POLITECNICO DI MILANO

DIPARTIMENTO DI
ELETTRONICA,
INFORMAZIONE
E BIOINGEGNERIA

Data, Web and Society (DWS) at Politecnico di Milano

OpenHouse @ DEIB 2015

Presented by Marco Brambilla

Data, Web and Society

- **The new "data revolution"**
  - From "business" to "social"
  - Data availability anywhere
  - Data deluge
    - Social content production
    - Pervasive systems
  - Big data challenge
    - Understanding and solving big mankind problems, including climate, energy, transportation, health
  - Examples of popular applications
    - Smart cities, disaster recovery, personalized medicine

DWS Research at DEIB

- Forty years of history
  - Start in the early 70ies with the birth of the relational model
  - Always at the forefront of international research
- Topics of excellence

  Distributed Databases, Conceptual Data Modeling, Deductive and Active databases, Context-Awareness and Personalization, Pervasive Databases, Web Application Design, Search Computing, Crowd-based applications, Medical and Genomic Applications, Beyond-the-Desktop Interaction

DWS Relevance

- Most of the industries marking the current ICT revolution (e.g. Google, Amazon) are focused on data management at the large scale
- Availability of open, socially produced and controlled data is changing the society.
- Data management and analysis methods are just dealing with the top of the iceberg: many problems remain open and unsolved.
- Our scientific community is constantly challenged to come up with radically new approaches that can change society (and have impact)

DWS Research

- **Common approach to DWS research:**
  - **Inventing new paradigms** for «data» design and management: models, query languages, interaction paradigms, methods for data analysis, mining and optimization.
  - While **adapting to:**
    - **platform evolution**: distributed & pervasive systems, cloud, Web, mobile,
    - **new interaction paradigms**: (multi-)touch, motion-based, touchless, wearable,
  - **Reaching out to new application domains**: education, social, health, tourism, cultural heritage…

Projects

In the areas of

- Web Technology
- Social Computing
- Adaptive & Pervasive Computing
- User Interfaces and Learning
- Genomic Data Management & Health

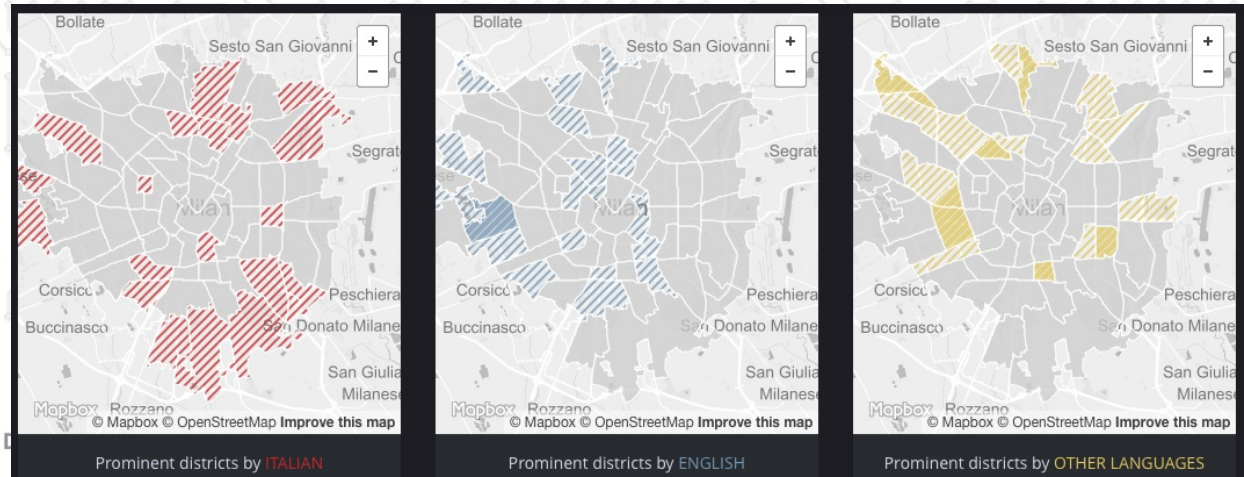
City Data Fusion & UrbanScope

- **Goals**: feel the pulse of our cities in real-time by fusing and making sense of information flows produced both by sensor networks and social networks.
- **Techniques**: semantic technologies, streaming databases, visual analytics, and crowd-sourcing techniques
- **Relevant features:** projects funded by EIT and Polimi, won IBM faculty award 2013.
- **Occupational outcomes**: Develop skills in Big Data Analytics, The fastest growing field in ICT
- **Web:**

