Une rétrospective de la recherche en bases de données au travers des 25 ans de publications BDA

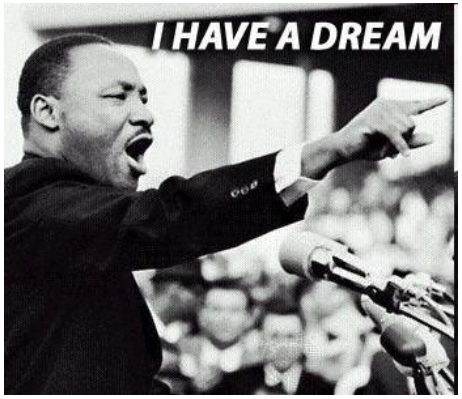
Georges Gardarin (Prism, UVSQ)
Michel Scholl (Cedric, CNAM)

Plan

- **Historique et organisation de la recherche en bases de données en France (1968-1988)**
  - M. Scholl
- **BDA: thèmes de recherche, résultats et perspectives**
  - G. Gardarin

La Préhistoire: 1968

Stanley Kubrick, « 2001, l'Odyssée de l'espace ».
HAL, le nom de l'ordinateur de bord



Printemps de Prague

Evènements de mai

Réalisation du système « Socrate » de gestion de bases de données
Jean-Raymond Abrial, (IMAG, Grenoble).

Années 1970
SOCRATE: SGBD navigationnel

- **Projet SOCRATE:**
  - ABRIAL, CAHEN, FAVRE, PORTAL, MAZARE,MORIN
  - Institut de Math. appliquées, Univ. de Grenoble1971
- **Projet IQS (Interactive Query System)**
  - Jean-Jacques Laclaverie, CII. Interface avec UFAS et IDS2 (Honeywell Bull), avec TDS, bien accueilli par les clients GCOS7
- **Socrate/Clio Syseca, fin des années 70-80**

Années 1970

- **Claude Delobel, Richard G. Casey:**
  - Decomposition of a Data Base and the Theory of Boolean Switching Functions. IBM Journal of Research and Development 17(5): 374-386 (1973)
- **François Bancilhon, Nicolas Spyratos:**
  - Protection of Information in Relational Data Bases. VLDB 1977: 494-500
- **Georges Gardarin, P. Lebeux:**
  - Scheduling Algorithms for Avoiding Inconsistency in Large Databases. VLDB 1977: 501-506

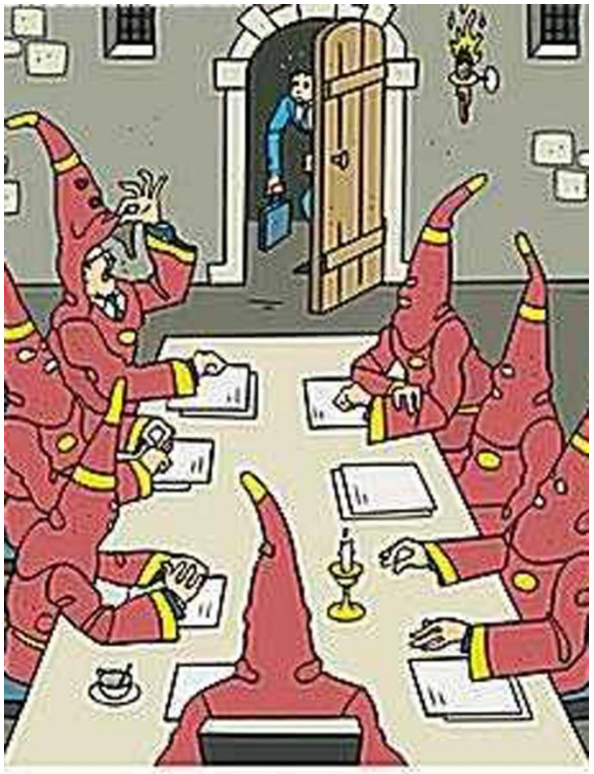
Années 70 SIRIUS

- **Projet pilote IRIA de Bases de Données réparties (1976)**
- **Jean Le Bihan Responsable**
- **Michel Adiba, Jean-Claude Chupin, Robert Demolombe, Georges Gardarin, Jean Le Bihan:**
  - Issues in Distributed Data Base Management Systems: A Technical Overview. VLDB Surveys 1978: 127-153
- **Jean Le Bihan, Christian Esculier, Gérard Le Lann, Witold Litwin, Georges Gardarin, S. Sedillot, L. Treille:**
  - SIRIUS: A French Nationwide Project on Distributed Data Bases. VLDB 1980: 75-85

