

# The Claremont Report on DB Research 2008

# Punto de Quiebre

- 1.Ubicuidad de “Big Data”
- 2.Analisis de Datos como centro de ganancia
- 3.Ubicuidad de datos estructurados y semi-estructurados
- 4.Expansion de demandas de desarrolladores
- 5.Cambios arquitecturales en la computacion

# Ubicuidad de datos

- Comunidades que trabajan con grandes cantidades de datos han crecido (Web, e-science, juegos, PLN, SNA, etc.)
- Usuarios de big data han crecido
- Desarrolladores en esta area marcaran esta area de la computacion

# Analisis de datos como negocio

- Barreras entre deptos. IT y Bussiness se estan eliminando
- Muchas empresas dond “data is the business”
- Captura, integracion y analisis de datos no son costos; son ganancias.
- Industria de analisis de datos se ha agigantado
- El data scientist: raw data + statistics

# Ubicuidad de datos estructurados y semi-estructurados

- Explosión de datos estructurados disponibles en la Web e intranets empresariales
- No viene de DB: extracción de info estructurada de texto, softw. Logs, sensores, crawls, etc.
- Tb. datos semi-estructurados: Blogs, comunidades, mensajería instantánea.
- Carrera por extraer info estructurada de texto

# Expansion de demandas de desarrolladores

- Adopcion de SABD por parte de programadores ha crecido notablemente
- Ha ayudado sistemas open source: MySQL, PostgreSQL, paquetes objeto-relacionales como Ruby on Rails
- Esto ha traído nuevas expectativas
- Oportunidades para nuevos modelos, nuevos componentes de sistemas para manipulacion y gestion de datos

# Cambios arquitecturales

- Macro: “Cloud computing” sugiere cambios fundamentales en la arquitectura de softw.
- Democratiza acceso a clusters paralelos
- Micro: de incremento de velocidad de procesamiento a numero de “cores” de procesadores y threads
- Tecnologias de almacenamiento cambian jerarquias de memoria
- Temas de energia

# Oportunidades

- 1.Reconsideracion de Motores de Bases de Datos
- 2.Programacion declarativa para plataformas emergentes (productividad programacion)
- 3.Relacion entre datos estructurados y semi-estructurados
- 4.Servicios de Cloud
- 5.Aplicaciones moviles y mundos virtuales
- 6.Temas de ambito del area, publicaciones, etc.