citydatafusion.org

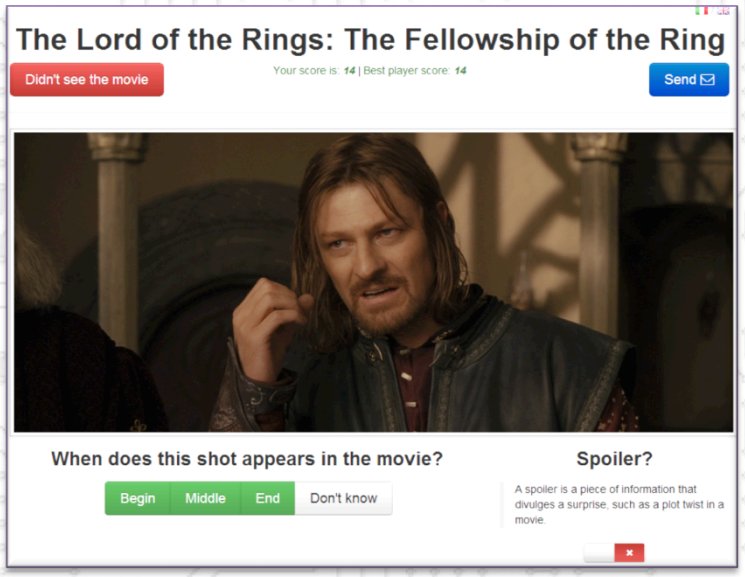
urbanscope.polimi.it

www.cityometers.com



Crowd Searching

- **Goals**: involving & mobilizing crowds for performing tasks, focus on answering data analysis queries.
- **Techniques**: Model & framework for building crowd-based applications over crowdsourcing platforms (AmazonTurk, CrowdFlowers) and social networks (Facebook, Twitter)



Uses reactive control to monitor applications that can span over multiple platforms and dynamically adapt to crowd behavior.

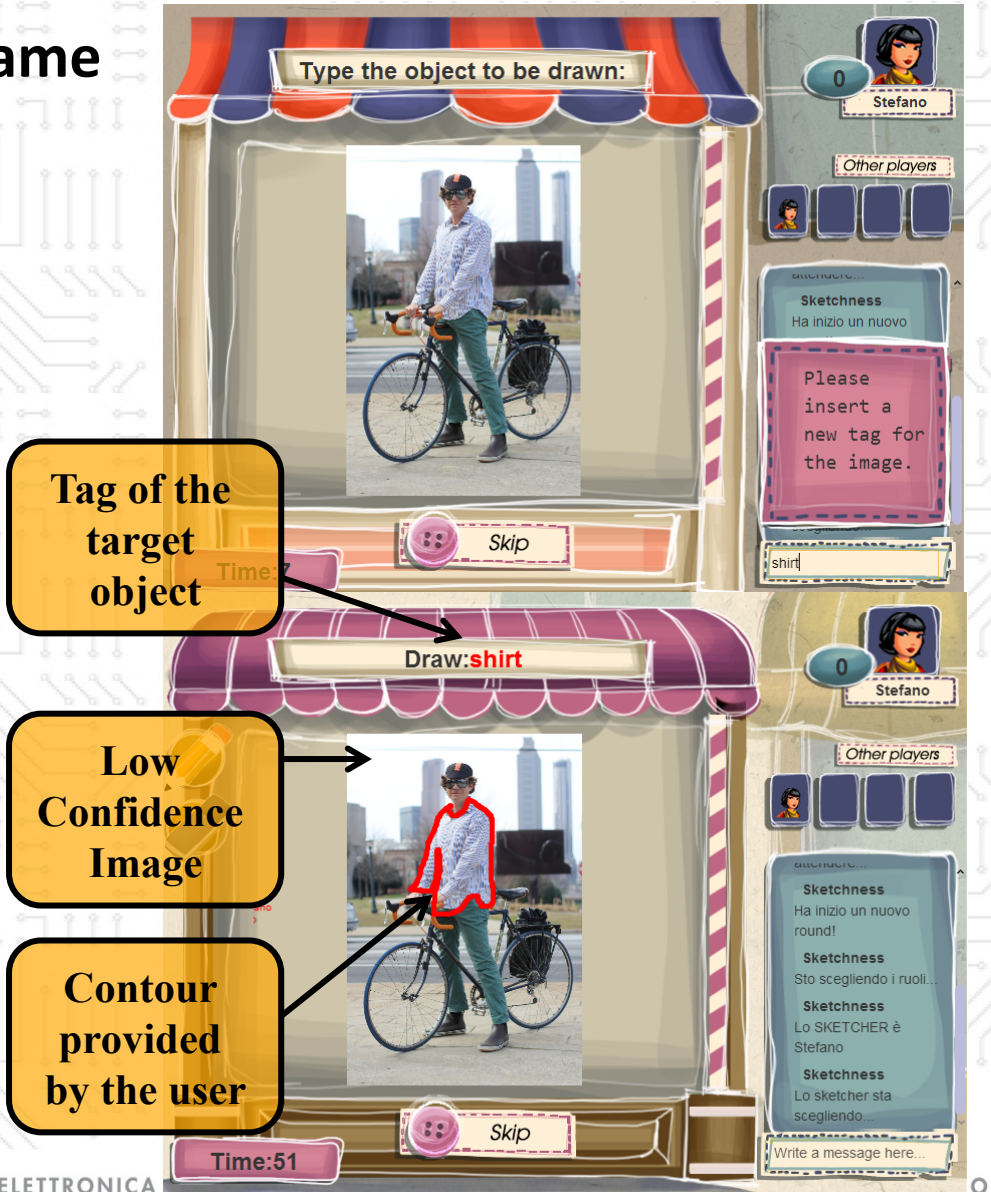
- **Publications:** US Patent, publications at WWW12, WWW13, EDBT13, IEEE Internet Computing 2015, JWE 2015.
- **Occupational outcomes**: Develop skills in design & deploy of crowd-based applications, an emerging field with >=100 new spinoffs