Années 80

- **1983 Rapport sur les axes de recherche en Bases de Données**
- **Auteurs :**

  M. Adiba,
  F. Bancilhon,
  G. Gardarin,
  H. Gallaire,
  J.M. Nicolas,
  R. Demolombe,
  M. Scholl



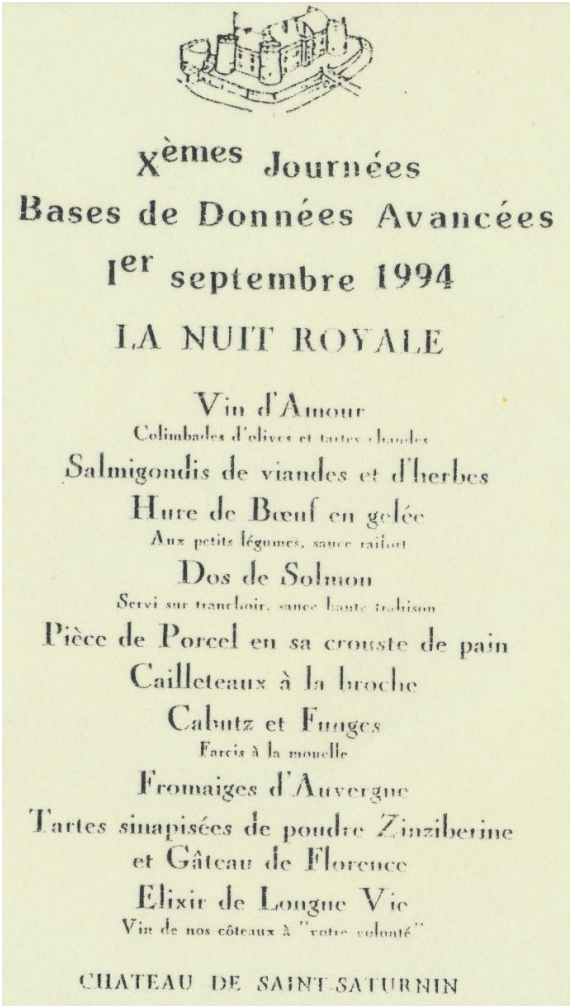
" Sorry, sir, only senior management are allowed in this room."

PRC BD3 1984-1985

- **Programme de recherches coordonnées**
- **Saut d'une génération**

→ Bases de données de 3e génération

BD3 (BDA)
Création de banquets annuels:
St Saturnin 1994



Xèmes Journées
Bases de Données Avancées
Ier septembre 1994

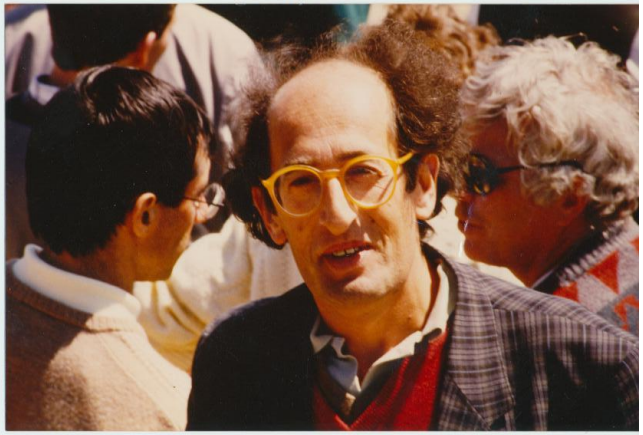
LA NUIT ROYALE

Vin d'Amour
Colimbades d'olives et tartes chaudes
Salmigondis de viandes et d'herbes
Hure de Bœuf en gelée
Aux petits légumes, sauce tailfort
Dos de Solmon
Servi sur tranchoir, sauce haute trahison
Pièce de Porcel en sa crouste de pain
Cailleteaux à la broche
Cabutz et Funges
Farcis à la moudle
Fromaiges d'Auvergne
Tartes sinapisées de poudre Zinziberine
et Gâteau de Florence
Elixir de Longue Vie
Vin de nos côteaux à "votre volonté"

