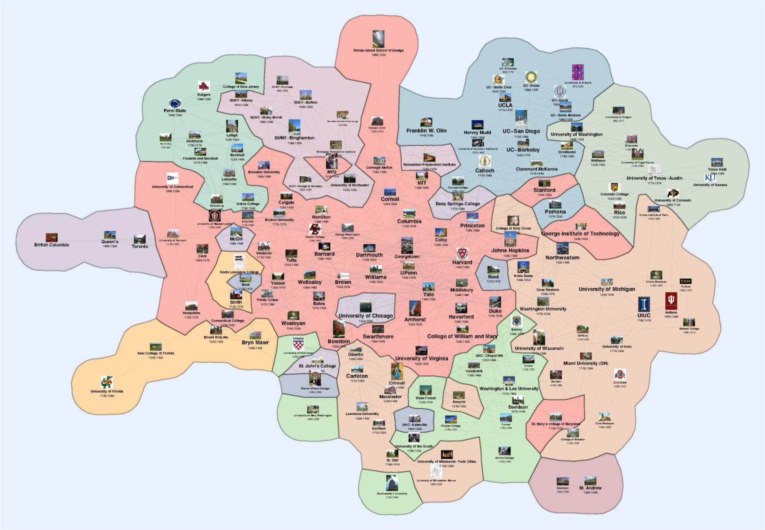
# Map of Academia



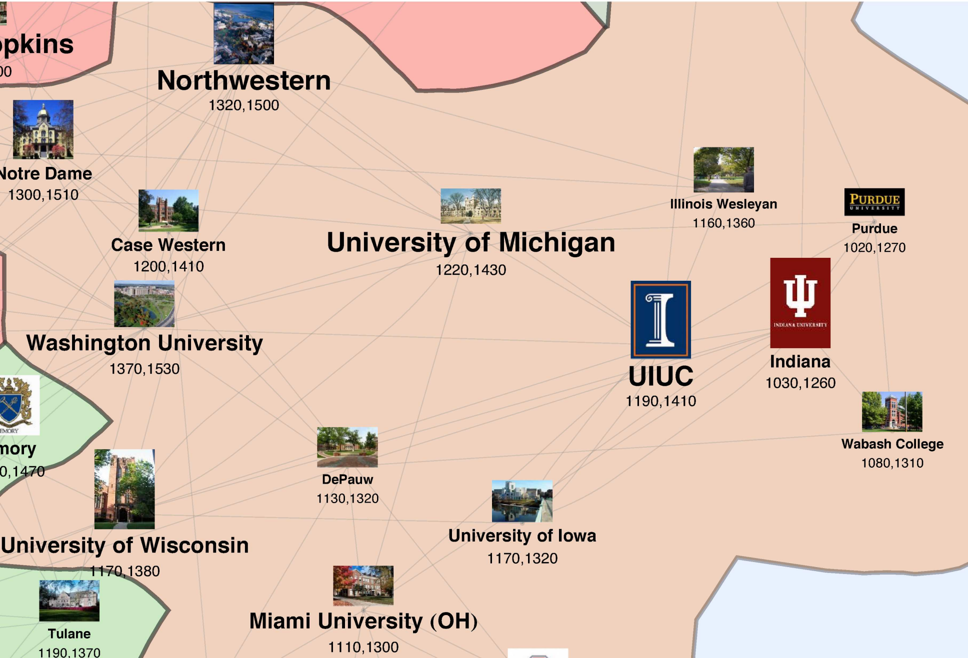
# Map of Academia



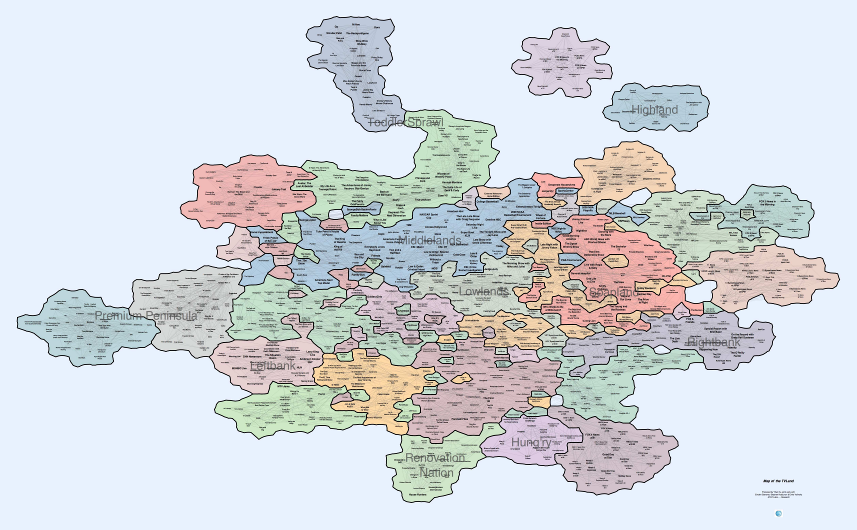
