

# Minería Visual de Grandes Volúmenes de Información

Dr. Alfredo Sánchez

<http://ict.udlap.mx>

Universidad de las Américas Puebla



**UDLAP**  
UNIVERSIDAD DE LAS  
AMÉRICAS PUEBLA

AMÉRICAS

**CUDI 2013**  
REUNIÓN DE OTOÑO  
CAMPECHE

Fecha



tematicas de

# Interacción Humano-Computadora





# Comunidad CUDI de Interacción Humano-Computadora





# Participantes - Taller IHC 2013

1. CIE-Mochis
2. Emerson Process Mgmt.
3. IMT
4. IT-Dgo (2)
5. IT-Saltillo
6. ITESM-Qro (2)
7. U Guanajuato (3)
8. UAA
9. UABC
10. UAM-C (3)
11. UANL (2)
12. UAQ
13. UASLP (7)
14. UATlax
15. UDLAP (9)
16. UNACAR (2)
17. UNCA
18. Unicauca
19. UPSLP (3)
20. UTM (3)
21. UV (5)



- Día Virtual de Sistemas Colaborativos
  - 18 de octubre 2013, 10:00 AM
  - Ponentes de México, España, Colombia, Argentina
- Día Mundial de la Usabilidad
  - 14 de noviembre
  - UANL, #WUDUANL
  - [usabilidad.uanl.mx](http://usabilidad.uanl.mx)

La Universidad Autónoma de Nuevo León a través de la Dirección General de Informática lo invita cordialmente a celebrar con nosotros et

## DÍA MUNDIAL DE LA USABILIDAD UANL 2013

CONFERENCIAS CON EXPERTOS NACIONALES E INTERNACIONALES.

SUSAN DRAY, Ph.D. | DRA. MÓNICA TENTORI | DR. CESAR CÁRDENAS

LIC. VERÓNICA TRAYNOR | LIC. NATALIA VIVAS | LIC. SELENE CASTILLA

14 DE NOVIEMBRE 2013

REGISTRO  
<http://usabilidad.uanl.mx>

EVENTO GRATUITO

#DíaMundialUsabilidad  
WUDUANL

SALA POLIVALENTE 1 Y 2, FACULTAD DE ORGANIZACIÓN DEPORTIVA, UANL  
HORA: 8:30 - 14:00 HRS. - INFORMES Y REGISTRO: USABILIDAD@UANL.MX  
#USABILIDAD@UANL.MX

Rosenfeld | axura | uxpa | Diseñando apps para móviles

DGI | UANL | World Usability Day



# Agenda

- Comprensión de colecciones
- Visualización de información
- Minería visual
- Visualización para la comprensión de colecciones
- Avances en el Laboratorio ICT (UDLAP)
- Perspectivas



# Colección digital

Conjunto de documentos u objetos digitales seleccionados y disponibles a comunidades de usuarios

## ACM Proceedings

### Dynamic Distributed Data-Intensive Applications, Programming Abstractions, and Systems

- 3DAPAS '11: Proceedings of the 2011 workshop on Dynamic distributed data-intensive applications, programming abstractions, and systems

### 3D Object Retrieval

- 3DOR '10: Proceedings of the ACM workshop on 3D object retrieval

### 3D Video Processing

- 3DVP '10: Proceedings of the 1st international workshop on 3D video processing

### Advances in Model-Based Testing

- A-MOST '07: Proceedings of the 3rd international workshop on Advances in model-based testing
- A-MOST '05: Proceedings of the 1st international workshop on Advances in model-based testing

### Amrita ACM-W Celebration on Women in Computing in India

- A2CWIC '10: Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India

### Advanced Architectures and Algorithms for Internet Delivery and Applications

- AAA-IDEA '06: Proceedings of the 2nd international workshop on Advanced architectures and algorithms for internet delivery and applications

### Automated analysis-driven debugging

- AADEBUG'05: Proceedings of the sixth international symposium on Automated analysis-driven debugging

### Autonomous Agents and Multiagent Systems

- AAIAS '12: Proceedings of the 2012 international conference on Autonomous agents and multi-agent systems





## Human-Computer Interaction Design in Product Design



Xianyi Yang ; Guo Chen

Education Technology and Computer Science, 2009. ETCS '09.

First International Workshop on

Volume: 2

Digital Object Identifier: 10.1109/ETCS.2009.359

Publication Year: 2009 , Page(s): 437 - 439

**IEEE CONFERENCE PUBLICATIONS**



Quick Abstract



PDF (436 KB)



HTML



## Research on human-computer interaction design for distance education websites



Peng Bian ; Yi Jin ; Nairen Zhang

Computer Science and Education (ICCSE), 2010 5th

International Conference on

Digital Object Identifier: 10.1109/ICCSE.2010.5593511

Publication Year: 2010 , Page(s): 716 - 719

**IEEE CONFERENCE PUBLICATIONS**



Quick Abstract



PDF (408 KB)



## Describing the interaction order: pre-theoretical practice for a science of human computer interaction



Frohlich, D.M.

Theory in Human-Computer Interaction (HCI), IEE Colloquium



## Collection Highlights



American Memory



Prints & Photographs



Historic Newspapers



Performing Arts



Veterans History



Sound Recordings

### Topics

[American History](#)

[Arts & Culture](#)

[Government, Politics & Law](#)

[Maps & Geography](#)

[News & Journalism](#)

[Religion & Philosophy](#)

[Science, Technology & Business](#)

[Sports & Leisure](#)

[World History & Cultures](#)

### Especially for

[Kids & Families](#)





*Buscar en las colecciones*

Junio 2013



12 Instituciones Mexicanas



18 repositorios



106,169

Libros | Tesis | Artículos  
Revistas | Videos | Imágenes



## Red Mexicana de Repositorios Institucionales

El objetivo general de REMERI es integrar una red federada de Repositorios de acceso abierto de las Instituciones Mexicanas de Educación Superior (IES) con la finalidad de integrar, difundir, preservar y dar



Para aprovechar mejor las colecciones,  
necesitamos comprender su contenido

(temas, autores, fechas,  
instituciones, inter-relaciones,...)

# Entendimiento “bottom-up”



¿Qué temas hay en las colecciones de ReMeRI?

- Métodos de recuperación de información
- Se requiere un esfuerzo significativo



# Entendimiento “bottom-up”

2, 3 Y 4 DE OCTUBRE

Google

¿Qué temas hay en las colecciones de ReMeRI?

Web

Images

Maps

Shopping

More ▾

Search tools

About 27,200 results (0.50 seconds)

Did you mean: [¿Qué temas hay en las colecciones de ReMeRAS?](#)

[Visualizando el universo « Investigación con Vida : UDLAP](#)

[blog.udlap.mx/.../2012/12/.../visualizando-el-universo...](#) ▾ [Translate this page](#)

Dec 9, 2012 - Teniendo en cuenta **que** la información es un recurso invaluable en las ... la cual **tiene** como objetivo crear un canal entre la información y ... para el caso concreto de la **colección ReMeRI** sería navegar por **temas** y subtemas.

[Investigación con Vida : UDLAP](#)

[blog.udlap.mx/investigacionconvida/](#) ▾ [Translate this page](#)

La tercera es la red Infotainment **que tiene** como objetivo brindar a los .... **que** para el caso concreto de la **colección ReMeRI** sería navegar por **temas** y ...

[Biblioteca Digital UA de C - Infoteca Central Saltillo](#)

[www.infosal.uadec.mx/cie/](#) ▾ [Translate this page](#)

**COLECCIÓN DE CD'S** para consultar en CIE Infoteca Saltillo. .... de dichas revistas están especializadas en **temas** relacionados con la Biblioteconomía, .... Índice de revistas latinoamericanas (alojamiento UNAM), **hay que** buscar el título de la .... Red Mexicana de Repositorios Institucionales (**REMERI** ), Red federada de ...

[Biblioteca Digital UAdeC - Infoteca Central Saltillo](#)

[www.infosal.uadec.mx/cie/new/menuind.php](#) ▾ [Translate this page](#)

# Entendimiento “bottom-up”

www.remeri.org.mx

Portal de la Red Mexicana de Repositorios Institucionales

Principal

Proyecto

Servicios

Capacitación

Participantes

Información



Indixe

Buscar en las colecciones

## Sistema de Consulta Indixe

Resultados de la búsqueda : temas (936) de 1 a 25

Lista de recursos ordenados por relevancia (relevancia, fecha, autor, institución, tipo, título )

ascendente - [descendente](#)

mostrando 25 (25, 50, 100, 200 )

[Anterior](#) - [Siguiete](#)



1.

Tesis de Maestría

[\\*ALGUNOS TEMAS CONTROVERTIDOS EN MATERIA DE IETU\\*](#)

GARCIA VIZCAYA, SAIRA IVONNE.

IPN (2010)

<http://tesis.bnct.ipn.mx/dspace/handle/123456789/7987>



2.

Tesis de Maestría

[\\*INVESTIGACION Y POSGRADO DE LA INGENIERIA HIDRAULICA EN MEXICO. PROPUESTA DE TEMAS DE INVESTIGACION E INDICADORES\\*](#)

VACA BADILLO, ISAAC.

IPN (2008)



- Busca cambiar el enfoque de la recuperación de objetos específicos a ofrecer una vista completa de cada colección

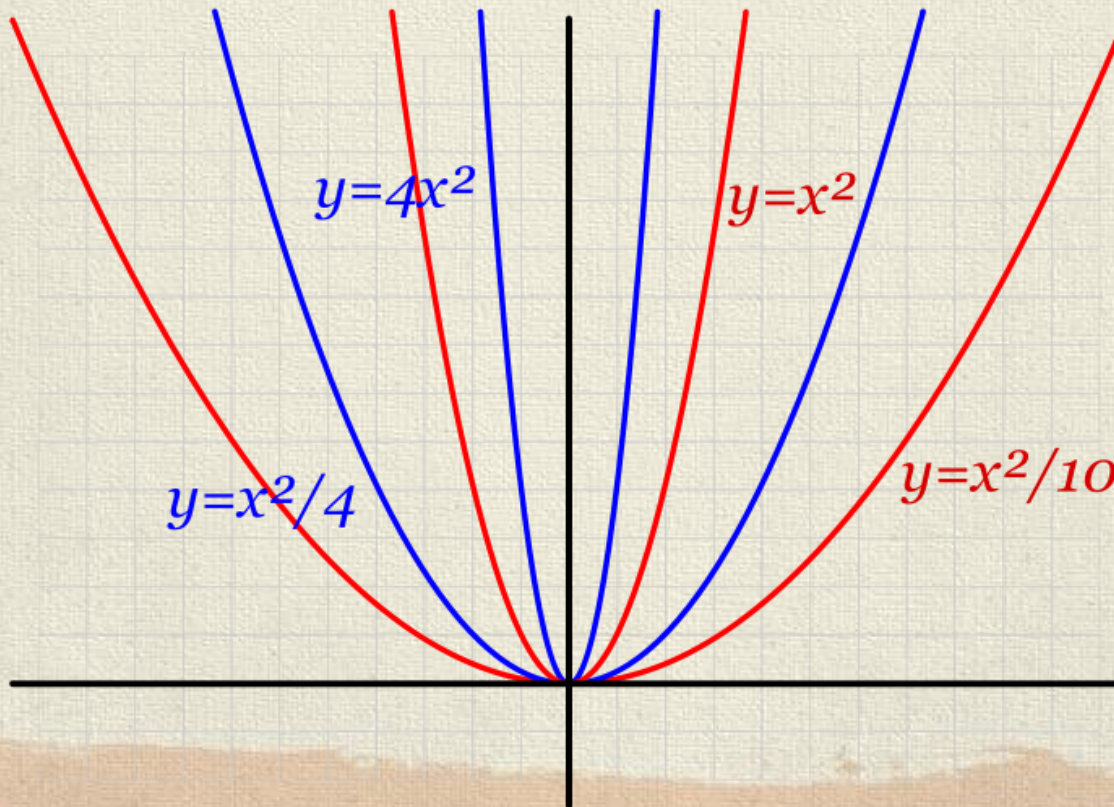
[Chang et al. 2004]



La comprensión de colecciones  
puede apoyarse en la  
**visualización de información**

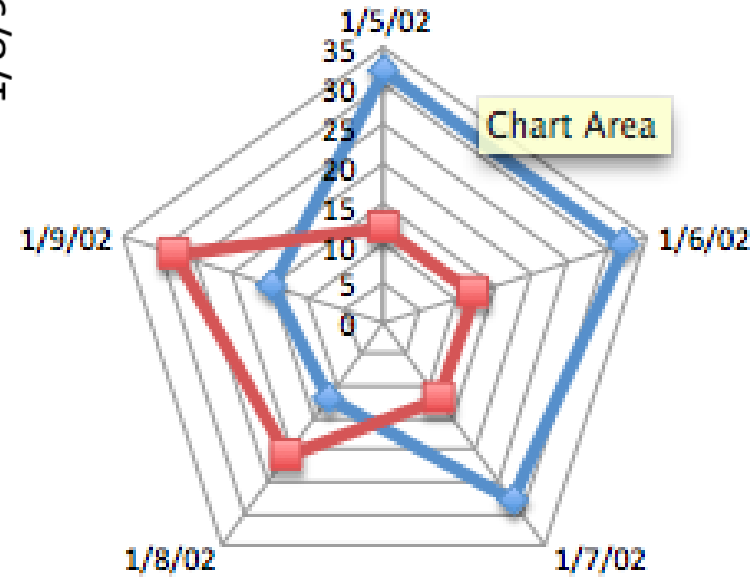
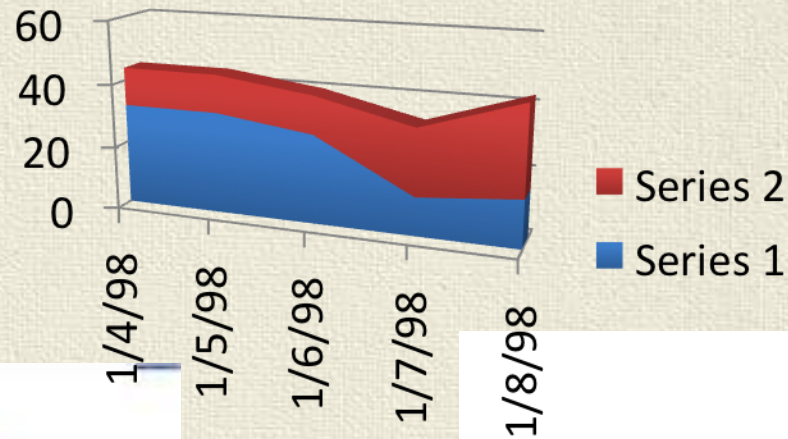


