

# Integración de anotación lingüística temporal en un SIG

Marta Guerrero Nieto  
Adolfo Urrutia Zambrana  
María José García Rodríguez  
Miguel Ángel Bernabé Poveda

*2 de octubre de 2009*

# ÍNDICE

---

- ▶ Motivación
- ▶ Objetivo
- ▶ Contexto
- ▶ Estado del arte
- ▶ Metodología
- ▶ Resultados
- ▶ Trabajos futuros
- ▶ Referencias



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

**gvsig**

"COOPERAR ES CRECER"



# MOTIVACIÓN

---

- ▶ Mermar el aislamiento de la dimensión temporal y fomentar la integración del tiempo y espacio.
- ▶ No todo el lenguaje natural puede incorporarse a una base de datos.
- ▶ Los Sistemas de Información Geográfica sólo aceptan expresiones temporales definidas, pero ¿cómo hacer que un SIG represente otro tipo de expresiones temporales?



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

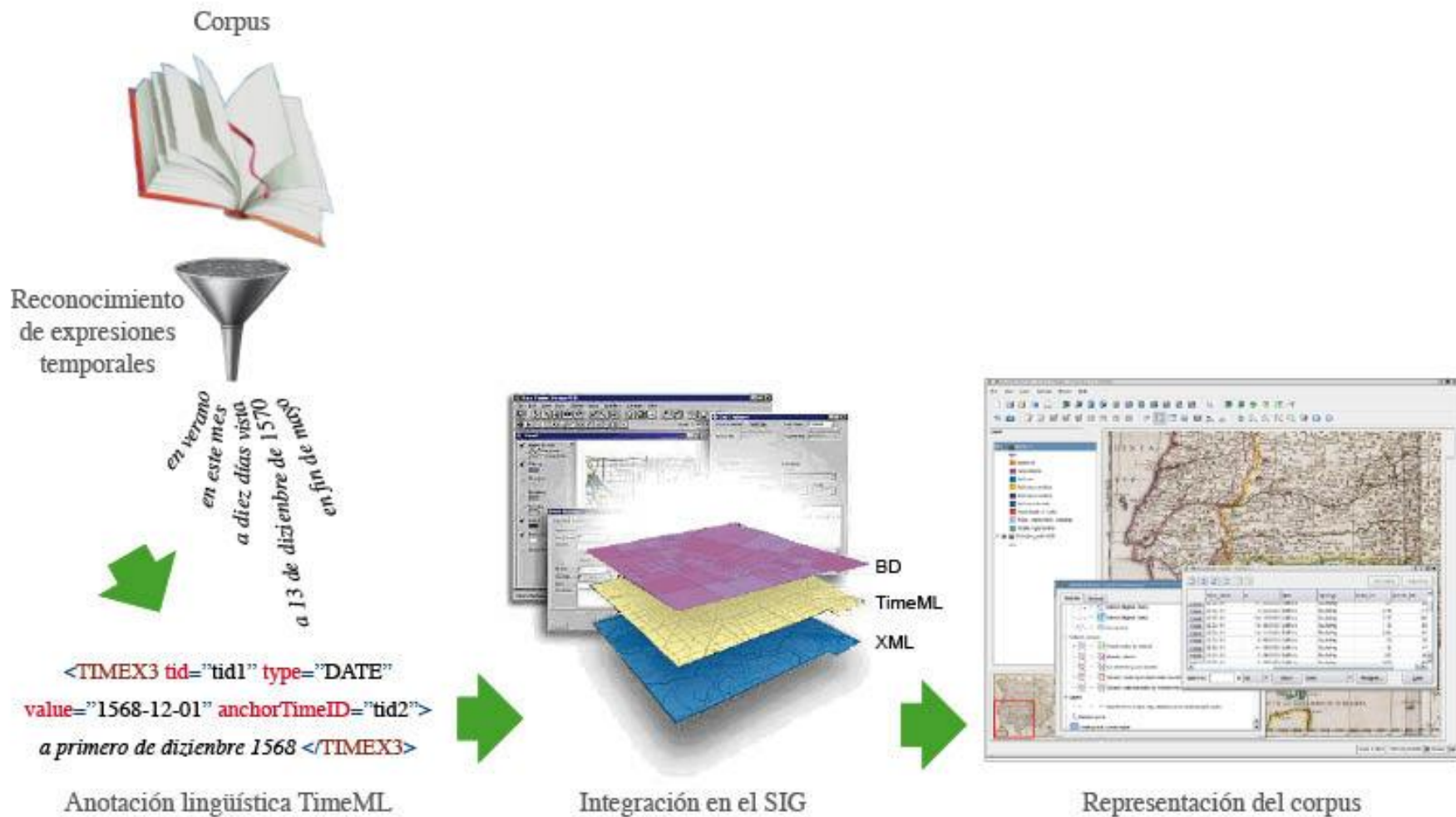
**gvsig**

"COOPERAR ES CRECER"



# MOTIVACIÓN

## Flujo de la integración de la anotación lingüística temporal y el SIG



# OBJETIVOS

---

- ▶ Anotación lingüística de expresiones temporales de un corpus histórico en español.
- ▶ Ampliar las fuentes de datos de un SIG para que incorpore corpus etiquetados con TimeML.
- ▶ Representación y análisis del corpus en un SIG.



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

**gvsig**

"COOPERAR ES CRECER"



# CONTEXTO

---

- ▶ Desde la creación del SIG se han venido dando una serie de transformaciones que han tenido relación con las necesidades de análisis y la representación de información.
- ▶ Para poder representar los cambios continuos y discretos de las entidades territoriales sea incluido la variable tiempo en el SIG.
- ▶ La incorporación del tiempo implica la modificación de la estructura de la base de datos espacial para poder gestionar ambas variables.

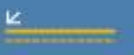


1<sup>as</sup>

Jornadas de Latinoamérica y Caribe  
USUARIOS DE

**gvSIG**

"COOPERAR ES CRECER"



