

CORSO DI BASI DI DATI E CONOSCENZA

BASI DI DATI E SISTEMI WEB-BASED
ELEMENTI DI DATA MANAGEMENT
GESTIONE DEI DATI E DELLA CONOSCENZA 1

Roberto Basili
a.a. 2021/22

Obiettivi Formativi

- **Scenario**

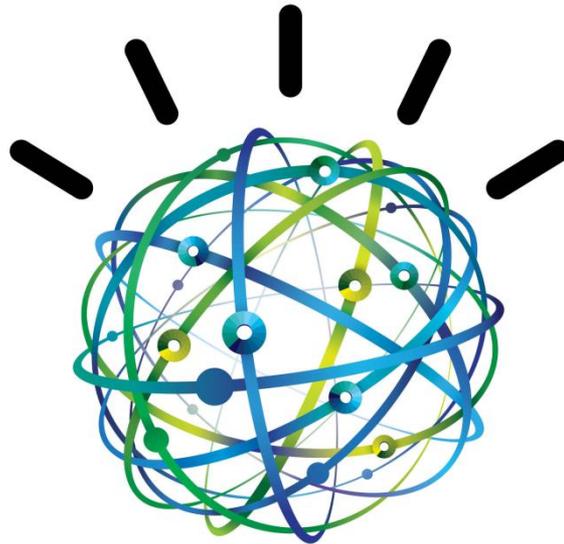
- Le grandi quantità di dati accumulate nelle reti di calcolatori costituiscono una risorsa enorme con requisiti critici rispetto alla efficienza dei metodi di **accesso**, **localizzazione** della informazione, **persistenza** e **riutilizzo** in applicazioni eterogenee
- Le applicazioni sono sempre più esigenti dal punto di vista della indipendenza tra logiche dei dati e loro integrazione nei processi di business
 - Intelligenza Artificiale
 - Big Data
 - Data Analytics

Obiettivi Formativi

- Approfondire i problemi della **modellazione concettuale** delle informazioni ed acquisire una significativa capacità di progettazione logica dei dati in applicazioni distribuite e complesse
- Conoscere le principali **metodologie e tecnologie di gestione e progettazione delle basi di dati** (Data Base Management).
- Applicare tali tecnologie nelle **applicazione delle BbD** relazionali in scenari tradizionali e Web.
- Introdurre la **relazione tra le Basi di Dati e la rappresentazione della conoscenza**.

