

Sixth Framework Programme  
Information Society Technologies  
Future Emerging Technologies



Project Acronym:

**GeoPKDD**

Project full title:

**Geographic Privacy-aware Knowledge Discovery and  
Delivery**

Project Number: FP6-014915

Instrument: Specific Target Research Project

<p><b>Publishable executive summary</b></p>
---

Date of preparation: M7

Revised: M7

Revision: final

Operative commencement date of Contract: 01/11/2005

Project Coordinator: Fosca Giannotti

Project institute coordinator: Knowledge Discovery and Delivery-LAB / ISTI-CNR

**Programme Name:** .....IST  
**Project Number:**.....014915  
**Project Title:** .....GEOPKDD

**Document Type:** .....Deliverable  
**Contractual Date of Delivery:** .....M7  
**Actual Date of Delivery:** .....M7  
**Title of Document:** .....Technical Report  
.....Publishable executive summary

**Author(s):** .....Fosca Giannotti (ISTI-CNR),  
.....Dino Pedreschi (ISTI-CNR)

**Dissemination Level**..... This document is public

### GeoPKDD consortium participants

Partic. Role	Partic. N.	Participant name	Short name	Country
CO	1	<b>KDD Lab.</b> joint research group of <b>ISTI-CNR</b> , Istituto di Scienza e Tecnologie dell'Informazione, Pisa. <a href="http://www-kdd.isti.cnr.it/">http://www-kdd.isti.cnr.it/</a> and <b>Univ. Pisa</b> , Dept. of Computer Science <a href="http://www.di.unipi.it">http://www.di.unipi.it</a>	<b>KDDLAB</b>	I
CR	2	<b>Univ. of Hasselt</b> , Theoretical Computer Science Group. <a href="http://alpha.uhasselt.be/research/groups/theocomp/">http://alpha.uhasselt.be/research/groups/theocomp/</a>	<b>HASSELT</b>	B
CR	3	<b>EPFL - Ecole Polytechnique Fédérale de Lausanne</b> , Lab. DB, Lausanne. <a href="http://lbdwww.epfl.ch/e/">http://lbdwww.epfl.ch/e/</a> and University of Milan - Computer Science and Communication Department <a href="http://www.dico.unimi.it/">http://www.dico.unimi.it/</a>	<b>EPFL</b>	CH
CR	4	<b>Fraunhofer Institute for Autonomous Intelligent Systems</b> , Sankt Augustin. <a href="http://www.ais.fraunhofer.de/">http://www.ais.fraunhofer.de/</a>	<b>FAIS</b>	D
CR	5	<b>Wageningen UR</b> , Centre for GeoInformation. <a href="http://cgi.girs.wageningen-ur.nl/">http://cgi.girs.wageningen-ur.nl/</a>	<b>WUR</b>	NL
CR	6	<b>Research Academic Computer Technology Institute</b> , Research and Development Division. <a href="http://www.cti.gr/">http://www.cti.gr/</a> and Univ. Piraeus, Dept. of Informatics <a href="http://www.unipi.gr">http://www.unipi.gr</a>	<b>CTI</b>	GR
CR	7	<b>Sabanci University</b> , Faculty of Engineering and Natural Sciences. <a href="http://www.sabanciuniv.edu/">http://www.sabanciuniv.edu/</a>	<b>UNISAB</b>	TK
CR	8	<b>WIND Telecomunicazioni SpA</b> , Direzione Reti Wind Progetti Finanziati & Technology Scouting. <a href="http://www.wind.it">http://www.wind.it</a>	<b>WIND</b>	I
CR	9	<b>Polytechnic University of Madrid</b> , Latingeo Laboratory <a href="http://www.latingeo.net">http://www.latingeo.net</a>	<b>UPM</b>	ES

## 1 MOBILITY, DATA MINING AND PRIVACY

The technologies of mobile communications and ubiquitous computing pervade our society, and wireless networks sense the movement of people and vehicles, generating large volumes of **mobility data**. Miniaturization, wearability, pervasiveness are producing traces of our mobile activity, with increasing positioning accuracy and semantic richness: Location data from mobile phones (GSM cell positions), GPS tracks from mobile devices receiving geo-positions from satellites, etc. The objective of the GeoPKDD project is to discover useful knowledge about human movement behaviour from mobility data, while preserving the privacy of the people under observation. Pursuing this ambitious objective, the GeoPKDD project has started a new exciting multidisciplinary research area, at the crossroads of **mobility**, **data mining**, and **privacy**.