http://crowdsearcher.search-computing.com

Games with a Purpose

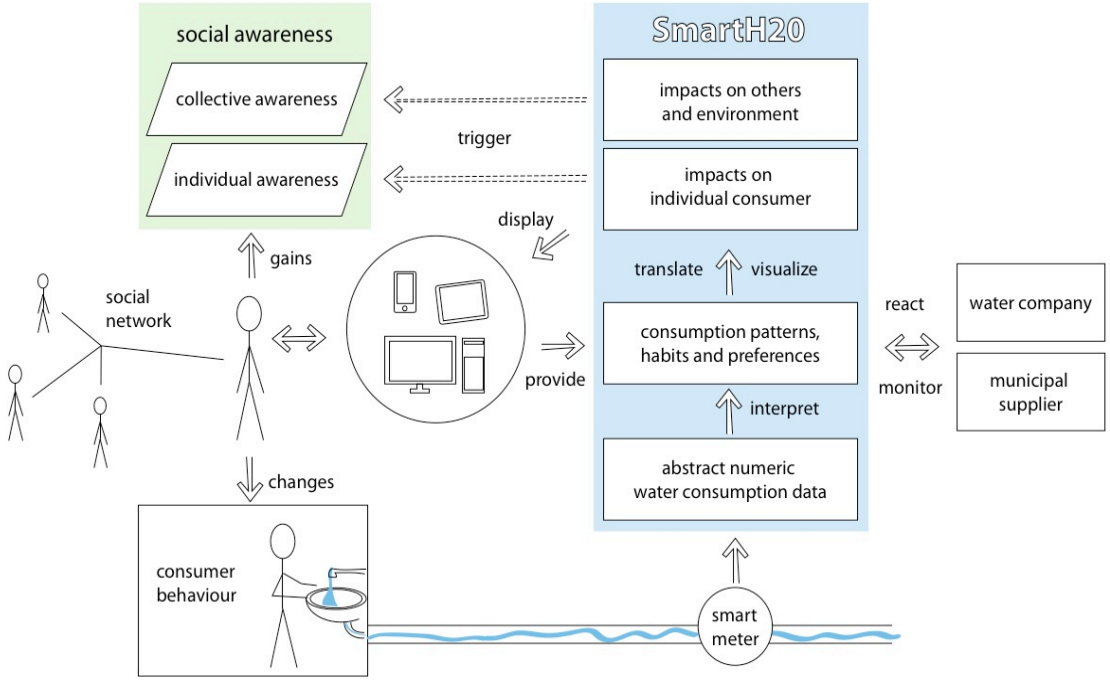
SKETCHNESS: object detection game

- CUbRIK: Human-enhanced time-aware multimedia search (2010-14)
- Human computation applied to multimedia search problems
- Development of games and crowdsourcing applications
- Total cost: 8 900 376 €
- EU funding: 6 834 400 €
- PoliMi funding: 900 000 €



Smart H20

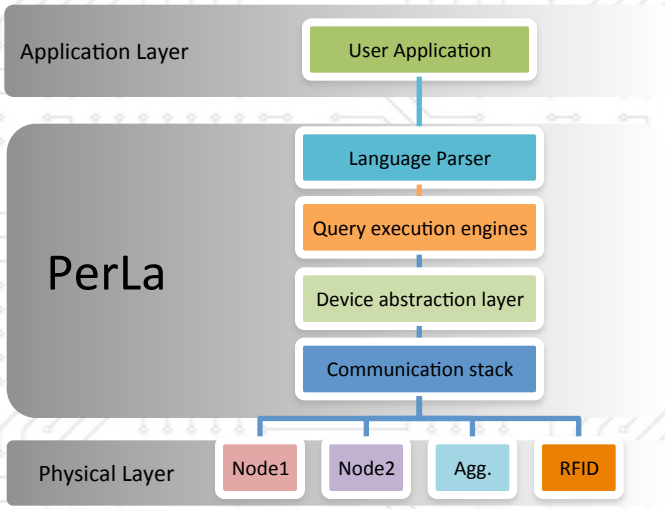
- FP7 STREP Project
- 2014-2017
- Helping people consume less water with persuasive games
- Helping water companies understand how water is consumed with smart meters and social network data
- "Playing" with water pricing models
- 15M users involved (Thames Water, the water company of London)