- Transformación de datos simbólicos a representaciones geométricas



# Visualización de información

2, 3 Y 4 DE OCTUBRE



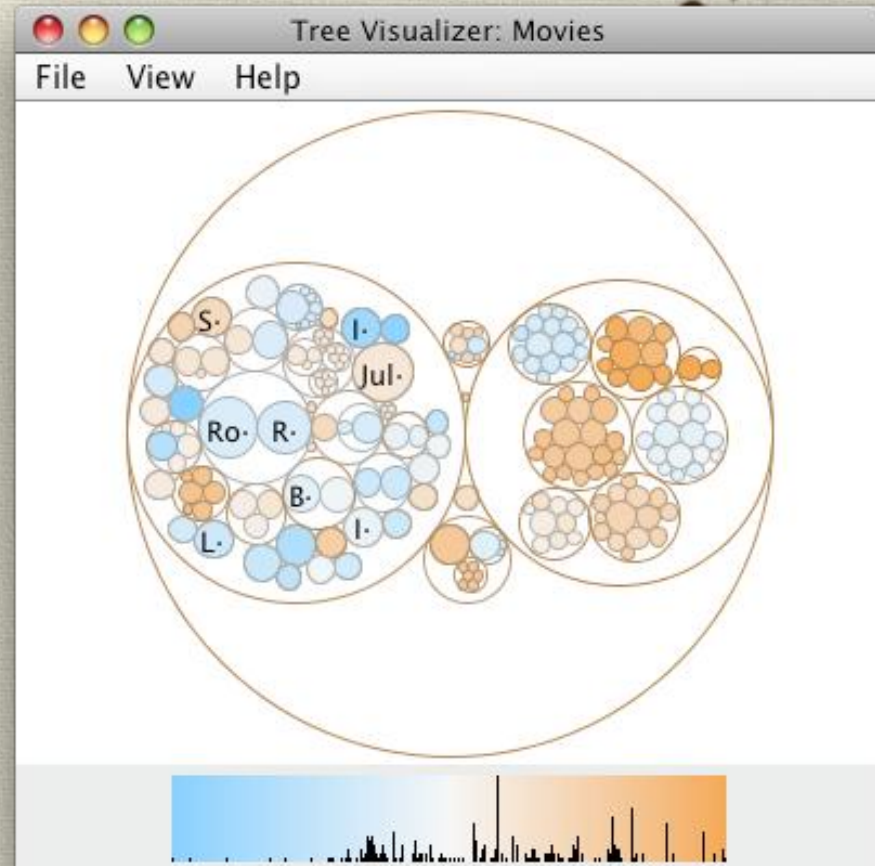


# Visualización de grandes volúmenes de información



# TreeMaps

2, 3 Y 4 DE OCTUBRE



[Shneiderman 1992]






# Campos de estrellas

Popularity

**Witches of Eastwick, The**  
 Director: **Miller, George** Year: 1987  
 Country: USA Language: English  
 Actors: **Pfeiffer, Michelle**  
 Nicholson, Jack  
 Jenkins, Richard  
 Joakum, Keith  
 Struycker, Carel  
 Cher  
 Sarandon, Susan  
 Cartwright, Veron



Title : ALL  
 ABCDFGHLMPNSTWZ

Actor : ALL  
 ABCDFGHJKLM PRS TWZ

Actress : Pfeiffer, Michelle  
 ABCDFGHJKLM PRS TWZ


Director : Miller, George  
 ABCDFGHJKLM PRS TWZ

105 Length 231

0 450

Ratings  G  PG  
 PG-13  R

Films Shown: 210



Copyright (C) 1993 HCIL

Year of Production

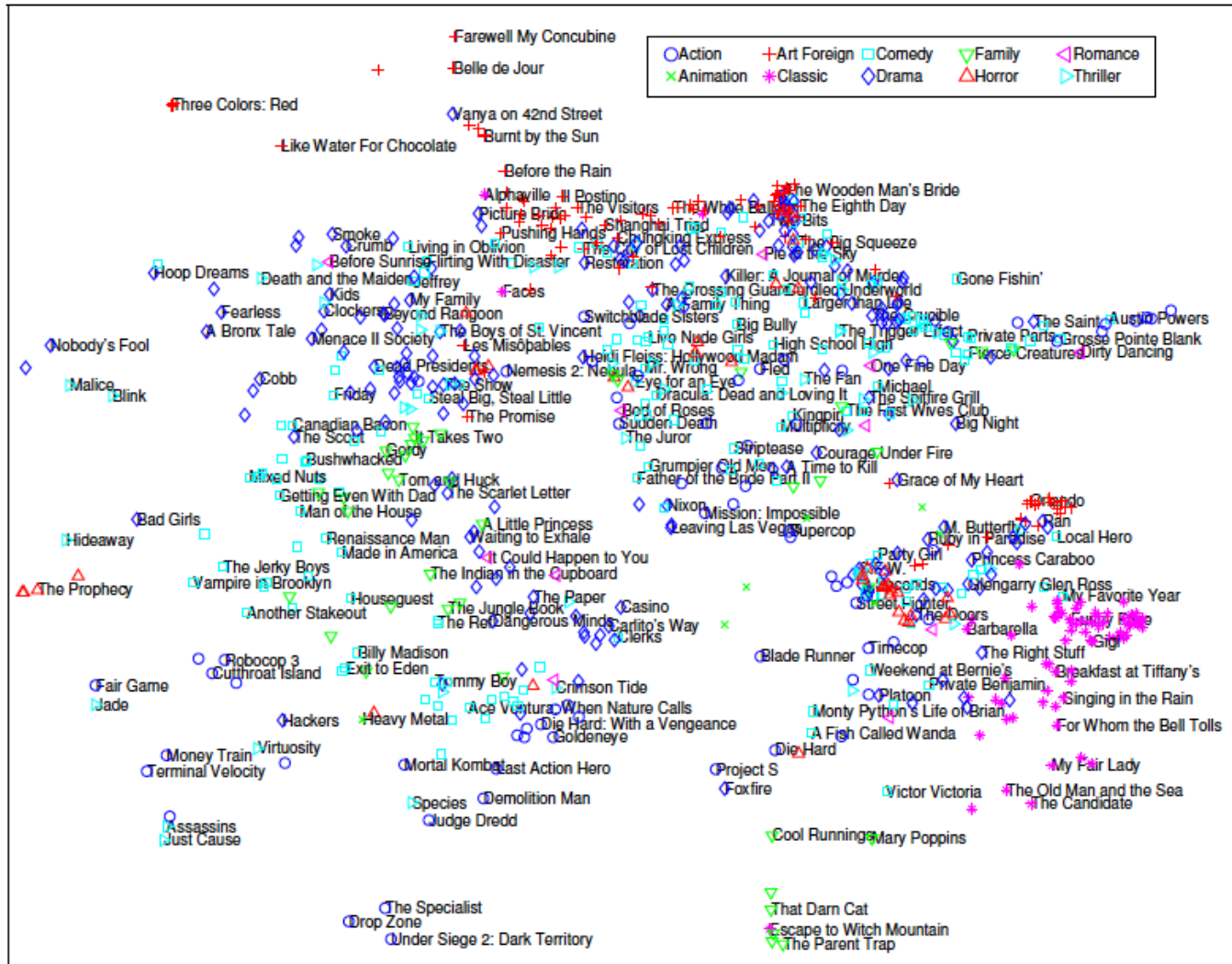
TUBRE

[HCIL UMD]



# Puntos

2, 3 Y 4 DE OCTUBRE

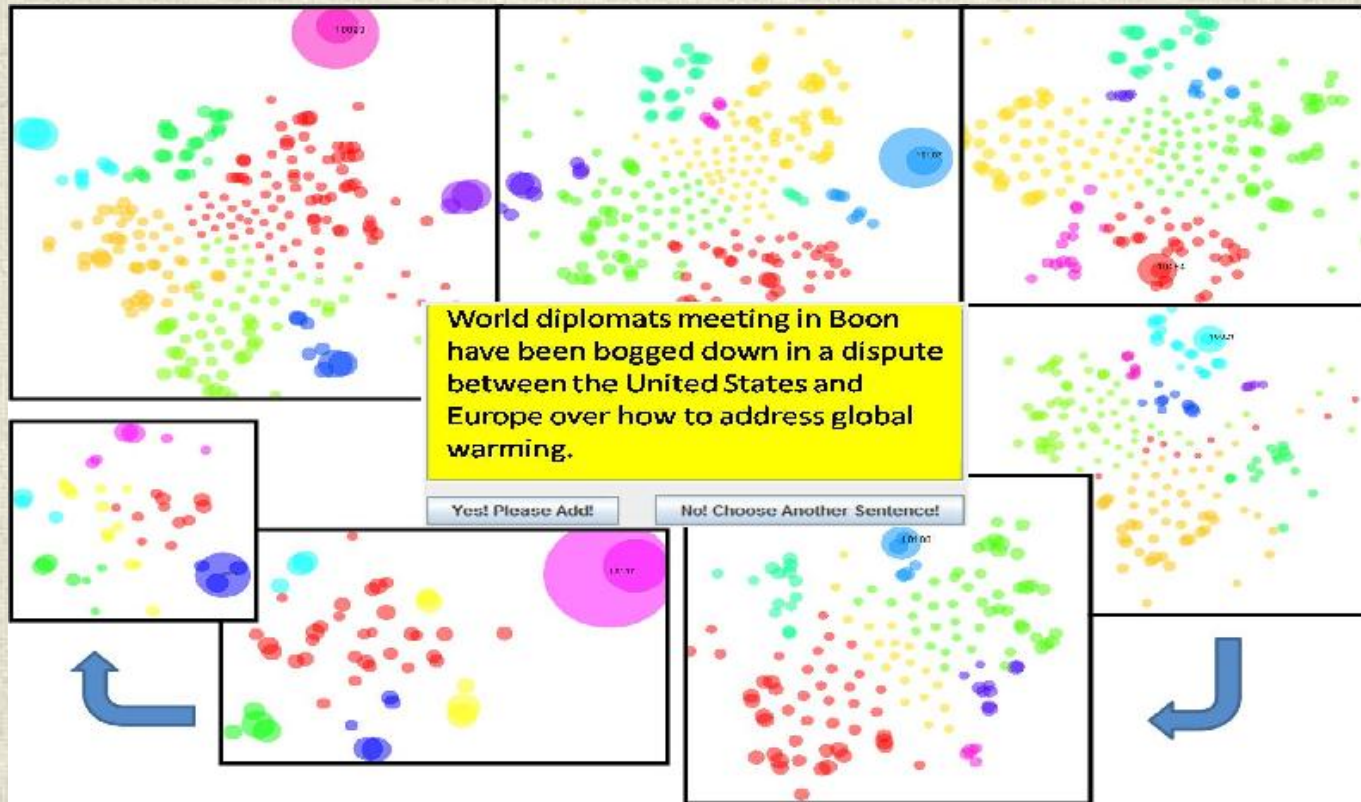


(Iwata et al., 2008)





# Grupos



(Zhang et al., 2010)

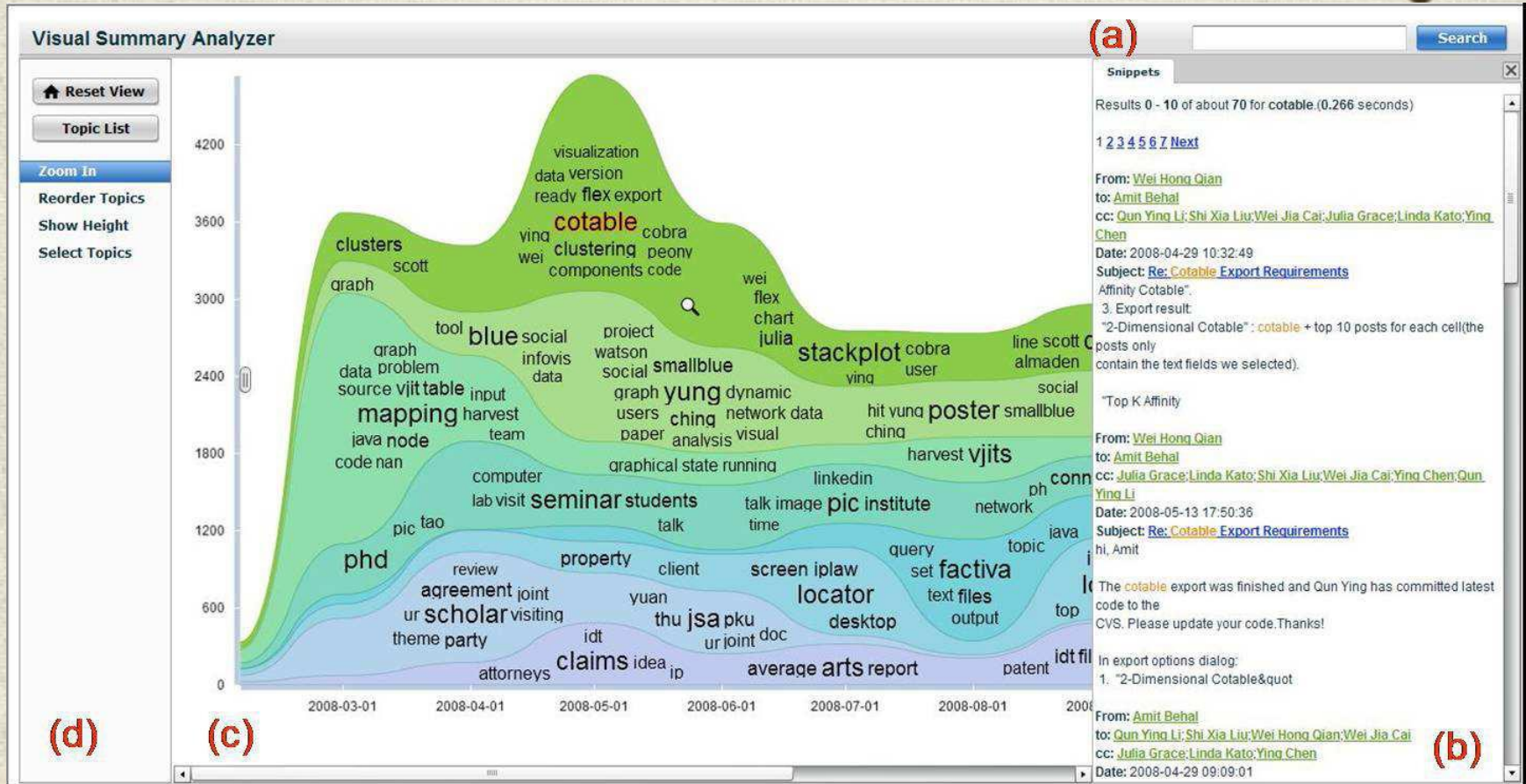






# Líneas de tiempo

2, 3 Y 4 DE OCTUBRE





# Paisajes





# Estado del arte

<b>Tipo de visualización</b>	Puntos	Red	Grupos	Jerarquía	Temporal	Paisaje	Tópicos
<b>Concientización</b>	Contenido		Estructura		Cambios		Relaciones
<b>Funciones</b>	Vista panorámica		Zoom y filtrado		Detalles a demanda		
<b>Tareas de usuario</b>	Búsqueda	Monitoreo	Navegación	Exploración	Entendimiento	Planeación	
<b>Tipo de reducción</b>	Geométrica				Semántica		

