

Siete pronti per un pianeta più intelligente?

Pietro LEO

Il sistema WATSON e la nuova frontiera dei sistemi intelligenti in grado di sfidare l'uomo

Executive Architect
Member of IBM Academy of Technology

Email: pietro_leo@it.ibm.com

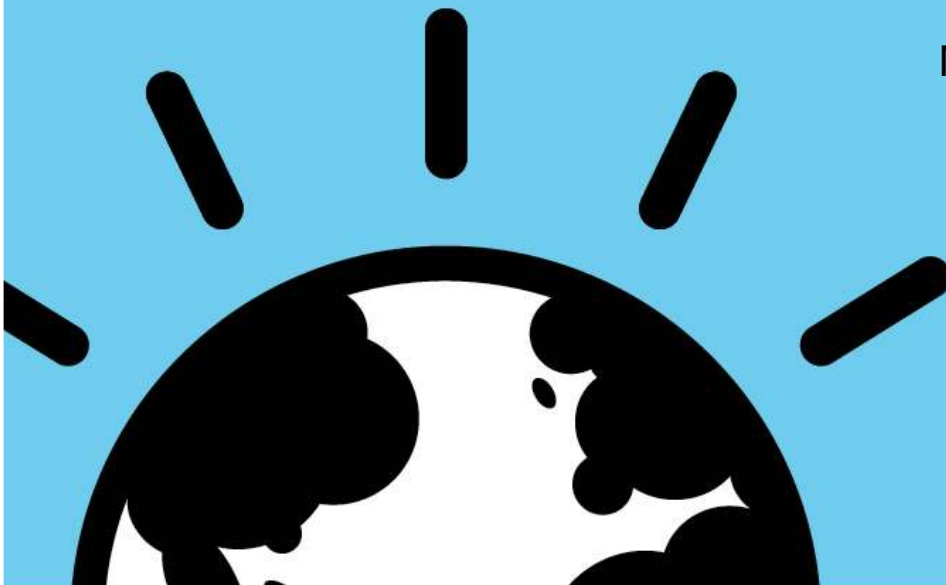
Twitter: [@pieroleo](https://twitter.com/pieroleo)

Linkedin: <http://it.linkedin.com/in/pieroleo>



FORUM PA 2011

Roma, 9 - 12 maggio



Siete pronti per un pianeta più intelligente?



interface with information
Information-based Intelligence
artificial intelligence Deep Q&A Jeopardy!
natural language processing THINK
human vs computer
grand challenge
evidence

Siete pronti per un pianeta più intelligente?



A Grand Challenge



Watson

Vuoi giocare o parlare?

• Scacchi

- Uno spazio di ricerca matematicamente ben definito
- Limitato numero di mosse e stati del gioco disponibili
- Tutti i simboli, i pezzi, le mosse possibili, e il loro significato e ruolo sono definiti nelle regole del gioco descrivibili matematicamente



• Il linguaggio umano

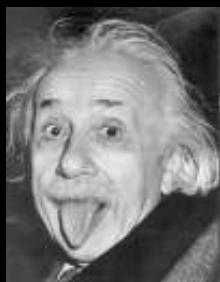
- Ciascuna parola presa isolatamente ha un significato (quando disponibile) limitato
- La Comprensione del linguaggio umano è un task profondamente cognitivo e specifico della nostra specie
- Combinando parole in frasi generiamo infiniti spazi di significato
- I computer non sono in grado di comprendere le parole e l'uso delle stesse e derivarne significati alla stessa stregua di quanto è in grado di fare l'uomo



Domande difficili?

Person	Birth Place
A. Einstein	ULM

Strutturato



• Dove è nato X?

– One day, from among his city views of Ulm, Otto chose a water color to send to Albert Einstein as a remembrance of Einstein 's birthplace.

Non Strutturato

Person	Organization
J. Welch	GE



• X dirige?

– If leadership is an art then surely Jack Welch has proved himself a master painter during his tenure at GE.

Siete pronti per un pianeta più intelligente?



Watson e e La sfida a Jeopardy!



Come si gioca a ?



- 3 concorrenti che si sfidano a rispondere a delle domande in una partita basata su 3 turni
- Uno dei 3 concorrenti sceglie una “categoria di domande” tra 6 possibili e non note a priori e sceglie il livello di difficoltà (tra 5) delle domande
- Il conduttore formula la domanda e risponde chi si prenota per primo: vinci il denaro associato al livello di difficoltà della domanda se c’azzechi o perdi il denaro se sbagli la risposta
- Vince il concorrente che alla fine delle 3 puntate ha accumulato più soldi

Come si gioca a ?