Dati e Applicazioni software

- IBM's Watson

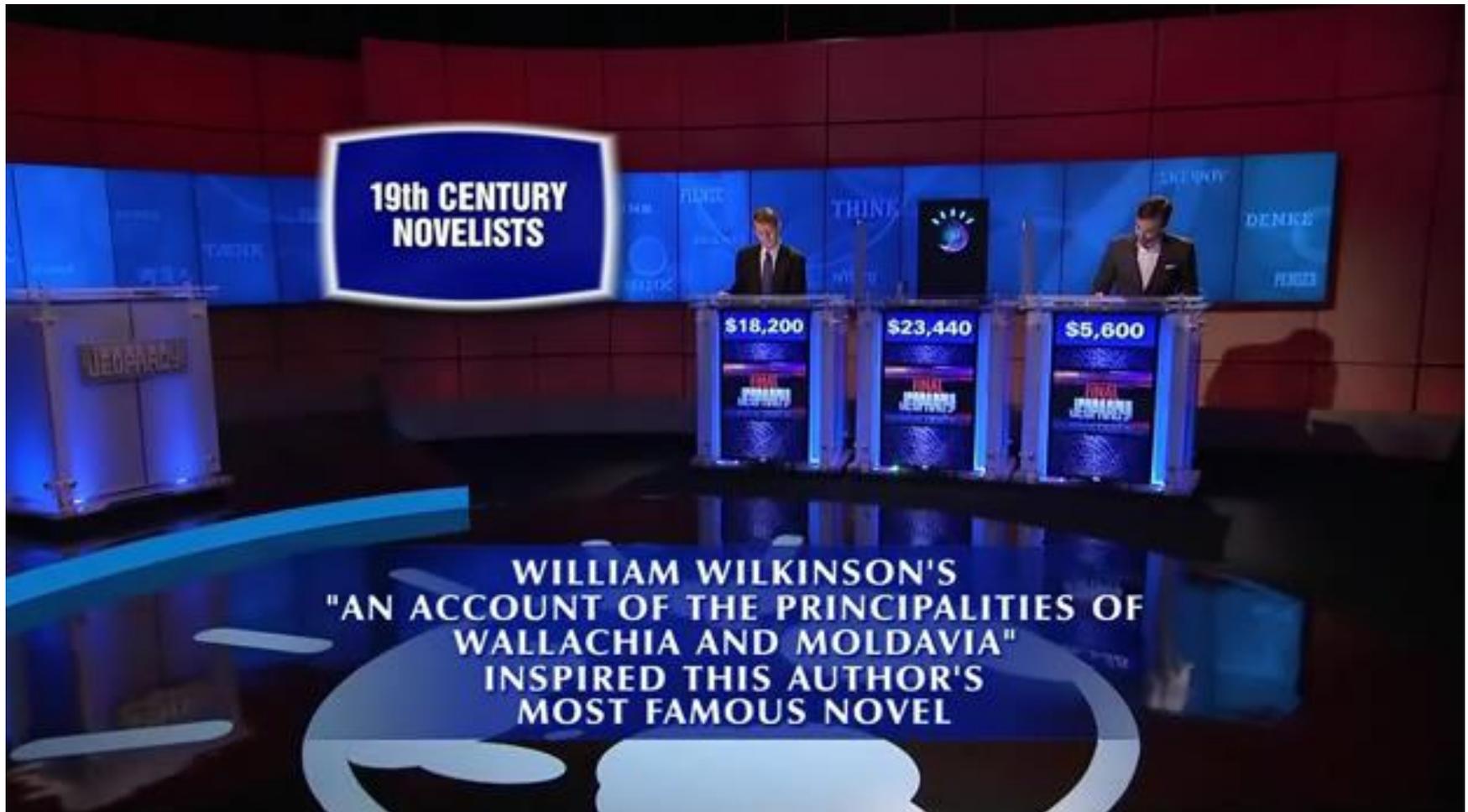


- http://www-03.ibm.com/innovation/us/watson/science-behind_watson.shtml

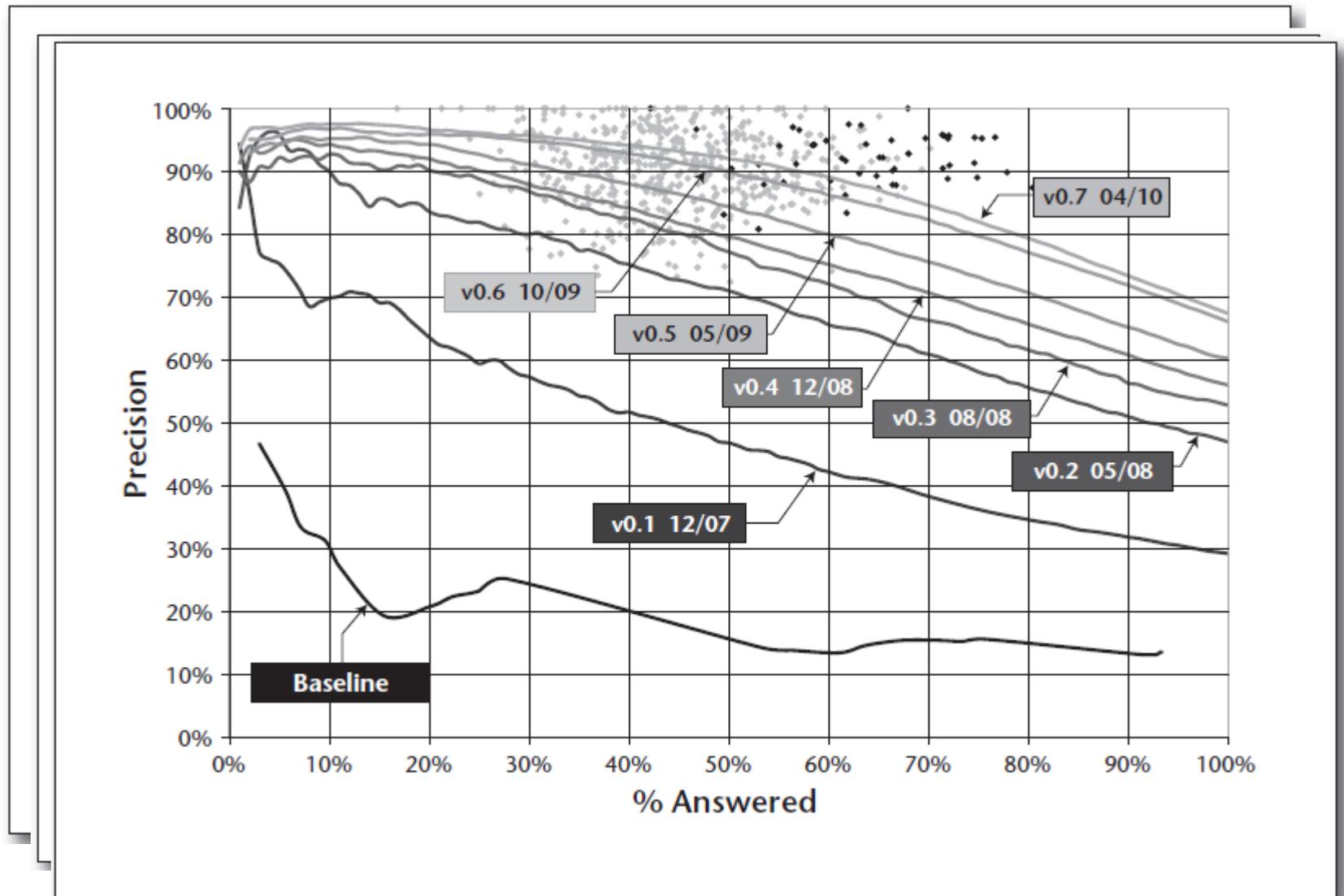
Domande



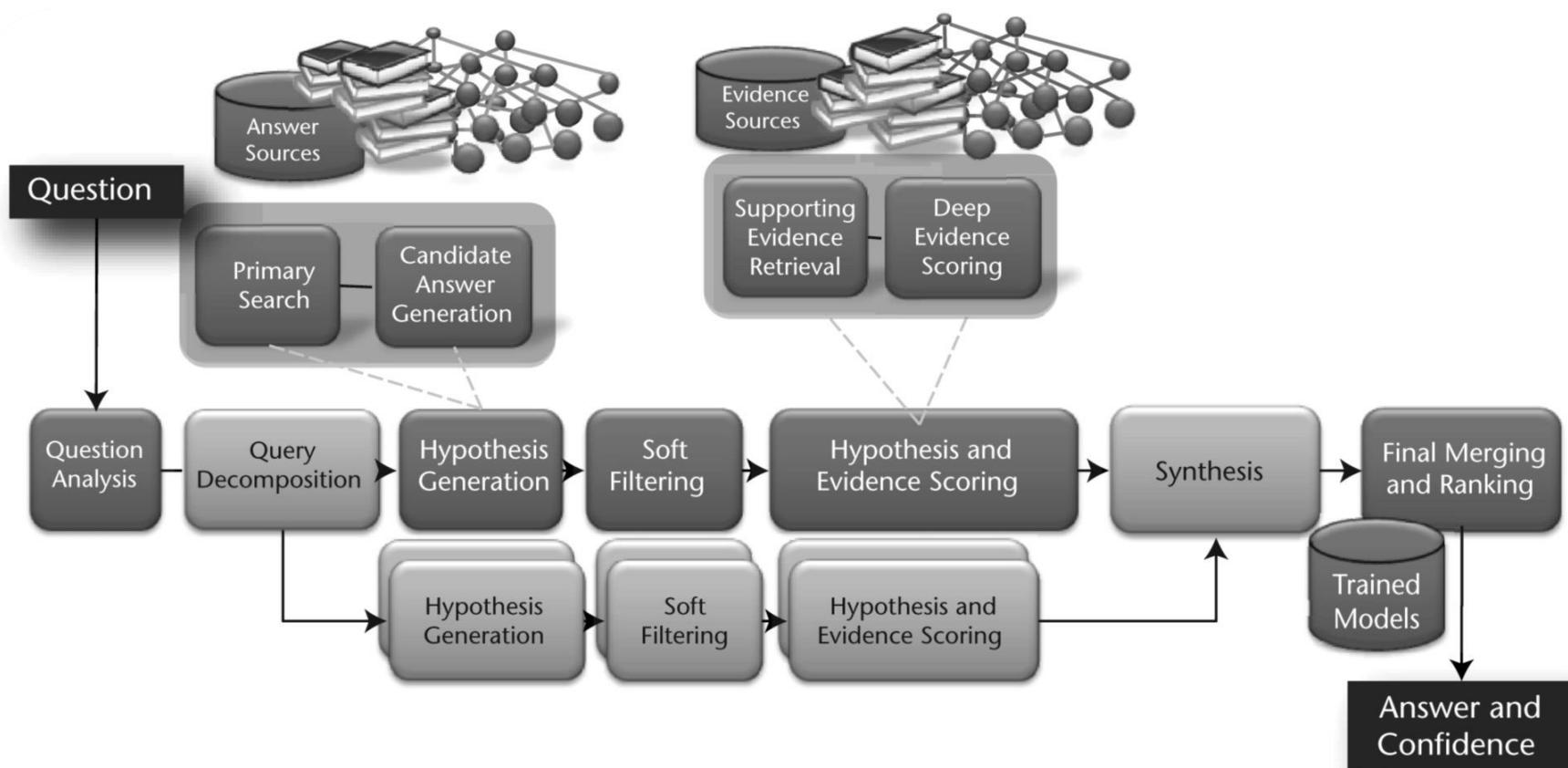
I successi di Watson



