

Visualização de Informação

Parte III

Multi-dimensional Visualization:
Visual Mining of Text , Images
and other Multi-dimensionoanl
entities.

*Rosane Minghim +
The team*



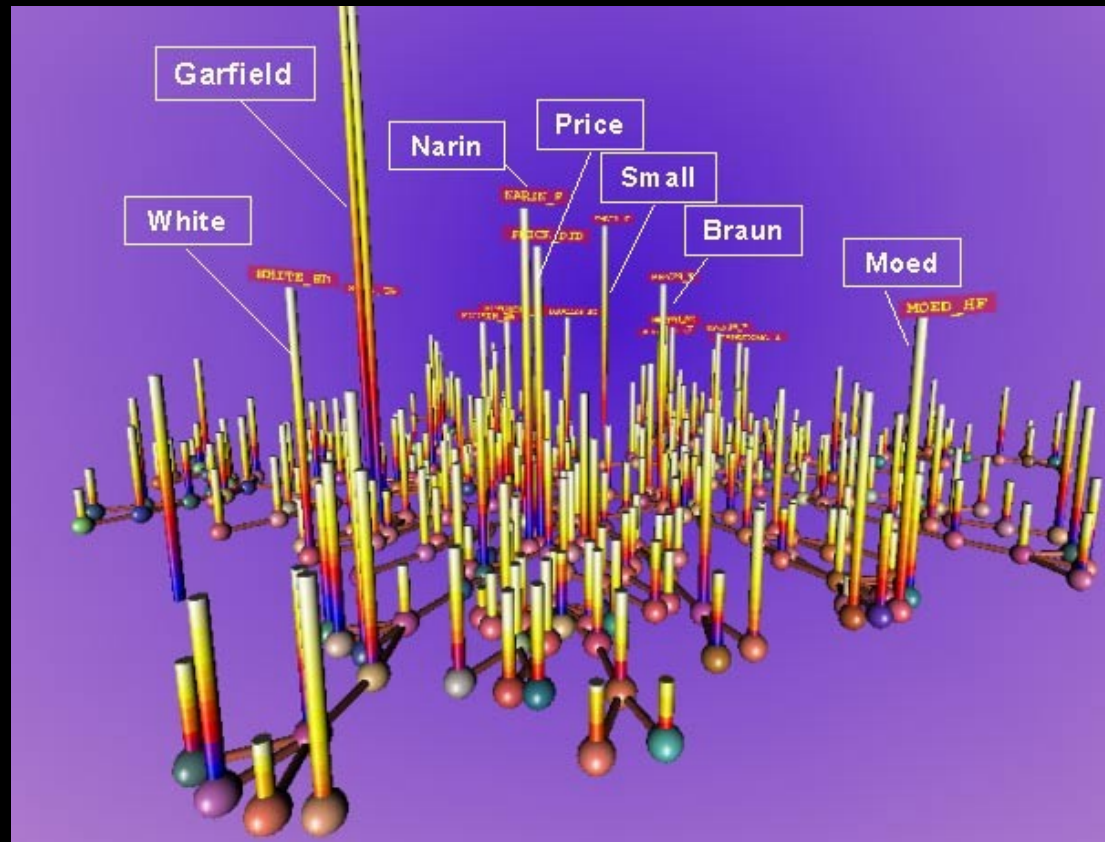
Instituto de Ciências Matemáticas e de
Computação
USP-São Carlos

Visual Text Mapping

- Approach 1: Relationship Based (Metadata)
- Approach 2: Content based



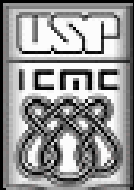
Relationships : Citation and Co-citation



(Borner)
(2003)

Content-based Text Mapping

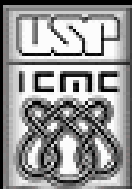
- Approach 1: Dimension reduction (Projections)
- Approach 2: Point Placement (PP)



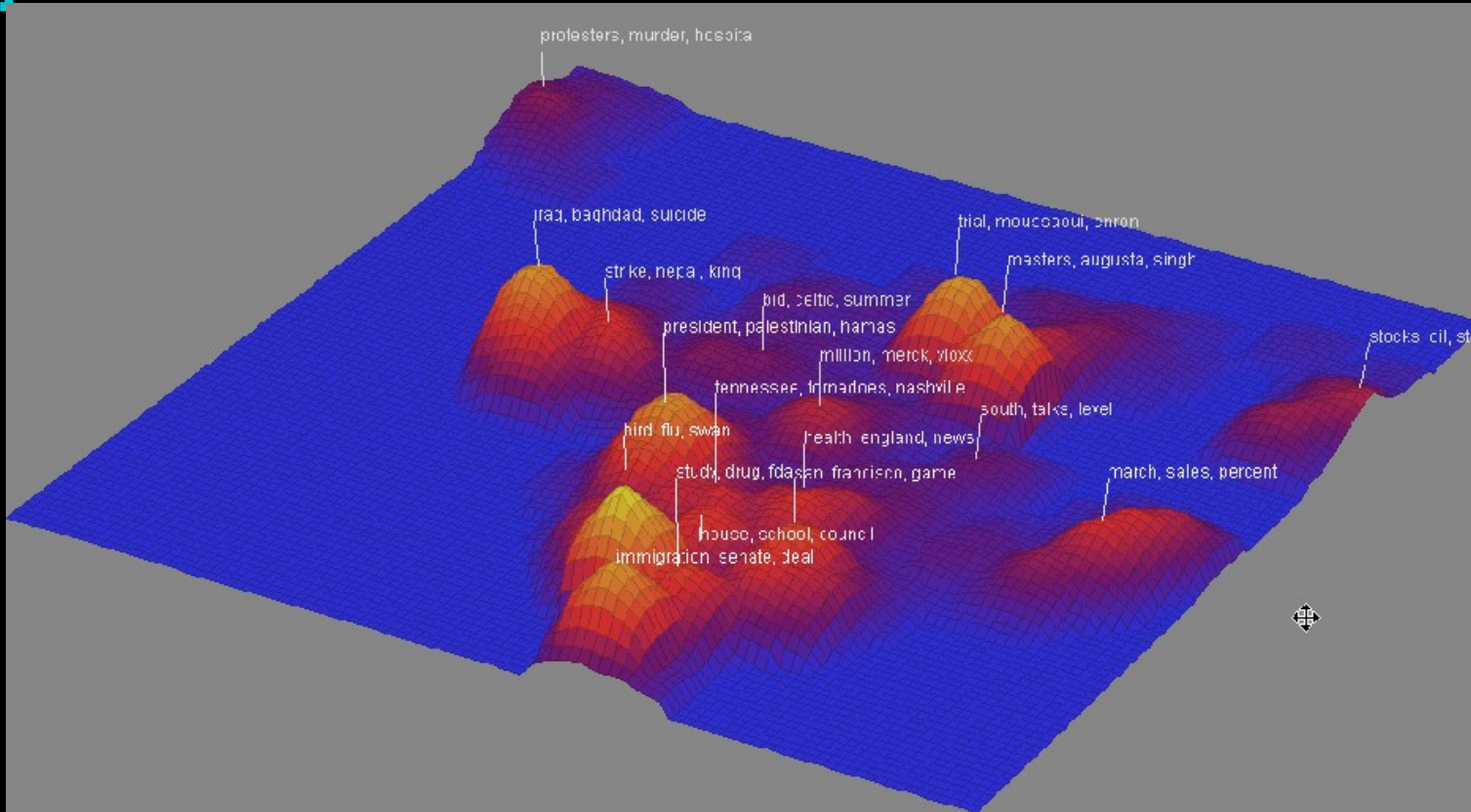
Content - based



(Dimensional
Reduction)
News flash
IN-SPIRE
(PNL)



Content - based

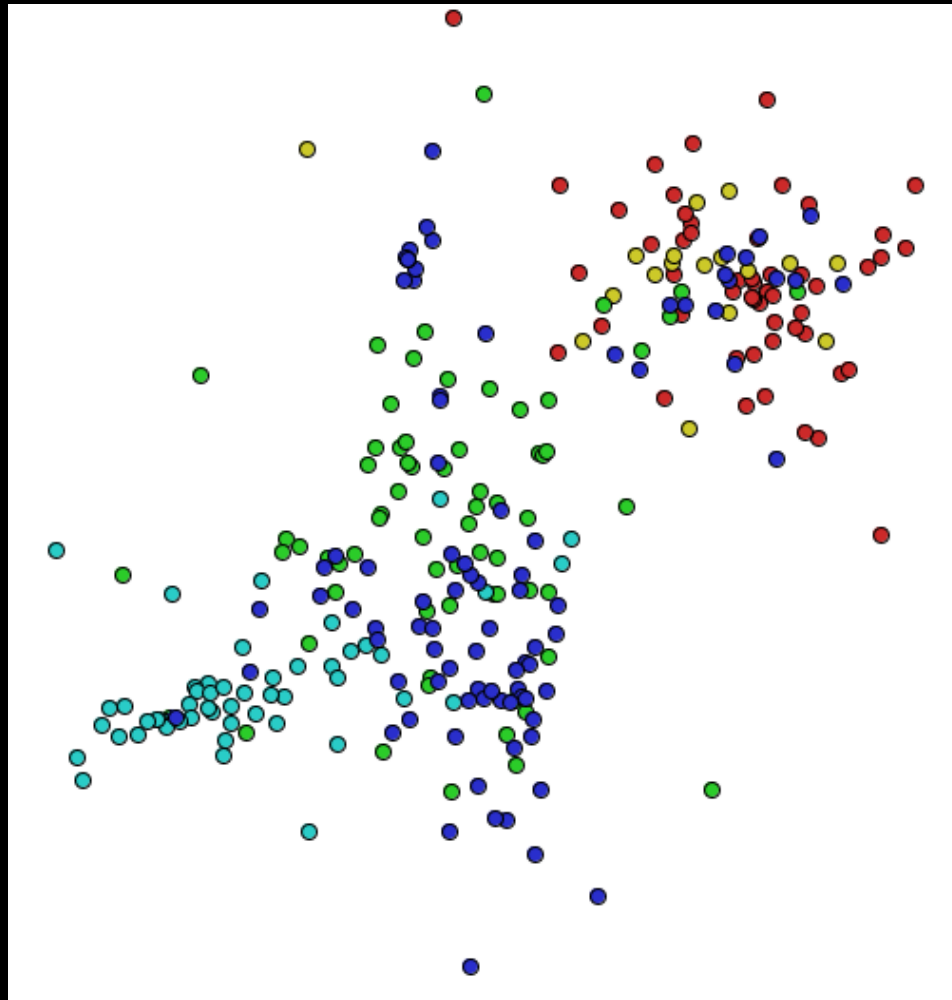


(Surface
View)
IN-SPIRE

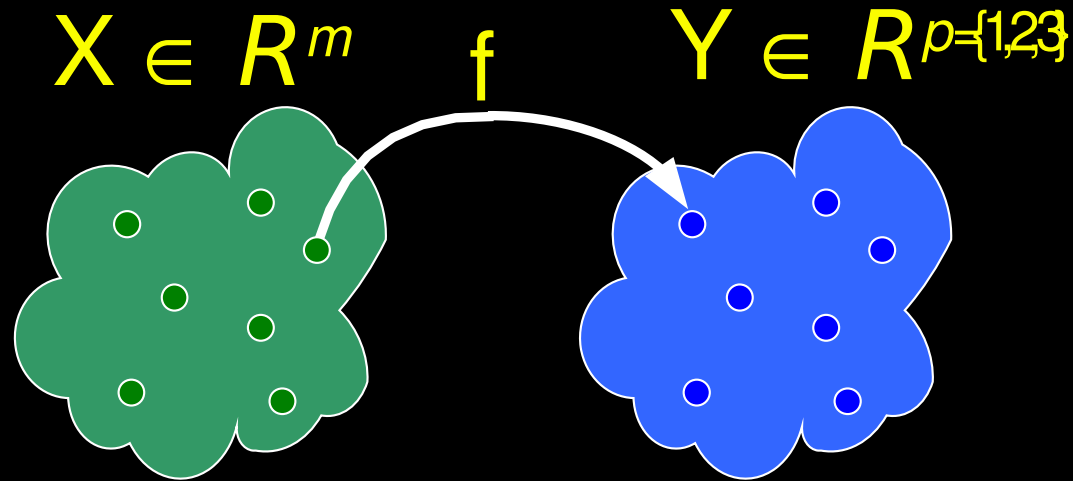


Mapeamento para o plano permitindo a exploração.