PerLa

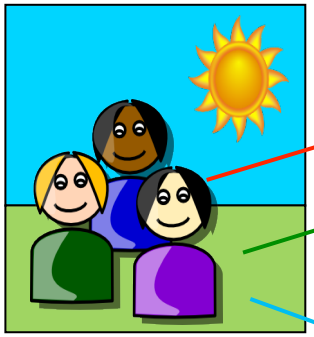
A Language for Managing Data in Pervasive Systems
http://perlawsn.sourceforge.net/index.php

- Use of the DB abstraction:
  - defines a user friendly language, as similar as possible to SQL, to handle pervasive systems
- Heterogeneity
  - deploy-time
  - run-time
- Context management
- Middleware
  - makes the support for new devices easy
  - reduces the amount of the needed low level code



- Projects using PerLa
  - Prometeo (PoliMi): rockfall monitoring in Lecco
  - ArtDeco (MURST-Firb): wine production process monitoring
  - GreenMove (Regione Lombardia): Electric car-sharing management system
  - SeNSori (Industria 2015): Energy Monitoring and saving in buildings

ADaPT: Automatic Data Personalization Based on Contextual Preferences

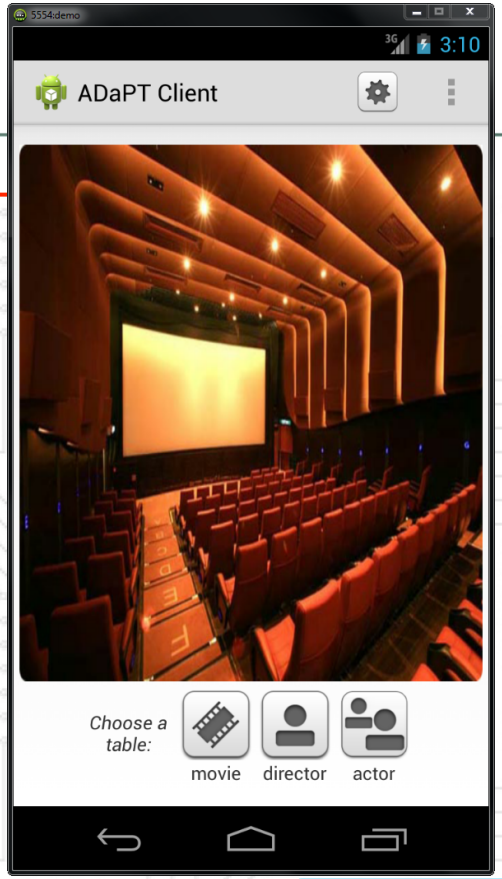


In daytime with friends

What do we want to do this afternoon?

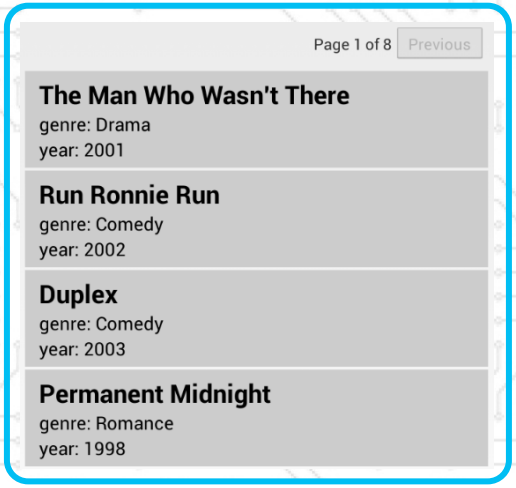
Why don't we watch a movie?

Let's look on the ADaPT app on my mobile phone



At night with friends

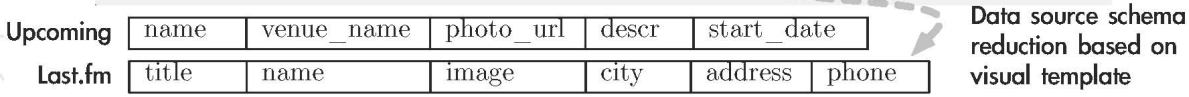
Movie list has changed!



http://tanca.dei.polimi.it/index.php?option=com_content&view=article&id=60:premine&catid=37:progetti-in-corso&Itemid=59

End User Development of Mashups

- Composition paradigms, models and tools for the development, **by the end users**, of multi-device, collaborative mashups



Main ingredients:

- Abstraction from technical details: a platform speaking the **user language** (functionality and terminology), based on UI-centric composition paradigms

- Continuous feedback: Immediate visual feedback → immediate mashup execution, without distinction between design-time and run-time