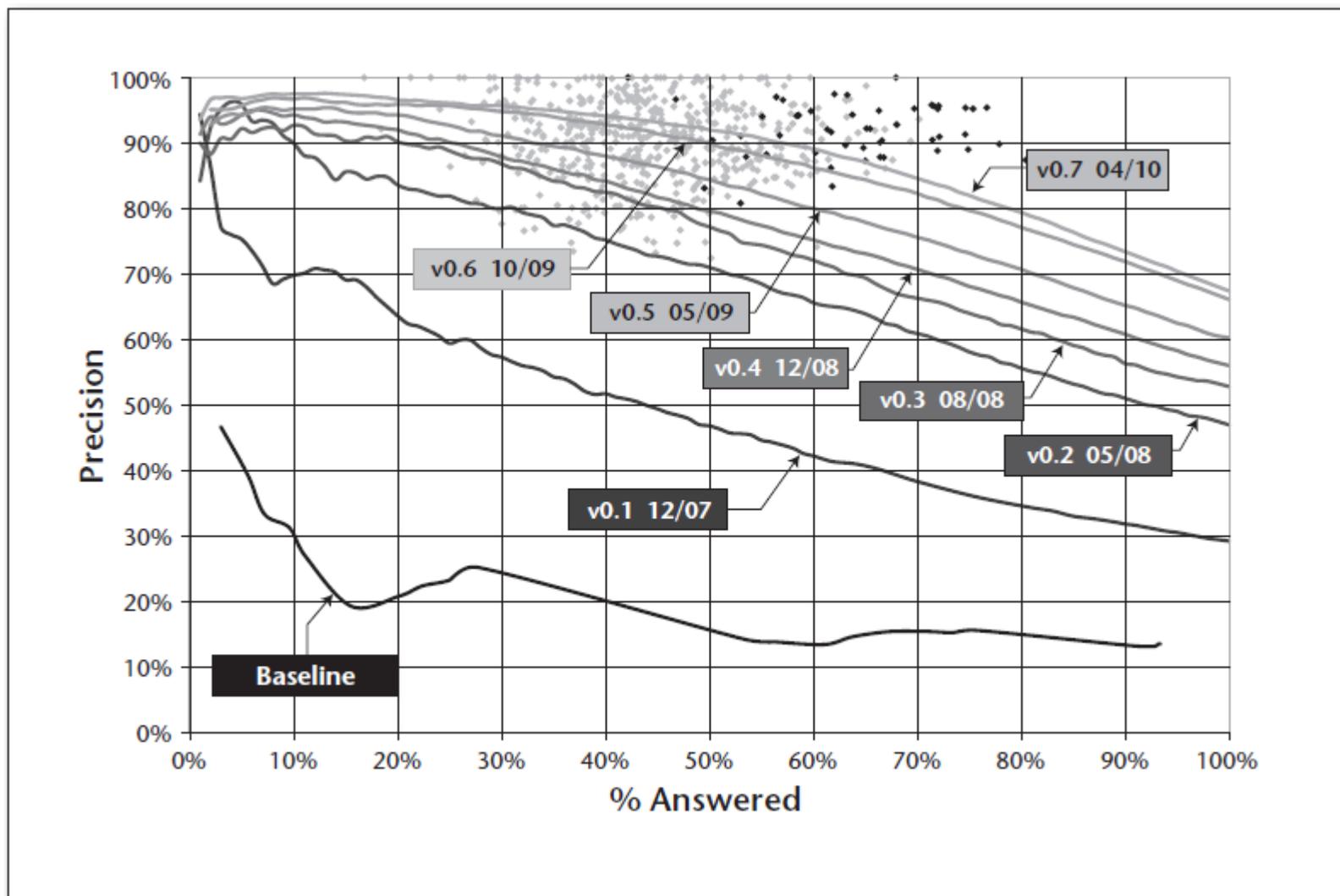
Watson ed il software “intelligente”



Watson: a DeepQA architecture



Pronti per Jeopardy!



Un altro task complesso risolto in modalità *data-driven: image captioning*

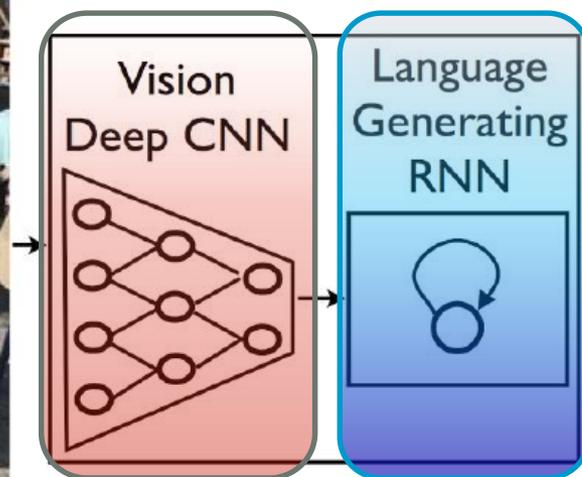
- Una **caption** (o **didascalia**) è una breve descrizione in linguaggio naturale di una immagine che descriva i suoi **contenuti**, in modo sintatticamente e semanticamente appropriato:
 - Oggetti e relazioni tra loro
 - Azioni
 - Proprietà degli oggetti



"A chef preparing food inside of a kitchen near a window."

