

Statistics Jobs

Enrico Aiello 11/11/2013

La mia carriera accademica

- Università di Bologna: Laurea in scienze statistiche demografiche e sociali;
- Tesi in statistica per la ricerca sperimentale;
- Università di Bologna: Master in Biostatistica;
- Oracle University: Databases & Datawarehousing;

Il Consorzio www.cineca.it

❖ è un Consorzio Interuniversitario senza scopo di lucro formato da **70 Università italiane**, due **Enti di Ricerca Nazionali** (OGS e CNR) e il **Ministero dell'Istruzione dell'Università e della Ricerca**.

❖ è il più importante centro di supercalcolo in Italia e uno dei più importanti a livello mondiale, al **13° posto** nella classifica mondiale **Top500** (<http://www.top500.org>) con il super calcolatore FERMI (IBM BG/Q) installato al Cineca nel giugno 2012

❖ 3 sedi operative: Milano – Bologna - Roma



Principali attività del CINECA

- ✓ Supercalcolo (HPC) – Supporto alla ricerca scientifica
- ✓ Visualizzazione scientifica e ambienti virtuali interattivi
- ✓ Trasferimento tecnologico
- ✓ Sistemi Informativi per:
 - Università
 - MIUR
 - Sanità
 - altre PA e imprese
- ✓ Editoria elettronica e biblioteche digitali, Knowledge management, portali, multimedia, e-learning
- ✓ Partecipazione a progetti Unione Europea (*Human Brain Project, PRACE, MONT-BLANC, V-MusT.net, VOA3R, ...*)
- ✓ Servizi di Data Center

Il CINECA e la Sanità: aree di applicazione

Servizi ICT per la ricerca clinica

- ✓ Clinical trial e registri epidemiologici
 - Registri di monitoraggio patologia/farmaco
 - Clinical trial (profit, no profit)
- ✓ Workflow management systems a supporto dei processi documentali
- ✓ Reti di servizi
 - Sistemi di sorveglianza e allerta epidemiologica
 - Network scientifici
 - Servizi per le CRO
 - E-learning accreditato ECM

Sorveglianza epidemiologica ed economica nel SSN

- ✓ Sistemi di supporto alle decisioni per il governo clinico (Osservatorio ARNO)
- ✓ Data Warehouse socio-sanitari
- ✓ Studi epidemiologici
- ✓ Osservatori territoriali GIS

Micronet

Il Progetto

Il progetto nasce da una sperimentazione congiunta fra CINECA e Istituto Superiore di Sanità.

Micronet è un sistema di sorveglianza delle infezioni e malattie da agenti microbici basata sulla rilevazione e trasmissione automatica e tempestiva dei risultati di accertamento etiologico infettivo e della loro resistenza agli antibiotici, tramite ben definiti eventi sentinella.

La sperimentazione ha portato a realizzare una **piattaforma automatica integrata** tra sistemi informativi locali ed eterogenei e un sistema centralizzato web-based di acquisizione e analisi dei dati.

L'integrazione dei dati provenienti dai diversi laboratori è resa possibile dall'uso di thesauri centralizzati, realizzati espressamente dall'Istituto Superiore di Sanità e dal network di microbiologi partecipanti alla sperimentazione.

Micronet

Obiettivi

- Sorveglianza e controllo delle malattie infettive, con particolare enfasi sull'antibiotico resistenza
- Sistema di allerta in tempo reale
- Supporto nella sorveglianza attiva (es.: infezioni nosocomiali, gruppi di rischio, prevenzione epidemie, ...)
- Allerta precoce nuovi patogeni
- Integrazione con gli altri flussi informativi ospedalieri relativi al paziente
- Base informativa utilizzabile a fini di ricerca

Micronet

Per Approfondire

- <https://micronet.cineca.it>
- D'Ancona, F., Rizzo, C., Alfonsi, V., & Ciofidegli Atti, M. L. (2007). P1815 Micronet: an Italian automatised laboratory based surveillance and early warning system for infectious diseases. *International Journal of Antimicrobial Agents*, 29, S518.
- Aiello, Enrico, et al. "Micronet. Rete informatica per la raccolta multicentrica di dati epidemiologici da laboratori di microbiologia." *Notizie dal cineca* (2008).
- Sisto, A., et al., A. (2012). Carbapenem non-susceptible *Klebsiella pneumoniae* from Micronet network hospitals, Italy, 2009 to 2012. *Euro surveillance: bulletin européen sur les maladies transmissibles - European communicable disease bulletin*, 17(33).