Ex: Patents **surgery**, **drugs**, **molecular**
bio

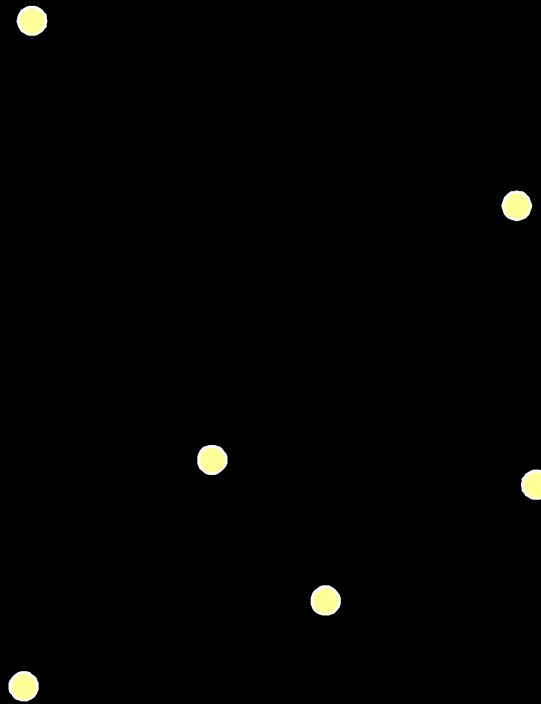


Projection Techniques

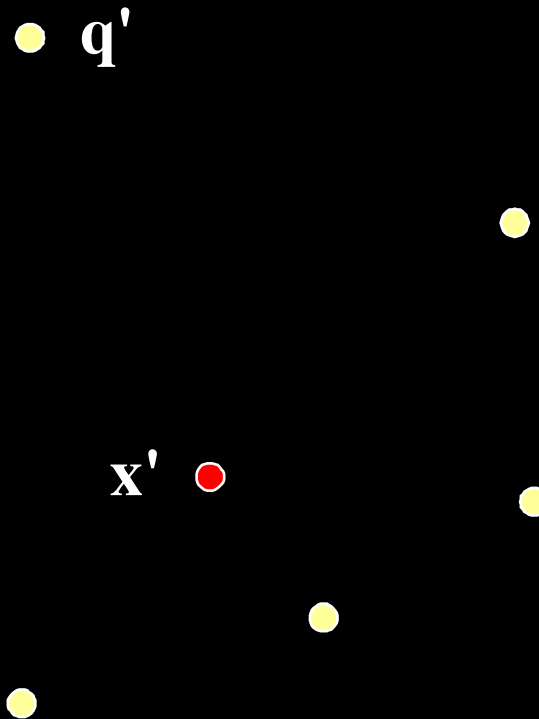


- $\delta: x_i, x_j \rightarrow R, x_i, x_j \in X$
- $d: y_i, y_j \rightarrow R, y_i, y_j \in Y$
- $f: X \rightarrow Y, |\delta(x_i, x_j) - d(f(x_i), f(x_j))| \approx 0, \forall x_i, x_j \in X$

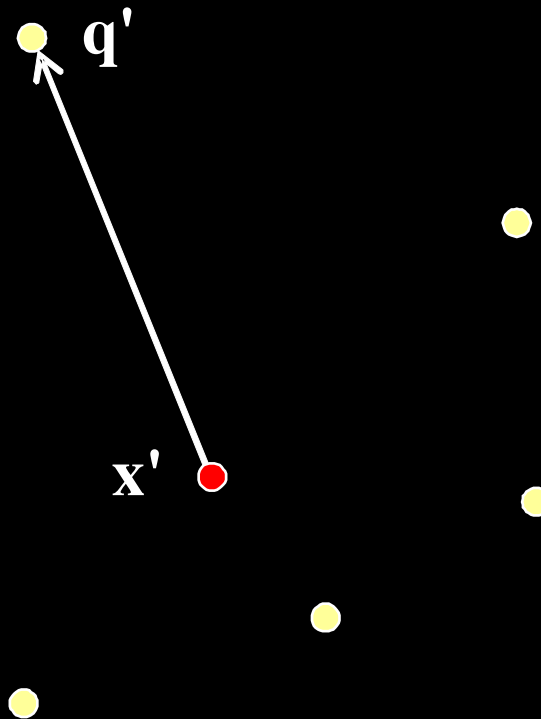
Force Based Point Placement



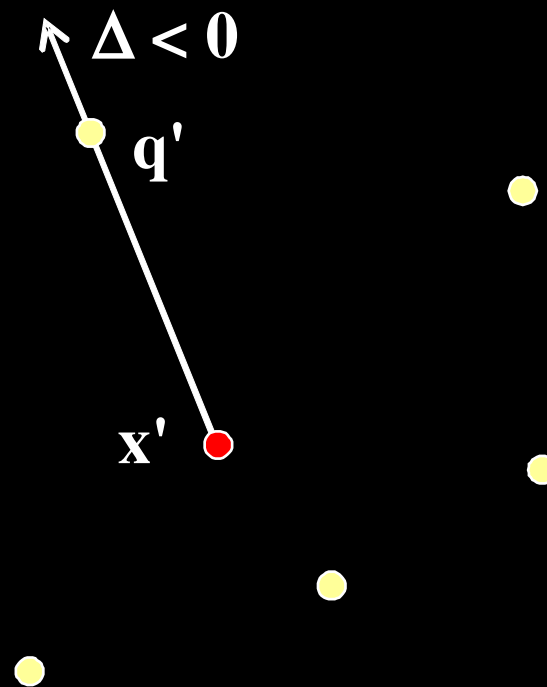
Force Scheme [Tejada et al., 2003]



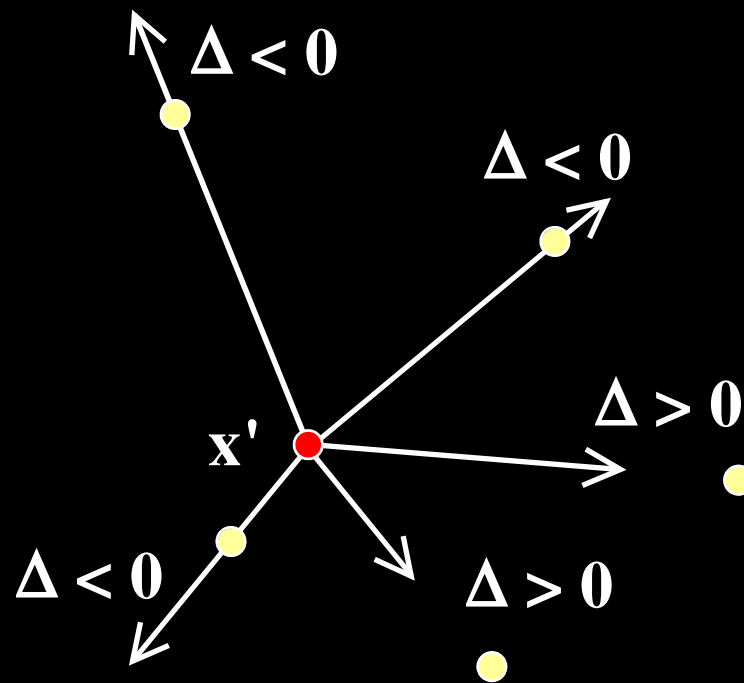
Force Scheme [Tejada et al., 2003]



Force Scheme [Tejada et al., 2003]

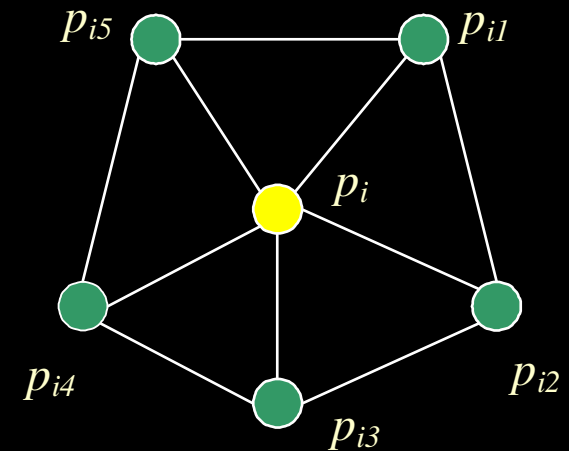


Force Scheme [Tejada et al., 2003]



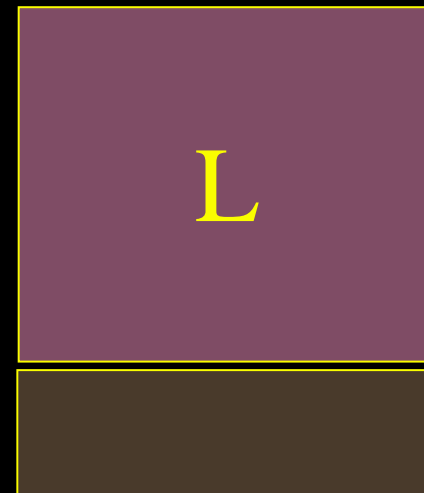
LSP: Matriz Laplaciana

- Seja $V_i = \{p_{i1}, \dots, p_{ik_i}\}$ a vizinhança de um ponto p_i e seja c_i as coordenadas de p_i em \mathbb{R}^p