## Minería visual:

Proceso de exploración de colecciones basado principalmente en la **percepción humana** para detectar patrones, estructuras o comportamientos *interesantes*

Normalmente, el resultado de la minería visual no se conoce de antemano



# Estado del arte

2, 3 Y 4 DE OCTUBRE

<b>Tipo de visualización</b>	Puntos	Red	Grupos	Jerarquía	Temporal	Paisaje	Tópicos	
<b>Concientización</b>	Contenido		Estructura		Cambios		Relaciones	
<b>Tareas de interfaz</b>	Vista panorámica		Zoom y filtrado		Detalles a demanda			
<b>Tareas de usuario</b>	Búsqueda	Monitoreo		Navegación	Exploración	Entendimiento		Planeación
<b>Tipo de reducción</b>	Geométrica				Semántica			



CUDI 2013  
REUNIÓN DE OTOÑO  
CAMPECHE



2, 3 Y 4 DE OCTUBRE

# Comprensión de colecciones en el Laboratorio ICT de UDLAP



# Campos de estrellas escalables (EVA2D)

**EVA2D**

File Edit Window Options Help

Starfield

Category

Building O  
Electric Me  
Engineerin  
Congress  
Medical R  
Astrophysi  
Egyptian,  
History an  
Haiti [KGS  
History [JV  
Criminal J  
Express S  
Maps [G:  
North Ame  
Early Histo  
Collected  
History [B

Allen, Ivor Brinker - G Davies, R. Garcia Cu Hofland, B Lefevre, H Mittelman, Povina, Alfr Selman, R Valdez Bec

Author

Filters

Title  
A 50 anos... Zwanzig j...

Author  
Abate Lar... Zysman, ...

Subject  
Abad Que... Zweig, St...

Category  
Collection... Bibliograp...

Edition  
1 Ed. 198 Ed.

Publisher  
A & C Bla... Zvonimir ...

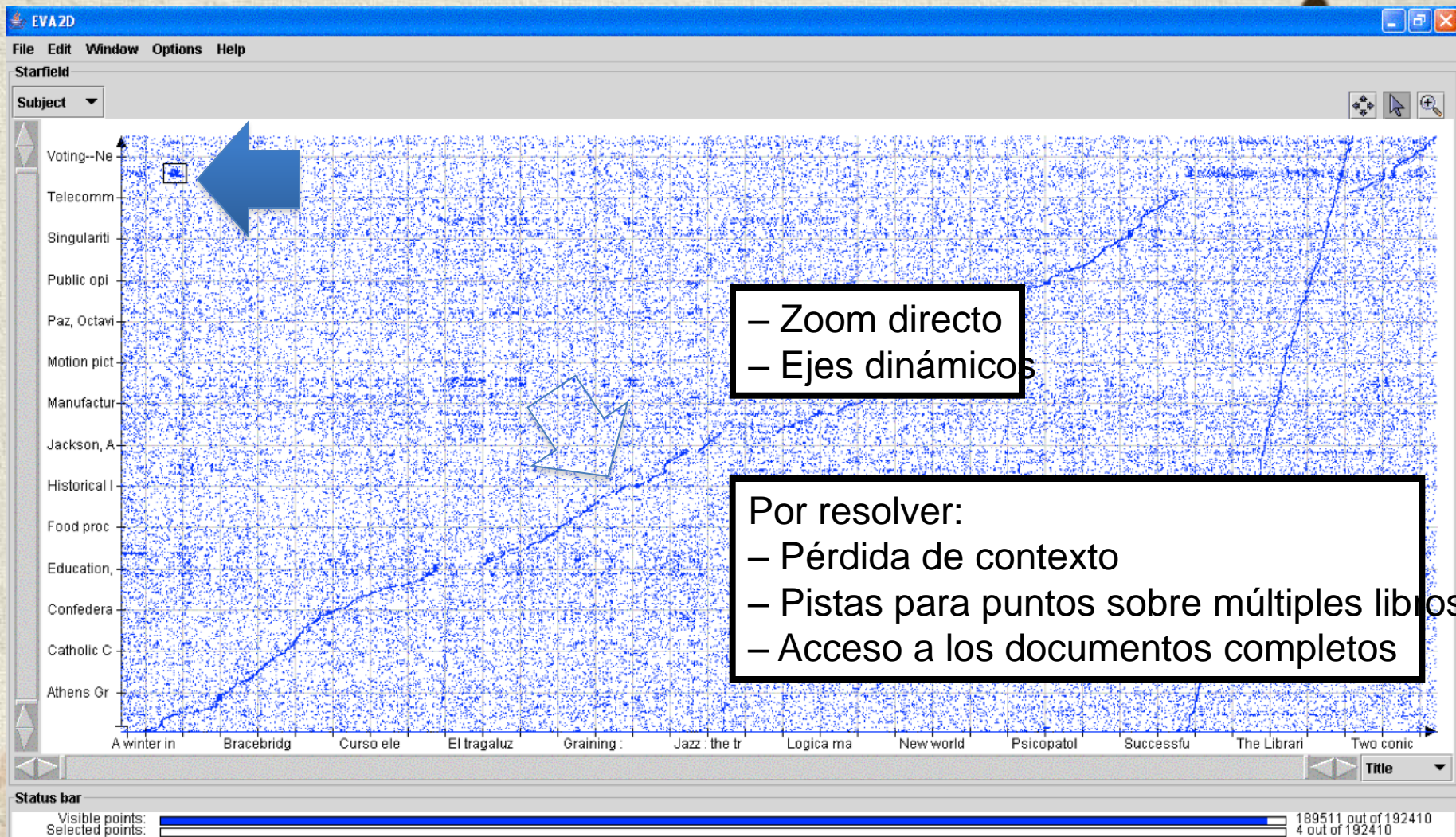
Year  
1774 2000

Place  
Abilene, T... Zutphen, ...

Status bar  
Visible points: 30249 out of 192410  
Selected points: 4 out of 192410

- Más de 200,000 libros
- Zoom directo
- Feedback cuasi-inmediato
- Filtros de precisión







# Ojos de pescado (fisheye views)

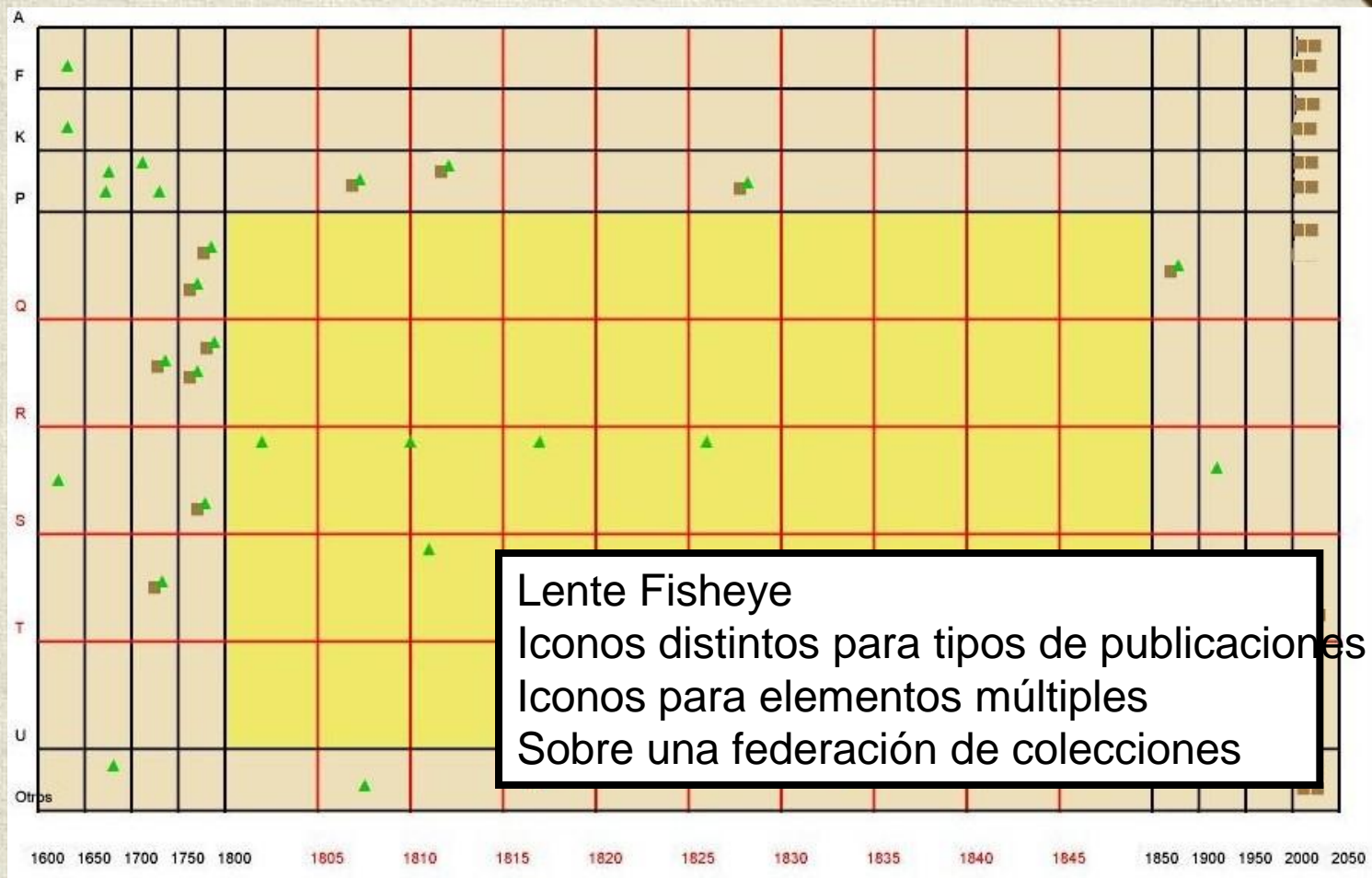
2, 3 Y 4 DE OCTUBRE



©2008, Erlend Robaye



# Star-fish: Contexto+foco



Lente Fisheye

Iconos distintos para tipos de publicaciones

Iconos para elementos múltiples

Sobre una federación de colecciones



# Star-fish: Contexto+foco

The screenshot shows a web browser window displaying a digital collection interface. The main area features a grid of digital collections with a 'Star-fish' visualization overlay. The grid has columns representing years from 1600 to 2005 and rows labeled A through Otrps. A yellow star-shaped area highlights a specific collection. To the left, there is a sidebar with a legend and a document details table. An inset window shows a PDF document cover page.

**Legend:**

- ▲ Fondo Antigo de UAEH
- Tesis de UAEH
- ▲ Fondo Antigo de UDLAP
- Tesis de UDLAP
- ▲ Fondo Antigo de USON

Detalles << Ver documento >>	
Título	Extraccion y acetilacion de los componentes de la grana cochinilla ( <i>Dactylopius coccus</i> COSTA)...
Autor	Gomez Hernandez, Belem
Editorial	Universidad Autonoma del Estado de Hidalgo
Colaborador	Universidad Autonoma del Estado de Hidalgo
Fecha	2006
Tipo Doc	Tesis de Licenciatura
Formato	application/pdf
Tema	Quimica

**Inset Window Content:**

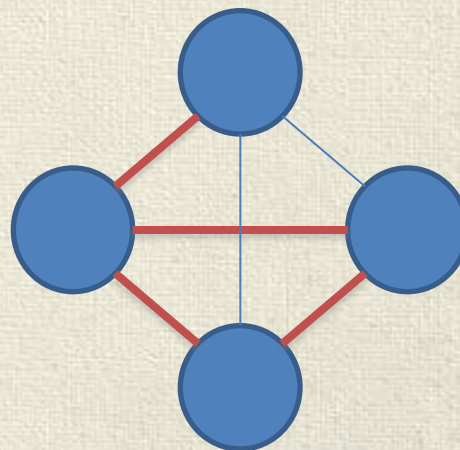
UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO  
INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA  
LICENCIATURA EN QUÍMICA  
"EXTRACCION Y ACETILACION DE LOS COMPONENTES DE LA GRANA COCHINILLA (*Dactylopius coccus* COSTA)"  
TRABAJO DE INVESTIGACIÓN  
QUE PARA OBTENER EL TÍTULO DE:  
QUÍMICO  
P R E S E N T A:  
BELEM GOMEZ HERNANDEZ  
Asesores: Dr. José Roberto Villagómez Ibarra  
M. en C. Julio Rodríguez Baños

# Visualización de relaciones



# Redes de colaboración

- Estructuras que representan grupos inter-conectados de personas que trabajan juntos con niveles variados de acoplamiento
- Particularmente interesantes las redes de colaboración implícitas