# Map of Academia



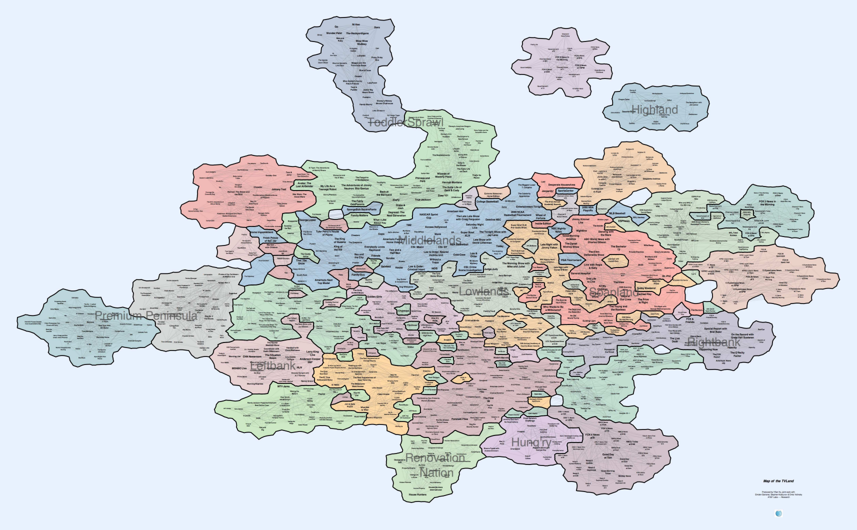
# Map of Academia



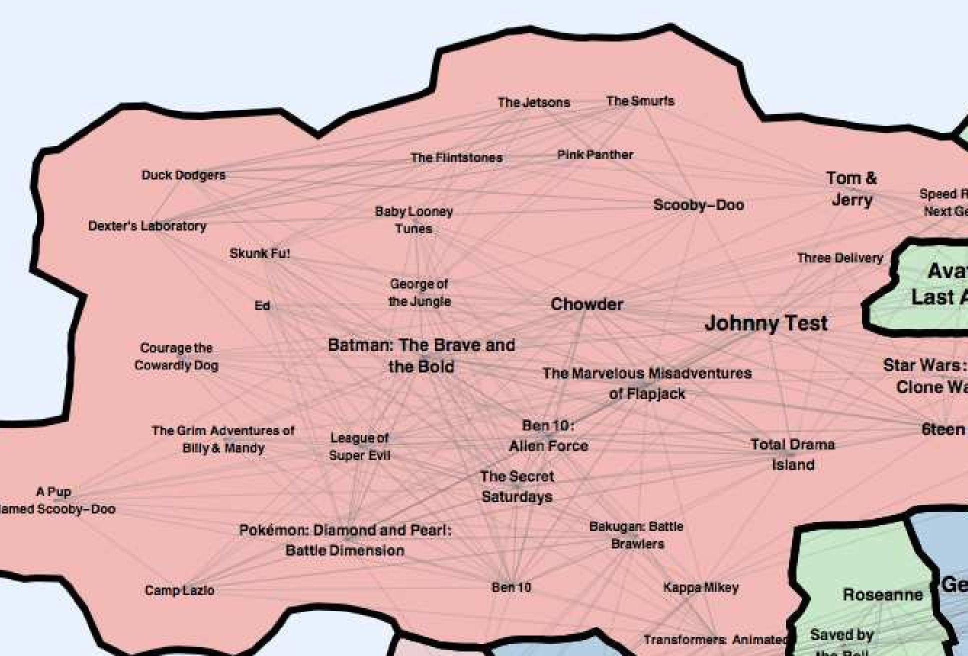
# Map of TV: TVLand



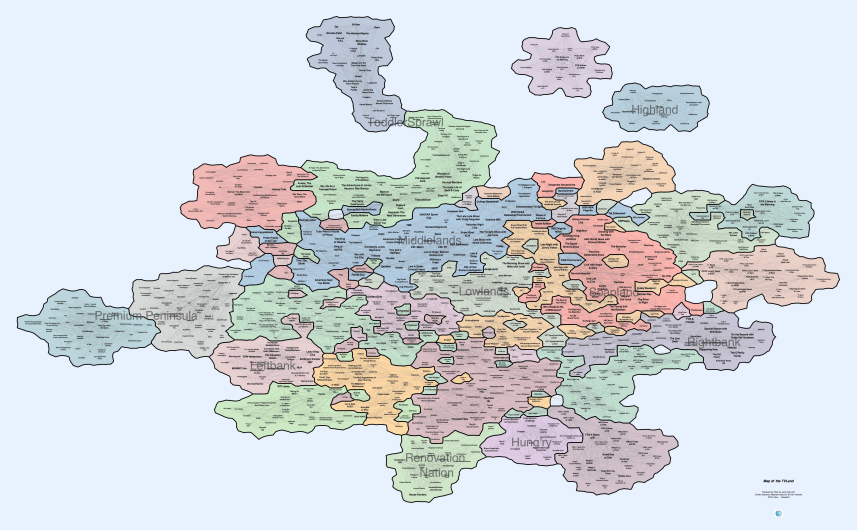