**Vasti domini di
conoscenza**

**Linguaggio
complesso**

Popular culture	\$600 Twins watching this young man's video "baby" on youtube helped it to a record 260 million views as of summer 2010	\$400 One type of this aquatic animal gives milk that's 65% fat; pups are weaned in 4 days, the least of any mammal	Mammals
	\$200 New Orleans Saints kicker Tom Dempsey's record 63-yard field goal was kicked in this school's stadium	\$800 ...this poet, whose "idylls of the king" started an Arthurian craze that inspired "a Connecticut Yankee"	
National Football League			Mark Twain Shall Meet

Precisione



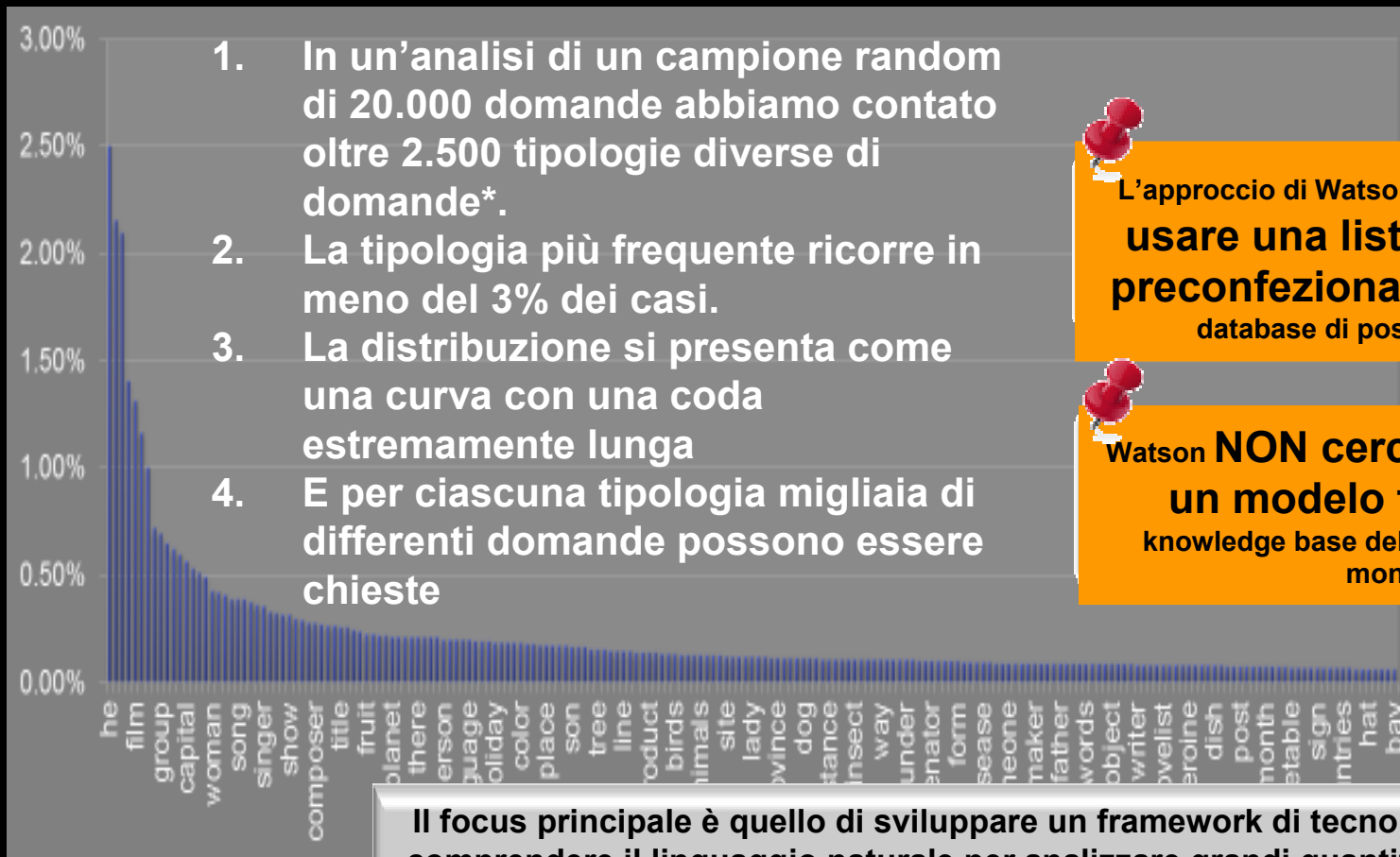
Velocità



Strategia



Perchè Jeopardy! è sfidante?