# ¿A quién interesan las redes de colaboración?

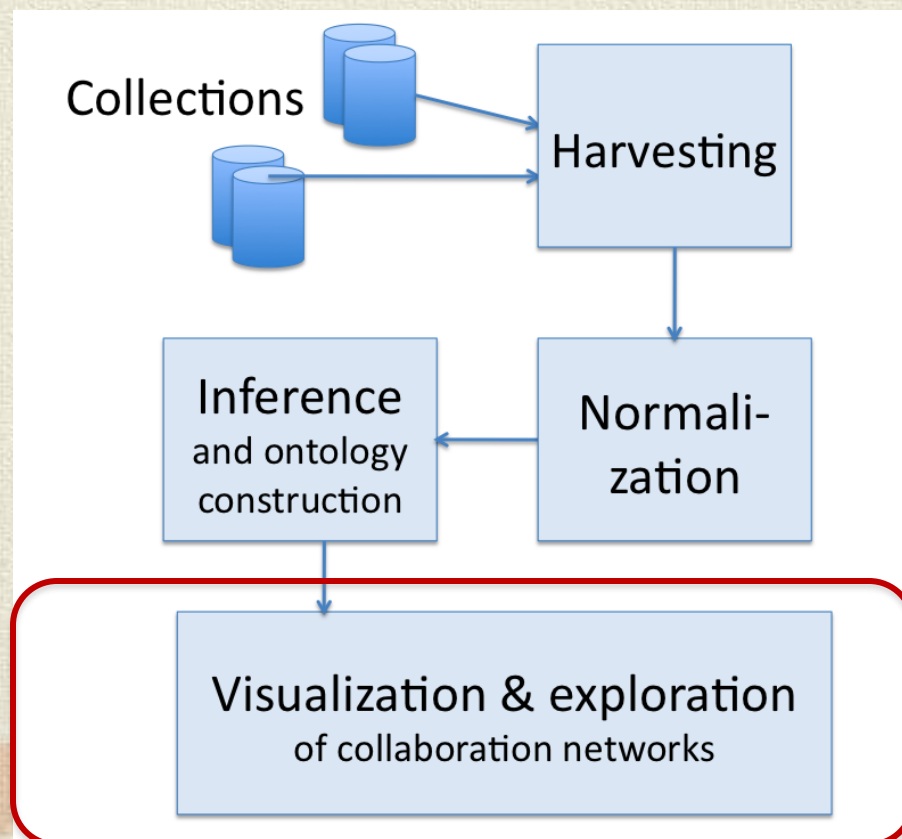
- Investigadores:
  - Para conocer sobre su propia comunidad
  - Para encontrar socios de proyectos multi-disciplinarios, multi-institucionales, multi-nacionales
- Agencias de financiamiento
  - Para definir beneficiarios potenciales de convocatorias
  - Para validar grupos de trabajo propuestos



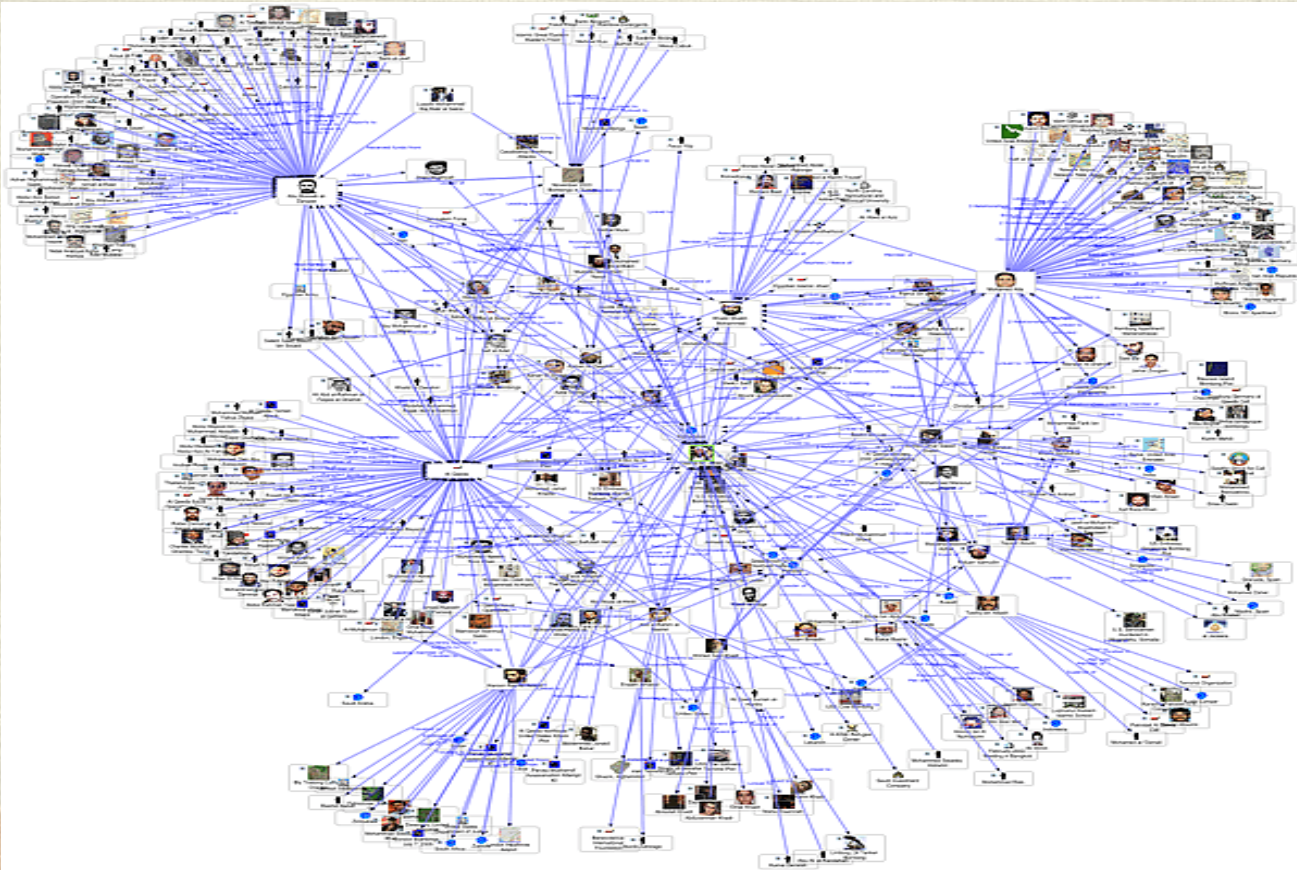
# Cosecha y minería de metadatos

2, 3 Y 4 DE OCTUBRE

- Meta: Derivar, visualizar y explorar redes de colaboración a partir de colecciones digitales

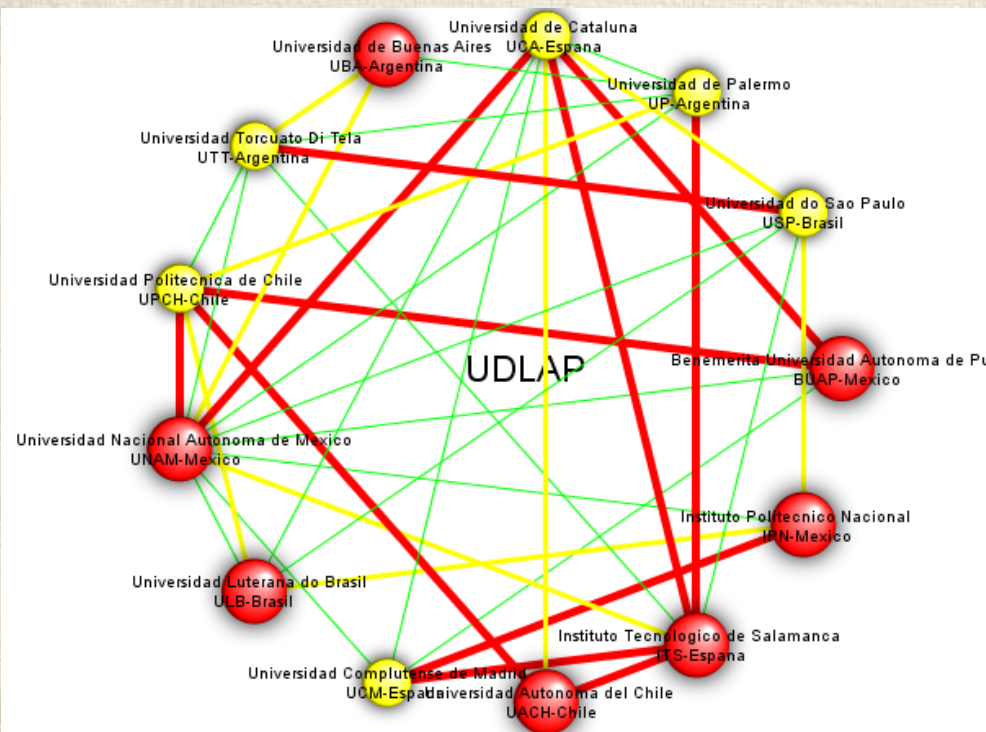


- Enfoque tradicional: nodos y aristas

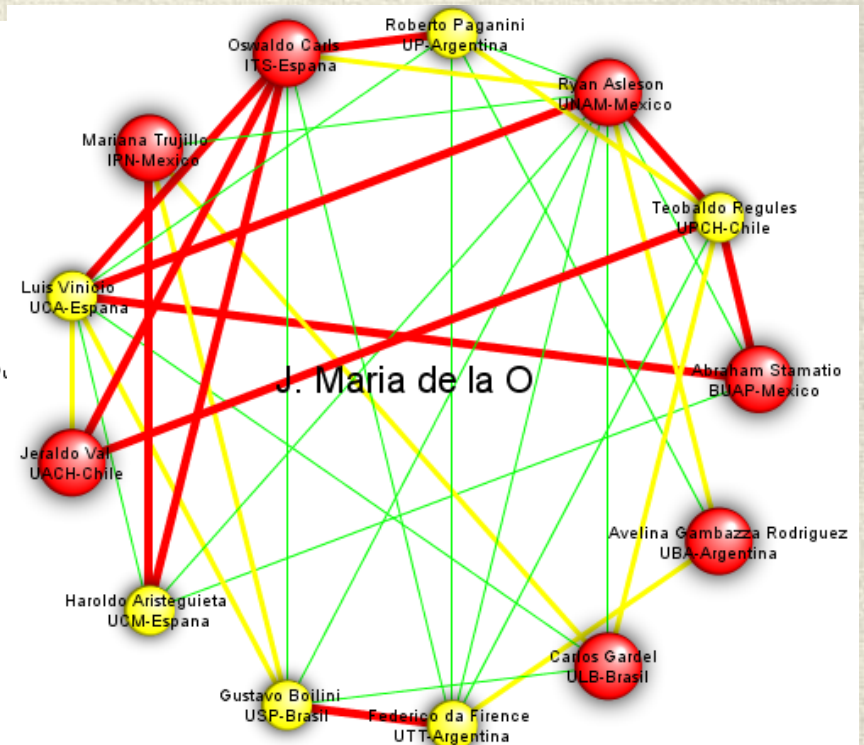




## Instituciones



## Autores









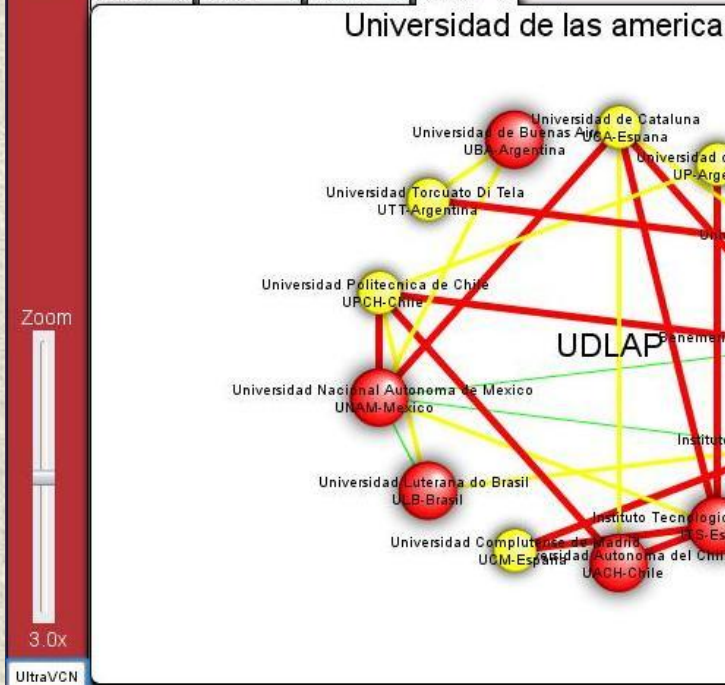
# Grafos egocéntricos + filtros

**Institución** Autor  
Institución: Universidad de las americas puebla  
País: mexico **Consultar**

Colaboradores Potenciales  
Todos 3.0 2.0 1.0  
Relaciones entre Colaboradores  
Todas 3.0 2.0 1.0