### 1.1 The Geographic Privacy-aware Knowledge Discovery Process

Research on moving-object data analysis has been recently fostered by the widespread diffusion of new techniques and systems for monitoring, collecting and storing location aware data, generated by a wealth of technological infrastructures, such as GPS positioning and wireless networks. These have made available massive repositories of spatio-temporal data recording human mobile activities, that call for suitable analytical methods, capable of enabling the development of innovative, location-aware applications. This is a scenario of great opportunities and risks: on one side, mining this data can produce useful knowledge, supporting sustainable mobility and intelligent transportation systems; on the other side, individual privacy is at risk, as the mobility data contain sensitive personal information. The GeoPKDD project, since 2005, investigates how to discover useful knowledge about human movement behaviour from mobility data, while preserving the privacy of the people under observation. GeoPKDD aims at improving decision-making in many mobility-related tasks, especially in metropolitan areas:

especially in metropolitan areas:

- Monitoring and planning traffic and public transportation systems
- Localizing new facilities and public services
- Forecasting/simulating traffic-related phenomena
- Geo-marketing and location-based advertising
- Innovative info-mobility services
- Detecting changes in collective movement behaviour.

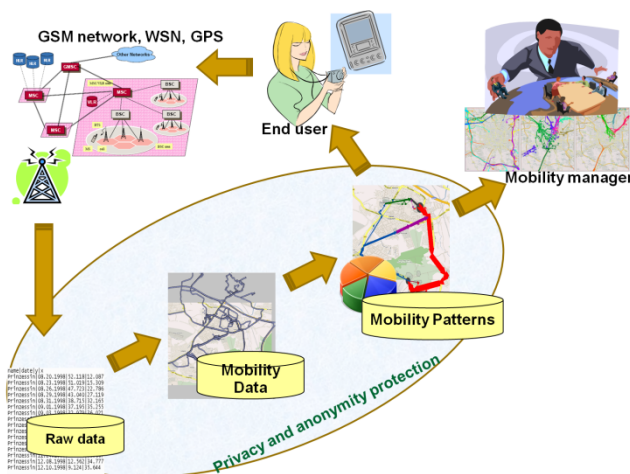


Fig. 1 The GeoPKDD process

### 1.2 Mobility Data Management and Warehousing

The first achievement of GeoPKDD consists in a *trajectory data base management system and warehouse*. A *trajectory*, the basic form of mobility data, is a sequence of time-stamped locations, sampled from the itinerary of a moving object. Our trajectory data base provides efficient means for reconstructing trajectories from raw location data, as well as storing and querying massive trajectory data; the basic analytic framework is also provided: the trajectory warehouse.

**Trajectory Warehouse** The T-Warehouse is a spatio-temporal data cube representing various aggregated measures of the moving objects, such as presence and speed. Our T-OLAP engine supports exploratory analysis, drilling up and down the space and time dimensions.

### 1.3 Mobility Data Mining

The second achievement of GeoPKDD consists in a toolkit of breakthrough analytical methods for mining from massive trajectory dataset, within a privacy-aware framework. Among a wide repertoire of patterns, models and privacy-preserving techniques, we mention the following:

**Trajectory Patterns.** A T-pattern is a sequence of locations that are frequently visited in the specified order with similar transition times; thus, a T-pattern reveals a frequently followed itinerary. Our mining algorithms automatically discover T-patterns in trajectory data.

**Trajectory Clustering.** A T-cluster is a set of similar trajectories, according to a repertoire of trajectory similarity functions; thus, a T-cluster reveals a group of objects sharing a common movement behaviour, e.g., home-work-home commuting. Our density-based clustering algorithms discover T-clusters in trajectory data.

**Trajectory Classification and Location Prediction.** Predictive models for trajectory data include: a classification method for inferring, e.g., the transportation means associated to a trajectory (private car, public transportation, pedestrian, etc.), and a predictor of the next location of a moving object given its past trajectory.

**Trajectory Anonymity.** A  $k$ -anonymous trajectory dataset is one where the itinerary of each person is indistinguishable from that of other  $k-1$  persons – anonymity viewed as *hiding in the crowd*. Our T-anonymity methods transform a trajectory dataset into a new,  $k$ -anonymous dataset, such that the key analytical properties are preserved, together with users' privacy.

### 1.4 Mastering the GeoPKDD Process

The third achievement of GeoPKDD consists in two analytical platforms, supporting the interactive, iterative, combined usage of the various tools to the purpose of discovering mobility knowledge. GeoPKDD developed two prototype platforms: a semantic-based query & reasoning systems, and a visual analytic environment.

**Semantic-based query & reasoning system.** This system allows the user to describe the entire knowledge discovery process using a set of primitives, based onto a **Data Mining Query Language**:

- The *spatio-temporal query primitives* support selection and pre-processing of trajectory data w.r.t. geographic background knowledge, as well as anonymization.
- The *trajectory mining primitives* (see Sect. 2) allow extracting and validating mobility patterns and models.
- A *reasoning component* allows to specify domain-driven *ontologies*, inferring types of trajectories and patterns.

**Visual Analytics.** The aim of this system is to help the analyst navigate through mobility data and patterns and visually drive the analytical process. Some key features:

- *Visualization of T-patterns* to support the navigation of the extracted patterns in the spatial and temporal dimensions.
- *Progressive refinement of T-clusters*: A user-driven exploration and evaluation of the discovered T-clusters, based on a step-wise iterative method.
- *Visual exploration of the T-Warehouse* to browse aggregated measures of moving objects, such as density and speed.