L'approccio di Watson è quello di **NON usare una lista di risposte preconfezionate** attingendo da un database di possibili risposte

Watson **NON cerca di costruire un modello formale**, una knowledge base delle conoscenze del mondo

Il focus principale è quello di sviluppare un framework di tecnologie in grado di comprendere il linguaggio naturale per analizzare grandi quantità di dati testuali **così come sono**

Sorgenti strutturate (Database e Knowledge Base) forniscono ulteriore background knowledge per interpretare il testo

*13% sono tipologie non distintive

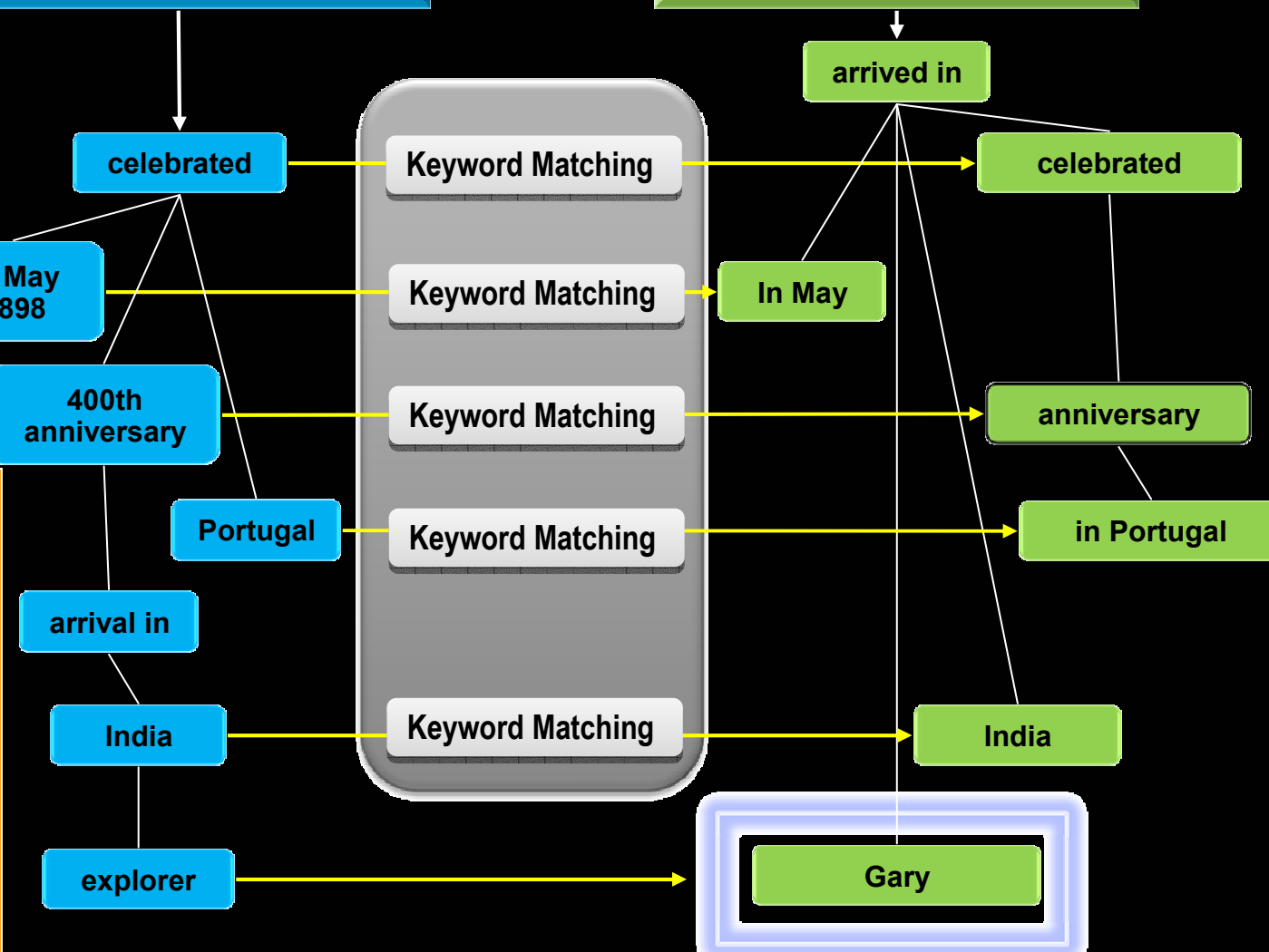
Domanda

In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

Evidenza

In May, Gary arrived in India after he celebrated his anniversary in Portugal.