## ESTADO DEL ARTE

---

- ▶ Modelo de instantánea (Langran, 1988)
- ▶ Compuesto espacio-temporal (Langran, 1988)
- ▶ Estampado temporal simple (Hunter, 1990)
- ▶ Modelo de tres dominios (Yuan, 1994)
- ▶ Modelo espacio-temporal orientado a objetos (Worboys, 1994)
- ▶ Modelo orientado a eventos (Peuquet, 1995)
- ▶ Modelo entidad-relación espacio-temporal (Tryfona, 1999)
- ▶ Modelo objeto-relacional (Parent, 1999)
- ▶ Objetos en movimiento (Güting, 2000)



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe  
USUARIOS de

**gvSIG**

"COOPERAR ES CRECER"



## ¿QUÉ ES TimeML?

---

- ▶ TimeML (Pustejovsky, 2005) es un lenguaje de marcado de expresiones temporales, eventos y sus relaciones.
- ▶ Las relaciones temporales están basadas en los conceptos temporales de Allen (1984): *before, after, included, at, starts, finishes, etc.*
- ▶ Tiene dos objetivos principales:
  1. Representación de los acontecimientos y las expresiones temporales
  2. Organización de los eventos en el tiempo teniendo en cuenta el anclaje temporal.



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

gvsig

"COOPERAR ES CRECER"





## ¿QUÉ ES TimeML?

---

- ▶ TimeML fue desarrollado en 2002, en un workshop denominado TERQAS.
- ▶ En 2003 se desarrolló más en el workshop TANGO.
- ▶ TimeML está siendo consolidado como estándar ISO (ISO WD 24617-1:2007).
- ▶ Aprobado como idioma de anotación para TempEval, una de las tareas del International Workshop on Semantic Evaluations (Verhagen, 2007, 2009).



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

**gvsig**

"COOPERAR ES CRECER"



## ¿QUÉ ES TimeML?

---

- ▶ A partir de la necesidad de manejar la información temporal surgen la anotación semántica temporal (en español TERSEO).
- ▶ Quiere solventar el problema de extracción de información temporal basada en pregunta-respuesta.
- ▶ Los primeros documentos analizados fueron del ámbito periodístico.
- ▶ Nace en el contexto del Procesamiento de Lenguaje Natural (PLN).