# Map of TV: TVLand



# Map of TV: TVLand

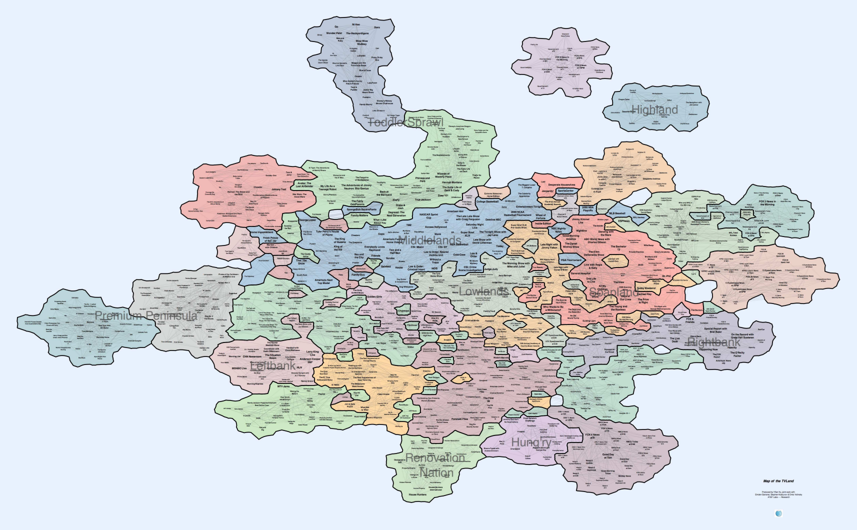


# Map of TV: TVLand

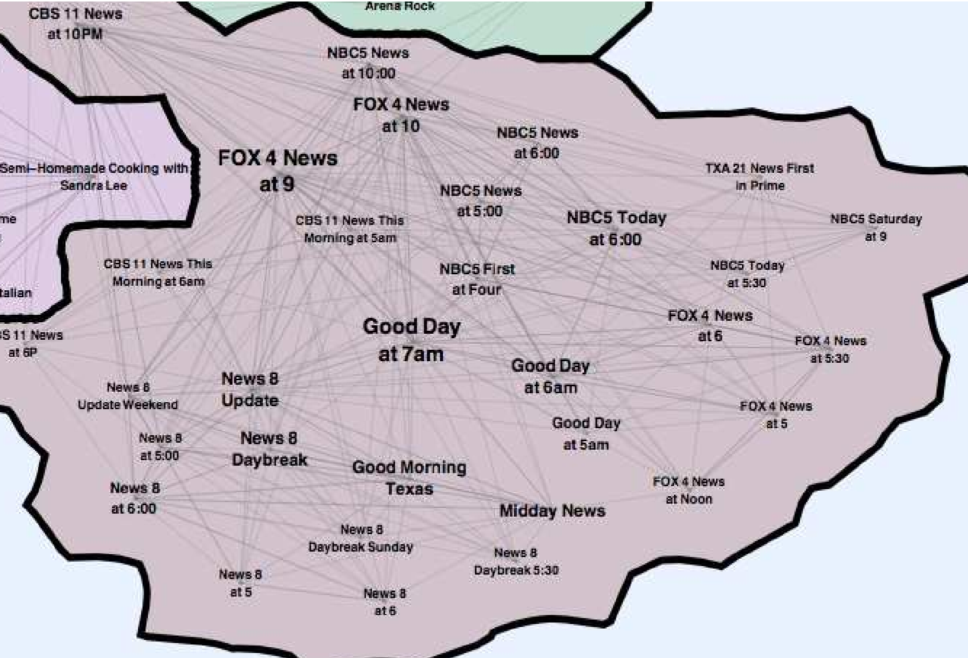