L'evidenza suggerisce che "Gary" potrebbe essere la risposta MA il sistema deve comprendere che la semplice "corrispondenza" delle parole conta poco rispetto ad altri tipi di verifiche, prove ed altre evidenze



Domanda

In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

celebrated

Portugal

May 1898

400th anniversary

arrival in

India

explorer

Differenti tipi di evidenze portano a credere che la risposta sia "Vasco da Gama"

L'ipotesi fatta non è comunque certa al 100%!

- Ricerca ampia e approfondita
- Esplorazione di molte ipotesi
- Diversi algoritmi di inferenza
- Valutazione delle evidenze

Temporal Reasoning

Statistical Paraphrasing

GeoSpatial Reasoning

Data math

Paraphrase

GeoKB

Evidenze!

On the 27th of May 1498, Vasco da Gama landed in Kappad Beach.

landed in

27th May 1498

Kappad Beach

Vasco da Gama

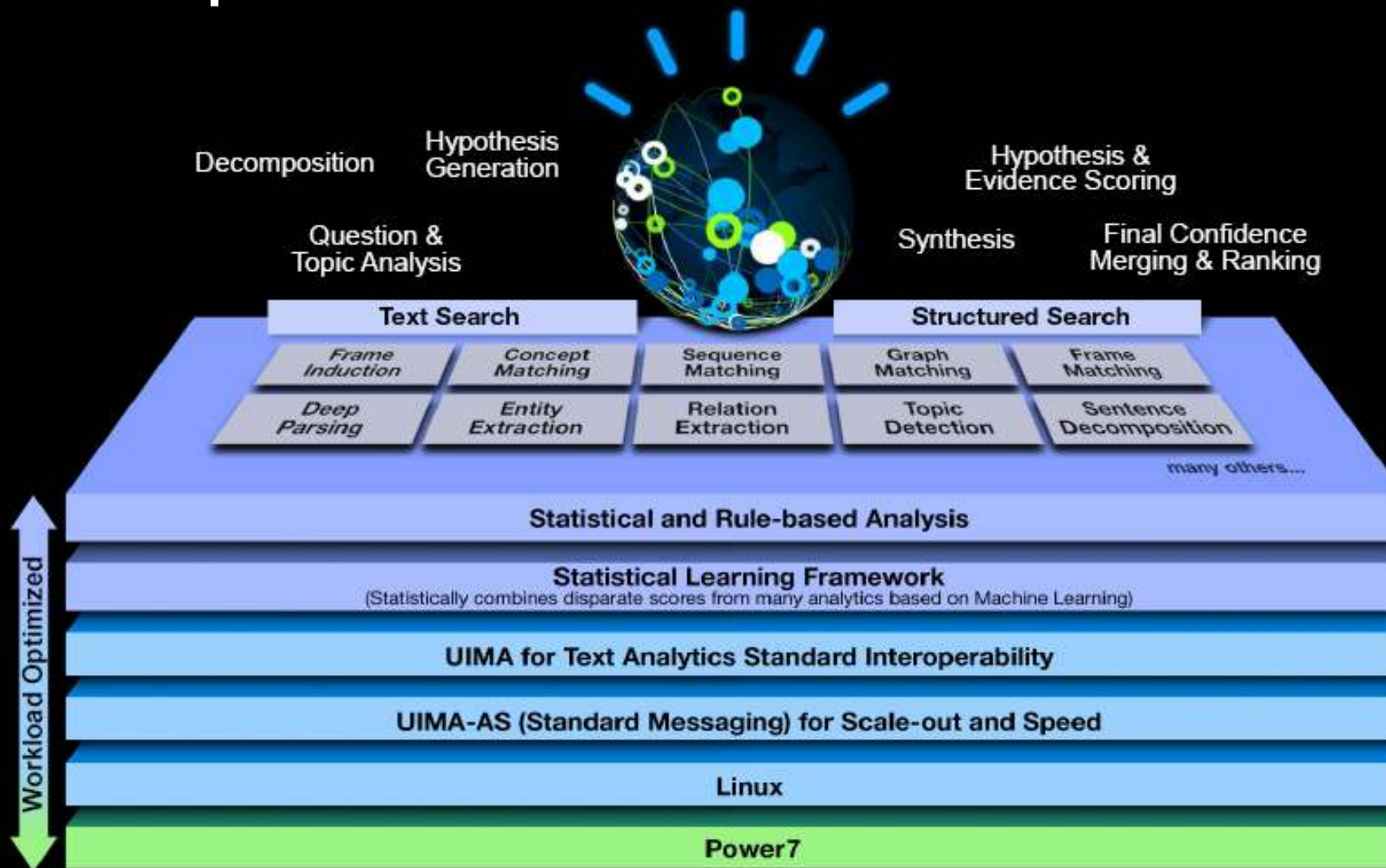
*Deep Q&A (Watson)
e i sistemi di Business Analytics*