## 1.5 Demonstrating the GeoPKDD Process

The final achievement of GeoPKDD consists in two application demonstrators in the industrial and societal domains:

- a set of analytical services for **mobility analysis and traffic management** in the city of Milano;
- an **educational mobile gaming** service in the city of Amsterdam.

In particular, within the transportation case, we demonstrated how the various methods and systems developed in the project support the creation of novel analytical services for mobility management, such as:

- the automated construction of **origin/destination matrices** from mobility data in a timely, reliable and objective manner, in order to analyze users' flows between geographical areas, overcoming the limitations of the current survey-based approach;
- the detailed analysis and discovery of **systematic movement behaviours**, i.e., the movements that repeat periodically during the week, with particular emphasis to cases of home-to-work and work-to-home commuting patterns.

The educational mobile game demonstrator tackles the problem of empowering the **recreation planners** with knowledge about the actual behavioural patterns of players within a recreation site. Examples include:

- the discovery of Points of Interest in a recreational site, given that game activities are not always performed in planned recreational zones;
- the discovery of similar players' interactions in a recreational site, given that the behaviour of groups of players partly depends on interaction with others (e.g., group competition vs. cooperation).

## 2 PROJECT EXECUTION

GeoPKDD was structured into three *core research* activities:

- i. *Reconstruction of trajectories from streams of raw mobility data, and construction of a privacy-aware trajectory warehouse*: the goal of this activity was to develop a *trajectory warehouse*, a data warehouse specialized for trajectories, also capable to deal with privacy issues. To this aim, it is needed to develop a *trajectory data model*, able to support multiple granularities in space and in time and uncertainty, and to devise OLAP mechanisms for multi-dimensional analysis of trajectory data. These activities were mainly pursued in Workpackage 1 (WP1).
- ii. *Spatio-temporal privacy-preserving data mining and knowledge extraction algorithms, yielding spatio-temporal patterns*: the goal of this activity was to design, implement and experiment algorithms for *spatio-temporal data mining* for mobility data. Such algorithms must be provided with scalable versions for massive datasets, and native methods for provably and measurably protecting *privacy*. These activities were mainly pursued in Workpackage 2 (WP2).
- iii. *Geographic knowledge interpretation and visual analytics techniques to deliver meaningful patterns to end users*: the goal of this activity was driving the analytical process and the interpretation of the resulting spatio-temporal patterns, by means of: i) mechanisms for *reasoning* on the background geographic knowledge and on the extracted patterns; ii) mechanisms to express complex analytical queries into a *data mining query language*, in which the data mining tasks can be formulated; ii) *visual analytics* methods, capable of driving the GeoPKDD process towards interesting spatio-temporal patterns and models. These activities were mainly pursued in Workpackage 3 (WP3).

The work in these research parts was accompanied by other two orthogonal activities:

- iv. *Application demonstrators*, aimed at integrating the achievements of the three core research activities into a coherent framework capable of demonstrating the usefulness of the GeoPKDD process in constructing two application demonstrators in selected societal and industrial domains; These activities were mainly pursued in Workpackage 4 (WP4).
- v. *Harmonization and dissemination*, aimed at making known the potential societal and economic impact of GeoPKDD, and at promoting the harmonization with European and national privacy regulations and the possible collaboration with privacy authorities. An important project goal is to activate a *Privacy Regulation Observatory* able to create relationships with the various organizations related to privacy issues, including non governmental organizations. These activities were mainly pursued in Workpackage 5 (WP5).

**SIG – Special Interest Groups:** 8 Special Interest Groups have been activated and successfully operated to boost inter-WP and inter-partner research.

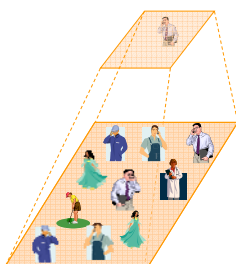
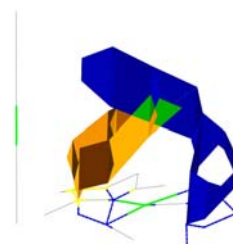
## 2.1 WP1: Privacy-aware trajectory warehouse (Coord.: CTI)

**Semantic-enriched trajectory data model** (EPFL (Leader), CTI, HASSELT): a conceptual data model for trajectory data, which supports: (i) trajectory semantics, based on the concept of stops and moves; (ii) new constructors representing interesting patterns computed by data mining.

**Trajectory reconstruction** (CTI): a technique for trajectory reconstruction, which transforms sequences of raw sample points into meaningful trajectories according different filters (temporal gaps, spatial gaps, maximum speed, tolerance distance).

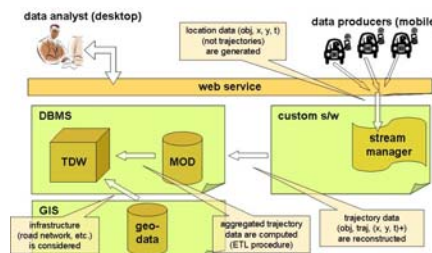
**Trajectory Data Warehouse** (CTI (Leader), EPFL, KDDLAB, HASSELT): the Trajectory Data Warehouse (TDW) has been implemented and tested. Further investigations addressed the definition of adequate aggregate functions for OLAP operations; a new method to compute holistic functions has been developed, such as the aggregate *presence* measure (number of distinct trajectories in a cell).