- Assisted composition: e.g., by means of recommendations

- Different prototypes covering:
  - Lightweight paradigms for data integration
  - Synchronization of widgets
  - Collaborative composition and sharing of interactive Web dashboards and mobile apps

http://home.deib.polimi.it/matera/peudom/index.html

Multimedia Authoring

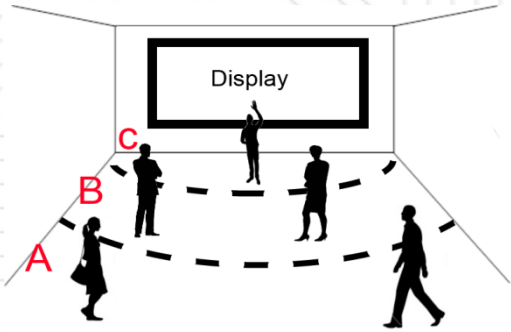
- **Goals**: Tools' development:
  - Authoring environment for multichannel interactive storytelling and hypermedia application development; delivery over PC, smartphone, multi-touch tables, app 4 Apple, YouTube…
  - Innovative portals for rich data
- **Relevant features:**
  - Projects and services (not just prototypes) used by thousands of real users – in schools, cultural institutions, public places
  - Partners: Comune di Milano, Italian Ministry for Education, EXPO2015, Accenture, national and international museums, Ministries for education of 18 European countries…
- **Occupational outcomes**: Develop skills in design & deploy of multimedia multichannel content-rich applications
- **Web:** hoc.elet.polimi.it; www.policulturaportal.it/

Beyond-the-desktop Interaction

Goals: tools and applications involving touchless motion-based interaction with large & small displays, in various application domains: tourism, culture, education, health, domotic

Relevant features:

- integration with the cloud;
- dynamic personalization & adaptation to user profile;
- benefits for users' with special needs (e.g., disabled children)



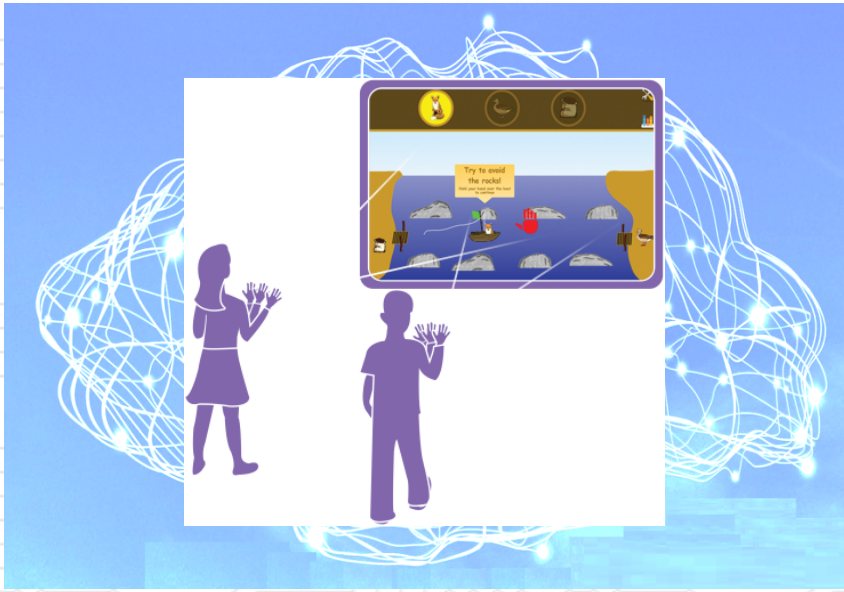
- Collaborations with Therapeutic/education Institutions & Research centers in Spain, Finland, Germany, The Netherlands, Hungary, US

Occupational outcomes: Develop skills in design & deploy of innovative interactive applications

Web: hoc.elet.polimi.it; http://www.m4allproject.eu/

Multimedia & Beyond the Desktop Projects

- **Playful Learning on the Cloud**
  - EIT ICT LAB program 2013; «cloud for people» approach for education and therapy (demos at CEBIT Hannover 2014)

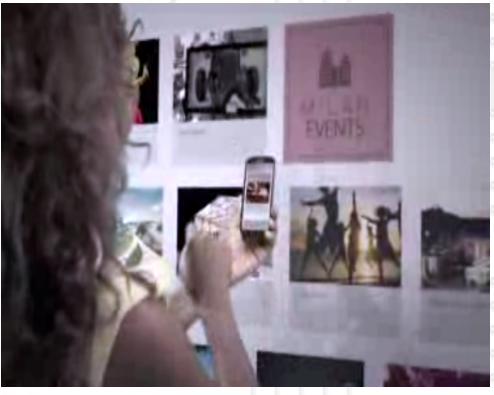


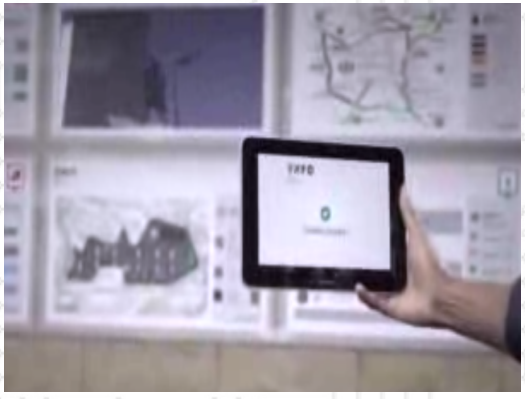
- **Motion-Based Interaction for All (M4ALL)**
  - EC Long Life Learning program; Development of Educational games for disabled Children