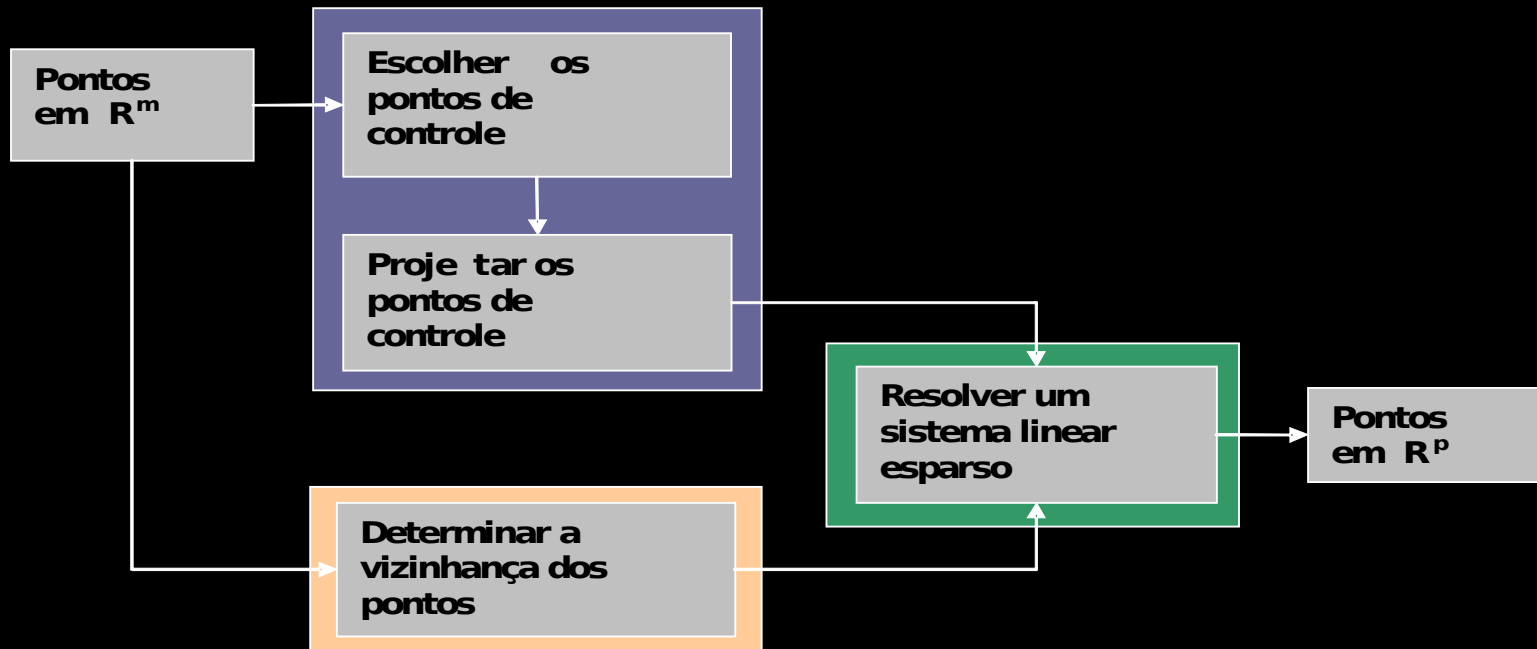


- Cada p_i será o centróide dos pontos em V_i

LSP: Adicionando os Pontos de Controle



LSP: Visão Geral



Choosing the Control Points

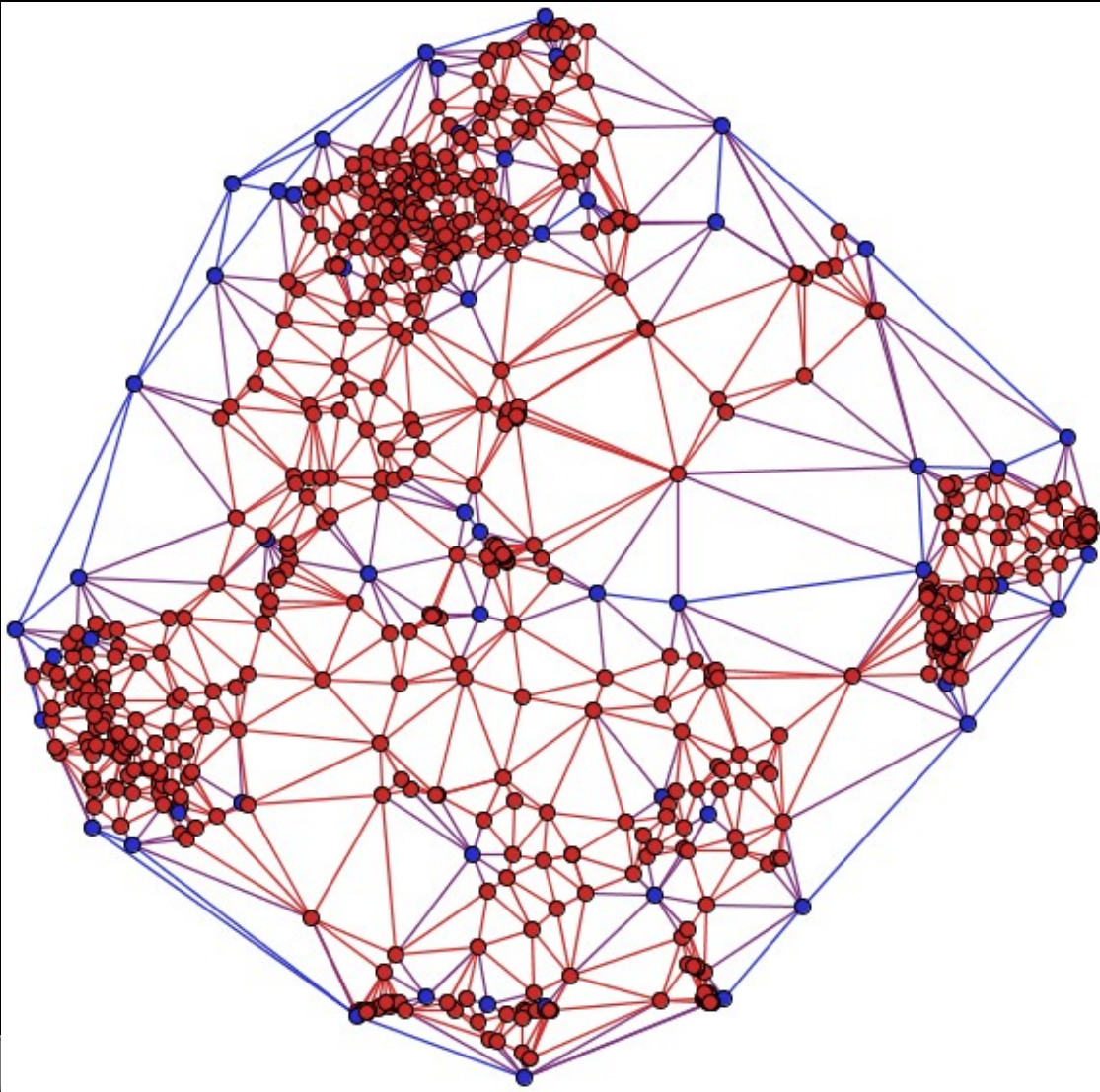
- In order to select the control points
 - the space R^m is split into nc clusters using k-medoids.
 - the control points are the medoids of each cluster



Choosing the Control Points

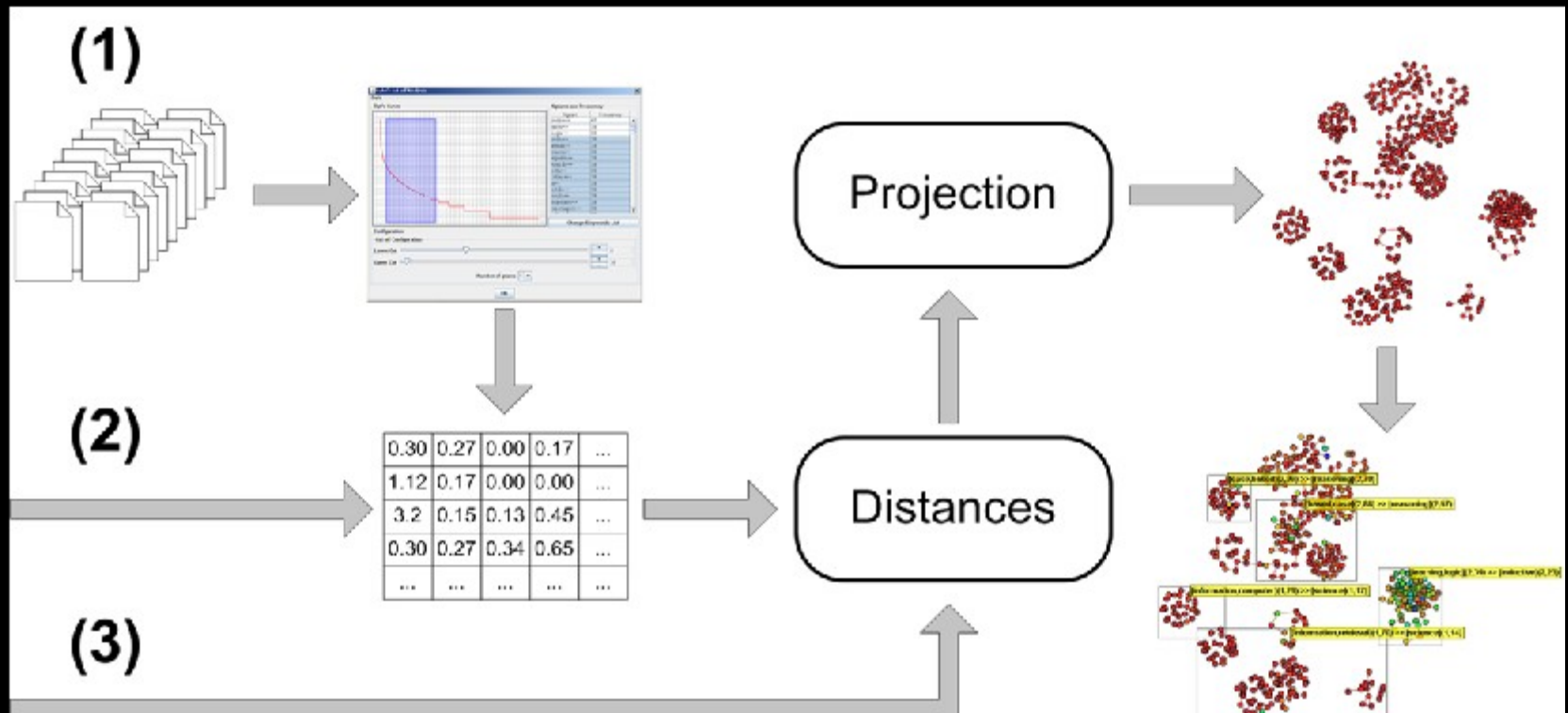
- Once the control points are chosen, these points are projected onto R^d through a fast dimensionality reduction method
 - Fast Projection (Fastmap or NNP)
 - Force Placement