**Uncertainty management** (HASSELT (leader), CTI): a data model for trajectory data dealing with *uncertainty* in trajectory positions using the concept of bead, or space-time prism, equipped with a query language. This notion has also been applied to the problem of *map matching*, i.e., mapping the points of a trajectory sample to match a road network, taking into account noise and outlier data points, and measurement gaps (e.g., due to tunnels or GPS-warm-up phase).



**Privacy and Security** (CTI (leader), EPFL): K-anonymity for trajectory data: a privacy component for the trajectory database, integrated with the WP2 toolkit of privacy-preserving methods; extension of the GEO-RBAC role-based location-based access control model to protect information available in an enterprise mobile context.

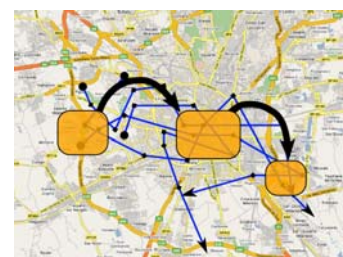
Globally, WP1 delivered a comprehensive **trajectory database management system and warehouse**, capable of acquiring, transforming, storing, querying and analyzing massive trajectory datasets.



## 2.2 WP2 Privacy-aware Spatio-Temporal Patterns and Algorithms (Coord.: KDDLALB)

**Methods for spatiotemporal mining of trajectories data** (KDDLALB (leader), UNISAB, HASSELT, CTI): define the forms of spatio-temporal local patterns and global models to be mined from trajectory data, and design and implement efficient algorithms for extracting such patterns.

**Trajectory Pattern Mining Discovery:** a tool based on the notion of *trajectory pattern* developed during the first year and consolidated in the second, resulting in a solid pattern mining tools made available within the DMQL. During year 3, this method has been applied to various datasets, and was experimented in the *demonstrators*

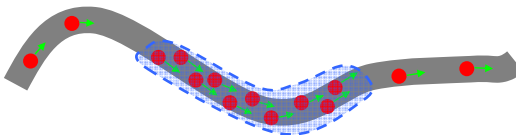


**Semantic Trajectories Mining:** geographic semantics incorporated in spatio-temporal pattern discovery, based on the concept of stops and moves.

**Sequential pattern mining over trajectory:** sequential pattern mining over trajectory data has been developed, introducing the concept of *temporal support of a sequential expression*.



**Similarity notions for trajectories:** a vast repertoire of trajectory similarities, supporting the expressiveness and the efficiency of trajectory clustering.



**T-Optics, the trajectory clustering tool,** which extracts a set of density-based clusters from an input dataset of GPS-like trajectories, has been re-interpreted as a scalable, incremental method within the visual analytics environment, in order to support a user-driven analysis of massive trajectory dataset (result of a tight collaboration with WP3.)

**Location prediction:** the WhereNext method for predicting the future location of a moving object has been developed, based on T-patterns, coherently with the idea that global models can be built out of a collection of local patterns.

**Classification models:** The problem of assigning the category of a trajectory (e.g., the type of vehicle, such as pedestrian, bike, car) has been studied, resulting in a **trajectory classification** framework that is based on the analysis and detection of T-patterns describing how trajectories interact with each other. A scalable algorithm that learns **Bayesian networks** from trajectory data has been designed, which discovers location dependencies.

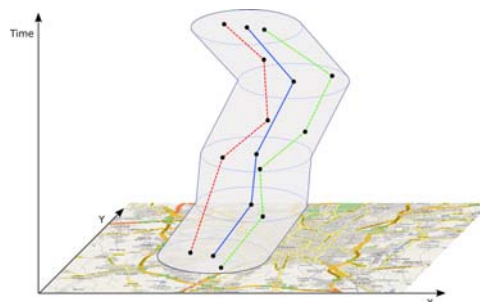


**Methods for privacy-preserving mining of trajectory data** (UNISAB (leader), KDDLALB, CTI): we characterized many subtle **re-identification attacks**, thus bringing further evidence to the fact that a robust notion of k-anonymity is needed for mobility data.

**Trajectory anonymity tools:** TrajKAnon and Never Walk Alone, developed during the first two years of the project, have been thoroughly validated, showing their complementary strong and weak points, so that we now have a methodology to select the most appropriate method depending on the specific GPS trajectory dataset.

**Anonymization methods for sequential data and GSM cell sequences** have been studied and thoroughly validated over real GSM data provided by WIND.

**Privacy-preserving distributed data mining:** efficient crypto protocols for distributed privacy preserving clustering of trajectories have been created, based on secret sharing, which enable cooperative mining without sharing the data.