CHATEAU DE SAINT-SATURNIN

Programmes de recherche coordonnées

- **1984-1985 (MRES)**
- **450 000/600 000 €/an**
- **PRC BD3 (1985-1989)**
- **géré par l'INRIA**
- **20 équipes de recherche**
- **Organisées en 3 pôles**
- **80 chercheurs**



PRC BD3 Conseil Scientifique

- **Président C. Delobel**
- **16 membres**
- **Dont 7 industriels (Bull,Intertechnique, Copernique,Telesystemes, Syseca, Rnur, ERLII)**

PRC BD3 Comité de pilotage

- **Les 2 directeurs (GG, MS)**
- **Les 4 responsables de Pôles (M. Adiba, N. Bidoit, C. Rolland, L. Cholvy)**
- **F. Bry (ECRC), J. Kouloumdjian (INSA), F. Velez (ALTAIR)**

Pôle Interface et outils (C. Rolland)

- **Interfaces, mise à jour et conception**
- **9 équipes**
- **LGI (Nguyen-Chiaramella)**
- **MASI (M. Bouzeghoub)**
- **CRIM (Boxenbaum,Pichat)**
- **Cerfia(Chrsiment, Zurfluh)**
- **Crin (Foucaut, Créhange)**
- **Lyon 3 (Flory)**
- **LISAN (Miranda)**
- **Paris 1(C. Rolland)**
- **Bourgogne(Parent,Spaccapietra)**

Pole BD Déductives : 6 équipes (N. Bidoit, L. Cholvy)

- **LRI (N. Bidoit)**
- **CERT/DERI (R. Demolombe)**
- **INRIA/SABRE (E. Simon)**
- **LIUC (Y. Hanatani)**
- **INSA (J. Kouloumdjian)**
- **INRIA/Verso (S. Abiteboul)**

Pole Objets Complexes Multimedia (Adiba/Collet)

- **LGI (M. Adiba)**
- **CRIM (Ferrié)**
- **CERFIA (Chrisment, Zurfluh)**
- **GRTC (J. Le Maitre)**
- **INRIA/SABRE (G. Gardarin)**
- **INRIA/Verso (S. Abiteboul, M. Scholl)**
- **GIP ALTAIR**

Les Journées BDA de 1985 à 2008

Une sélection non exhaustive et partisane de thèmes et d'articles

Référence : voir DBLP et les souvenirs

Actes pas toujours trouvables

1. BDA 1985: St-Pierre de Chartreuse

- **Delobel Pdt, Gardarin CP accepte tout !**

BDA 1985: St-Pierre-de-Chartreuse

- **Thèmes essentiels**
  - Extensions des SGBDR (Sabre, Pépin, Campus …)
    - Déduction : logique, intégrité, règles
    - Influence de Syntex (Toulouse)
    - Documents structurés et multimédia (SGML)
  - Outils de conception, Interfaces
    - SECSI versus OICSI, Relation universelle, SQL++
  - Architecture de machines BD
    - Filtres matériels (Verso versus Schuss)
- **Invités**
  - Dennis Tsichritzis: Outils de la Bureautique
  - Michael Stonebraker: Expert Database Systems
  - François Bancilhon: Règles récursives dans les BD déductives

2. BDA 1986: Giens

BD DEDUCTIVES :

ANC(x,y) ← PAR(x,z), ANC(z,y)

ANC(x,y) ← PAR(x,y)

?ANC(toto,y)

Pb_x(toto)
Magic^{10}(toto)