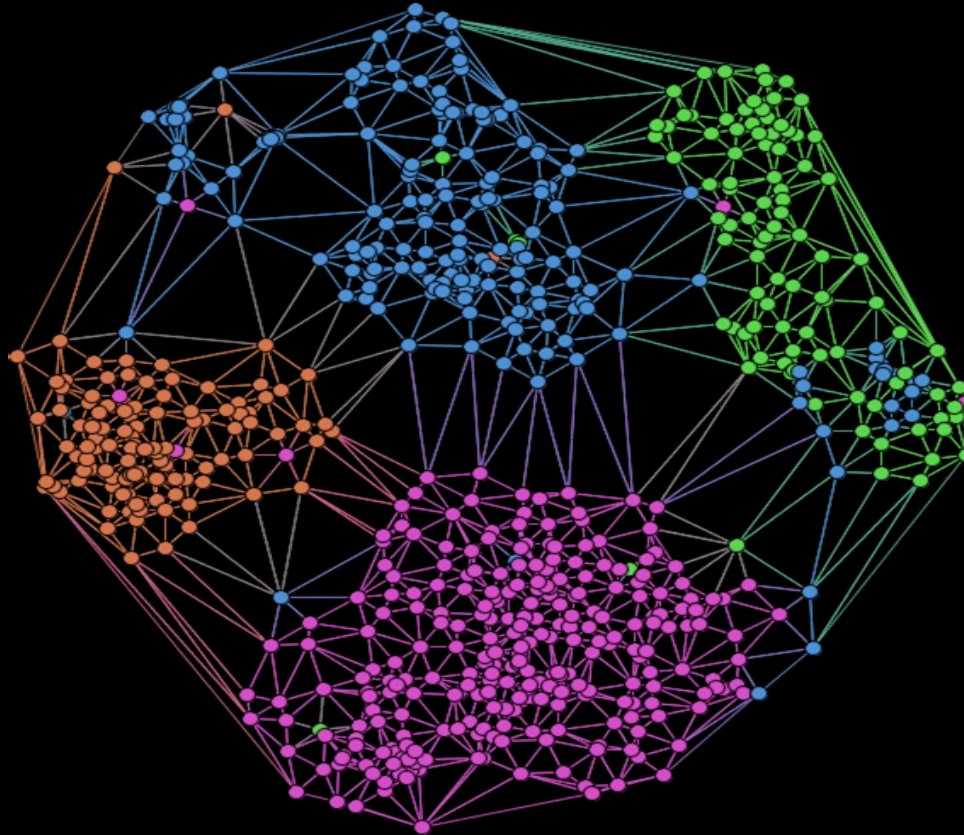


Control
points in
blue

Content - based by Projections

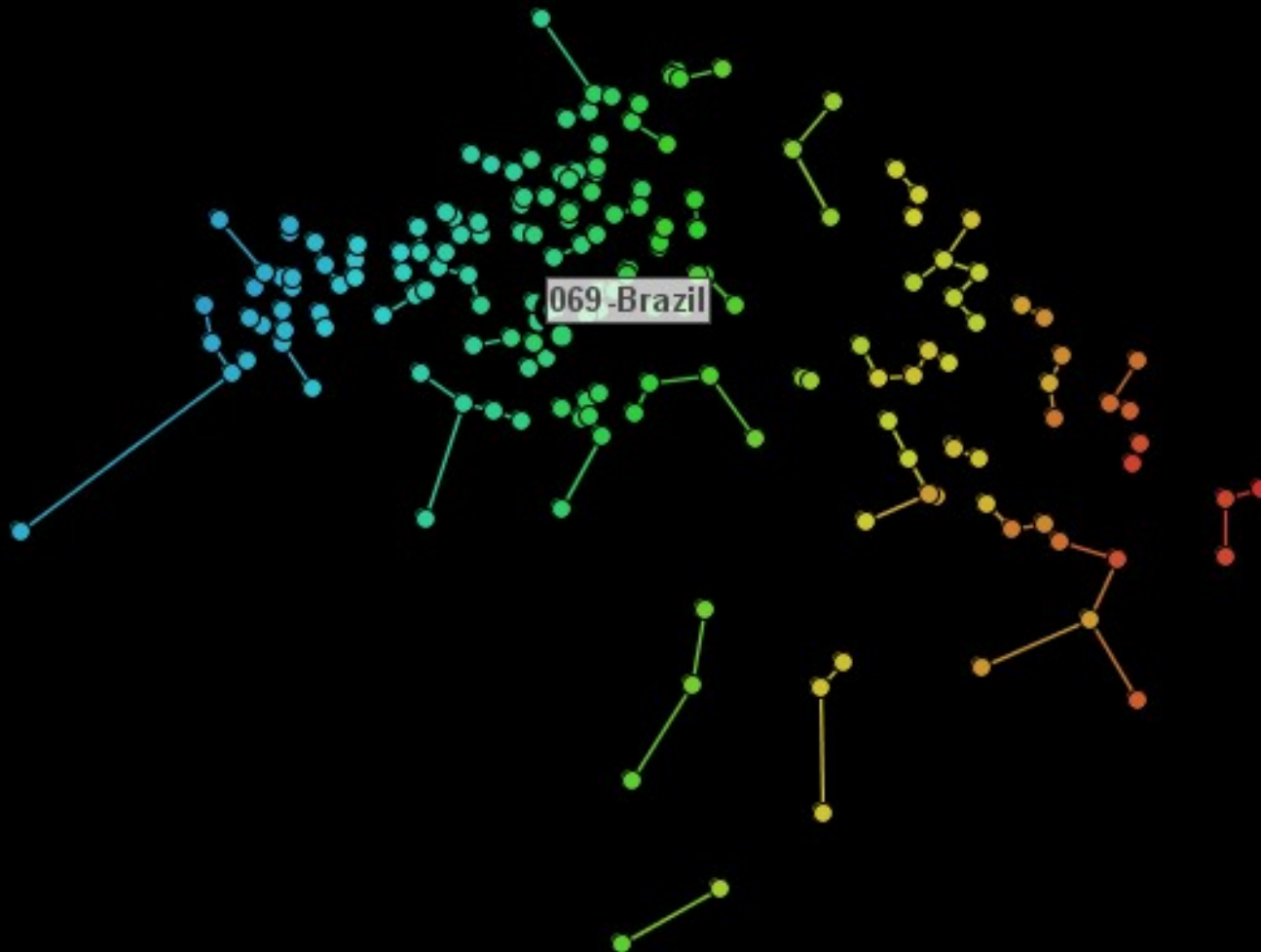


Example



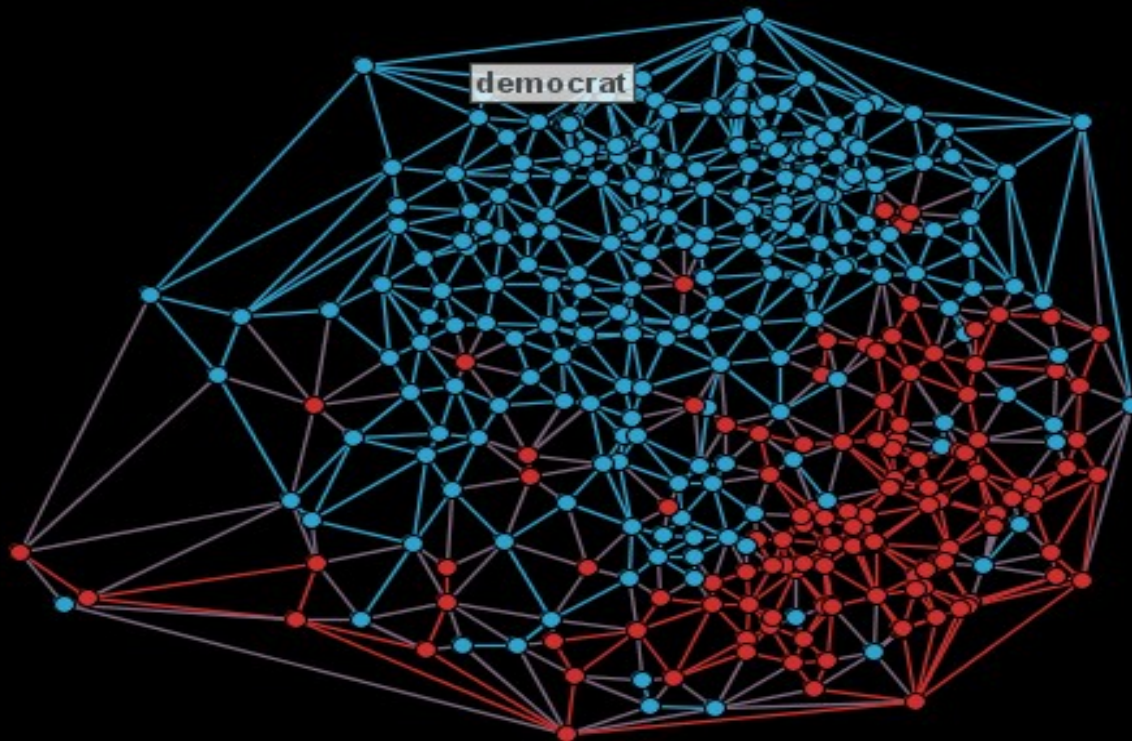
Exemplo de Projeção: IDH

Min  Max

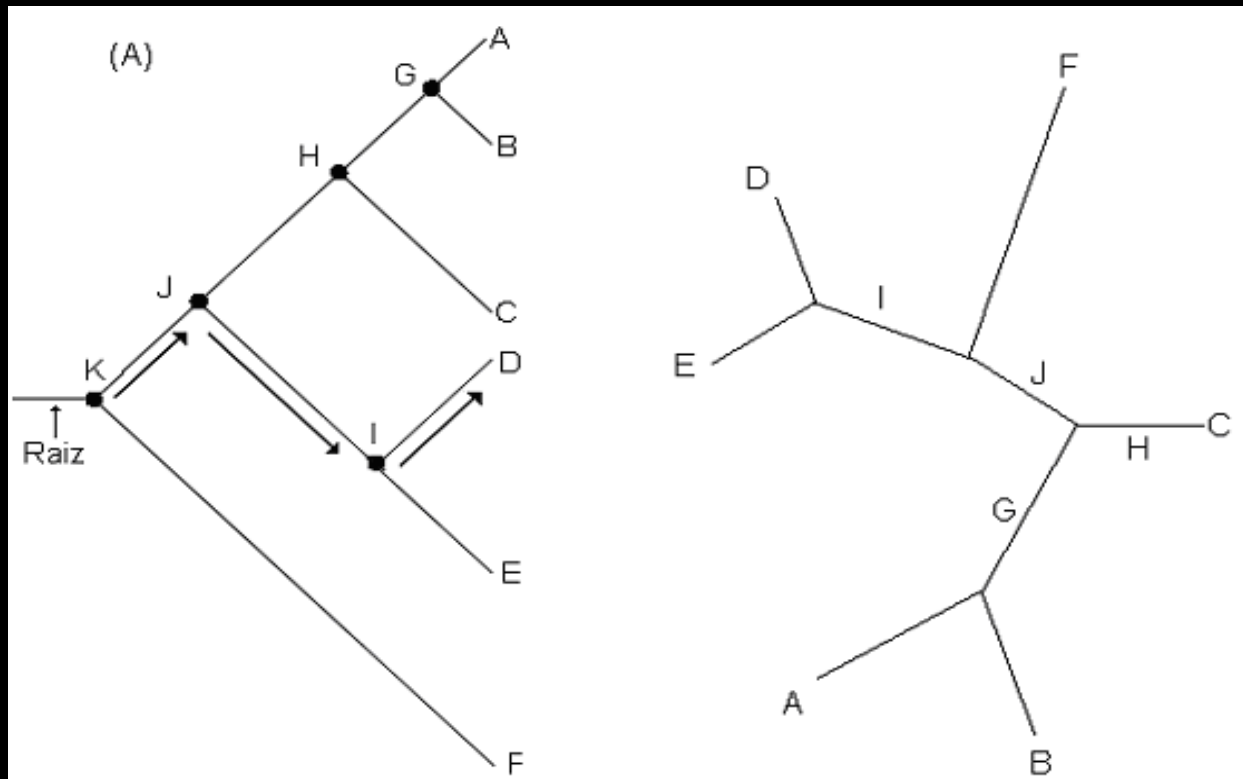


Exemplo de Projeção: Votação

Min  Max

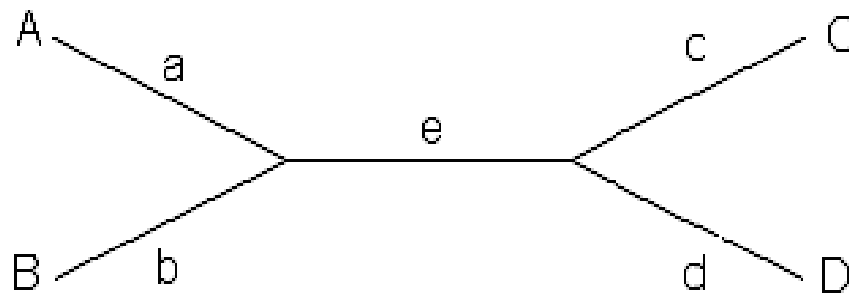


Point Placement by Phylogenetic Tree Construction Algorithms (N-J Trees)



Point Placement by Phylogenetic Tree Construction Algorithms (N-J Trees)