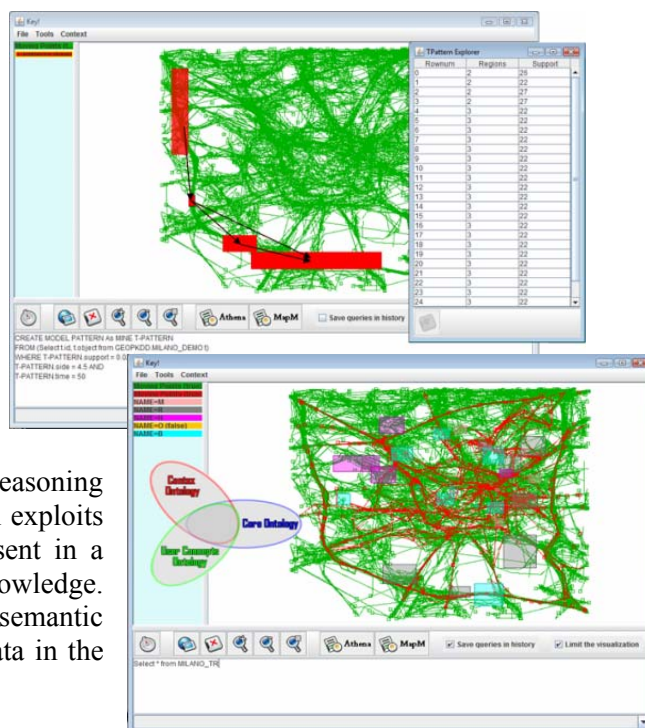
Globally, WP2 delivered a comprehensive **toolkit for privacy-aware mobility data mining**; most such tools have been integrated into the analytical frameworks provided by WP3, namely the DMQL and the visual analytics component.

### 2.3 WP3 Geographic Knowledge Interpretation and Delivery (Coord.: FAIS)

**Data Mining Query Language – Dedalus:** (HASSELT (leader), KDDLALB). A comprehensive query & mining framework has been realized, which extends the object relational model with the mining and entails operators. The modular architecture supported the integration of most algorithms developed in Wp2. The system has been extensively used to support the demonstrators.

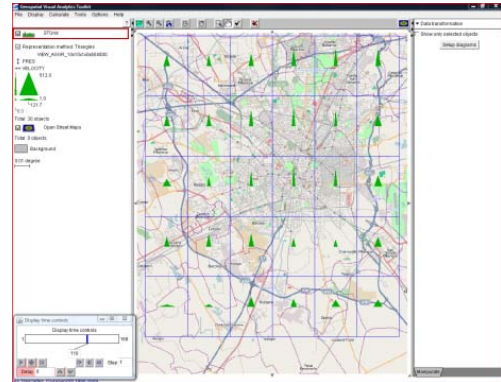
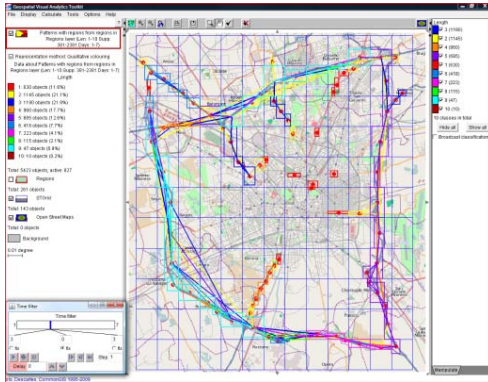
#### Reasoning component – Athena

(KDDLALB (leader), WUR, EPFL): a reasoning module on top of the Dedalus DMQL, which exploits a multi-tier ontological framework to represent in a uniform way data, patterns and domain knowledge. The framework allows the combination of semantic ontology-driven query and reasoning over data in the moving object database.



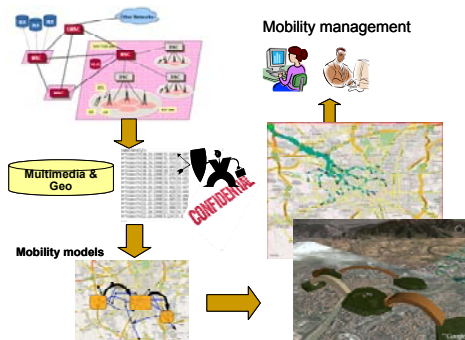
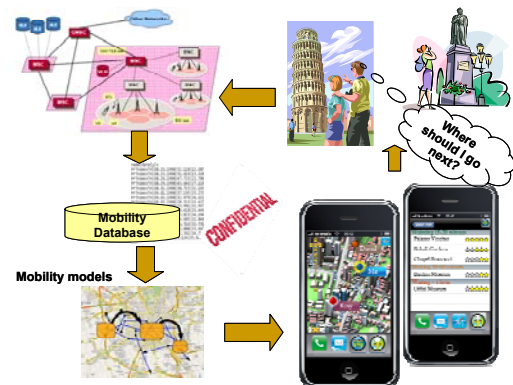
**Visual analytics environment** (FAIS (leader), KDDLALB): suite of tools for visual exploration and analysis of movement data. The toolkit incorporates a variety of data base processing, cluster analysis, and interactive visualization methods for the analysis of trajectories. A visual analytics perspective is adopted: the interactive geographic visualization of trajectory clusters is a useful tool to drive the analytical exploration and progressively refine the queries towards useful results. The system tightly integrates visual analytics with trajectory clustering (from WP2) and Hermes (from WP1). The visual analytics toolkit has been successfully applied to several trajectory datasets.