Automatic image captioning: il ruolo delle reti neurali

- Il processo di *captioning* è decomposto in passi successivi, in cui, tramite i dati, **due reti neurali** vengono addestrate a gestire indipendentemente (1) il **riconoscimento visivo** e (2) la **generazione di (uno o più) testo(i)**



A group of people shopping at an outdoor market.

There are many vegetables at the fruit stand.

... una rete neurale addestrata per il *captioning* in Italiano



*Uno scuolabus giallo
parcheeggiato sul lato
della strada.*

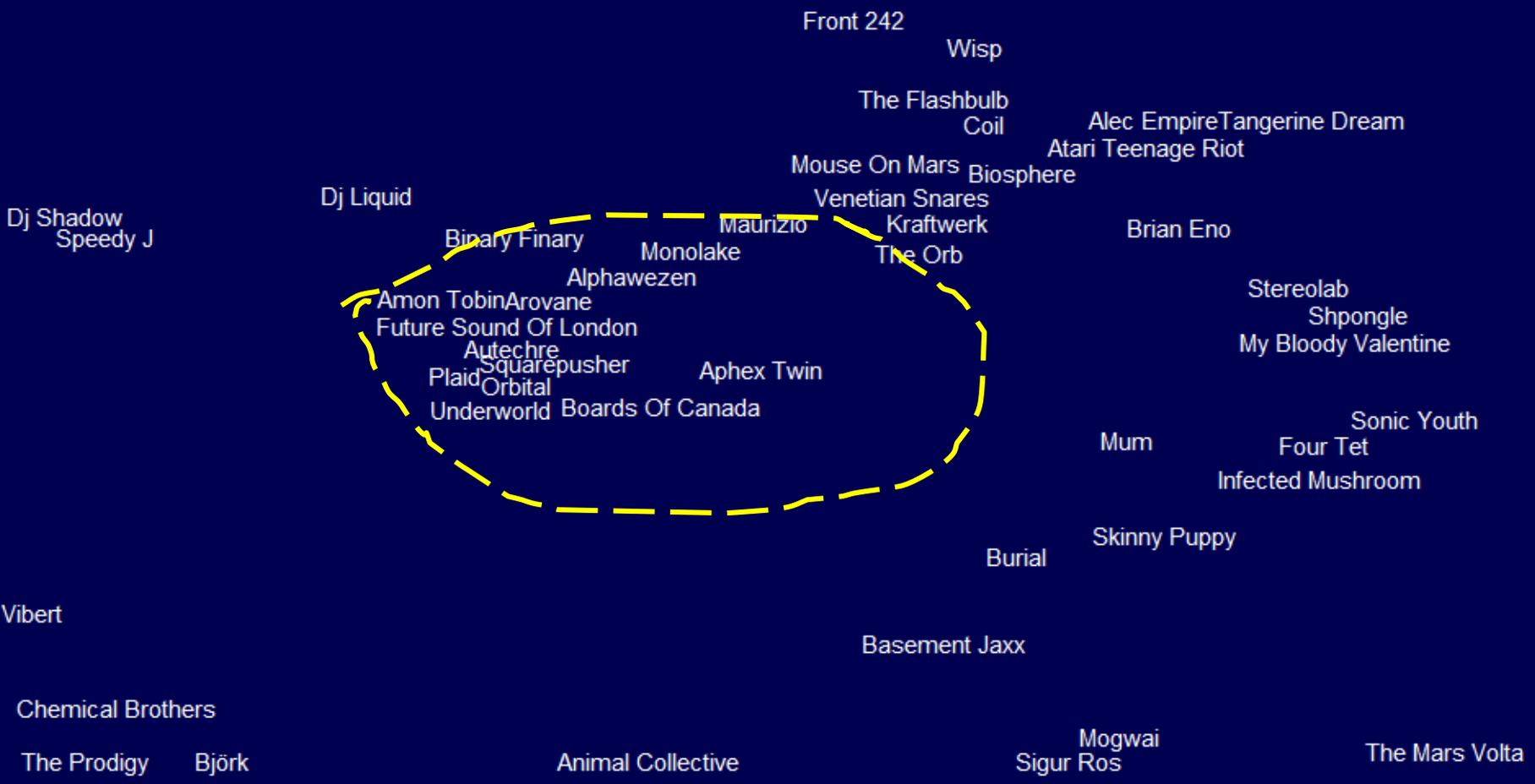


*Un uomo che
cavalca un cavallo
su una strada
cittadina.*

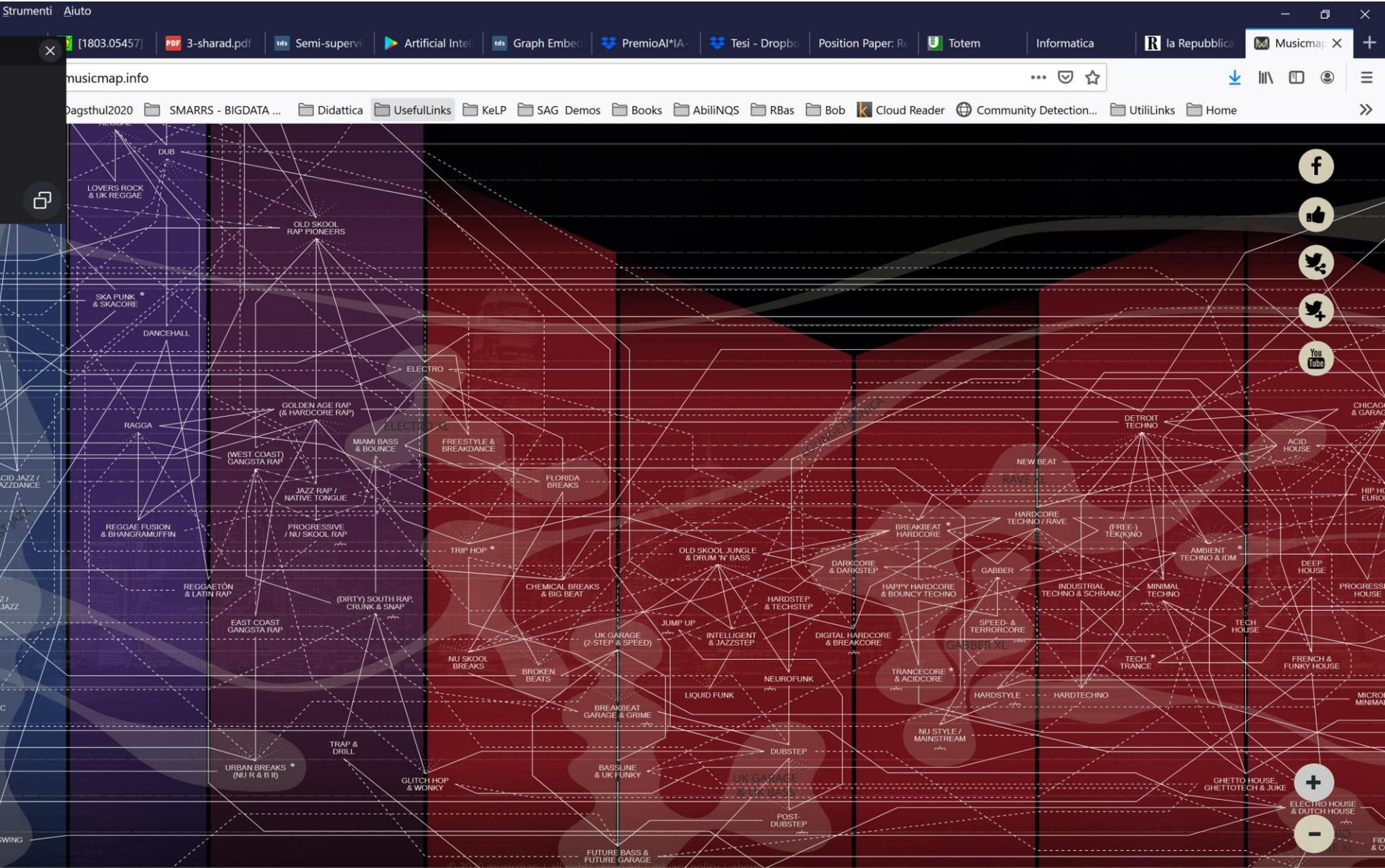


*Un segnale di stop
che si siede su un
angolo di strada.*

map



Music maps: 2020



Es. di Applicazioni fondate sui dati

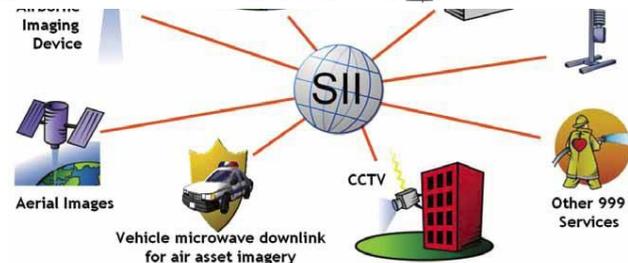
- Motori di Ricerca
- Sistemi di predizione della affidabilità economico finanziaria dei clienti
- Sistemi di recommending
 - Musica
 - Servizi avanzati di e-commerce (e.g. Amazon)
 - Promozioni merci e servizi (ad es. Google Ads)
 - Servizi turistici (e.g. Trip Advisor)
- Brand Reputation Management
 - Monitoraggio della comunità dei clienti
 - Gestione delle crisi reputazionali
 - Personalizzazione dell'Offerta

Livelli e Indipendenza



- Livello Logico

- Livello fisico



- SFIDE:
 - Content-based Logics
 - Interoperabilità semantica
 - Natural Language Processing



Web, dati e informazioni: le Lingue

I contenuti del Web, caratterizzati da ricche caratteristiche multimediali, sono per lo più **opachi dal punto di vista semantico**



The screenshot shows a Mozilla Firefox browser window displaying the website www.takungpao.com.hk. The browser's address bar shows the URL, and the page title is "www.takungpao.com.hk". The website header features the Takungpao logo and a banner for the "2011 中国証券金紫荆獎 Golden Bauhinia Awards". Below the header, there is a search bar and navigation links. The main content area displays a news article titled "胡總語特首:防範經濟金融風險" (Hu Jintao's special message to the Chief Executive: Preventing economic and financial risks). The article includes a photograph of two men shaking hands and text discussing the APEC summit. The browser's status bar at the bottom shows the URL: http://www.takungpao.com.hk/news/11/11/13/2011_apec_xgbd-1423309.htm.