BDA 1986: Giens

- **La récursion en pointe :**
  - François Bancilhon: Comparaison des performances des méthodes d'évaluation de règles récursives
  - Georges Gardarin & CdM: Evaluation des programmes logiques récursifs par des fonctions récurrentes
  - Jean Rohmer: Une extension de la méthode d'Alexandre pour optimiser le traitement des axiomes récursifs dans les bases de données déductives
- **Le temps et les versions**
- **Les outils de conception**
- **Les interfaces et langages**

3. BDA 1987: Port-Camarque

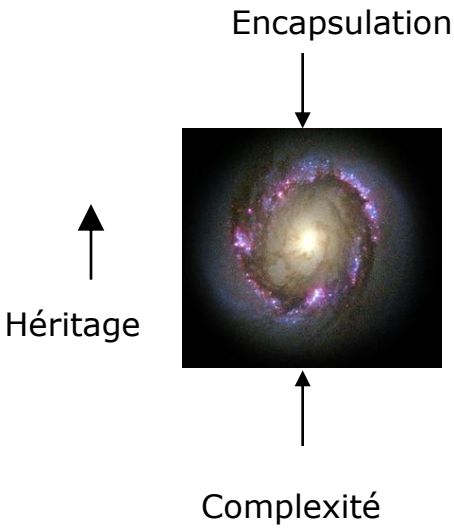
- **Datalog++: les règles avec fonctions, négations, mises à jour (Serge Abiteboul)**

BDA 1987: Port-Camarque

- **Emergence des objets (complexes)**
  - Peter Pistor: Advanced information management prototype
  - Contributions de Verso (langage) et Sabre (Objet-Relat.)
- **Le déductif devient bases de connaissances**
  - Matthias Jarke: Knowledge base management for database software
  - G. Gardarin et Ph. Pucheral se trainent avec l'optimisation de règles récursives
  - Sémantique des requêtes et mises à jour de vues par R. Demolombe et N. Spyratos
- **Abiteboul propose un pont O-R via Datalog++**
- **Les thèmes classiques survivent**
  - Peter Buneman: Langages de programmation et BD
  - Conception, Interfaces, Dynamique

4. BDA 1988: Bénodet

- **Comme un guetteur mélancolique, j'observe la montée de l'objet … (inspiré d'Apollinaire)**



BDA 1988: Bénodet

- **L'envahisseur est là : l'objet**
  - Christophe Lécluse, Philippe Richard, Fernando Vélez: O2 un modèle de données orienté-objets
  - Christine Collet: Gestion d'objets complexes au travers de formulaires dynamiques
  - Qiming Chen, Georges Gardarin: Un modèle d'implantation pour raisonner avec des objets complexes
  - Sophie Gamerman, Fernando Vélez: Comparaison de langages de programmation en utilisant des applications base de données
  - Christophe Damier, Bruno Defude: Un modèle de données pour les informations géographiques
  - Benoit David: Le modèle Spatiarel - Une extension du modèle relationnel pour gérer les données spatiales surfaciques.
- **Les BD logiques s'étendent : Raisonnement, Intégrité**

5. BDA 1989: Genève

- **M. Léonard, …**
- **Datalog intègre les mises à jour**
  - Serge Abiteboul, Eric Simon: Fundamental properties of deterministic and non deterministic extensions of DATALOG
  - Serge Abiteboul, Stéphane Grumbach, Agnès Voisard, Emmanuel Waller: An extensible rule-based language with complex objects and data-functions
  - Laurence Cholvy: Mises à jour dans les bases de connaissances
- **O2 est en plein développement**
  - Christophe Lécluse, Philippe Richard: Langages orienté-objet et bases de donées: l'expérience O2

6. BDA 1990: Montpellier

- **Objet, Déduction**
  - Stefano Ceri: Object orientation and logic programming for databases: a season's flirt or long-term marriage?
- **Multibases (Litwin)**
  - Serge Abiteboul: Virtuality in object-oriented databases.
  - Stefano Spaccapietra: Intégration des vues et relativisme sémantique.
  - C. Bréant: SABRINA-STAR: a cooperation system for heterogeneous and pre-existing databases.
- **Optimisation de requêtes**
  - Béatrice Finance, Georges Gardarin: An extensible query rewriter for databases with objects and rules.
  - Pascale Borla-Salamet, N. Lefèbvre: Une technique de reformulation des constraintes d'intégrité exprimées sur des relations déduites.
  - Rosana S. G. Lanzelotte: Modelling search strategies in an extensible query optimizer.