Particularly, analysis of Milano car tracks and Amsterdam children tracks are parts of the GeoPKDD demonstrators. Analysis of simulated RFID data has been awarded at IEEE Visual Analytics Science and Technology conference Challenge, IEEE VAST 2008.



#### 2.4 WP4: Application Demonstrators (Coord.: WUR-UPM)

**Application demonstrators** (WUR (leader), WIND, KDDLAB, FAIS, CTI): two application demonstrators were selected in the industrial and societal domains: (1) a set of analytical services for traffic managing in the city network of Milano; (2) an Educational Mobile Gaming service in the city of Amsterdam. For the transportation management demonstrator, WIND conducted a small-scale survey addressed to urban planning departments, mobility agencies, traffic managers, and other public administrations responsible for mobility both at urban and regional scale. As a result of several meetings with experts of the Mobility Agency of Milano, the following problems were identified as relevant in the decision making process:



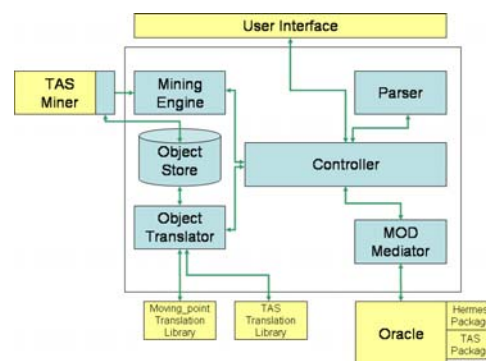
- Provide insights about how the **flow of people** between some given **origin and destination** develops and is distributed along the possible paths.
- Provide **O/D matrices** in a timely, reliable and objective manner, overcoming the current limitations of the survey-based approach.
- Highlight **systematic movement behaviours**, loosely defined as movements that repeat during the week, with particular emphasis to cases of home-to-work and work-to-home.
- Provide information about how the movement of people and the traffic **change** in presence of **exceptional events**, both on large scale (e.g., concerts and demonstrations) and on small scale (e.g., closure of a road segment, or changes of billing rules for the use of some roads, such as the ecoPass introduced in Milan at the end of 2007).

In the educational game case, the problem tackled is how to help recreation planners to understand the *actual behavioural patterns of players within a recreation site*, i.e.,:

- To understand the actual trajectories followed by players;
- To estimate the flow between points of interest;
- To understand the interactions and movement of players and foresee possible new riddles;
- To identify possible players interactions and how the players respond when some unexpected event happens;
- To discover alternative places for the recreational game;
- To discover dangerous and suspicious behaviour in the recreational game.

**Demonstrator platform** (KDDLAB (leader), all partners): Various scenarios of integration between individual modules have been developed and used for demonstrating the above selected services.

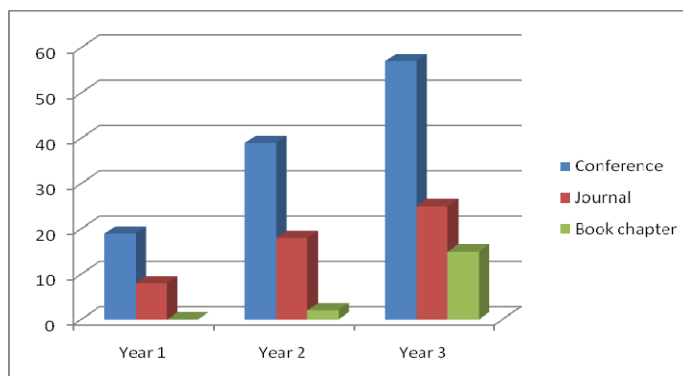
The available GeoPKDD platforms have been experimented on two different data sets in the transportation case: GPS data (17K vehicles with on-board GPS receivers for insurance purposes moving for one week in Milano urban area, approx 200K trajectories); GSM data in Milano area related to people moving during their phone calls. Also, two GPS datasets were used in the recreational demonstrator.



## 2.5 WP5: Harmonization and Dissemination (Coordinator: KDDLAB)

### Dissemination (all partners):

- **GeoPKDD Workshops:** 14 workshops on specific themes of the project have been held, 7 in the third year, among which: The First Interdisciplinary Workshop on *Mobility, Data Mining and Privacy* - Preserving anonymity in geographically referenced data. Rome, Italy - 14 February 2008; The Dagstuhl Seminar 08471 *Geographic Privacy Aware Knowledge Discovery and Delivery*, Dagstuhl 16-21 November 2008; *Workshop on GeoVisualization of Dynamics, Movement and Change* at the AGILE 2008 Conference (Girona, Spain, May 2008).
- Papers presented at **international conferences:** Year 1: **19**, Year 2: **39**, Year 3: **57**.
- Papers published or accepted for publication on **international journals:** Year 1: **8**, Year 2: **18**, Year 3: **25**.
- **Book chapters:** Year 2: **2**, Year 3: **15**.
- **29 invited presentations** and **tutorials** of the project ideas have delivered in research institutions and international events.
- **17 papers** still under evaluation were published within the series of **GeoPKDD technical reports**.