Sólo Importantes  
Sólo Fuertes

Paleta de colores para las relaciones

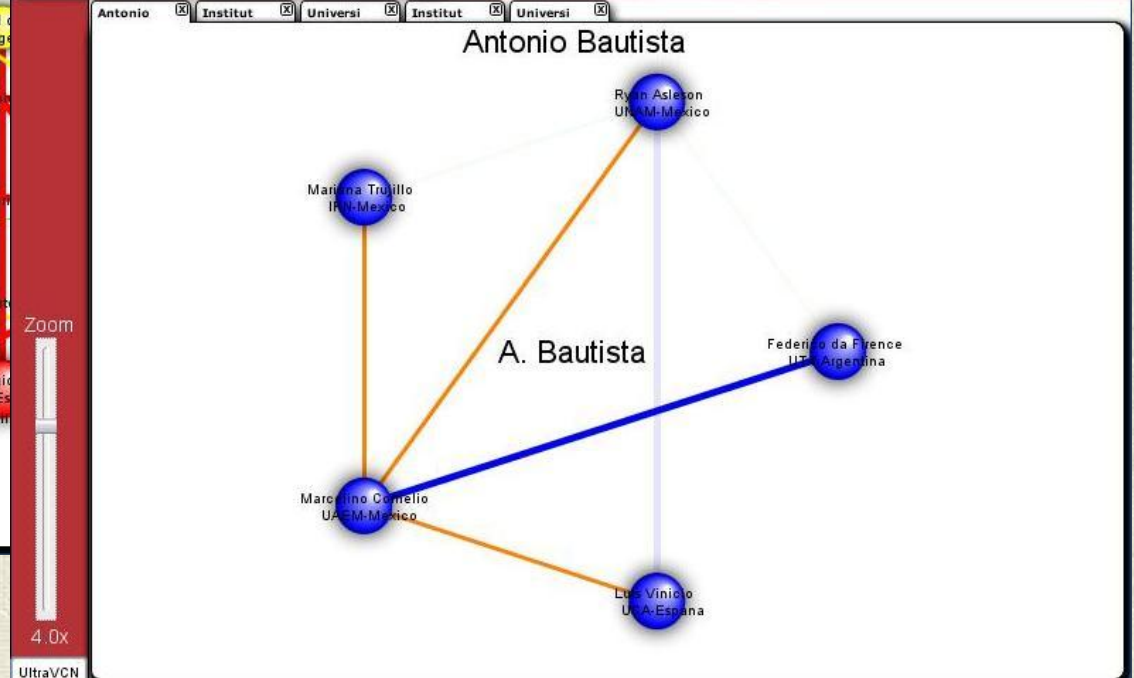


**Institución** Autor  
Nombre: Antonio Bautista País: mexico  
Institución: Instituto Tecnológico de Estudios Superiores de Occidente **Consultar**

Colaboradores Potenciales  
Todos 3.0 2.0 1.0  
Relaciones entre Colaboradores  
Todas 3.0 2.0 1.0