- **Multichannel Interaction for Autistic Children**
  - POLISOCIAL Program 2013-2014;
  - Combination of robotic and motion based human-technology interaction

Multimedia & Beyond the Desktop Projects

- **StreetSMART**
  - EIT ICT LAB program 2014; Personalized ambient interaction with public and personal displays



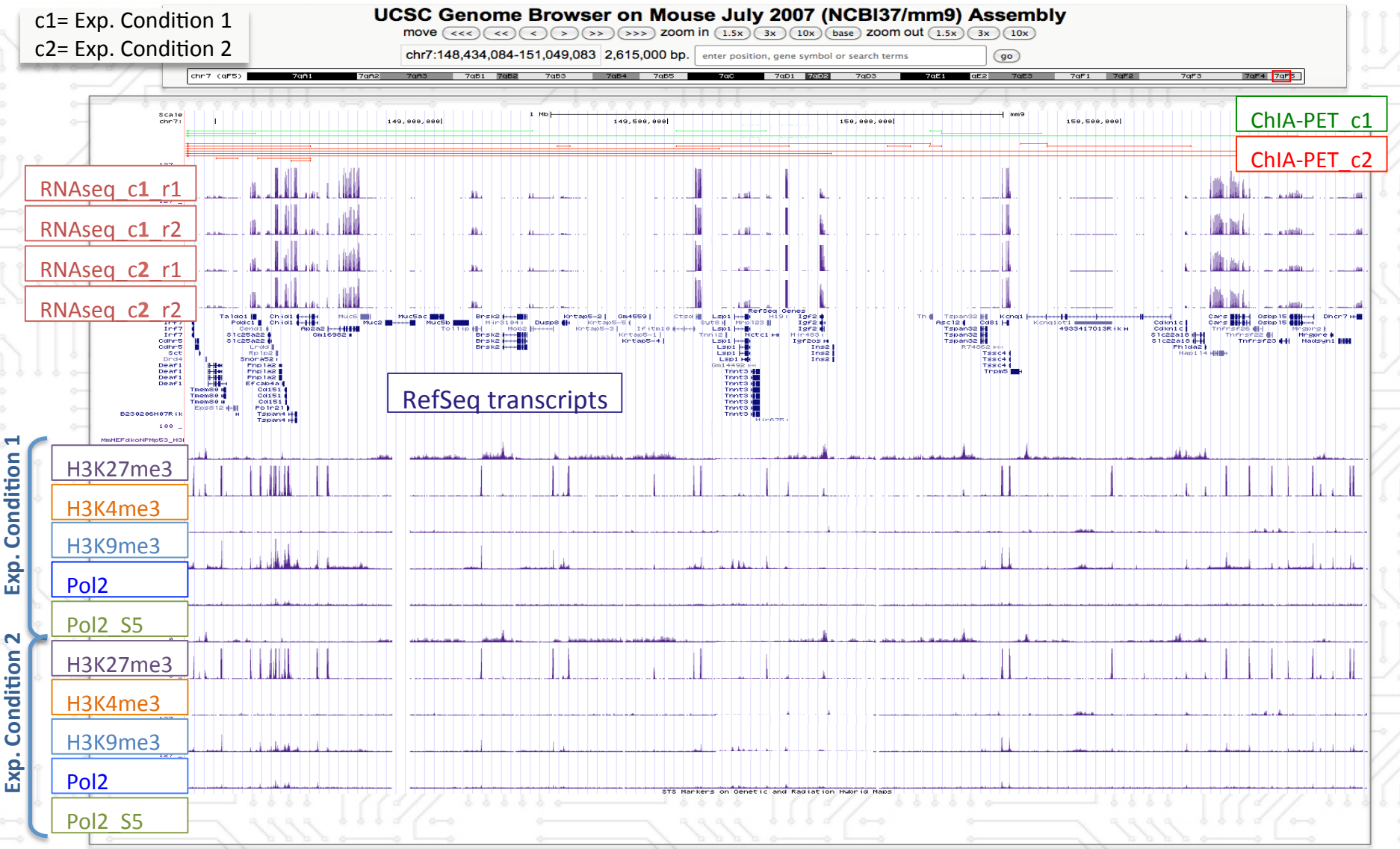


- **COMFIT**
  - EC InterReg Program; tools for integrated multimedia communication for Tourism