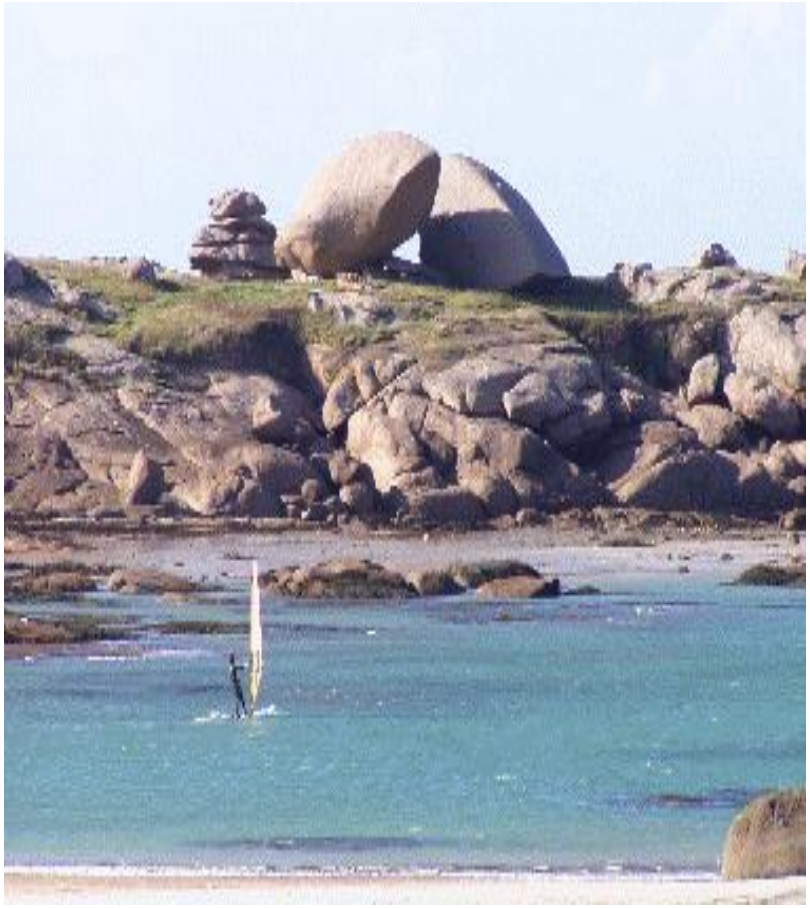
7. BDA 1991: Lyon

- **Dynamique des Schémas**
  - M. Léonard; E. Waller; D. Rieu …
- **Extension des langages de requêtes**
  - Retour au classique + sémantique Datalog
  - G. Gardarin, P. Valduriez: ESQL2: an object-oriented SQL with f-logic semantics
- **Conferences invitées**
  - Dennis Shasha: Promises versus assumptions in database fault tolerance.
  - Jean-Pierre Cheiney: Maintien différentiel ensembliste de fermetures transitives

8. BDA 1992: Trégastel

- **Interfaces et langages (9)**
- **Performance et optimisation**
  - Requêtes (B. Finance, R. Lanzelotte, …)
  - Contrôles (V. Benzaken, A. Doucet, B. Traverson, …)
- **Conception de BDO**
  - Pascal Poncelet, Maguelonne Teisseire, Lotfi Lakhal: IFO2, modèle et principe pour la conception de bases de données avancées.
  - Michel E. Adiba, Christine Collet, Pascal Dechamboux, Bruno Defude: Object oriented persistent application development: model, language and integrated tools.
  - Isabelle Comyn-Wattiau, Mokrane Bouzeghoub: A deductive process to consider constraints during view integration.
- **Applications géographiques**
  - A. Voisard, M. Scholl, R. Laurini, B. Amann

La Côte de Granit Rose



9. BDA 1993: Toulouse

- **Session 4: Perspectives et Transferts en Bases de Données: Dix Ans Après**
  - Table-Ronde animée par Nicole Bidoit, Claude Delobel et Laurent Vieille
- **INRIA O2**
- **ECRC QsQ**
- **Sabre pas invité !**

- **Chevaliers du Taste-vin**
  - Clos de Vougeot, Bourgogne ???