$$d_{AB} + d_{CD} \leq \max(d_{AC} + d_{BD}, d_{AD} + d_{BC})$$



$$d_{AB} = a + b$$

$$d_{AC} = a + e + c$$

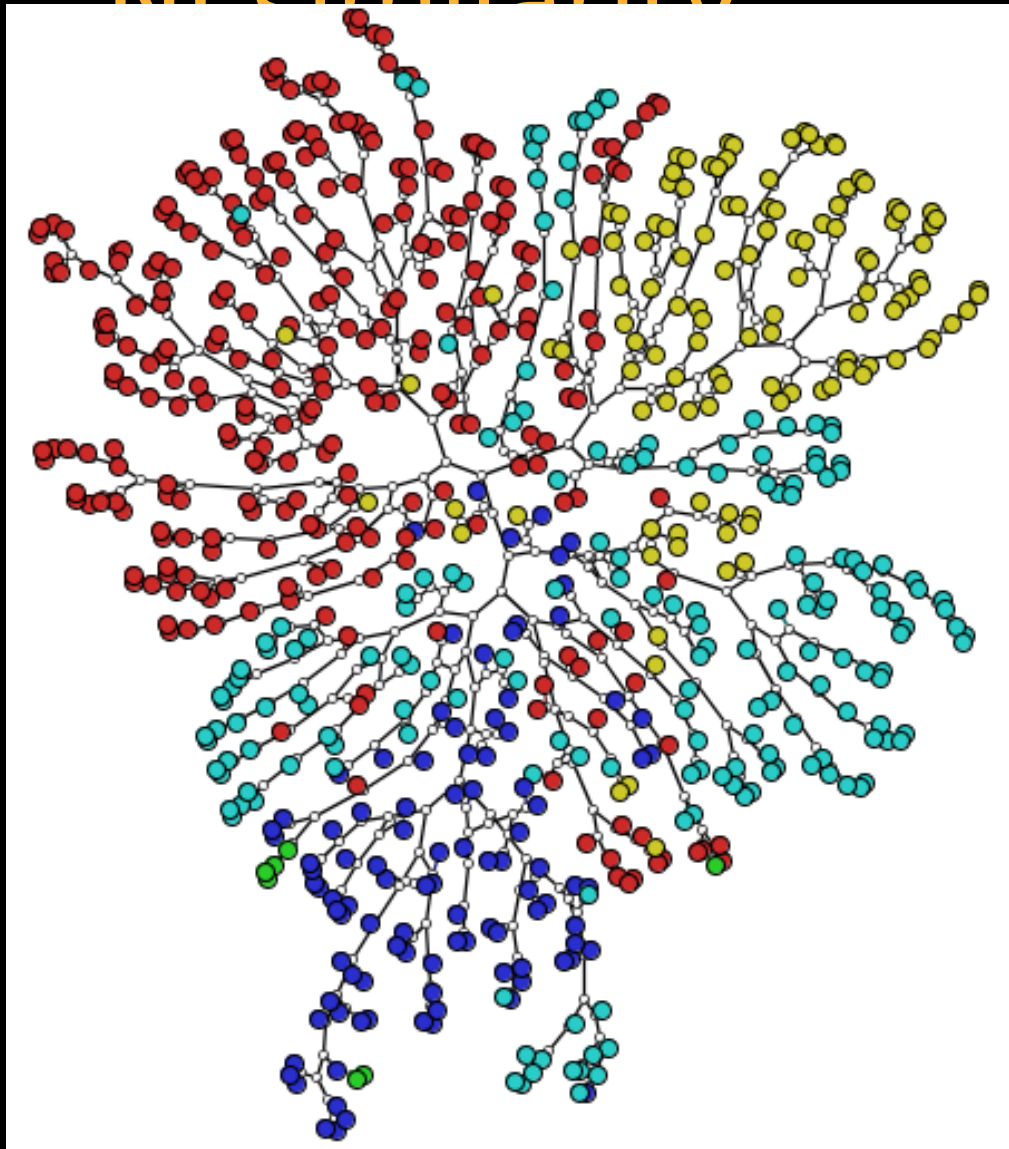
$$d_{AD} = a + e + d$$

$$d_{CD} = c + d$$

$$d_{BD} = b + e + d$$

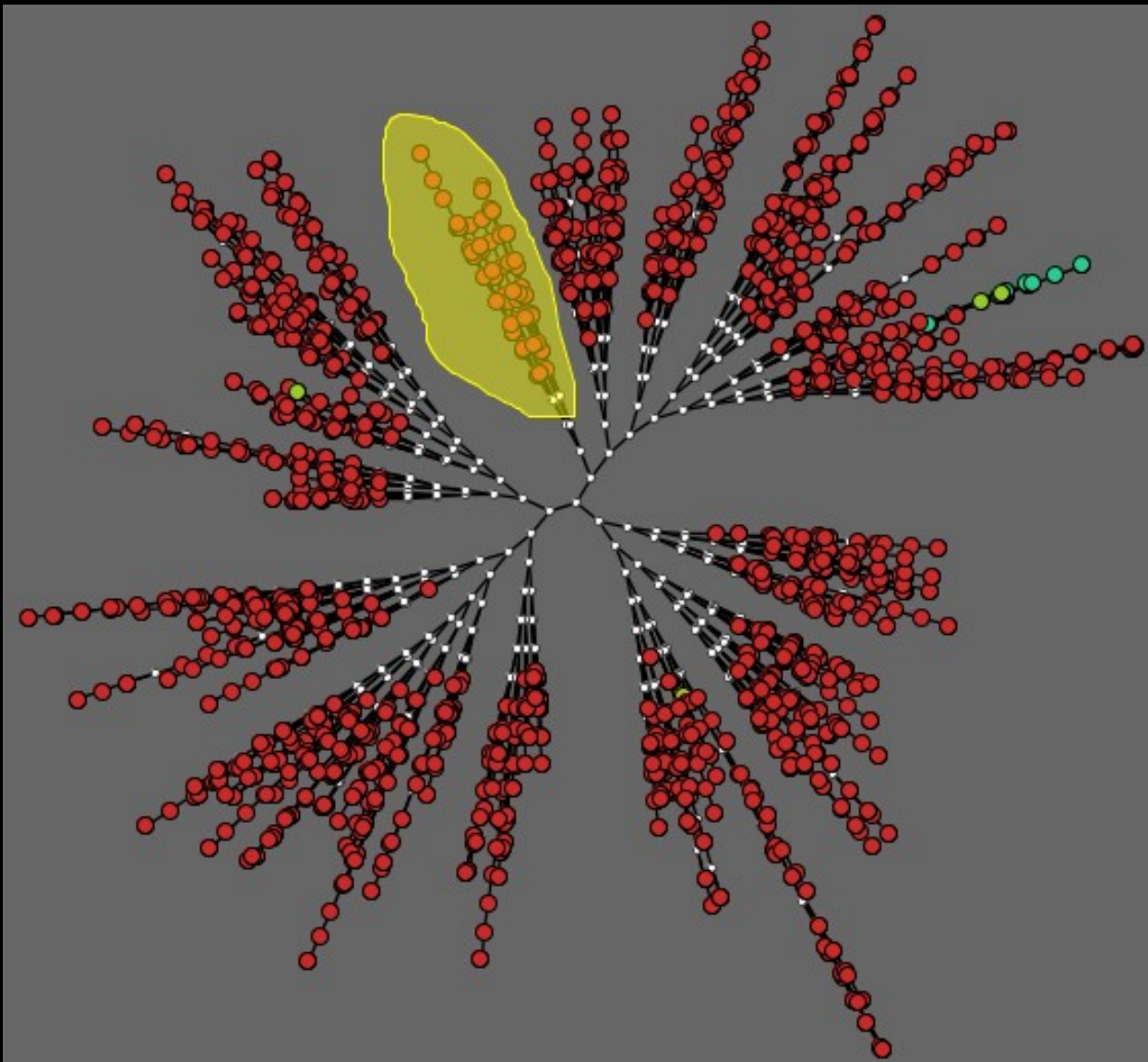
$$d_{BC} = b + e + a$$

MI similarity

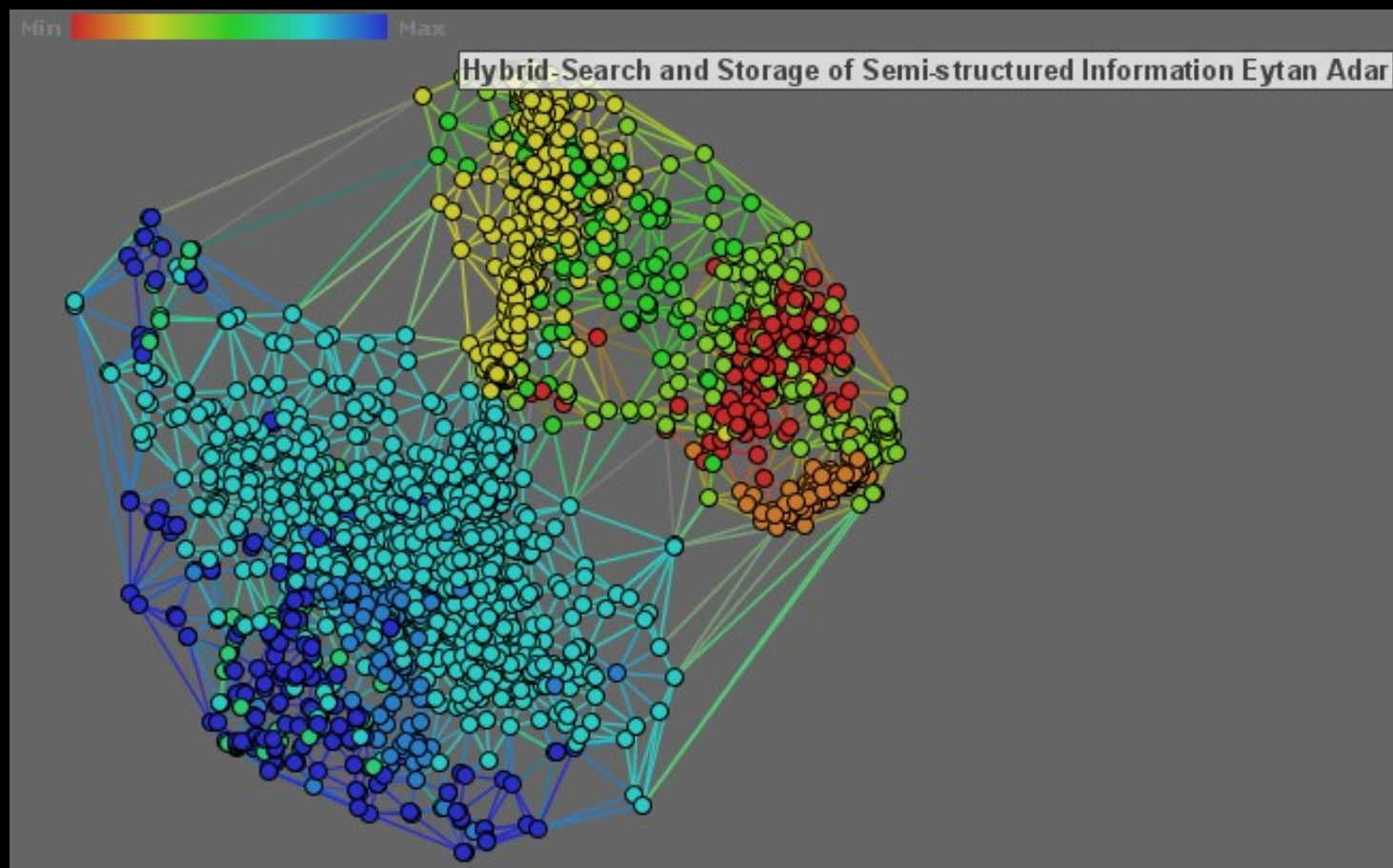


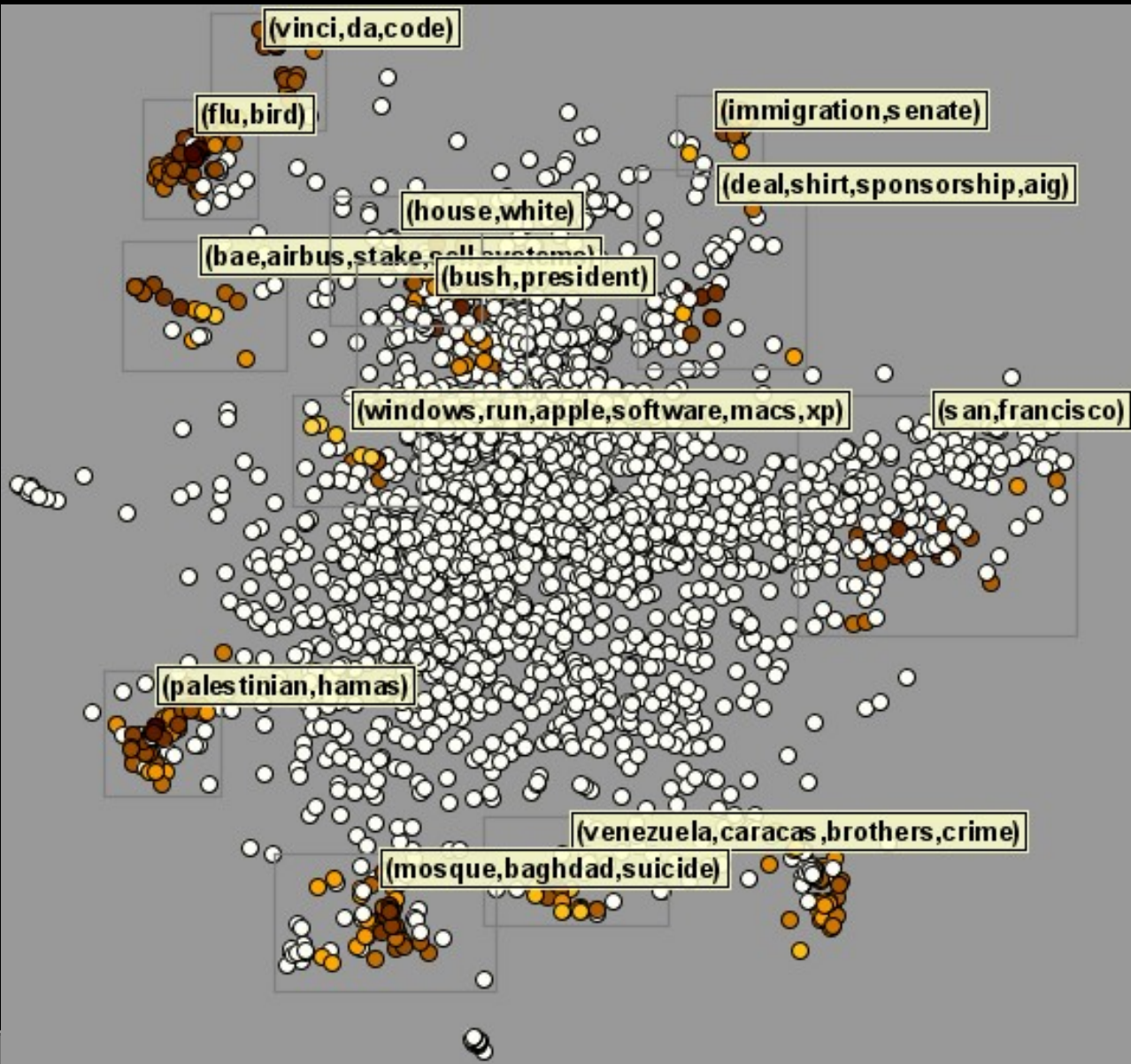
- Alternate view (N-J Tree)





Exploration





- Detailing topics



bird:[flu]<-[bird],(61/81%;97%):swan:
flu:[flu]<-[bird],(46/82%;100%):swan, dead:

jeffrey:[enron]<-[jeffrey],(13/10%;100%)::

francis co:[san]<-[francis co],(13/11%;100%)::

england:[day]<-[england],(8/8%;62%):cochin, india