Genomic Computing

- **Scenario:** With the cost of whole genome sequencing going down to about 100$ in the next 5 years, genomic data management will soon become the biggest "big data" problem of mankind.
- **Goal**: Support biological and clinical research on genomic data.
- **Technique**: Use of a genomic data model (GDM) and genometric query language (GMQL) to support massive queries over thousands of samples with cloud platforms (Hadoop)
  - Getting ready for massive sequencing of the genome of each individual, leading to personalized medicine.
- **Relevant features:** Joint PoliMi-IEO-IIT project - Strong interaction with European Institute of Oncology, Prof. Veronesi & Pelicci; new PRIN Project "GenData 2020" involving nine excellent Italian research centers; 9 PhD students hired in 2012-2015.
- **Occupational outcomes**: bio-informatics research.
- **Web:** http://www.bioinformatics.deib.polimi.it/genomic_computing/

Genomic Computing
3D Chromatine structure on UCSC Genome Browser



Aging w/o Losing Mobility and Autonomy

In collaboration with the AIR and Architectures groups

- **Goals**: Supports the mobility and autonomy of users through assisted navigation or automatic navigation of wheelchairs.
- **Techniques**: Combines data from indoor localization systems & environment monitoring to plan users' paths tailored to the specific needs of the users. A user-friendly mobile interface guides and monitors the user through the plan.
- **Relevant features:** EU AAL Joint Program, involving 8 partners from 4 EU countries (CH, I, UK, D), 2 end users and 3 SMEs.
- **Occupational outcomes**: Develop multidisciplinary skills applied to the Ambient Assisted Living field, characterized by many challenges, but also opportunities for the citizens, the social and healthcare systems as well as industry.
- **Web**: http://www.alma-aal.org/

DWS in the Bachelor

- **Core Course: Database 1 (3° y, 1° s)**
  - Classic Course on Data Management, teaching two fundamental skills:
    - How to design a database
    - How to query a database
  - Includes foundational aspect of database theory
- **Optional Courses: (3° y, 2° s)**
  - HyperMedia Applications (Web and Multimedia)

DWS in the Master

- Provides core contents to the methodological area:
  - Information Management and Communication
- Offers key ingredients for the specialization tracks:
  - **Big Data**
  - Interactive Applications
  - Bio-informatics and e-Health
  - Pervasive Systems
  - Internet Engineering

Core: Database Systems 2

Follow-up to Database Systems 1, focus on:

- **Technology**:
  - Physical Data Structures
  - Transactions
- **Architectures**
  - Distributed, parallel, replicated databases
- **Language Paradigms**
  - XML Databases and XQuery
  - Object-Oriented and Object-Relational Databases
  - Active Databases
  - NoSQL Databases

Textbook: Atzeni, Ceri, Fraternali, Paraboschi, Torlone:

Basi di dati: Architetture e Linee di Evoluzione, Ed. 2 and Volume Unico (in Italian, in print) complemented by books of Navathe-Elmasri and Widom et. Al. (in English).

Track: Big Data

Objective: enable informed decision-making in enterprises based on accurate data processing and analysis. Appropriate data manipulation upon large-scale repositories generates knowledge, which in turn gives competitive advantage to companies. Focus of the track is on processing, organizing and analyzing data in order to derive knowledge.

Track held in collaboration with **Mathematical Engineering** and of **Management Engineering**

Typical courses: Business Information Systems, Model Identification and Data Analysis, Data Mining and Text Mining, Technologies for Information Systems, Data Management for the Web, Pervasive Data Management, Distributed Systems, Computer Systems Performance Evaluation.

Advanced programming for scientific computing, Digital Business Innovation, and others from Math and Mgmt.

Track: Pervasive Systems

Objective: Designing and deploying systems which are time and location sensitive, responsive to events and event streams in real time, often very small in size.

This Track presents students with the methodologies and technologies that provide the basis for pervasive computing, emphasizing:

- in the methodological line "ICT Management", the aspects related to the management of information and knowledge generated and used by those systems;
- in the line "Software Methodologies", the development of middleware and services;
- in the "Architectures" line, issues of system design from hardware architectures to the design of the system software required to build applications and services.

All lines share both the cultural foundations of computer engineering and the necessary interdisciplinary aspects related to communication networks, modeling and information theory.

Typical courses: Middleware technologies for distributed systems, Distributed systems, Internet of things, Multidisciplinary project, Formal methods for concurrent and real-time systems, Multimedia internet applications, Embedded Systems, Dependable systems,…

Track: Internet Technology

- Objective: creating Internet professionals, jointly designed with TLC.
- New Course: **Data Management for the Web**
  - Topics: Web Information Retrieval, Semantic Web and Open/ Linked Data, Web Design Methods and Tools, Crowd-based Applications.
- **Other ICT Courses:** Advanced User Interfaces, Design and Implementation of Mobile Applications, Middleware Technologies for Distributed Systems, Service Technologies 1 and 2
- **Other Telecom Courses**: Internet of Things, Wireless Networks, Wireless Internet, Multimedia Internet Applications.