10. BDA 1994: Clermont-Ferrand

- **Un repas mémorable**
  - Château de Saint-Saturnin, 63

BDA 1994: Clermont-Ferrand

- **Typage, Cohérence, Objets, ...**
- **Bases de Données Spatiales**
  - Jan Paredaens: On the Foundations of Spatial Databases.
- **DB Graph**
  - A.Gutiérrez, Ph. Pucheral, H. Steffen, JM. Thévenin: Database Graph Views: A Practical Model to Manage Persistent Graphs.
- **Parallélisme et Transactions**
  - Laurent Daynès, Olivier Gruber, Patrick Valduriez: On the Cost of Lock Inheritance in Lock Managers Supporting Nested Transactions.
  - Mohamed Zaït, Patrick Valduriez, Daniela Florescu: On the Validation of a Parallel Query Optimizer
  - Patrick Casadessus, Benoît Dageville, Luc Bouganim: Performances du SGBD Parallèle DBS3 sur la machine KSR

11. BDA 1995: Nancy

- **Le dorénavant classique**
  - Objet, déduction, concurrence
  - Optimisation de requêtes
- **BD actives et triggers**
  - Christine Collet, Javam Machado: Optimization of Active Rule Processing with Parallelism.
  - Françoise Fabret, Eric Simon: Optimizing Repetitive Computations of Database Triggers within a Transaction.
  - Bruno Defude, Hervé Martin, Anne Picault: Un système de contrôle de l'activité de transactions basé sur des exceptions.
- **Le web devient présent**
  - Peter M. G. Apers: Database-support for the Web
  - J. Le Maître, E. Murascio: SGMLQL
- **Enfin une session applications**
  - Mais ce n'en est pas (interfaces) sauf géo-spatial
- **Et des démonstrations (5)**

12. BDA 1996: Cassis

- **Du classique : Conception, Intégrité, Sémantique, Optimisation …**
  - Georges Gardarin, Fei Sha, Zhao-Hui Tang: Calibrating the Query Optimizer Cost Model of IRO-DB, an Object-Oriented Federated Database System.
  - Georges Gardarin, Jean-Robert Gruser, Zhao-Hui Tang: Algorithmes pour le Traitement des Expressions de Chemin dans les Bases de Données Orientées Objet.
  - Luc Bouganim, Daniela Florescu, Patrick Valduriez: Répartition Dynamique de la Charge dans un SGBD Parallèle Hiérarchique.
- **Et des versions :**
  - Stéphane Gançarski, Geneviève Jomier: Vers un Langage de Manipulation pour Bases de Données Multiversion.
  - Anne Doucet, Stéphane Gançarski, Geneviève Jomier, Sophie Monties: Mantien de la Cohérence dans une base de Données Multiversion.

BDA 1996: Cassis

13. BDA 1997: Grenoble

- **Beaucoup d'aspects système**
  - Sacha Krakowiak: Avancées récentes en systèmes répartis et leur impact sur les SGBD (invité)
- **Des aspects moins classiques :**
  - M. Tamer Özsu: Issues in Multimedia Data Management (conf. invitée).
  - Stéphane Grumbach, Philippe Rigaux, Michel Scholl, Luc Segoufin: DEDALE, A Spatial Constraint Database.
  - Patrick Bosc, Ludovic Lietard, Olivier Pivert: Gradualité, Imprécision et Dépendances Fonctionnelles.
  - Fadila Bentayeb, Dominique Laurent: Inversion de l'Algèbre Relationnelle et Mises à jour.
  - Monique Picavet, Marie-Pierre Van Hoecke: La carte à microprocesseur : une base de données individuelle et mobile dans un système très réparti.
  - Hubert Naacke, Georges Gardarin, Anthony Tomasic: Leveraging Mediator Cost Models with Heterogeneous Data Sources.