seventh:[inning]<-[seventh],(7/7%;88%)::

england:[season]<-[england],(5/5%;38%)::

hamas:[palestinian]<-[hamas],(35/56%;90%):governm

victory:[night]<-[victory],(5/5%;56%)::

israeli:[israeli]<-[gaza],(7/29%;78%):palestinians:

night:[night]<-[hit],(5/5%;36%):hits:

form:[form]<-[quickly],(4/27%;100%):ehud, katsav, olm

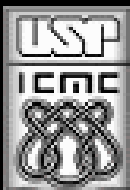
night:[game]<-[night],(5/5%;33%)::

couric:[couric]<-[katie],(7/10%;100%)::

gagne:[gagne]<-[dodgers],[elbow],[nerve],[pitchin

happily:[stevens]<-[happily],(3/5%;100%)::

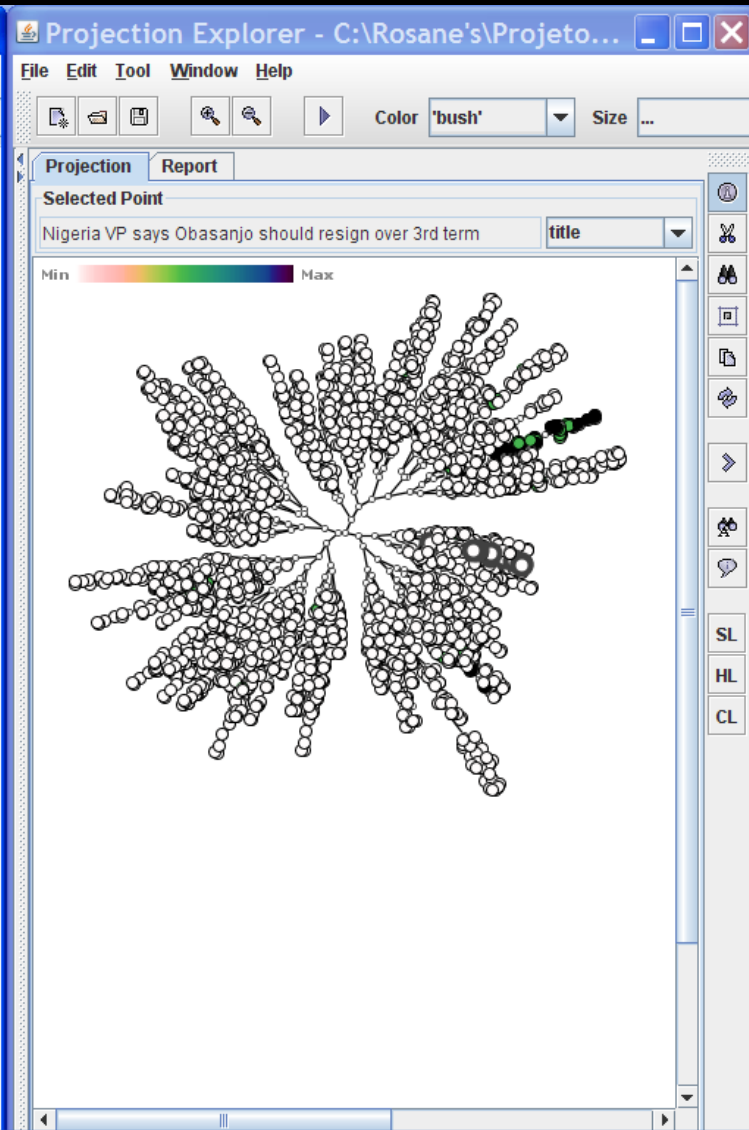
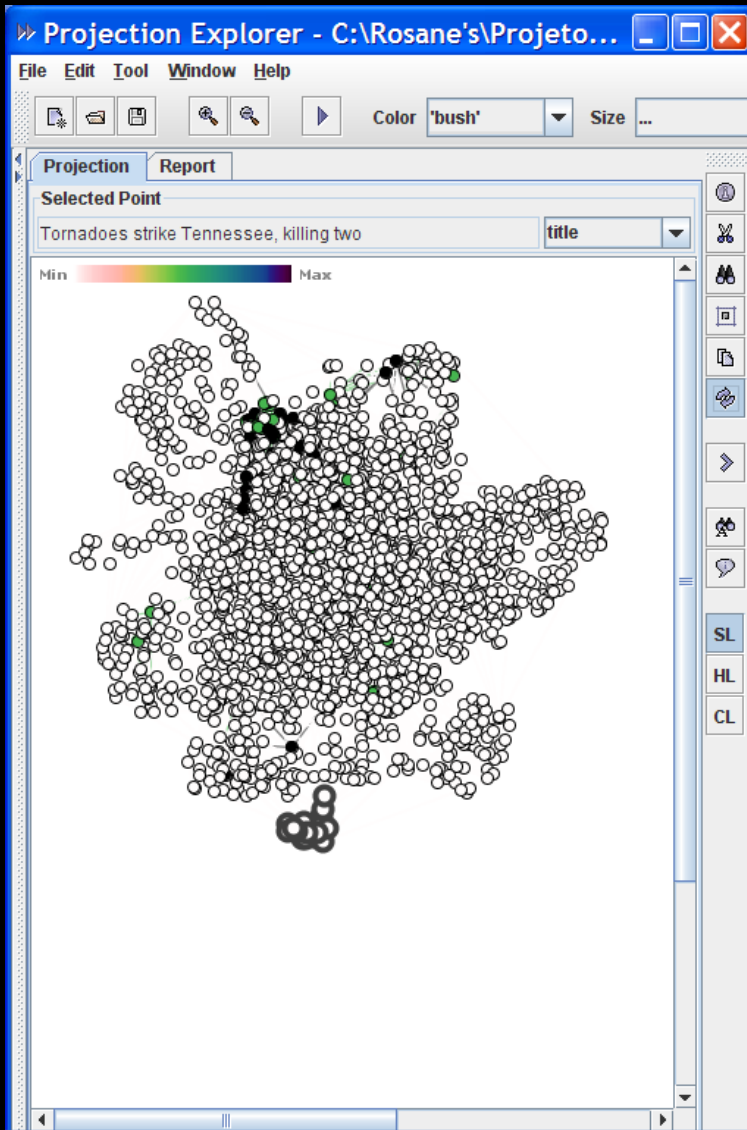
olympics:[begins]<-[olympics],(3/5%;100%)::