Track: Interactive Applications

Objectives:

- Fostering technological and methodological skills needed to design and develop (innovative) interactive applications using conventional and non conventional devices
- Focus: interfaces & interaction paradigms + underlying technology and algorithms;
- Case of study in different domains (e.g., entertainment, learning, commerce, cultural heritage, tourism, health)

Key courses: Advanced User Interfaces (10 CFU), HyperMedia Applications (Web and Multimedia), VideoGames Design and Programming

Typical Additional Courses: Computer Graphics, Pervasive Data Management, Design and Implementation of Mobile Applications, Internet of things, Mathematical Game Theory

Track: BioInformatics & e-Health

- **Objective**: Application of computer science principles to biology and medicine to increase biomedical knowledge and improve health care
- Area Course: **Genomic Data Management & Analysis**
  - Topics: Genetic and molecular biology concepts, Bioinformatics techniques, Bio-ontologies & Bio-databases, Bioinformatics tools, Medical information and standards, Interoperability in healthcare
- **Main Courses**: ICT for Health Care and Life Sciences, Bioinformatics and Computational Biology for Molecular Medicine
- **Other ICT Courses**: Data Mining and Text Mining, knowledge Engineering, Data Management for the Web, Sistemi Informativi Sanitari (Bio), Elaborazione di Segnali e Immagini Biomediche [2] (Bio)

Technologies for Information Systems

Information Systems in the Big Data era:

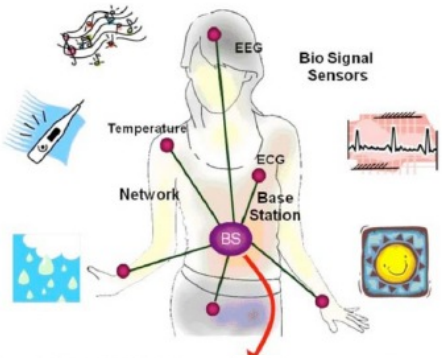
HOW DO WE MAKE SENSE OF THIS DATA FLOOD?

Challenges

- Internet of Things
  - Autonomic data-sources
  - Data and Users Mobility
  - Embedded systems Databases
  - Uncertainty and Lineage Management
- Multimodality
- Ambient and body intelligence
- Information noise

Problems

- Past problems are not completely solved.
- Missing or expensive infrastructures.
- Sound and complete data integration in open world is unpractical.
- 85% of interesting data is unstructured!



- **Modern Information System Architectures and Heterogeneous Data Integration (12 hrs lectures, 8 hrs exercises):**
- **Data Warehousing and Analysis (10 hrs lectures, 8 hrs exercises):**
- **Time Representation and Management in Information Systems (4 hrs lectures):**
- **Advanced topics (6 hrs lectures): Big Data analysis techniques, introduction to data exploration, intensional data representation, personalization and context-awareness**

Advanced User Interfaces

- **Learning goals:** providing methodological and technological background for designing and developing innovative interfaces exploiting beyond-the-desktop interaction and personalization features
- **Topics:** Interaction Paradigms: Motion-Based, (multi-) touch, tangible, wearable; Personalization; Recommender Systems; Interaction Design Process; Empirical Rsearch Methods; on-line evaluation; off-line evaluation
- **Application domains:** e-learning; e-tourism, e-shopping, iTV, e-health
- **Teaching Model:**
  - learning-by-doing (project-based); case-based
- **Synergies with other areas/courses**:
  - Video-game design and programming; Design and Implementation of Mobile Applications; Soft Computing; courses in the e-health track

Other Courses

- Data Mining and Text Mining
  - **Learning goals:** providing the basic methods for analyzing massive amount of data
  - **Topics:** data representation, data preprocessing, classification, clustering, and association; evaluation of models;
  - **Application domains:** business intelligence, marketing, health, user modeling, etc.
- Videogame Design and Programming
  - **Learning goals:** to learn how to design and develop games for traditional and mobile platforms
  - **Topics:** game design, game programming, formal and dramatic elements, artificial intelligence, procedural content generation, etc.
  - **Application domains:** video game development, gamification, etc.

Pointers

firstname.lastname@polimi.it

- Crowdsearching: Marco Brambilla, Stefano Ceri
- Genomic Computing: Stefano Ceri, Marco Masseroli
- City data fusion, UrbanScope: Emanuele Della Valle
- Games with a purpose, SmartH20: Piero Fraternali
- Perla: Fabio Schreiber
- Adapt: Letizia Tanca, Elisa Quintarelli
- EUD of mashups: Maristella Matera
- Multimedia & Interaction Projects: Franca Garzotto, Paolo Paolini
- Aging: Sara Comai

POLITECNICO DI MILANO

DIPARTIMENTO DI
ELETTRONICA,
INFORMAZIONE
E BIOINGEGNERIA

Thanks