- The web site [www.geopkdd.eu](http://www.geopkdd.eu) received more than 30K hits in the third year of the project, and more than 90K hits overall.
- **GeoPKDD Book:**
  - o *Mobility, Data Mining and Privacy – Geographic Knowledge Discovery*. Fosca Giannotti, Dino Pedreschi (Eds.) Hardcover, 430 pp., Springer, December 2007. ISBN 978-3-540-75176-2.
  - o **500** copies sold by Springer, **200** distributed by the project, **>1000** downloads from Springerlink
  - o 2 further books on the project topics in preparation.
- Communication to general public: project brochure, poster and presentation slides have been produced. A video for the project has been created and presented at the Future of internet conference. An interview with Italian RAI UNO TV News has been realized.
- **Young researchers:** **18** PhD students and **9** post-doc were trained in the GeoPKDD project.

**Privacy Regulation Observatory** (all partners): In the third year, PRO organized a workshop, held in the congress facility of the Italian Data Protection Commission, jointly with the second year review: The First Interdisciplinary Workshop on *Mobility, Data Mining and Privacy - Preserving anonymity in geographically referenced data* Rome, Italy - 14 February 2008 <http://wiki.kdubiq.org/mobileDMprivacyWorkshop/> (105 participants).

As a consequence of this workshop, a network of juridical and legal experts has been consolidated around the many techno-juridical issues of privacy and anonymity. Examples of such liaisons are: Giusella Finocchiaro, professor of Diritto Privato, Univ. Bologna, and Bert-Jaap Koops, Tilburg Institute for Law, Technology, and Society.

Speakers from the project have been invited to the symposium *In search of privacy*; [http://www.ru.nl/ds/ifl/agenda/symposium\\_'in\\_search/](http://www.ru.nl/ds/ifl/agenda/symposium_'in_search/) to be held on June 25th in Radboud University (NL).

An interdisciplinary project proposal on Anonymity funded by the Italian Ministry of Research was to set up and funded. The project is a joint initiative of PRO with a group of jurists of University of Bologna.

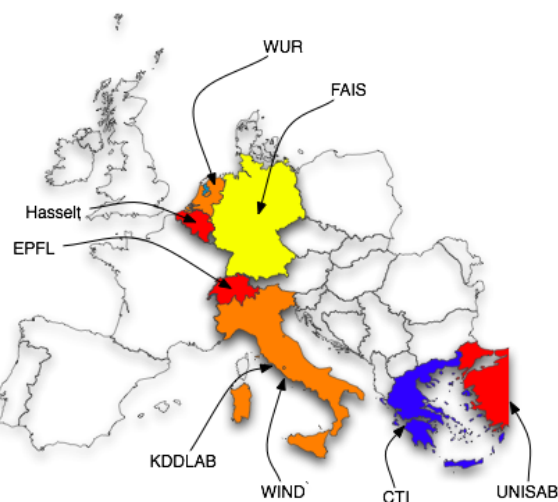
A **Google Research Award** has been granted to Dino Pedreschi for his follow-up research project *Anonimo – An alliance of IT and law for the protection of privacy and anonymity of personal data*.

**Exploitation and Follow up initiatives.** A consolidated analysis of potential applications in the area of urban transportation management was carried out by WIND expanding work of the second year. Such preliminary market analysis turned out to be solid and drove the service definition and development of the demonstrator in transportation analysis/management. The other partners have contributed to share their vision on exploitation, and the overall outcome is reported in Deliverable D5.6 (*Exploitation plan*).

During the third year the consortium put much effort in exploiting GeoPKDD results, ideas and open problems within other projects proposals, both at national and European level, and both industrial projects (exploiting the analytical services invented in GeoPKDD) and frontier research (further exploring the open challenges). Remarkably, as many as **8 GeoPKDD-follow-up projects started during the third year:**