- Finding Relationships

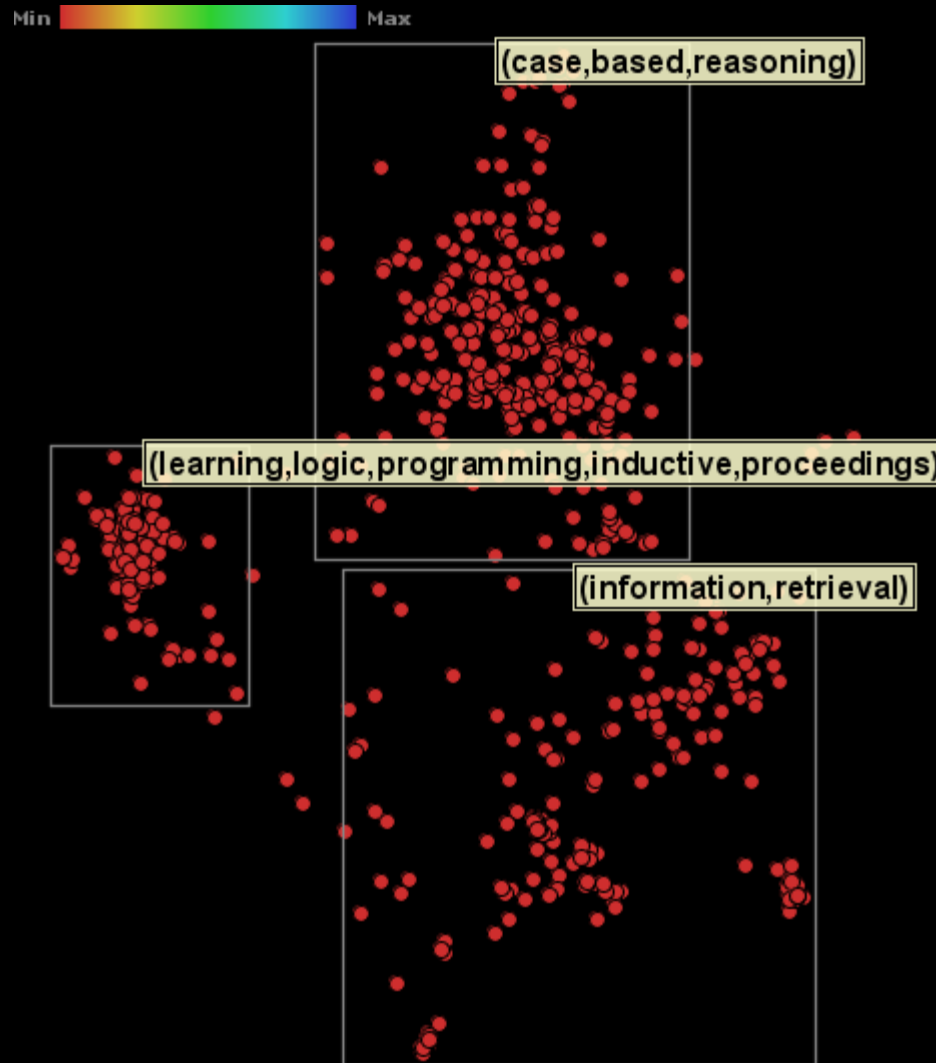




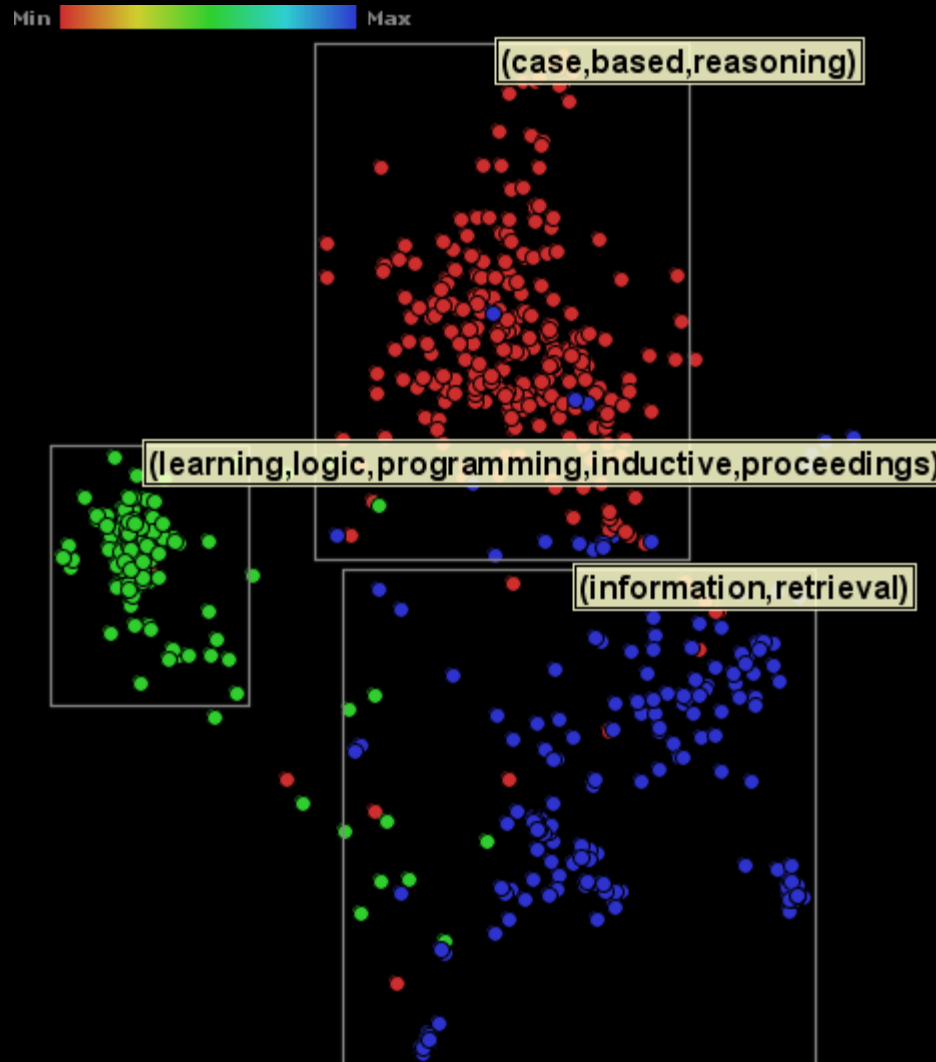
- Building a Surface



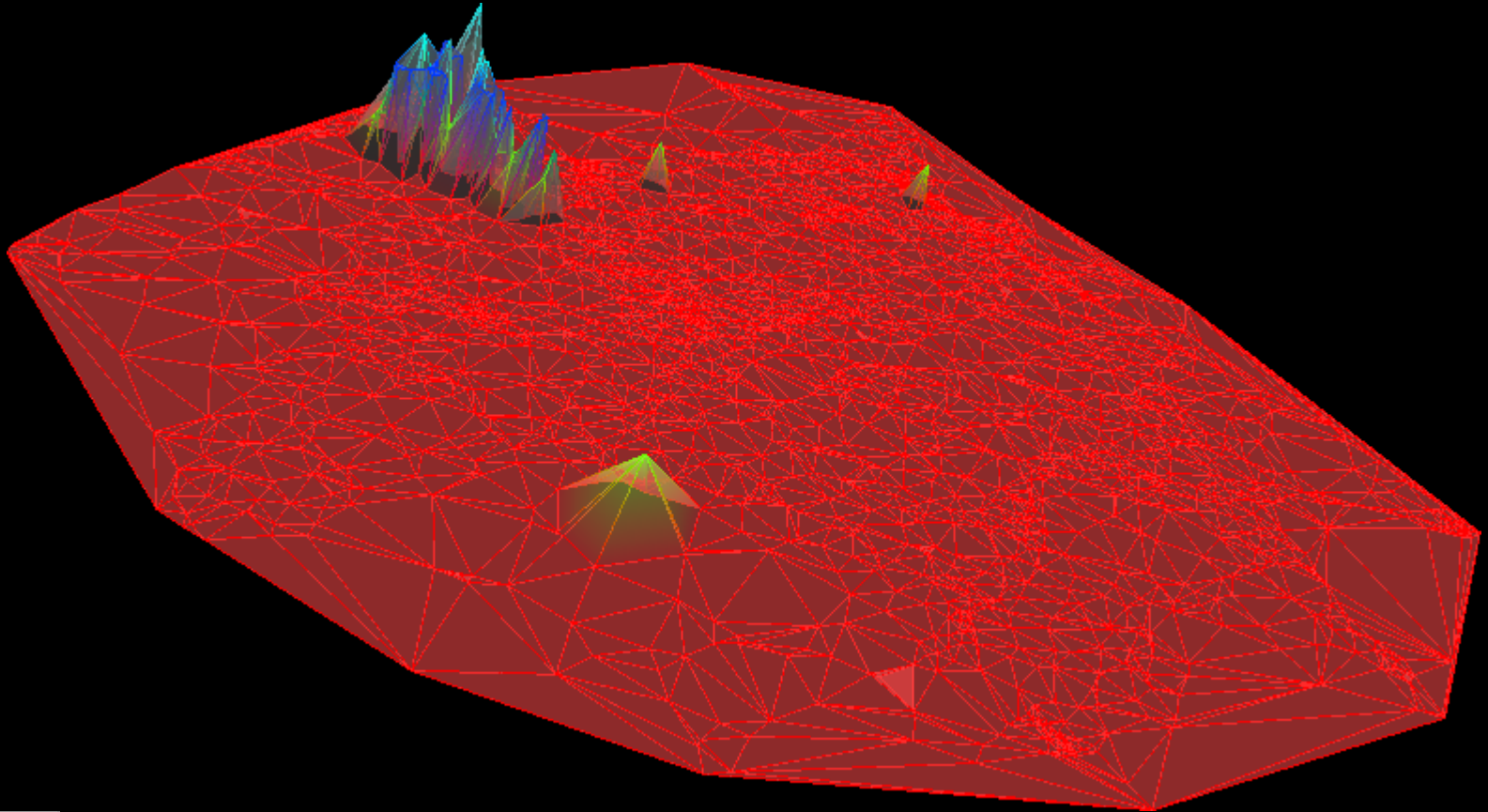
Exemplos de Mapas



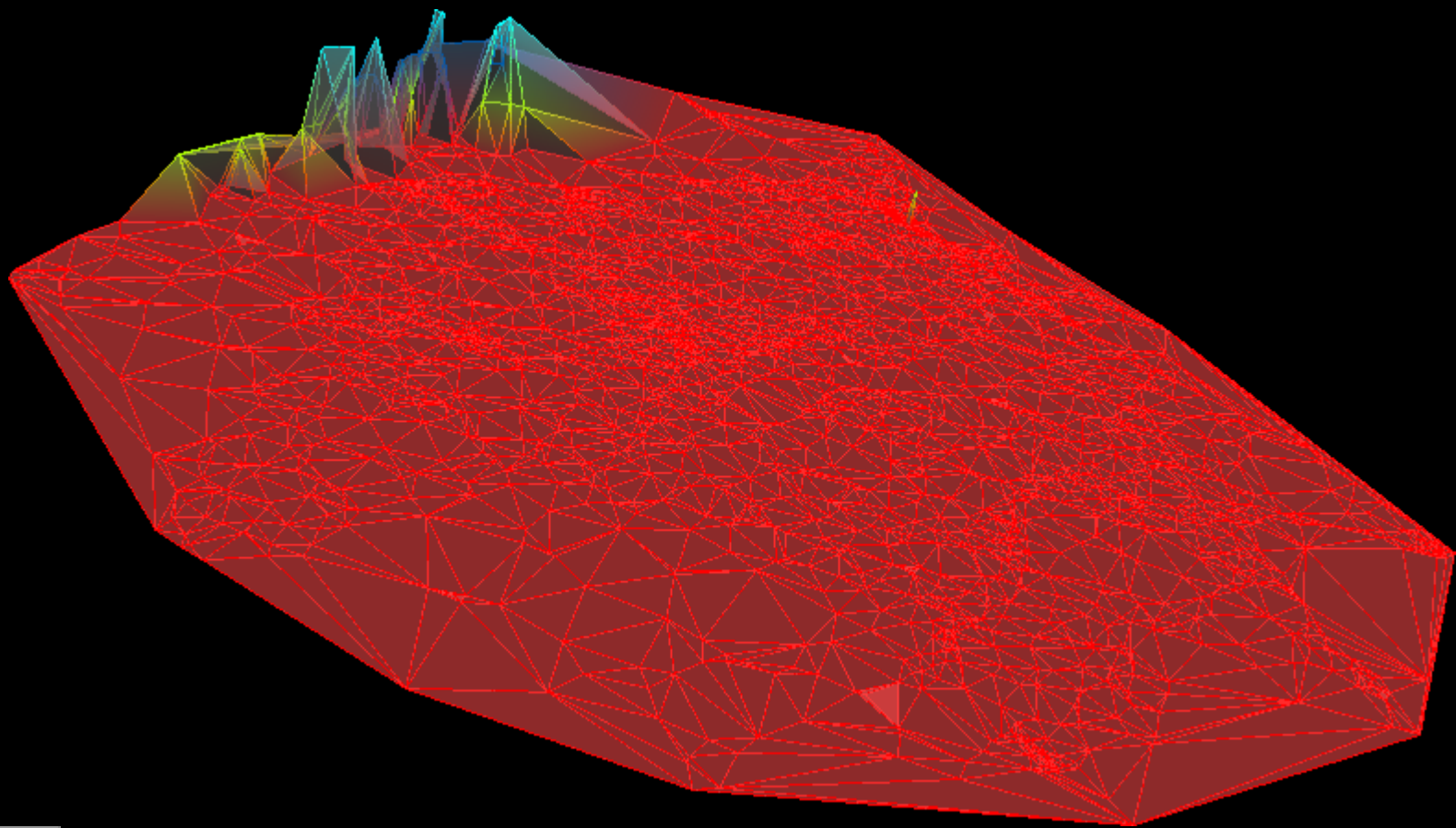
Exemplos de Mapas



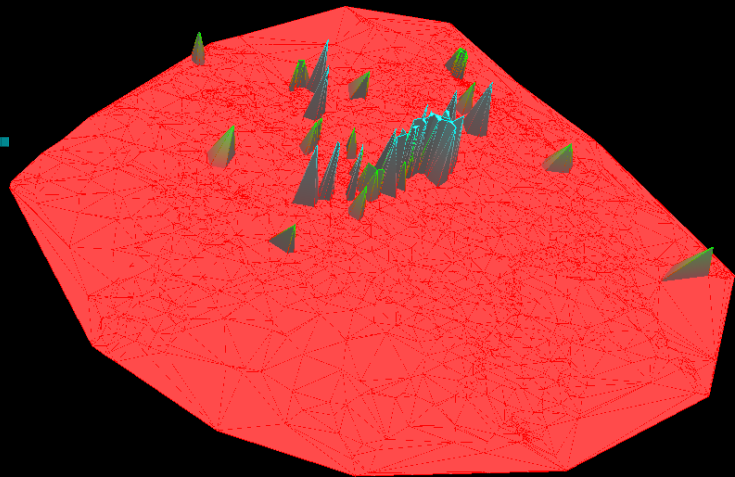
RSS News Flash



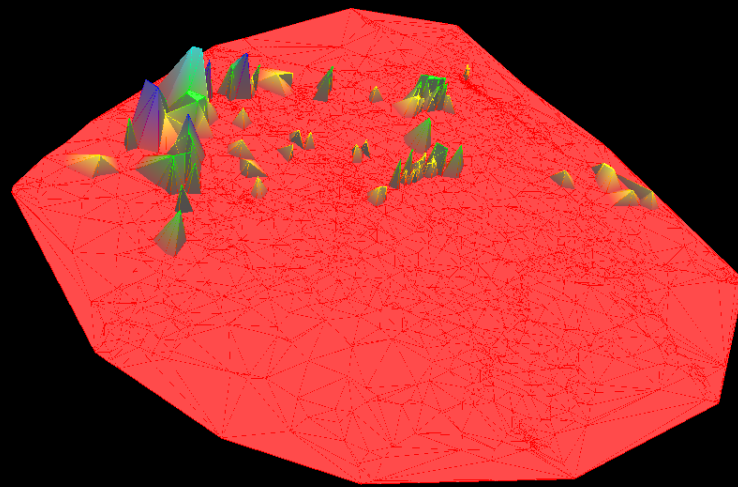
Bird and Flu



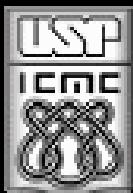
Palestinian

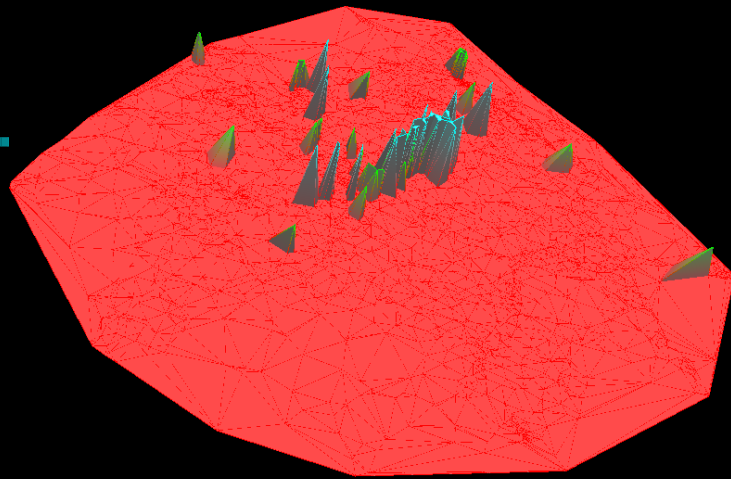


Bush

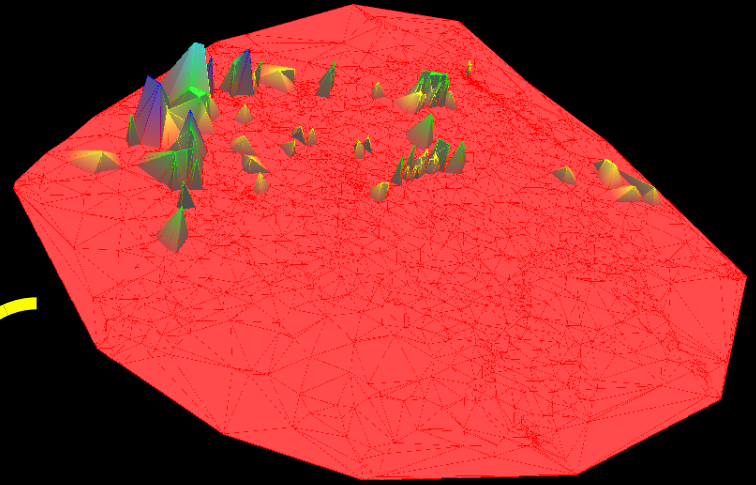


Iraq

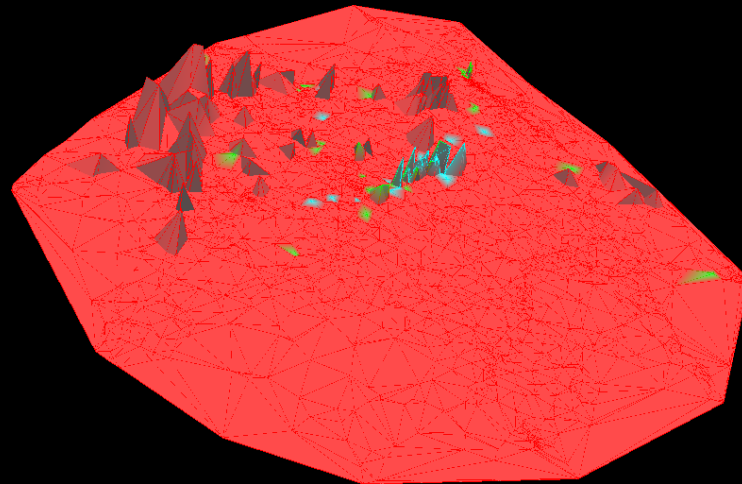
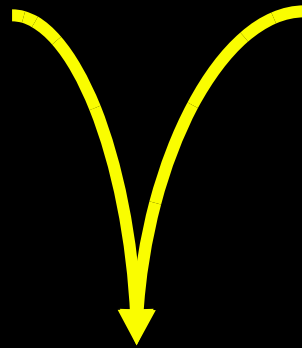