Sólo Importantes  
Sólo Fuertes

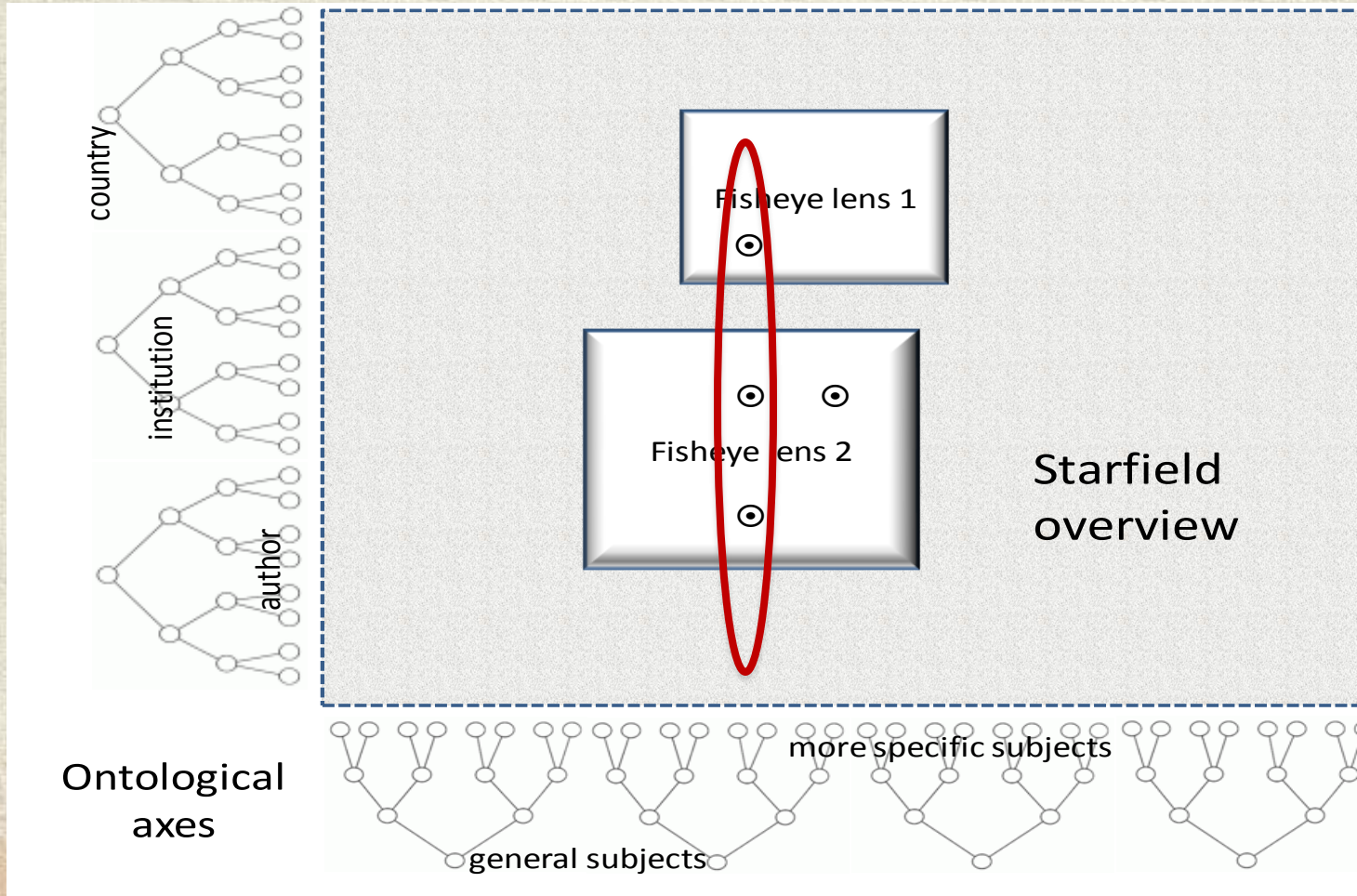
Paleta de colores para las relaciones



# OntoStarFish:

## Campos de estrellas ontológicos + *Fisheyes*

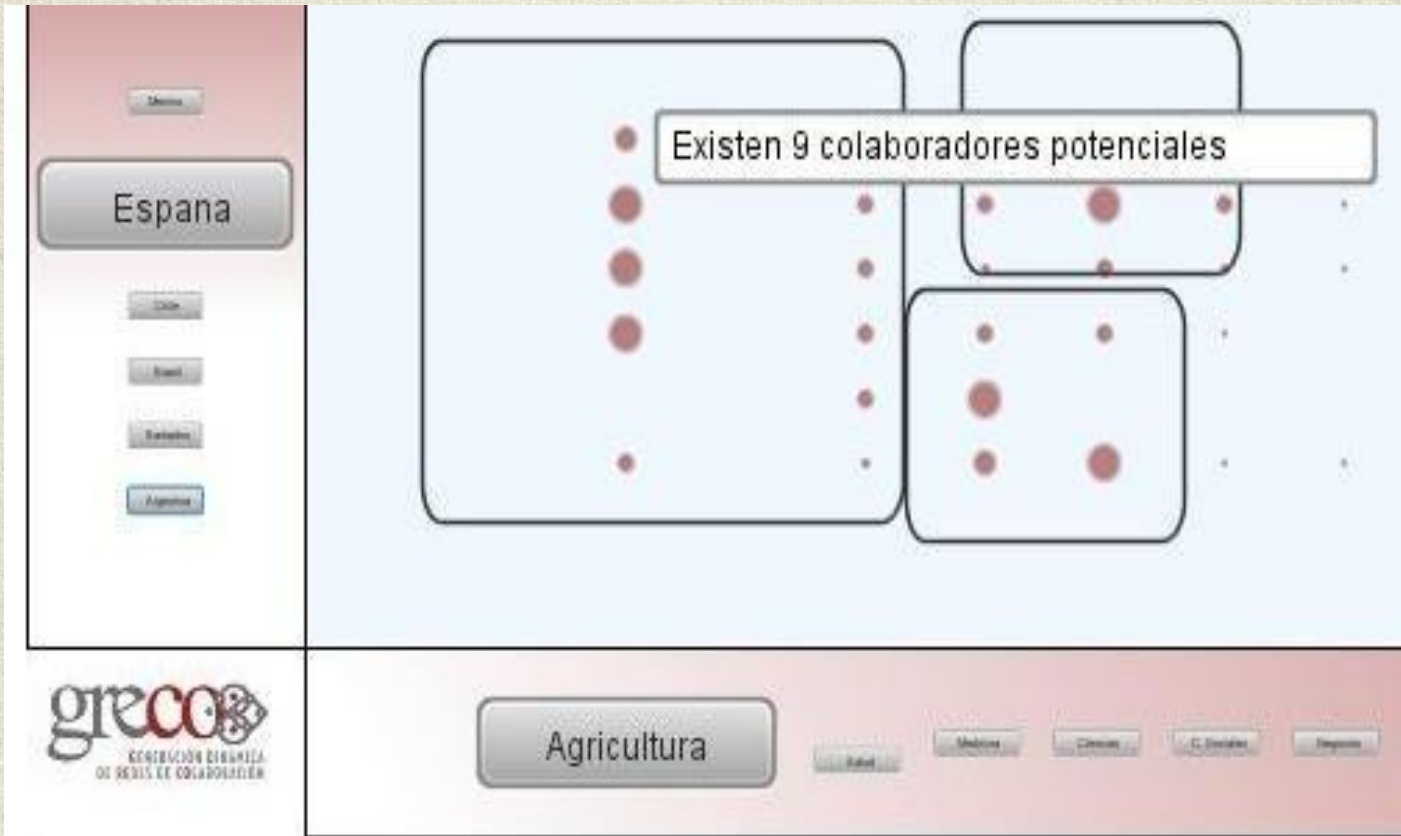
2, 3 Y 4 DE OCTUBRE



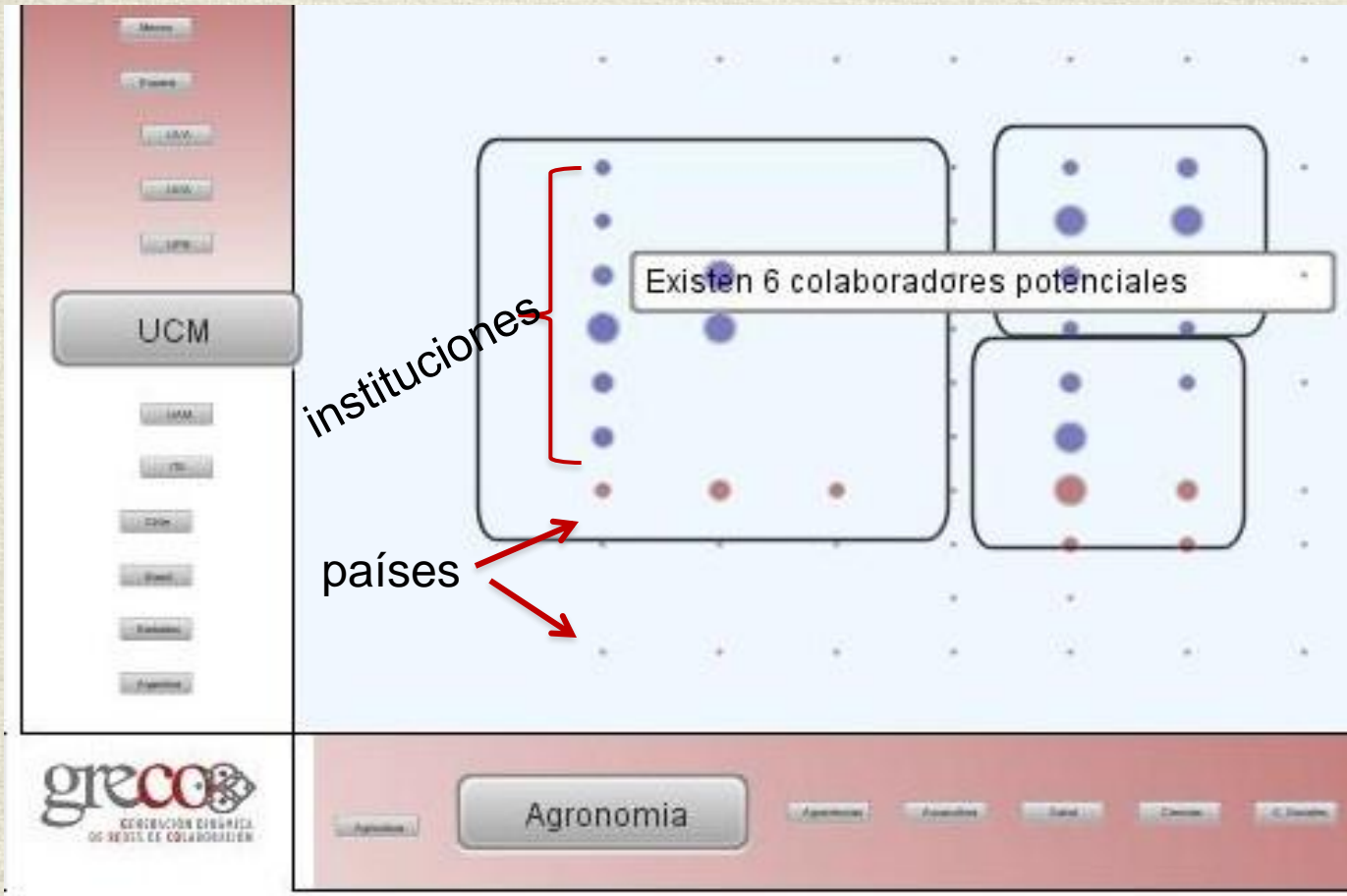


# Localizando investigadores

2, 3 Y 4 DE OCTUBRE

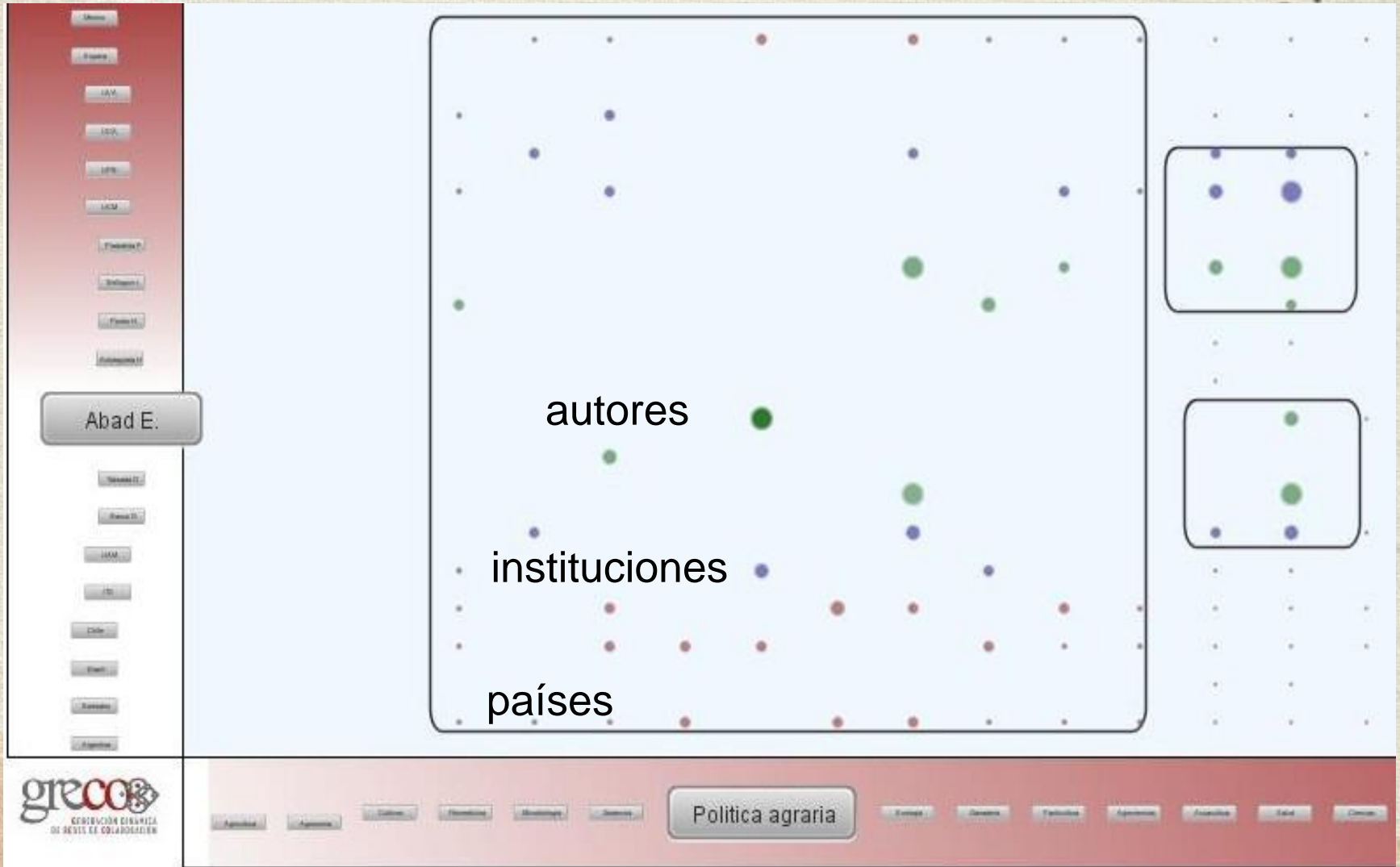


# Localizando colaboradores potenciales

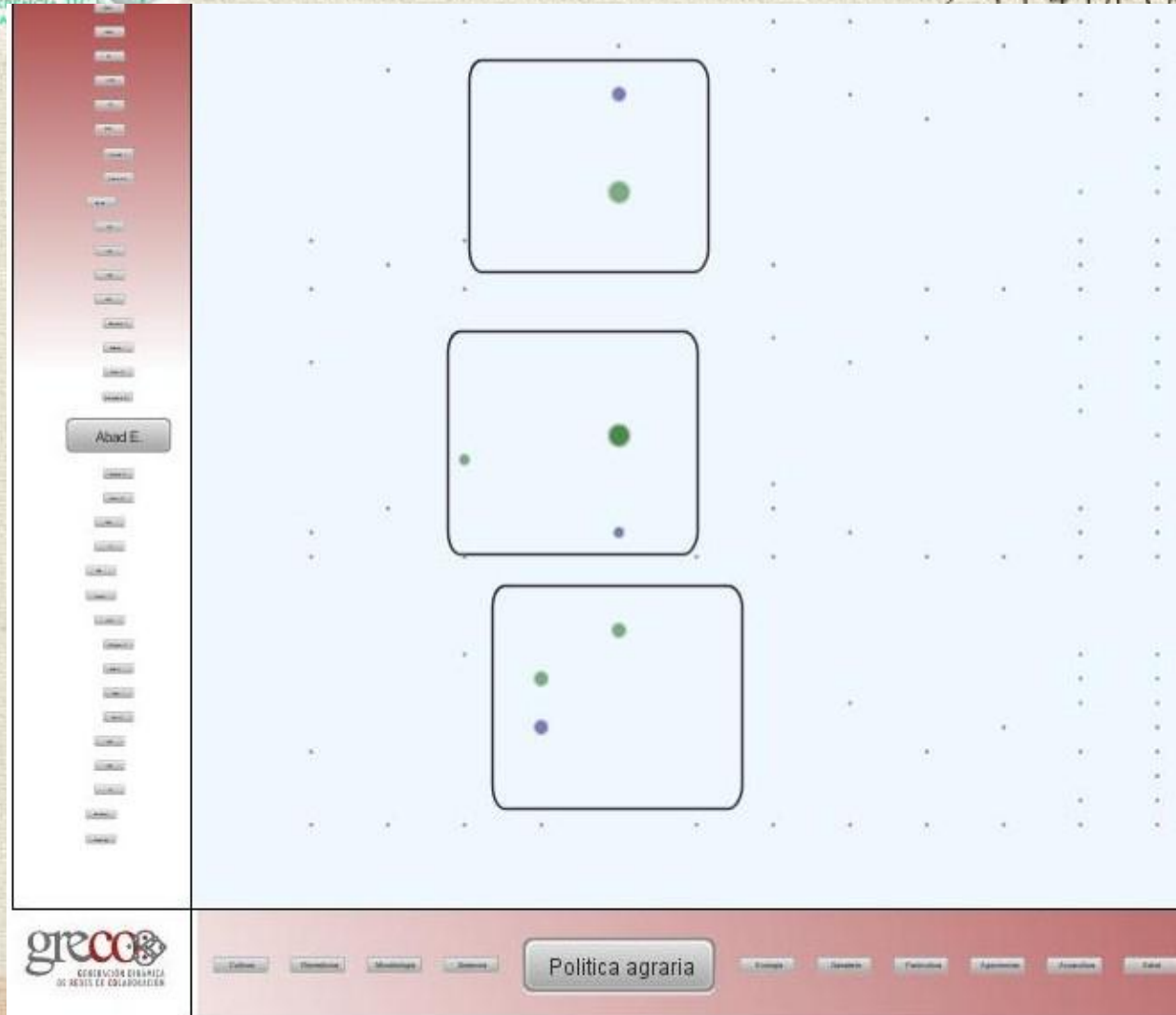




# Expansión de ambos ejes



# Uso de múltiples lentes *fisheye*

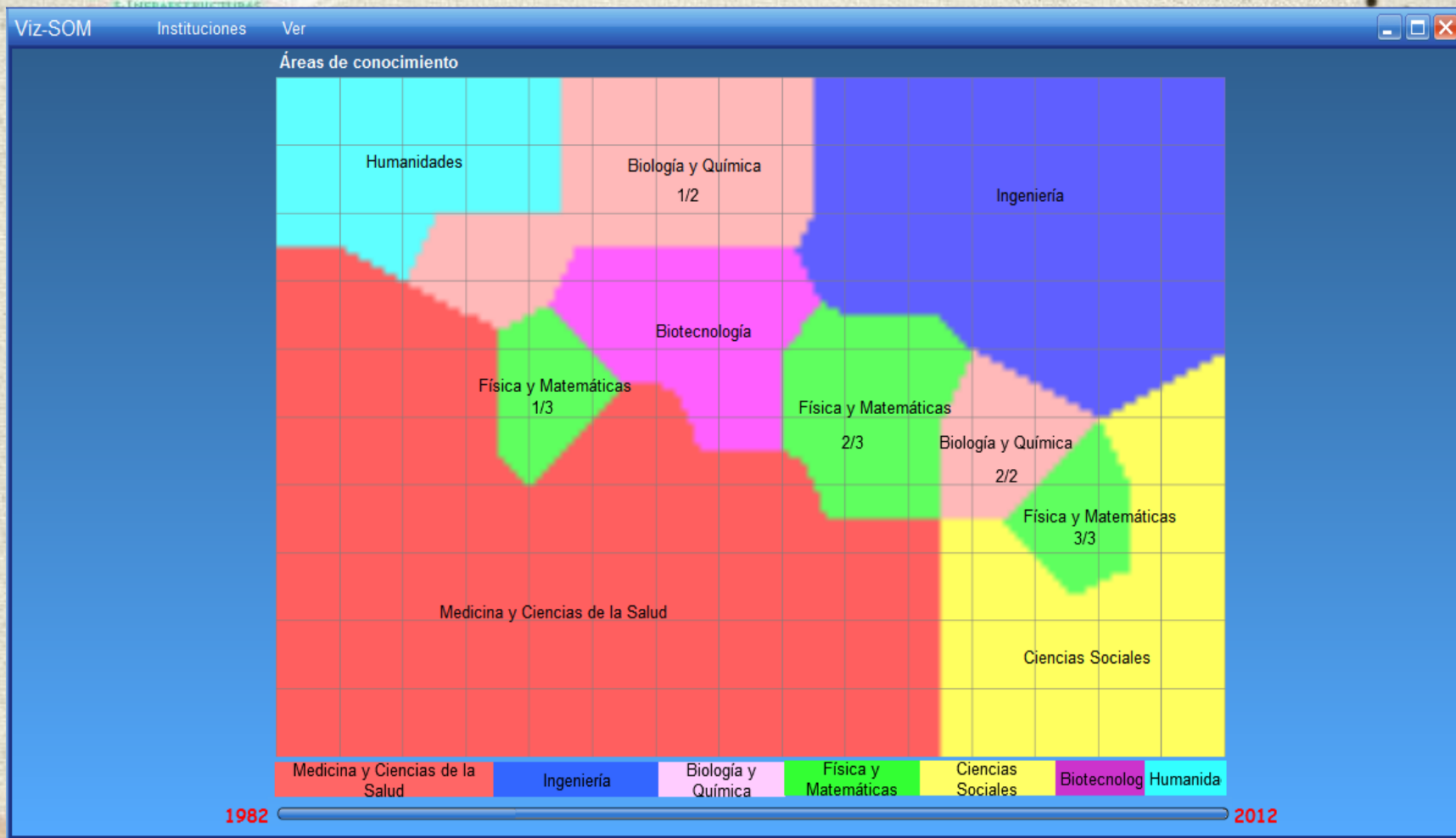


[Sánchez et al. 2011]