Information-based Intelligence

- **“Metodi statistici e di Analytics”** → abbinati alla forza bruta computazionale, per analizzare un numero incredibilmente grande di dati (BIG Data) attraverso algoritmi e computer ad alte prestazioni
- **“Scalare”** → più informazione hai, più potenti sono i computer, più sofisticati sono gli algoritmi.... Migliori risultati hai!
- Originata dagli **approcci empirici** adottati per lo studio e l'interpretazione di enormi moli di dati generati nell'ambito degli esperimenti di fisica delle alte energie
- **Data mining** durante gli anni '90
- **IBM Deep Blue** (1997)
- **Deep Q&A - IBM Watson** (2011)

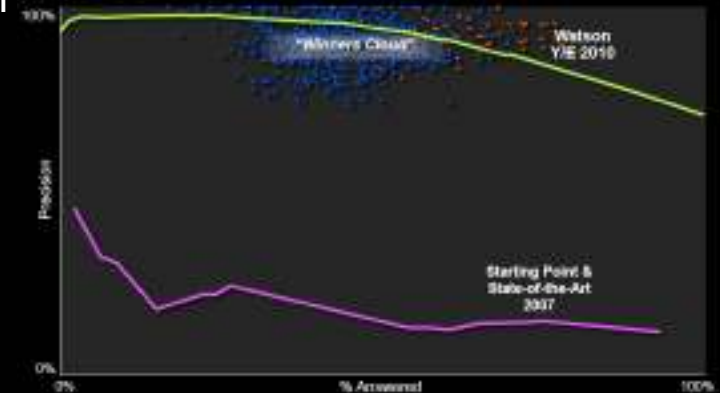
***Business
Analytics and
Optimization....
(Content analytics)***

DeepQA: L'architettura di Watson



Velocità, Precisione, Strategia

- **Deep Analytics** — Combinazione di un alto numero di strumenti di analisi in un'architettura di coordinamento abbiamo ottenuto un livello di precisione e confidenza nelle risposte a partire dall'analisi di una grande quantità di dati "as-is"



- **Velocità** — Ottimizzazione di un'infrastruttura hardware basata su processori POWER7 di 2880 core per ottenere un tempo di risposta in media di 3 secondi per domanda (l'elaborazione di una domanda da parte di un singolo core richiederebbe più di 2 ore di calcolo)
- **Risultati** — in 55 sfide con i campioni del gioco dello scorso anno Watson ha vinto nel 71% dei casi le sfide posizionandosi al primo posto, oltre ad aver vinto la sfida di febbraio con i campioni di sempre



Le ricadute applicative

Generazione di Ipotesi
Deep Q&A: Sistemi di domande e risposte
Supporto alle decisioni Analisi, decomposizione e ricostruzione di Fatti
Gestione della conoscenza

Information-based Tasks:

1. Analisi, decomposizione e ricostruzione di Fatti
2. Generazione di Ipotesi
3. Gestione della conoscenza
4. Supporto alle decisioni
5. Deep Q&A: Sistemi di domande e risposte....



Riconoscimento di Frodi



Supporto alla Diagnosi



Analisi e correlazione tra fatti investigativi



Supporto al ritrovamento e all'analisi di evidenze



Customer Service



Gestione della Conoscenza

Grazie!!!

Executive Architect
Member of IBM Academy of Technology

Email: pietro_leo@it.ibm.com

Twitter: [@pieroleo](https://twitter.com/pieroleo)

Linkedin: <http://it.linkedin.com/in/pieroleo>