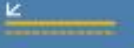


1<sup>as</sup>

Jornadas de Latinoamérica y Caribe  
USUARIOS DE

gvSIG

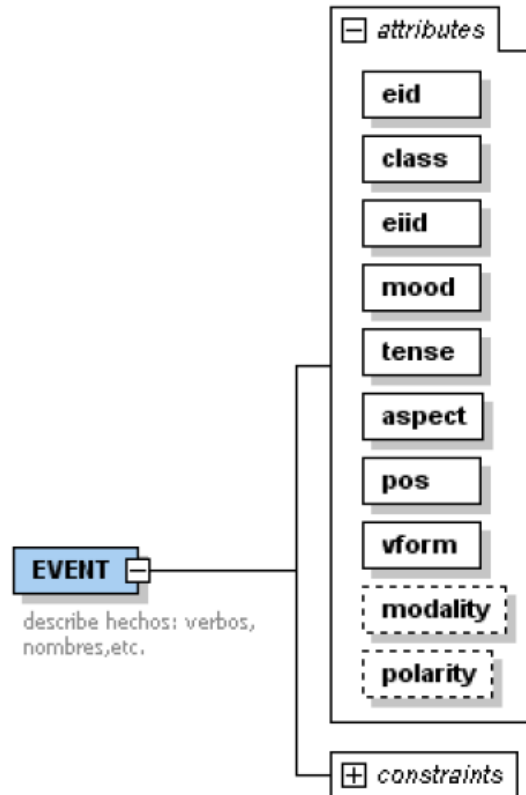
"COOPERAR ES CRECER"



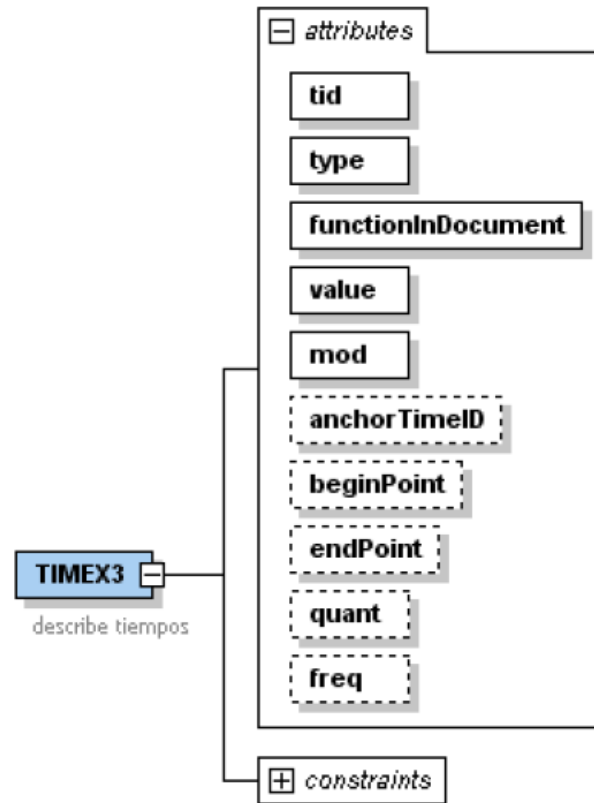
# ¿QUÉ ES TimeML?

## TimeML Etiquetas básicas

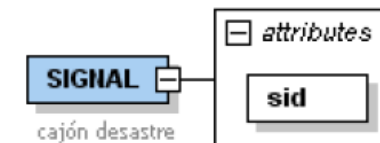
### EVENTOS



### EXPRESIONES TEMPORALES



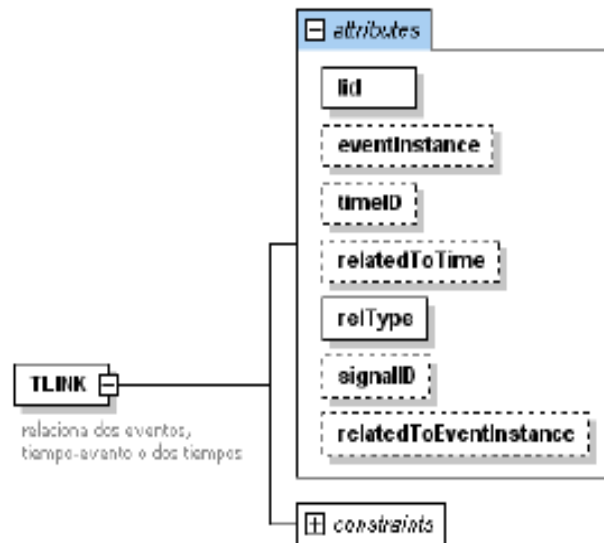
### "SEÑALES"