# Map of TV: TVLand

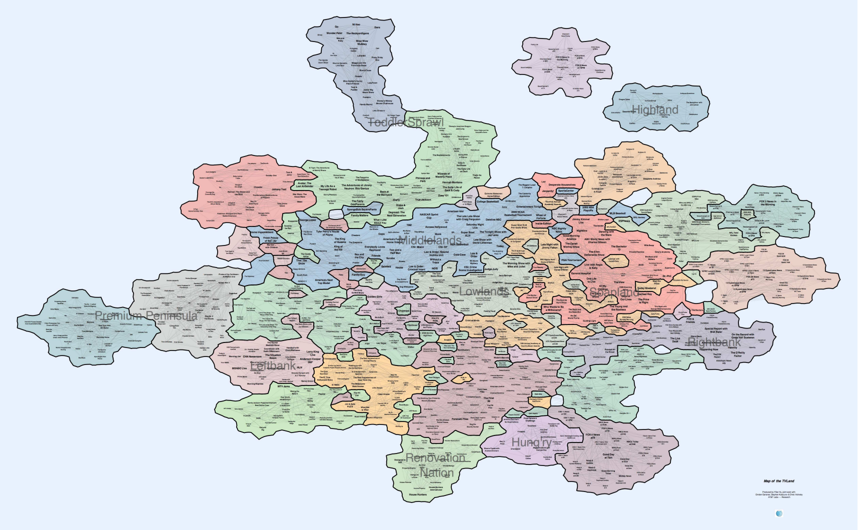
# Map of TV: TVLand



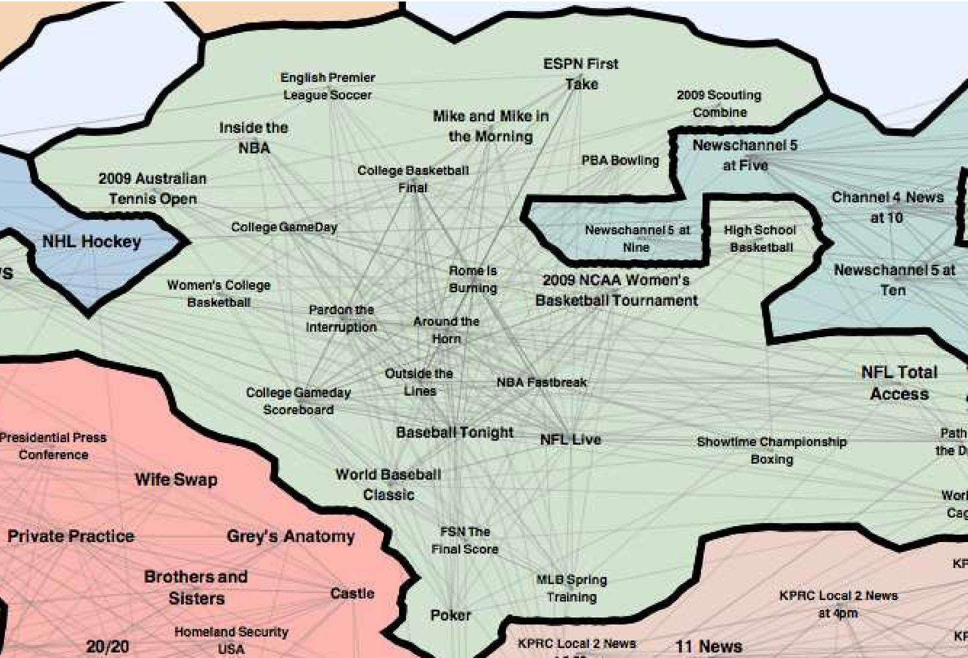
# Map of TV: TVLand



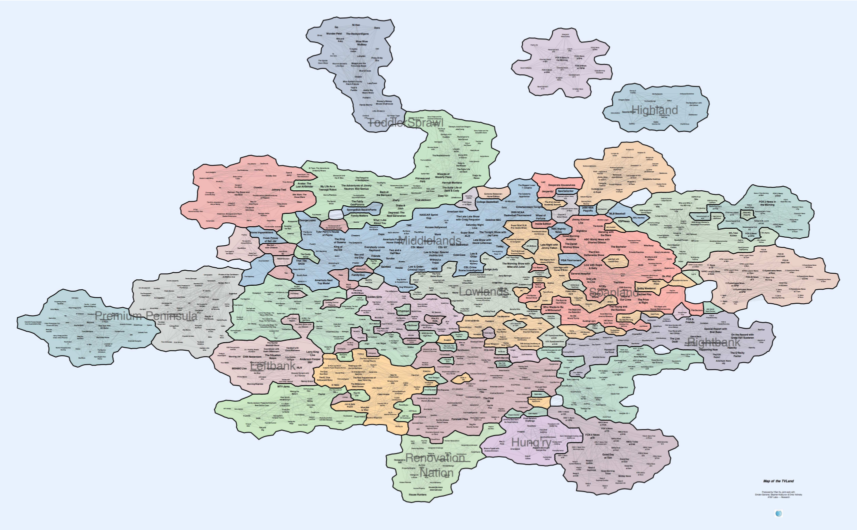