14. BDA 1998: Hammamet, Tunisie

Les vues revisitées

- **Alon Y. Levy: Answering queries using views: algorithms, applications and opportunities (invité)**
- **Serge Abiteboul, Bernd Amann, Sophie Cluet, Tova Milo, Victor Vianu: Active Views for Electronic Commerce.**

Le vent du Web

- **Mary F. Fernández, Daniela Florescu, Jaewoo Kang, Alon Y. Levy, Dan Suciu: Catching the boat with Strudel: experiences with a Web Site management system.**

La synthèse en transactions

- **C. Mohan: Transaction Processing and Distributed Computing in the Internet Age (invité)**

BDA 1998: Hammamet

- **Web ?**
- **Electronic commerce ?**

15. BDA 1999: Bordeaux

- **L'émergence (tardive) du Web et de XML**
  - Stefano Ceri: Model and Tools for Design of Data-Intensive Web Applications
  - Nicole Bidoit, Sofian Maabout, Mourad Ykhlef: Un langage imbriqué pour l'interrogation et l'intégration de données semistructurées.
  - Emmanuel Bruno, Jacques Le Maitre, Elisabeth Murisasco: Expression, contrôle et exploitation de la navigation hypertextuelle dans le langage SgmlQL.
  - Fei Sha, Georges Gardarin, Laurent Némirovski: Managing Semistructured Data in Object-Relational DBMS
- **L'intégration du Data Mining (peu réussie)**
  - Fei Wu, Georges Gardarin: Gradual Clustering Algorithms for Metric Spaces.
  - Jean-Marc Petit, Farouk Toumani: Discovering Inclusion and Approximate Dependencies in Relational Databases.
  - Florent Masseglia, Pascal Poncelet, Maguelonne Teisseire: Extraction efficace de motifs séquentiels : le prétraitement des données.
  - Nicolas Pasquier, Yves Bastide, Rafik Taouil, Lotfi Lakhal: Closed Set Based Discovery of Small Covers for Association Rules.

16. BDA 2000: Blois

- **Les BD pour le Web**

- **Serveur XML (e.g., Xylème)**
  - Laurent Mignet, Serge Abiteboul, Sébastien Ailleret, Bernd Amann, Amélie Marian, Mihai Preda: Acquiring XML pages for a WebHouse.
  - Amélie Marian, Serge Abiteboul, Laurent Mignet: Chance-centric Management of Versions in an XML Warehouse.
- **Extension des SGBDR (e.g., e-XML)**
  - Daniela Florescu, Donald Kossmann, Ioana Manolescu: Integrating Keyword Search into XML Query Processing.
  - Daniela Florescu, Valérie Issarny, Patrick Valduriez, Khaled Yagoub: Caching Strategies for Data-Intensive Web Sites.
- **La fouille de données : Règles associatives ++**
  - Stéphane Lopes, Jean-Marc Petit, Lotfi Lakhal: Discovering Agree Sets for Database Relation Analysis.
  - Florent Masseglia, Pascal Poncelet, Maguelonne Teisseire: Incremental Mining of Sequential Patterns in Large Databases.
- **Les SGBD pour carte à puce**
  - Christophe Bobineau, Luc Bouganim, Philippe Pucheral, Patrick Valduriez: PicoDBMS: Scaling Down Database Technique for the Smartcard

BDA 2000 : Les châteaux …

17. BDA 2001: Agadir, Maroc

Royal Palace, Rabat

BDA 2001: Agadir

- **XML/XQuery pour les BD (fédérées)**
  - Bernd Amann, Catriel Beeri, Irini Fundulaki, Michel Scholl, Anne-Marie Vercoustre: Rewriting and Evaluating Tree Queries with XPath.
  - Ioana Manolescu, Daniela Florescu, Donald Kossmann: Answering XML Queries over Heterogeneous Data Source
- **Le datacube à toutes les sauces**
  - Anne Laurent: Bases de données multidimensionnelles floues
  - Jean-François Boulicaut, Patrick Marcel, Christophe Rigotti: Query-Driven Knowledge Discovery via OLAP manipulations
- **Fouille de données : Web mining**
  - Florent Masseglia, Maguelonne Teisseire, Pascal Poncelet: Web Usage Mining Inter-Sites: Analyse du comportement des utilisateurs à impact immédiat