Il ruolo del testo per le Informazioni nel Web

Hu meets KMT honorary chairman in Hawaii - People's Daily Online - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Yahoo! Strumenti Aiuto

Hu meets KMT honorary chairman

Indietro Avanti Down

Chinese President Hu Jintao (R) shakes hands with Honorary Chairman of the Chinese Kuomintang (KMT) Lien Chan, in Honolulu, Hawaii, the U.S., Nov. 11, 2011.
(Xinhua/Huang Jingwen)

HONOLULU, United States, Nov. 11 (Xinhua) -- Hu Jintao, general secretary of the Central

Latest News: • Indonesia to host European Higher Education Fair

Beijing Sunny 15 / 1 City Forecast

Home >> China Politics

Hu meets KMT honorary chairman in Hawaii

(Xinhua)

11:10, November 12, 2011



Chinese President Hu Jintao (R) shakes hands with Honorary Chairman of the Chinese Kuomintang (KMT) Lien Chan, in Honolulu, Hawaii, the U.S., Nov. 11, 2011.

Selections for you



Miao ethnic group celebrates Miao's New Year in SW China



World's first Angry Birds exclusive shop opens in Helsinki

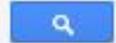


Most Popular

- 1 Hu reaffirms support to Hong Kong's sta...
- 2 Hu meets KMT honorary chairman in Hawaii
- 3 China in APEC: a mutually beneficial en...
- 4 Night life in Shanghai
- 5 China's 2011 foreign trade to grow 20 p...
- 6 Beijing house prices stumble 5.1 pct as...
- 7 Lama students start school in Tibet Col...
- 8 Police in central China crack phoney ca...



Hu Jintao



Ricerca

Circa 725.000 risultati (0,09 secondi)

- Tutto
- Immagini
- Mappe
- Video
- Notizie
- Shopping
- Più conte
- Tutti i ri
- Per argomento



- Qualsiasi dimensione
- Grandi
- Medie
- Icone
- Maggiori di...
- Dimensioni esatte...



- Qualsiasi colore
 - A colori
 - Bianco e nero
-



- Qualsiasi tipo
- Volti
- Foto
- Clip art
- Disegni



- Visual standard
- Mostra dimensioni

Una riflessione

- L'intermediazione che chiediamo alle tecnologie dell'informazione sembra richiedere:
 - **Capacità di calcolo** rispetto a volumi crescenti di dati
 - **Sensibilità verso i significati dei contenuti**, indipendenti dal formato dei dati
 - **Integrazione** di fonti diverse ed eterogenee
 - Aggregazione dei dati in processi complessi **la cui logica non dipende da**
 - eterogeneità di formati di *storage*,
 - linguaggi di comunicazione
 - protocolli di comunicazione

Scopi del Corso :

- Introduzione alla ***modellazione logico-concettuale dei dati*** nelle applicazioni del software
- Introduzione ai concetti di base dei ***Sistemi di Gestione delle Basi di Dati (DBMS)***.
 - Uso dei **DBMS relazionali** per la progettazione delle Basi di Dati
 - **Modelli dei dati** a livello logico e fisico
- Introduzione alle architetture software per lo sfruttamento dei DBMS nelle moderne applicazioni distribuite, e.g. **Multi-Tier Architectures**
- Introdurre le tendenze e le prassi avanzati nella **gestione dei dati nelle reti di comunicazione**

Requisiti

- Il corso e' diretto agli studenti del Corso di
 - Laurea in Ingegneria Gestionale (2 emiseimestri 12 CFU)
 - Laurea Ing. Internet – Telecomunicazioni (1 emiseimestre 6 CFU)
 - Alcuni hanno chiesto la erogazione di 12 CFU
 - Laurea di Scienze dei Media (2 emiseimestri 8 CFU)
- Prerequisiti:
 - Elementi di **algebra e logica**
 - Elementi di **algoritmica e programmazione** C++, PHP, Java o Python
 - Elementi di base dei **sistemi operativi**

Organizzazione

- Una particolare attenzione verrà dedicata ad aspetti tecnologici legati alla adozione delle tecnologie agenti sui dati:
 - Piattaforme di RDBMS software esistenti (es. MySQL o Postgres)
 - Integrazione tra *applicazioni sw e DBMS* (uso di linguaggi per la progettazione di applicazioni Web che usano le basi di dati)
 - Tecnologie *noSQL*
 - Tecnologie di *Visual Analytics*: dai dati all'infografica
- Esercitazioni dedicate
 - MySQL, PHP, Mongo DB