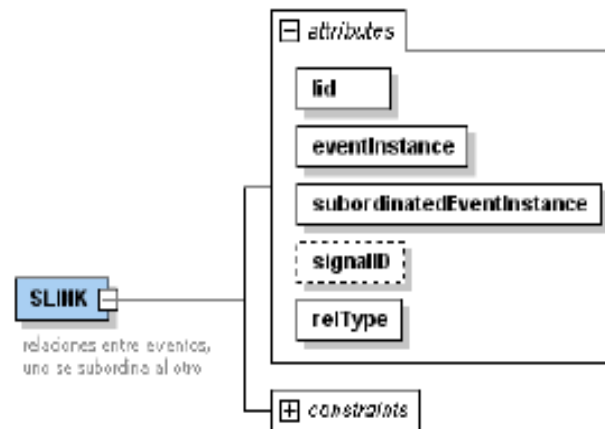
# ¿QUÉ ES TimeML?

## Etiquetas relacionales TimeML

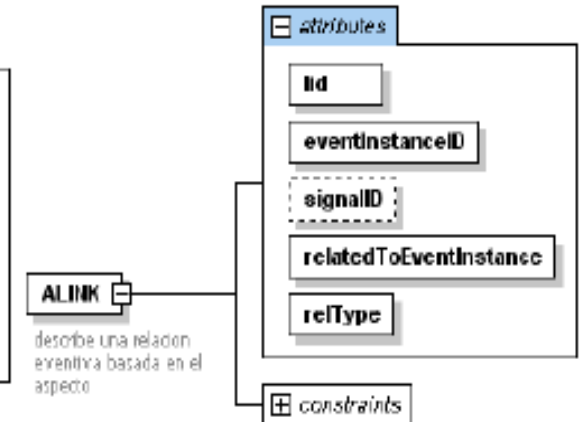
### RELACIONA TIEMPO Y EVENTOS



### RELACIONA DOS EVENTOS (Subordinación)



### RELACIONA DOS EVENTOS (Aspecto)



# METODOLOGÍA

---

- ▶ Anotación lingüística del corpus en TimeML (timex3, event, signal, tlink, alink, slink)
- ▶ TimeML posee una estructura equiparable a una base de datos, por lo que se puede ser incorporada a un SIG.
- ▶ La expresiones anotadas, extraídas del corpus, se almacenarán en esta base de datos.
- ▶ Una vez dentro de una base de datos se contará con herramientas de análisis mucho más potentes.