18. BDA 2002: Evry

- **XML actif et les Web Services**
  - Serge Abiteboul, Gregory Cobena, Julien Masanés, Gerald Sedrati: A First Experience in Archiving the French Web.
  - Serge Abiteboul, Omar Benjelloun, Ioana Manolescu, Tova Milo, Roger Weber: Active XML: A Data-Centric Perspective on Web Services.
  - Serge Abiteboul, Gregory Cobena, Benjamin Nguyen, Antonella Poggi: Construction and Maintenance of a Set Of Pages Of Interest (SPIN) using Active XML.
  - Mohand-Said Hacid, Alain Léger, Christophe Rey, Farouk Toumani: Dynamic discovery of e-services
  - Bernd Amann, Catriel Beeri, Irini Fundulaki, Michel Scholl: Querying XML Sources Using an Ontology-based Mediator.

- **La fouille de données et les applications**
  - Régis Saint-Paul, Guillaume Raschia, Noureddine Mouaddib: Database Summarization: Application to a Commercial Banking Data Set.
  - Marta Rukoz, Maude Manouvrier, Geneviève Jomier: Distances de similarité d'images basées sur les arbres quaternaires.
  - Eric Viara: Banques et bases de données en biologie moléculaire: de la donnée à la structure (conf. invitée).
  - Alex Aussem, Jean-Marc Petit: e-functional dependency inference: application to DNA microarray expression data.

19. BDA 2003: Lyon

- Manipulation de schemas
  - **Philip A. Bernstein: Generic Model Management: A Database Infrastructure for Schema Manipulation (conf. invitée)**
- La carte à puce rejoint la sécurité
  - **Nicolas Anciaux, Luc Bouganim, Philippe Pucheral: Memory Requirements for Query Execution in Highly Constrained Devices.**
- **Règles associatives, Séquences**
- **Systèmes répartis**

20. BDA 2004: Montpellier

- **XML : une valeur sûre**
  - Laurent Yeh, Georges Gardarin: Indexing XML Objects with Ordered Schema Trees.
  - Noureddine Chatti, Sylvie Calabretto, Jean-Marie Pinon: Vers un environnement de gestion de documents à structures multiples.
  - Luc Bouganim, François Dang Ngoc, Philippe Pucheral: Client-Based Access Control Management for XML documents.
  - Tuyet-Tram Dang-Ngoc, Georges Gardarin, Nicolas Travers: Tree Graph Views: On Efficient Evaluation of XQuery in an XML Mediator.
  - Ioana Manolescu, Andrei Arion, Angela Bonifati, Andrea Pugliese: Path Sequence-Based XML Query Processing.
  - Véronique Benzaken, Giuseppe Castagna, Cédric Miachon: CQL: a pattern-based query language for XML.

20-25. De 2005 à nos jours …

- **2005 : Saint-Malo**
  - XQuery, P2P, Confidentialité (BD Hypocratique)
- **2006 : Lille**
  - Services Web, Multidim., Clustering
- **2007 : Marseille**
  - Optimisation, Services Web, Web sémantique
- **2008 : Valence**
  - Ontologie, Fouille du cube, Top-k
- **2009 : Namur**
  - ???

Et demain ?

- **Les thèmes**
  - Polishing the ball ?
  - Very* large DB (e.g., Google)
  - BD mobiles spatio-temporelles
  - BD plus sûres (cartes à puces)
  - BD intégrées, Cloud computing
- **L'organisation**
  - Le mérite des grands projets
    - Socrate, Sirius, Sabre, Verso …
  - Théorie et pratique
    - Marier les deux avec harmonie