- **FP7 VisMaster FET/225429:** a FET-Open European Coordination Action Project on the Visual Analytics. **Coordinator:** Fraunhofer. **Partners:** 27 groups in industry, public sector and academic institutions, including KDDLAB, FAIS.
- **COST Action, MOVE:** Knowledge Discovery from Moving Objects (Under Evaluation, approved first step), a network for collaboration on knowledge extraction from massive amounts of moving objects data. **Coordinator:** Robert Weibel (Univ. Zurich, CH), **Partners:** 45 groups in industry, public sector and academic institutions, including most GeoPKDD partners.
- **Research Program of Regione Toscana special focus on Transportation: Mercato della Mobilità:** building a car-pooling social network of citizens on the basis of common mobility behaviours. **Coordinator:** Mirco Nanni (KDD Lab, Pisa); **Partners:** KDDLAB, Univ. Pisa, Libero Logico S.r.l.
- **Industria2015/Mobilità sostenibile. MOTUS:** a platform for integrating and managing mobility data acquired by heterogeneous sources to support the development of info-mobility services for tourists. **Coordinator (TELECOM - Italy).** **Partners: 24** Italian groups in industry, public sector and academic institutions, including KDDLAB.
- **ICT-2009.8.0 FET Open Coordinating Action: MODAP : Mobility, Data Mining and Privacy,** aims to stimulate the emergence of a new research community to address mobility data mining together with privacy issues. **Coordinator:** Yucel Saygin UNISAB. **Partners:** GeoPKDD consortium extended with selected research institutions from Dagstuhl Meeting. **Duration:** 30 months;
- **PRIN 2008 (Basic Research program of Italian Ministry of Research), ANONIMO** aimed at investigating the interdisciplinary aspects of data anonymization from the techno-legal perspective. **Coordinator** (Franco Turini, KDDLAB), **Partners:** KDDLAB, Univ. Milano, Univ. Bologna.
- **DFG, German Research Foundation, SPP-ViAMoD: Visual Spatiotemporal Pattern analysis of Movement Data,** Visual analytics methods to support spatiotemporal analysis of movement in a physical space, in particular in a geographical space. **Coordinator** Daniel Keim, Univ. Konstanz, **Partners:** Univ. Konstanz, FAIS.

### 3 GEOPKDD CONSORTIUM



Role	N.	Participant	Area of Expertise	Short name	Country
CO	1	<b>KDD Lab.</b> joint research group of <b>ISTI-CNR</b> , Istituto di Scienza e Tecnologie dell'Informazione, Pisa. <a href="http://www-kdd.isti.cnr.it/">http://www-kdd.isti.cnr.it/</a> and <b>Univ. Pisa</b> , Dept. of Computer Science <a href="http://www.di.unipi.it">http://www.di.unipi.it</a> Subcontractor: ICAR-CNR, Istituto di Calcolo e Reti ad Alte Prestazioni, Cosenza <a href="http://icar.cnr.it">http://icar.cnr.it</a>	Data mining, Spatio-temporal reasoning, Privacy-preserving data mining, Data Mining query languages.	KDDLAB	I
CR	2	<b>Univ. of Hasselt</b> , Theoretical Computer Science Group. <a href="http://alpha.uhasselt.be/research/groups/theocomp/">http://alpha.uhasselt.be/research/groups/theocomp/</a>	Database theory, constraint-based, point-based, and region-based models for spatio-temporal data	HASSELT	B
CR	3	<b>EPFL - Ecole Polytechnique Fédérale de Lausanne</b> , Lab. DB, Lausanne. <a href="http://lbdwww.epfl.ch/e/">http://lbdwww.epfl.ch/e/</a> Subcontractor: University of Milan - Computer Science and Communication Department <a href="http://www.dico.unimi.it/">http://www.dico.unimi.it/</a>	Multiple representation, spatial data warehousing, Ontology driven reasoning	EPFL	CH
CR	4	<b>Fraunhofer Institute for Autonomous Intelligent Systems</b> , Sankt Augustin. <a href="http://www.ais.fraunhofer.de/">http://www.ais.fraunhofer.de/</a>	Machine learning, knowledge discovery, and visual knowledge exploration, multi-relational data mining	FAIS	D
CR	5	<b>Wageningen UR</b> , Centre for GeoInformation. <a href="http://cgi.girs.wageningen-ur.nl/">http://cgi.girs.wageningen-ur.nl/</a>	Geographic knowledge discovery, land use GIS applications, geo-visualization	WUR	NL
CR	6	<b>Research Academic Computer Technology Institute</b> , Research and Development Division. <a href="http://www.cti.gr/">http://www.cti.gr/</a> and Univ. Piraeus, Dept. of Informatics <a href="http://www.unipi.gr">http://www.unipi.gr</a>	Moving Object Databases, Data Warehousing, Privacy-preserving clustering and other data mining methods	CTI	GR
CR	7	<b>Sabanci University</b> , Faculty of Engineering and Natural Sciences. <a href="http://www.sabanciuniv.edu/">http://www.sabanciuniv.edu/</a>	Data mining methods and privacy, data broadcasting and mobile data management	UNISAB	TK
CR	8	<b>WIND Telecomunicazioni SpA</b> , Direzione Reti Wind Progetti Finanziati & Technology Scouting. <a href="http://www.wind.it">http://www.wind.it</a>	Mobile Operator, Mobile Data Provider	WIND	I
CR	9	<b>Polytechnic University of Madrid</b> , Latingeo Laboratory <a href="http://www.latingeo.net">http://www.latingeo.net</a>	Geographic knowledge discovery, land use GIS applications, geo-visualization	UPM	ES

Coordinator: **Fosca Giannotti**, KDDLAB, ISTI-CNR, via Moruzzi n. 1, 56124 Pisa, Italy  
tel: +39 050 3152999 – fax: +39 050 3152811 – email: [fosca.giannotti@isti.cnr.it](mailto:fosca.giannotti@isti.cnr.it)

Webpage: <http://www.geopkdd.eu>