Organizzazione del Programma



- Corsi di Basi di Dati (Ing. Gestionale, Internet, Sc. Media):
 - Esame scritto alla fine del primo ciclo di lezioni (fine primo emisemestre)
- Progetto per gli studenti di Ing. Gestionale (secondo emisemestre)
- Basi di Dati e Sistemi Web-based
 - Esame scritto alla fine del primo ciclo di lezioni
 - Esame finale scritto (secondo emisemestre)
 - Presentazione Progetto

Orario del Corso

- **MAR 16:00-17:45, Aula 6 (MS Teams)**
- **MER 16:00-17:45, Aula 6**
- **VEN 11:30-13:30 , Aula C1**

- **Ricevimento: Mercoledì, dopo la lezione**
- **Sito Web:**
 - sag.art.uniroma2.it/didattica/basili/DB_21_22/
 - Avvisi, Slides (statiche), Bibliografia, esercizi proposti
 - Risorse
- **MS Teams: «BASILI-8039764-BASI_DI_DATI_E_CONOSCENZA»**

Pagina Web

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

repubblica.it/index.html?refres... Registration | Ricerca Internaz... Novità Compiti del Direttivo - Google Basi di Dati e Conoscenza (a.a. 201... Basi di Dati e Conoscenza (a.a. 201...

sag.art.uniroma2.it/didattica/basili/DB_19_20/ 90%

Clic2017 AIA2017 specIssue - G... Più visitati Ristoranti SMARRS - BIGDATA ... EPA Session Monday Didattica UsefulLinks KeLP SAG Demos Books AbiliNQS RBas Bob

Basi di Dati e Conoscenza (a.a. 2019/20)

Introduzione alle Basi di Dati ed alle Tecnologie dei Dati nel Web

Elenco dei File nel deposito

Sommario Contenuti

1. [Novità](#)
2. [Programma del Corso](#)
3. [Testi di Riferimento](#)
4. [Link Utili](#)
5. [Diapositive delle lezioni](#)
6. [Progetti ed Esercizi Proposti](#)

Novità

- *****Indirizzo WEB***** L'indirizzo della pagina Web del corso e' il seguente: http://sag.art.uniroma2.it/didattica/basili/DB_19_20/

TIMETABLE: Orario delle lezioni a.a. 2019-20:

Le lezioni del Corso si terranno secondo orario su tre giorni a partire da Martedì 24 Settembre 2019 :

- **MAR 14:00-15:45 aula 6, Edificio Aule Nuove, Macroarea Ingegneria.**
- **MAR 16:00-17:45 aula 6, Edificio Aule Nuove, Macroarea Ingegneria.**
- **MER 16:00-17:45 aula 6, Edificio Aule Nuove, Macroarea Ingegneria.**
- **VEN 11:30-13:30 aula 6 , Edificio Aule Nuove, Macroarea Ingegneria .**
- **Slides of the first lessons:** Sono disponibili nella sezione [Diapositive delle lezioni](#) i lucidi delle prime lezioni.
- **Registration:** E' necessario iscriversi al Corso (azione necessaria per prenotarsi a tutti gli appelli del Corso di Basi di Dati e Conoscenza)

Testi Consigliati

- Testo di Riferimento
 - Sistemi di Basi di Dati, di Raghu Ramakrishnan e Johannes Gehrke, Edizione Italiana, McGraw Hill, 2004
 - *BASI DI DATI*, Paolo Atzeni, Stefano Ceri, Piero Fraternali, Stefano Paraboschi e Riccardo Torlone, Edizione 5, 2018, pp. 832. ISBN: 9788838694455
- Materiali distribuiti dal docente, es.
 - Introduzione a MySQL, PL/SQL, Uso delle librerie di JDBC,
 - Introduzione al Datawarehousing
 - Introduzione alle tecnologie NoSQL
 - Introduzione al Machine Learning
- Slides delle lezioni (sito degli autori):
 - <http://www.cs.wisc.edu/~dbbook/>

Modalità d'Esame

- 1° Esonero a metà del primo emicorso (ca. 5-10 Nov 2019)
 - Test a Risposte Chiuse (TRC)
 - Emi1: Prova di Progettazione di un Db (PDb)
- 2° Esonero (entro 20 Dicembre 2019)
 - Test a Risposte Chiuse (TRC)
 - Emi1: Prova di Progettazione di un Db (PDb)
 - *Emi2: Domande Aperte (DA)*
- Prova Progettuale (studenti di Ing. Gestionale, 12 CFU)
- Prove Finali (e Recupero Settembre)
 - Emi1: TRC + PDb
 - Emi2: TRC+PDb+DA

Modalità Esame