# Map of TV: TVLand



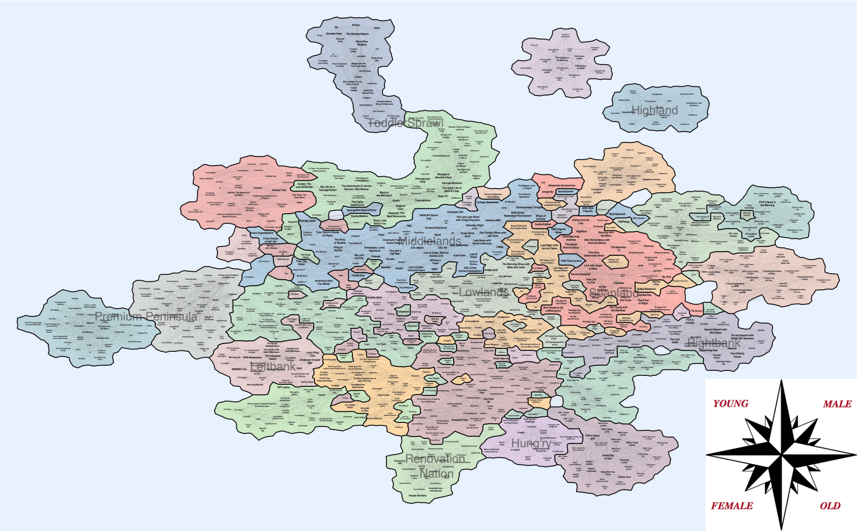
# Map of TV: TVLand



# TVLand Compass



# TVLand Compass



# TVLand Topography