1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

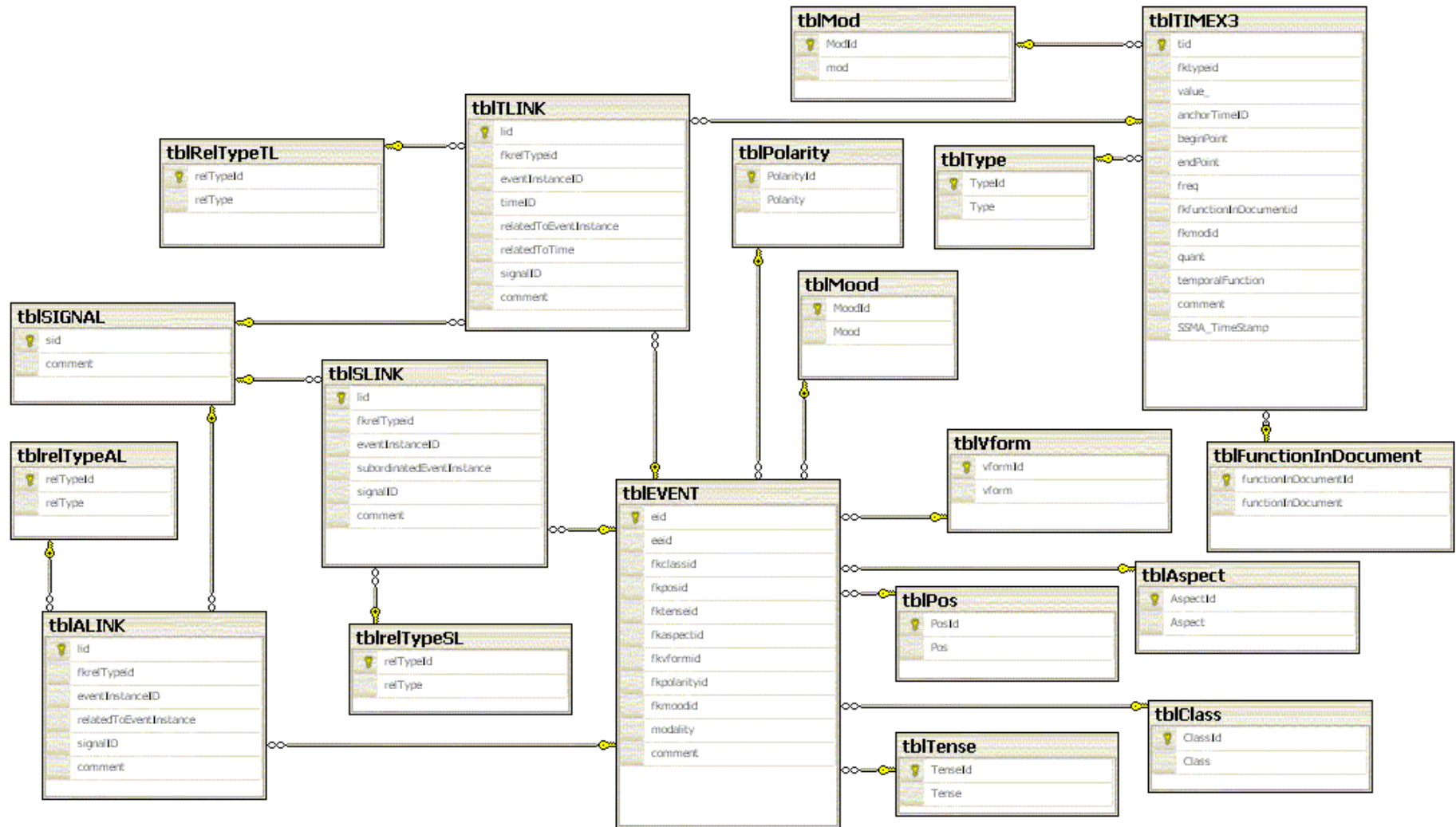
**gvSIG**

"COOPERAR ES CRECER"



# METODOLOGÍA

## Diagrama de la base de datos TimeML



# RESULTADOS

---

- ▶ Anotación lingüística del corpus escrito en español del S.XVI.
- ▶ Preparación de diccionario de expresiones temporales indefinidas.
- ▶ Creación de reglas para la identificación de expresiones temporales y eventos.
- ▶ Estudio de factibilidad para la modificación de herramientas de anotación automática al espa



# TRABAJOS FUTUROS

---

- ▶ Preparación para integrar el SIG con la anotación lingüística espacial.
- ▶ Preparación para enlazar en el SIG la anotación lingüística espacial y temporal.
- ▶ Visualización del corpus anotado en el SIG.



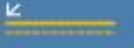
1<sup>as</sup>

Jornadas de Latinoamérica y Caribe

USUARIOS DE

**gvsig**

"COOPERAR ES CRECER"