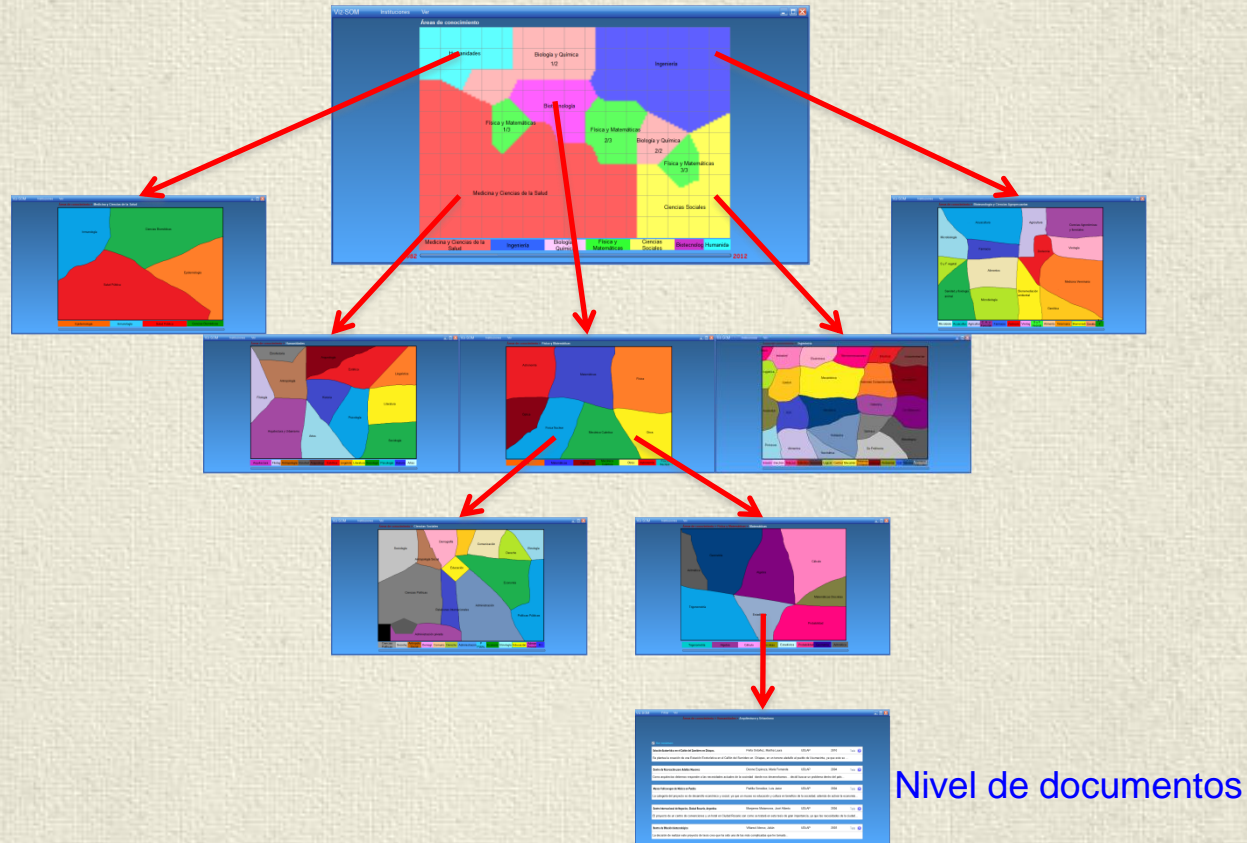


# Visualización de agrupamientos y relaciones usando mapas

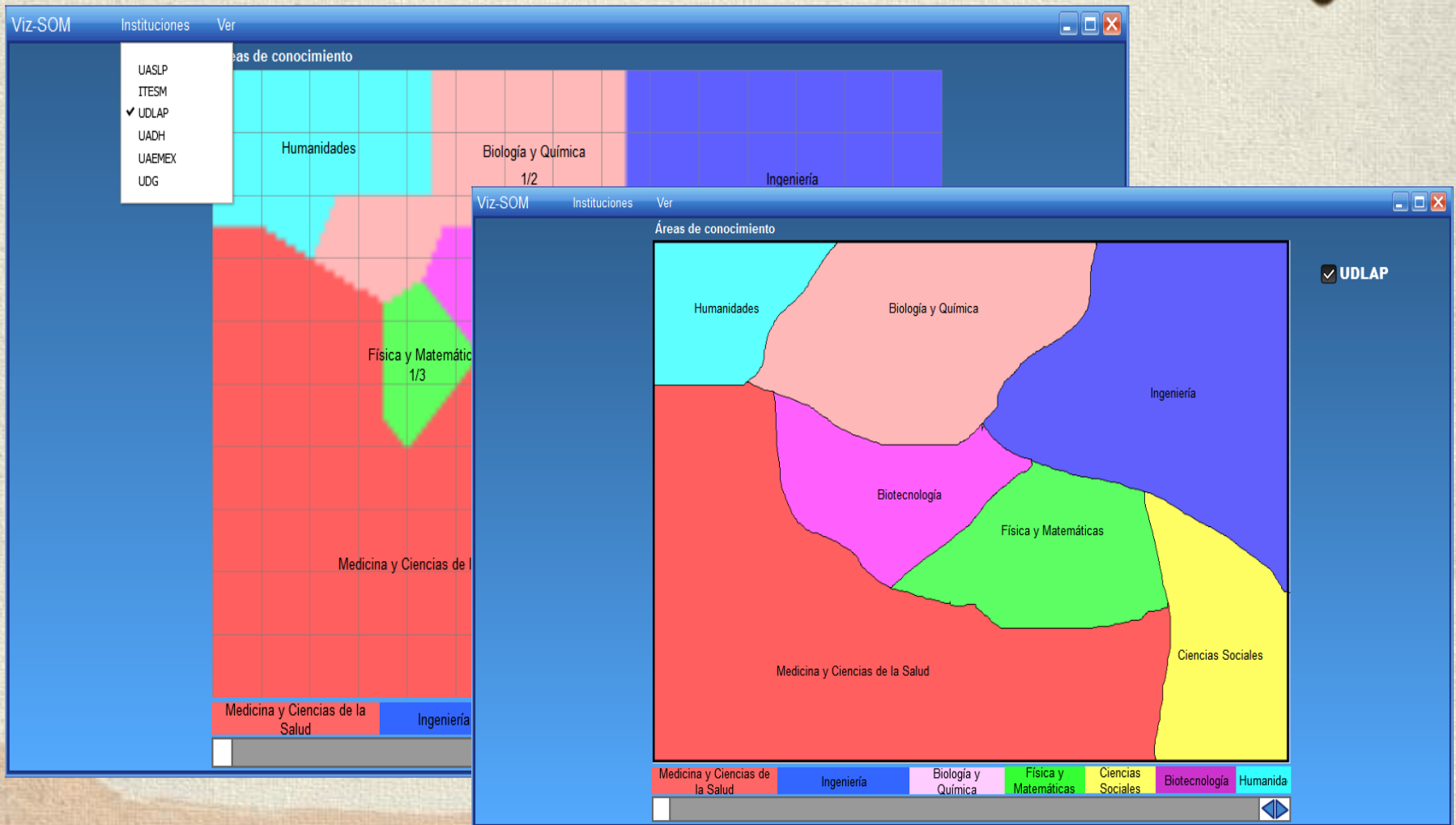
# Los mapas pueden ser una metáfora clave para comprensión de colecciones







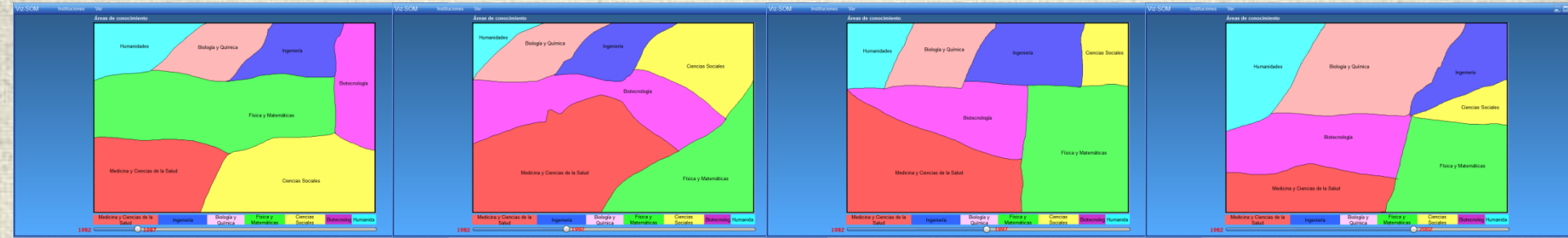
# Filtros por institución o tema



[Gutiérrez et al. 2012; Gutiérrez 2013]



# Evolución en el tiempo

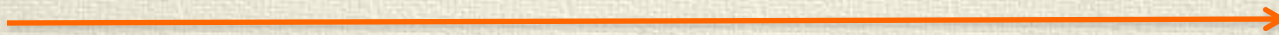


2004

2007

2010

2013



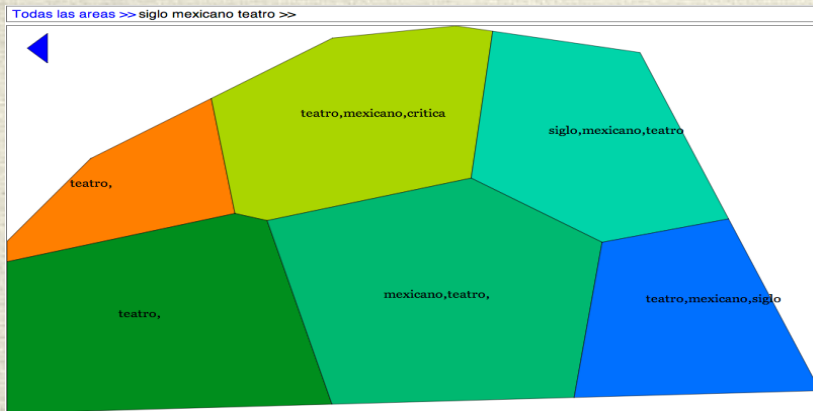
# VoroSOM: mapas auto-organizables + diagramas de Voronoi

- **Mapas auto-organizados (SOMs)**
  - Funcionan como clasificadores, produciendo jerarquías de categorías – ideal para navegar
  - Preservan propiedades topológicas de relaciones existentes en los metadatos
- **Diagramas de Voronoi**
  - Se pueden construir sobre una representación espacial 2D de las categorías SOM
  - Reflejan visualmente las relaciones semánticas de las categorías adyacentes



# Alternativas de representación

2, 3 Y 4 DE OCTUBRE



“Continentes”



“Islas”