Cruscotto direzionale delle prestazioni sanitarie

Il Progetto

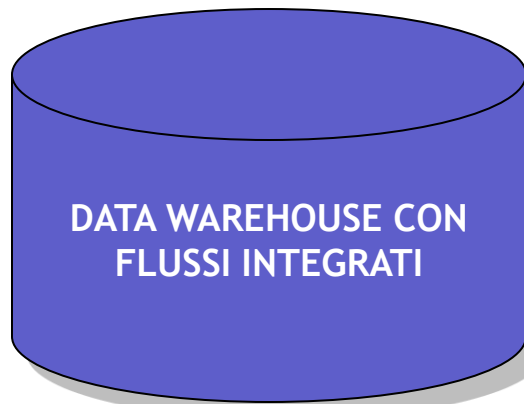
- Strumento che permette l'utilizzo di dati integrati attraverso percorsi di navigazione con visione sia economica che farmacoepidemiologica al fine di contenere la spesa senza diminuire l'efficacia della cura;

Cruscotto direzionale delle prestazioni sanitarie

VISIONI DI ANALISI


Epidemiologica
/clinica


Economica

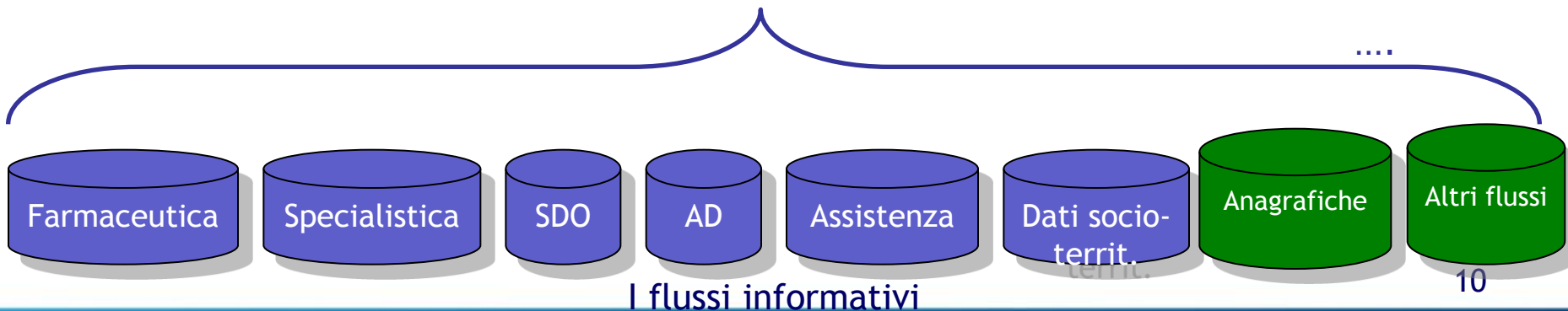


PROFILI DI ACCESSO AI DATI


Regione


ASL


MMG, PLS



Cruscotto direzionale delle prestazioni sanitarie

Obiettivi

- **Monitoraggio della spesa e delle prestazioni sanitarie** (prescrizioni farmaceutiche, diagnosi ospedaliere, prestazioni specialistiche)
- **Proiezione della spesa**
- **Monitoraggio degli indicatori di appropriatezza e di spesa per il raggiungimento degli obiettivi regionali/aziendali**
- **Benchmark tra aree territoriali**
- **Monitoraggio degli indicatori epidemiologici**

Cosa sono i Big Data?

- Ogni giorno creiamo 2,5 quintilioni di byte di dati e il 90% dei dati è stato creato solo negli ultimi due anni.
- Questi dati vengono registrati ovunque: sensori per la raccolta di informazioni sul clima, post su siti di social media, video e immagini digitali, record delle transazioni di acquisto e segnali GPS dei cellulari, solo per fare qualche esempio. Questi tipi di dati vengono definiti **big data**.
- I big data sono più di una semplice questione di dimensioni; rappresentano l'opportunità di estrapolare informazioni da tipi di dati e contenuti nuovi ed emergenti, consentendo di rendere il proprio business più agile e di rispondere a domande che prima erano considerate fuori portata. Finora, non esisteva un modo pratico per cogliere questa opportunità.

Cosa sono i Big Data?

I big data si sviluppano su tre dimensioni:

- **Volume:** le aziende sono sommerse da una quantità in costante crescita di dati di tutti i tipi e immagazzinano facilmente terabyte, in alcuni casi petabyte, di informazioni.
- **Velocità:** a volte 2 minuti sono troppi. Nei processi in cui il tempo è un fattore importante come, ad esempio nell'identificazione delle frodi, è necessario utilizzare i big data che fluiscono in azienda per massimizzarne il valore.
- **Varietà:** i big data sono dati di qualsiasi tipo - dati strutturati e non strutturati come, ad esempio, dati di testo, dati dei sensori, dati audio, dati video, flussi di clic, file di log e altri. Quando si analizzano questi tipi di dati insieme, emergono nuove informazioni di valore.

Data Scientist

“part analyst, part artist (...) a data scientist is somebody who is inquisitive, who can stare at data and spot trends. It's almost like a Renaissance individual who really wants to learn and bring change to an organization.”

[Anjul Bhambhri, vice presidente dei prodotti big data IBM]

Data Scientist

Background

- Informatica e applicazioni;
- Modellazione;
- Statistica;
- Matematica;

R

The R language has become a de facto standard among statisticians for developing statistical software and is widely used for statistical software development and data

analysis. [McKinsey Global Institute - Big data: The next frontier for innovation, competition, and productivity]

Contatti

- e.aiello@ceneca.it
- <http://www.cineca.it/it/content/invia-curriculum>