# REFERENCIAS

---

- ALLEN, J. A general model of action and time. *Artificial Intelligence* 23, 2, (1984).
- FERRO, L., GERBER, L., MANI, I., SUNDHEIM, B., & WILSONSON, G.. TIDES 2005 Standard for the Annotation of Temporal Expressions. The MITRE Corporation, (2005).
- GÜTING, R.H., BOHLEN, M.H., ERWIG, M., JENSEN, C.S., LORENTZOS, N.A., SCHNEIDER, M. AND VAZIRGIANNIS, M. 2000. A foundation for representing and querying moving objects. 25, 1-42.
- HUNTER, G.J. AND WILLIAMSON, I., The development of a historical digital cadastral database†. *International Journal of Geographical Information Science* 4, 169, (1990).
- LANGRAN, G. AND CHRISMAN, N., A Framework For Temporal Geographic Information. *Cartographica: The International Journal for Geographic Information and Geovisualization* 25, pp. 1-14, (1988)
- LANGRAN, G., Time in Geographic Information Systems, Taylor & Francis.. pp. 27-67, (1993).
- LAURINI, R., Thompson, D., Fundamentals of Spatial Information systems. Academic Press Limited, London, UK (1992).
- MANI, I., PUSTEJOVSKY, J., AND GAIZAUSKAS, R. (eds.) *The Language of Time*, Oxford University Press, (2005).
- MONTANARI, A., PUSTEJOVSKY, J., AND REVESZ, P., Eds. *TIME 2006: International Symposium on Temporal Representation and Reasoning*, Budapest, Hungary, (2006).
- MONTGOMERY, L., *Temporal Geographic Information Systems Technology and Requirements: Where we are today*, (1995).
- PARENT, C., SPACCAPIETRA, S. AND ZIMÁNYI, E., Spatio-temporal conceptual models: data structures + space + time. In 26, (1999).
- PEUQUET, D.J. AND DUAN, N., An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. *International Journal of Geographical Information Science* 9, 7, (1995).

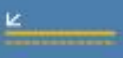


1<sup>as</sup>

Jornadas de Latinoamérica y Caribe  
USUARIOS DE

gvsig

"COOPERAR ES CRECER"



# REFERENCIAS

---

- PEUQUET, D., Time in GIS and Geographical Databases. In Longley et al (eds), Geographics Information Systems: Principles, techniques, applications and management, pp 91-103, (1998).
- PUSTEJOVSKY, J., KNIPPEN, R., LITTMAN, J., AND SAURÍ, R. Temporal and event information in natural language text. Language Resources and Evaluation 39, 2-3, 123-164. (2005).
- PUSTEJOVSKY, J., CASTAÑO, J., INGRIA, R., SAURÍ, R., GAIZAUSKAS, R., SETZER, A., KATZ, G., AND RADEV, D. TimeML: Robust specification of event and temporal expressions in text. In AAAI Spring Symposium on New Directions in Question-Answering (Working Papers), Stanford, CA, pp. 28-34., (2003).
- SAQUETE, E., MARTÍNEZ-BARCO, P., MUÑOZ, R., NEGRI, M., SPERANZA, M., SPRUGNOLI, R., Automatic resolution rule assignment to multilingual Temporal Expressions using annotated corpora, Proceedings of the Thirteenth International Symposium on Temporal Representations and Reasoning, pp.218-224, (2006).
- SAQUETE, E., FERRÁNDEZ, O., FERRÁNDEZ, S., MARTÍNEZ-BARCO, P., MUÑOZ, R., Combining automatic acquisition of knowledge with machine learning approaches for multilingual temporal recognition and normalization. pp. 3319-3332, (2008).
- TRYFONA, N. AND JENSEN, C.S., Conceptual Data Modeling for Spatiotemporal Applications. Geoinformatica 3, pp.245-268, (1999).
- VERHAGEN, M., MANI, I., SAURI, R., LITTMAN, J., KNIPPEN, R., JANG, S. B., RUMSHISKY, A., PHILLIPS, J., AND PUSTEJOVSKY, J. Automating temporal annotation with TARSQL. In 43rd Annual Meeting of the Association for Computational Linguistics, Ann Arbor, Michigan, (2005).
- WORBOYS, M.F., A Unified Model for Spatial and Temporal Information. The Computer Journal 37, pp.26-34 (1994).
- YUAN, M., Wildfire conceptual modeling for building GIS space-time models. In GIS/LIS '94, Phoenix, USA, (1994).



---

# Muchas gracias

**Miembros del proyecto DynCoopNet UPM**  
(*mguerrero, adolfo.urrutia, mjosegr, ma.bernabe*)@topografia.upm.es