# VoroSOM - Continentes

2, 3 Y 4 DE OCTUBRE

VOROSOM

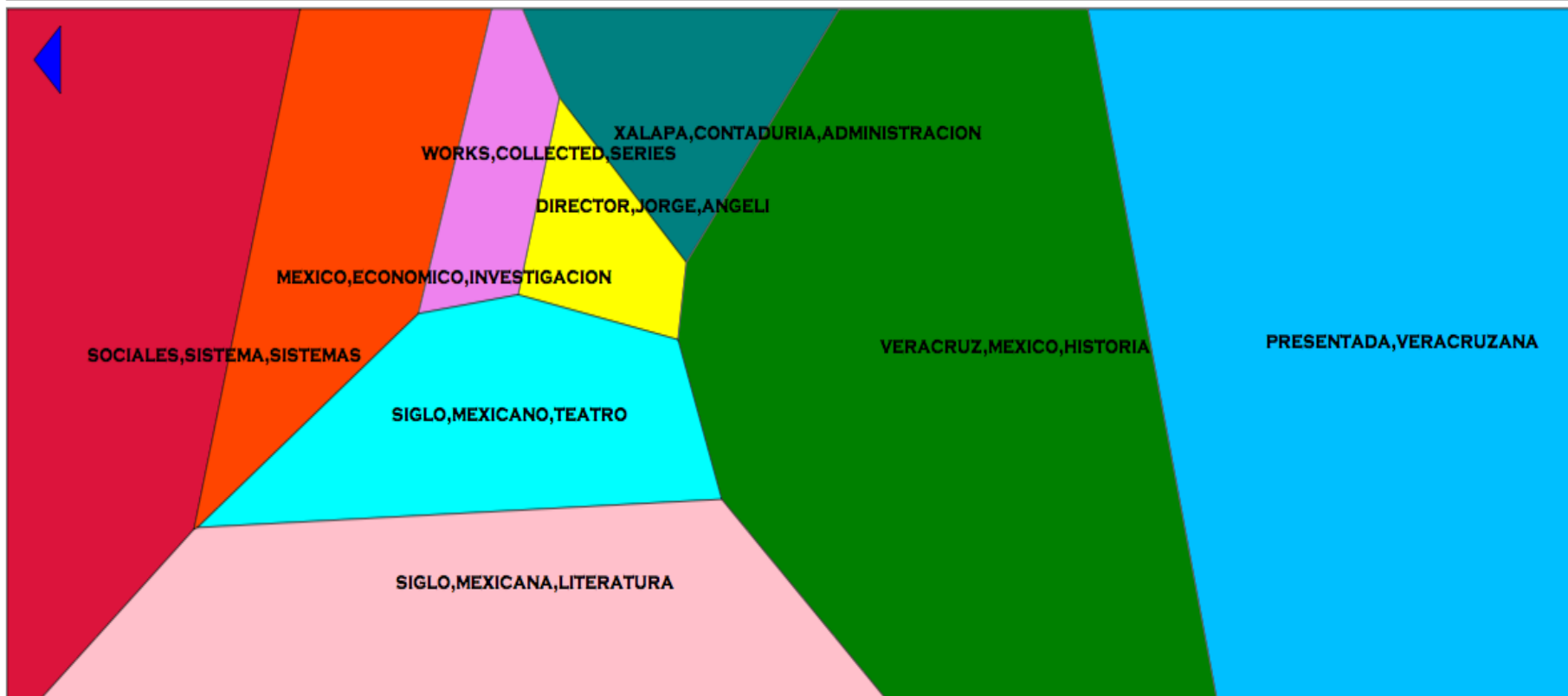
COLECCIONES POR INSTITUCIÓN

HERRAMIENTAS

AYUDA

CONTACTO

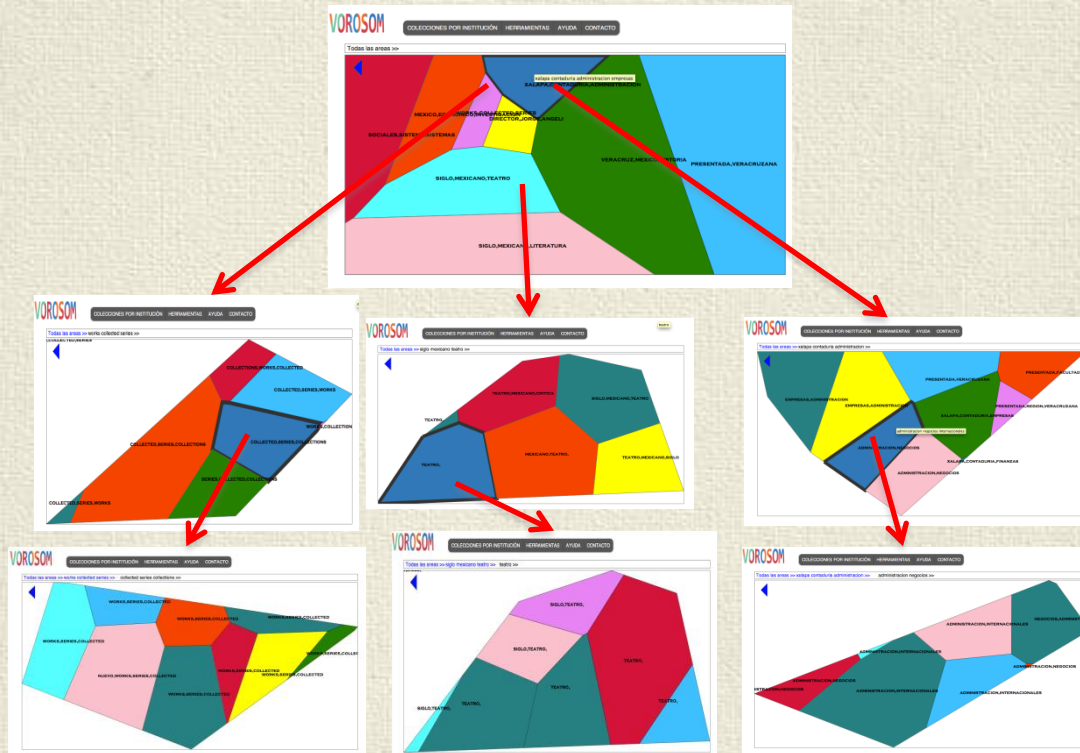
Todas las areas >>





# Navegación continentes

2, 3 Y 4 DE OCTUBRE



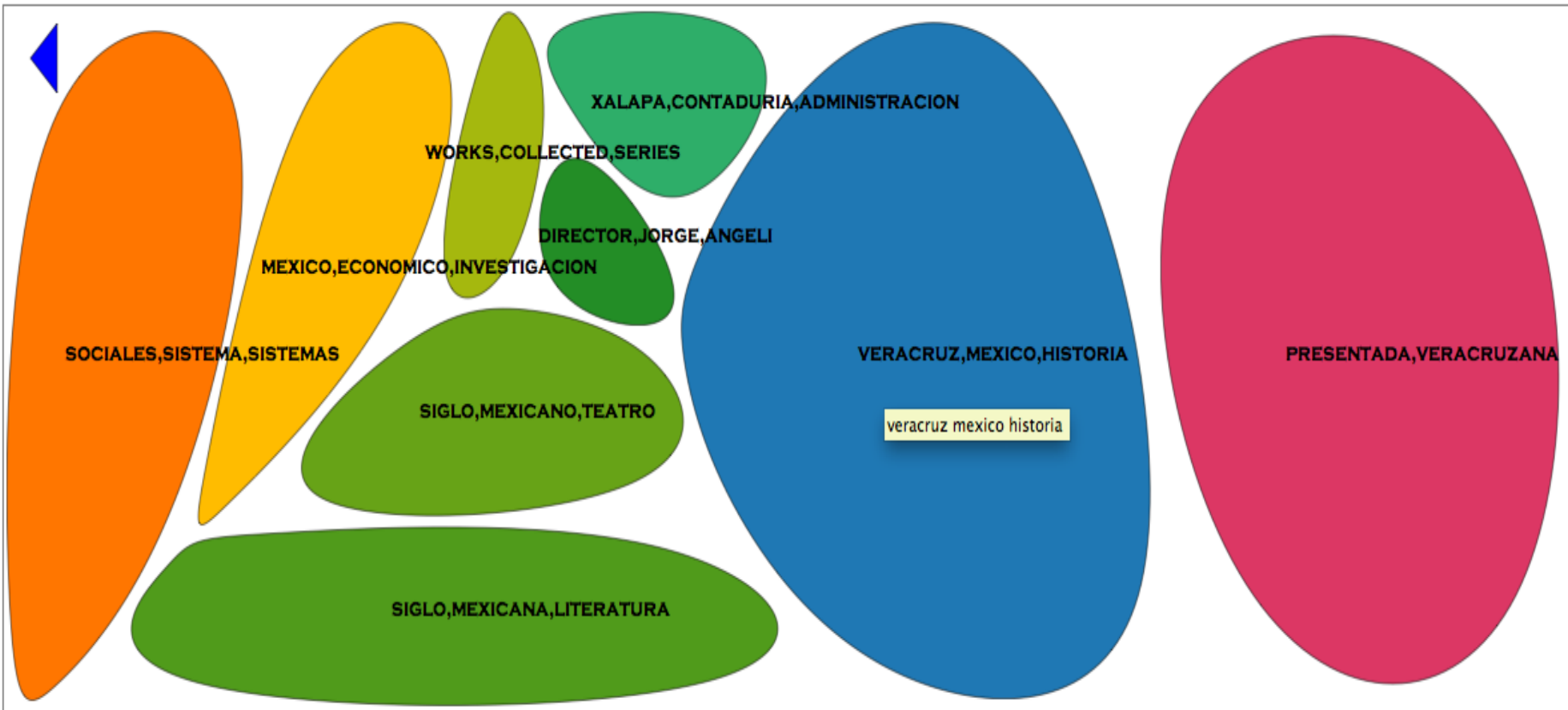
COLECCIONES POR INSTITUCIÓN

HERRAMIENTAS

AYUDA

CONTACTO

TODAS LAS AREAS >>









# Prototipos disponibles

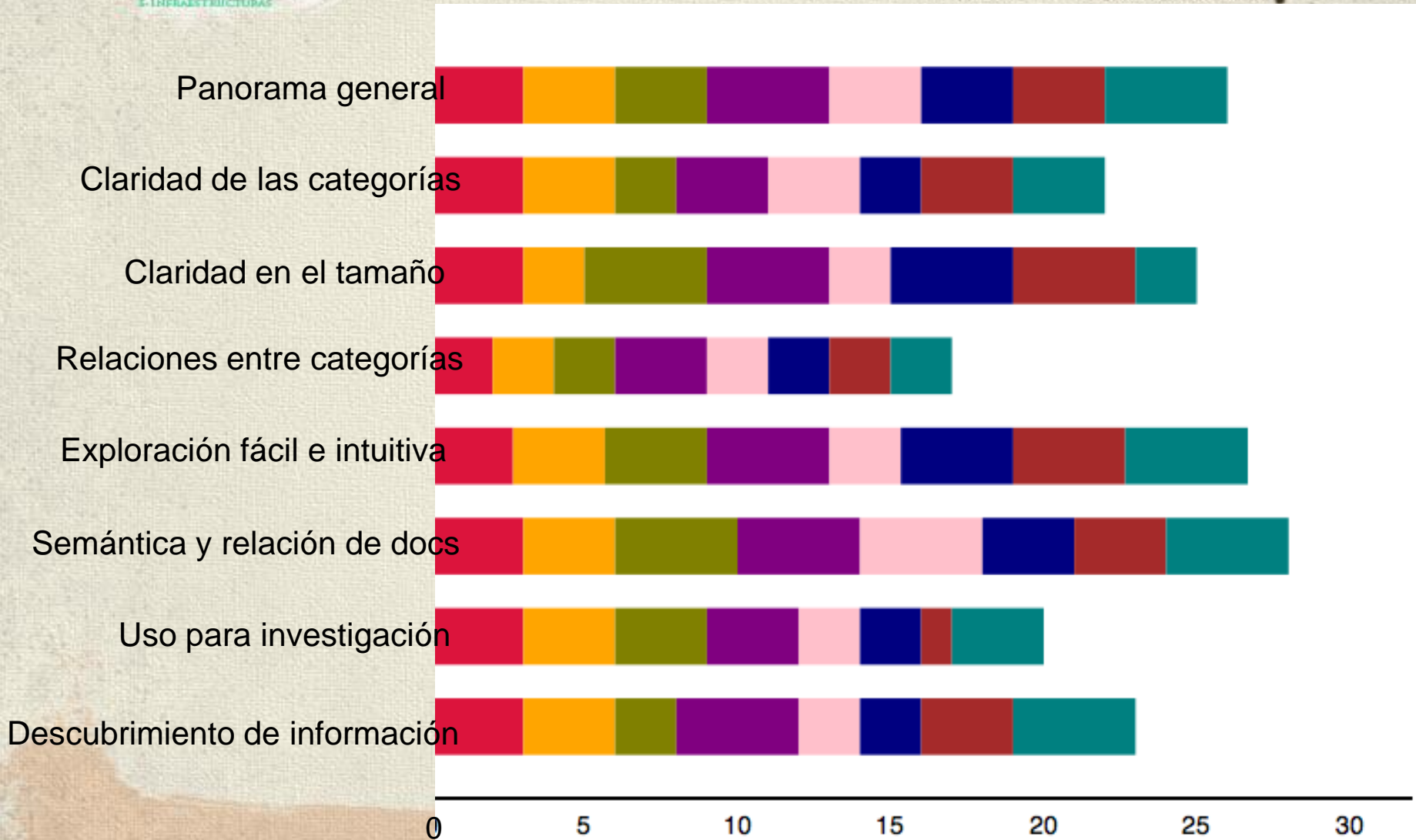
- VoroSOM – Continentes:
  - <http://ict.udlap.mx/remeri/vorosom1>
  - <http://ict.udlap.mx/remeri/vorosom2>
- VoroSOM – Islas
  - <http://ict.udlap.mx/remeri/vorosom3>



# Evaluación

(8 sujetos, verbalización constructiva)

2, 3 Y 4 DE OCTUBRE



# Representaciones alternas

2, 3 Y 4 DE OCTUBRE

**REMERI** Ayuda

Esquema de color: Degradado

Ruta y posición actual

Home

---

Filtros

DE TIEMPO

Mínimo: 1564

Máximo: 2013

INSTITUCIÓN

UASLP  UAEMEX

UDLAP  ITESM

UAEH  UDG

UV  UCSJ

TIPO

Tesis digital

Artículo

Documento

Elemento

Home

Ordenar por: Título

Ascendente Descendente

Realizado por Nilda Galán Hernández



# Representaciones alternas

**REMERI** [Ayuda](#)

Esquema de color: Degradado

**Ruta y posición actual**

Home  
arte mexicano mundo hombre mexicana cultura narrativa espacio texto lectura  
mexicano hombre historia

**Filtros**

**DE TIEMPO**

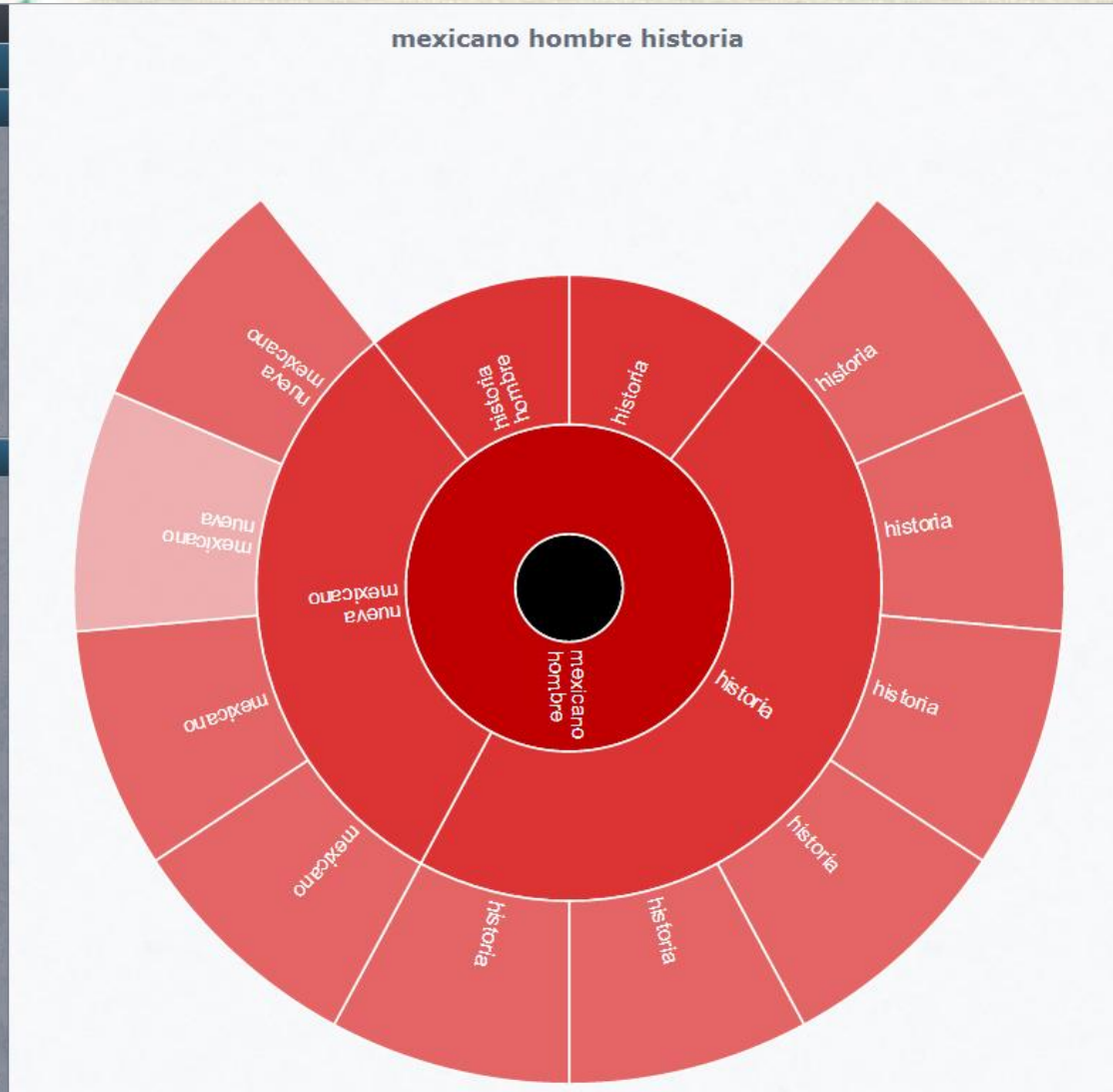
Mínimo: 1564  
Máximo: 2013

**INSTITUCIÓN**

UASLP  UAEMEX  
 UDLAP  ITESM  
 UAEH  UDG  
 UV  UCSJ

**TIPO**

Tesis digital  
 Artículo  
 Documento  
 Elemento



Ordenar por: Título

Ascendente Descendente

REMERI Ayuda

Esquema de color: Degradado

Ruta y posición actual

Home  
arte mexicano mundo hombre mexicana cultura narrativa espacio texto lectura  
mexicano hombre historia

Filtros

DE TIEMPO

Mínimo: 1564  
Máximo: 2013

INSTITUCIÓN

UASLP  UAEMEX  
 UDLAP  ITESM  
 UAEH  UDG  
 UV  UCSJ

TIPO

Tesis digital  
 Artículo  
 Documento  
 Elemento

## mexicano hombre historia



Ordenar por: Título  
Ascendente Descendente

### mexicano, nueva

Algunos elementos chamánicos entre los iluminados o alumbrados de la Nueva España

Lagarrita Attias Isabel  
1988 UV

Angel Rama y la nueva mirada

Lavín Cerda Hernán  
1985 UV

El "ensayo político sobre el reino de la Nueva España", de Humboldt

Brand Donald D  
1959 UV

El manierismo en Nueva España: letras y artes

Manrique Jorge Alberto  
1974 ITESM

La nueva concepción del personaje redentor

Aldana, Adolfo  
1920 UV

La Nueva crónica

Madden Lori  
1989 UV

La nueva lengua universal

Karon Garbe  
1874 UASLP



# Conclusiones

- La minería visual puede apoyar en la comprensión de grandes colecciones
- Para facilitar la minería visual, la visualización de información requiere:
  - Metáforas y representaciones accesibles
  - Interactividad
  - Mecanismos de inferencia y clasificación dinámicos



# Invitación

2, 3 Y 4 DE OCTUBRE

- Comunidad CUDI de IHC  
([www.cudi.edu.mx](http://www.cudi.edu.mx))
- Comunidad Latinoamericana de IHC
  - [www.sigchi.org/communities/laihc](http://www.sigchi.org/communities/laihc)
- Congreso Latinoamericano de IHC (CLIHC 2013),  
<http://clihc.org>, Costa Rica,  
4-5 Dic.
  - Orador invitado: Jonathan Grudin