Bush



Iraq



Generalizando o processo: Séries Temporais – Vazão em Hidrelétricas

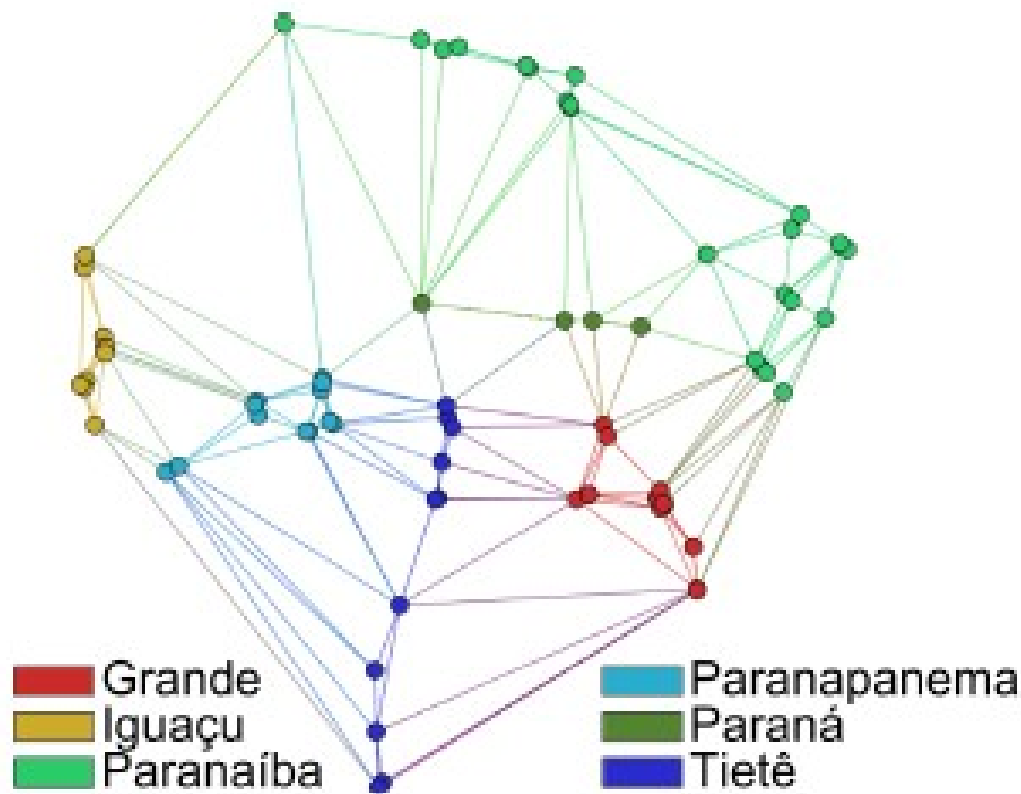


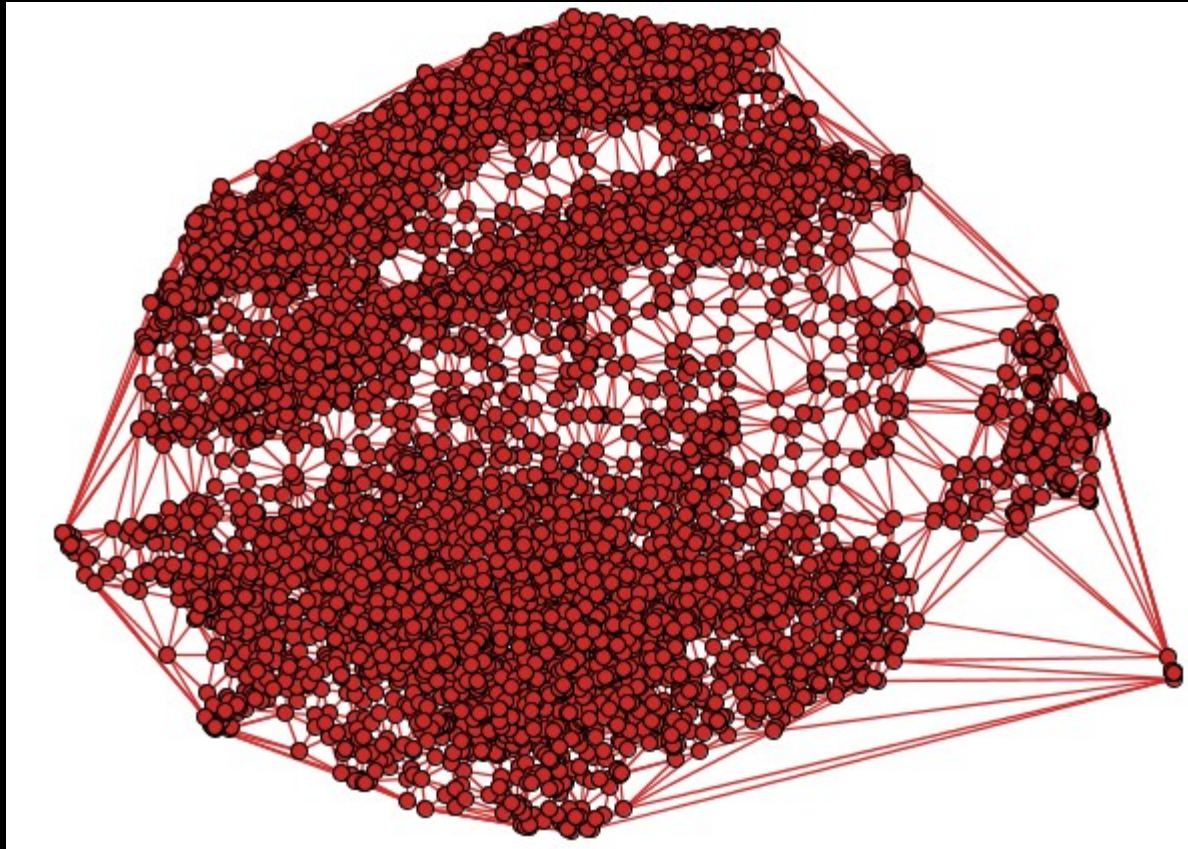
Figure 2. Power plants of the basin Paraná

Text from attributes

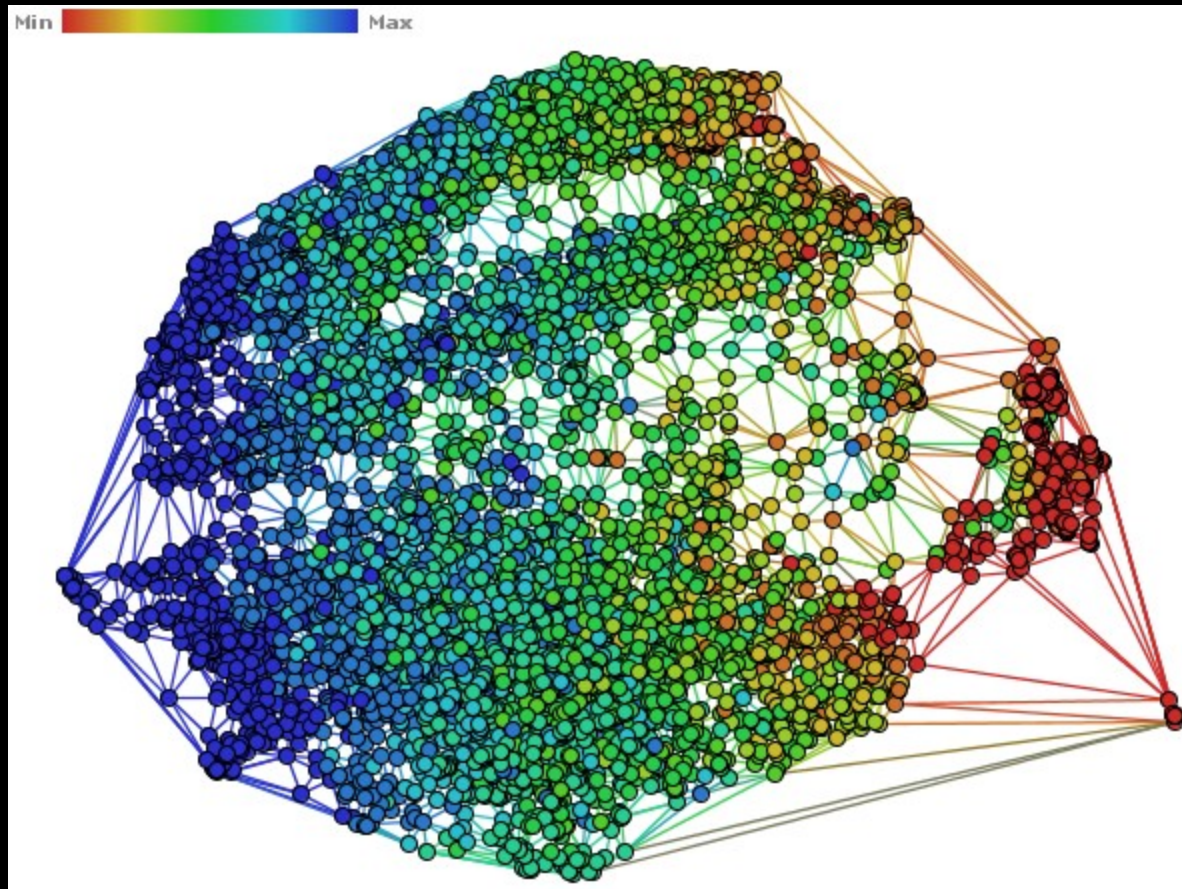
- Cattle performance data
 - Translated to text from categorical information, e.g.,
 - Ranges of weight to words such as:
{weight_below_fifty_percent;
weight_between_fifty_seventy_five; etc..}
 - 9135 individuals



Cattle performance data

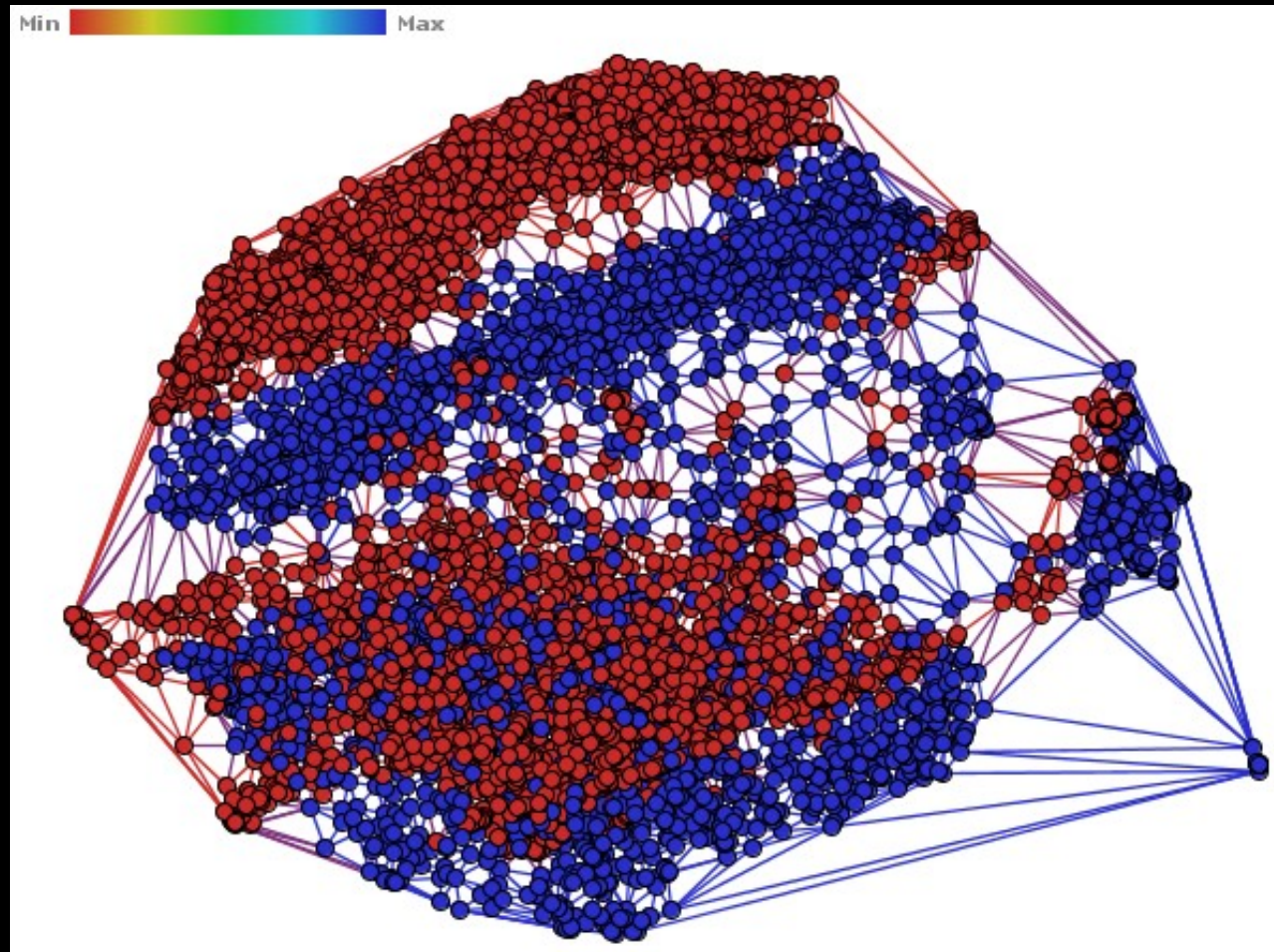


Cattle performance data



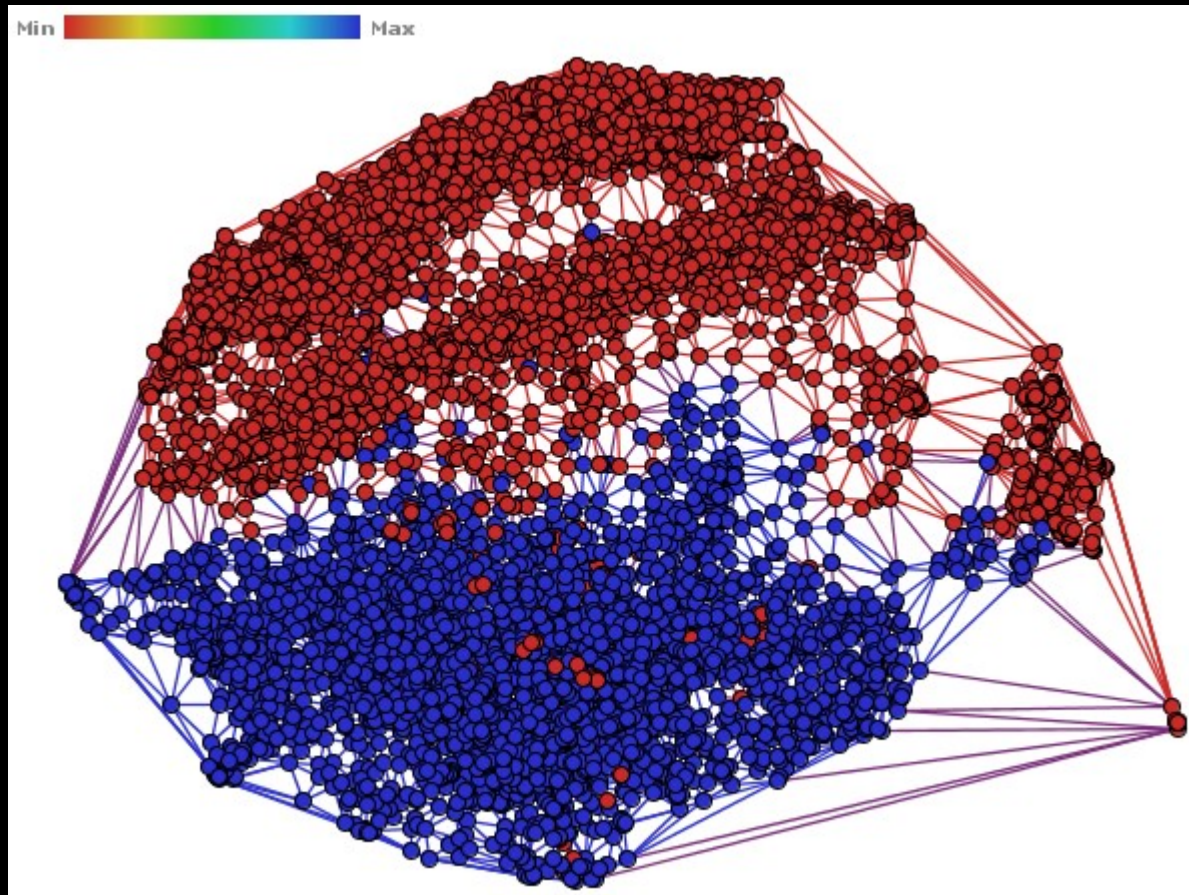
Colored
by word
'top'

Cattle performance data



Colored
by
female

Cattle performance data



Colored
by farm