



Regione Toscana

 **PENTOSCANA**



# Open & Big Data

## Le iniziative di Regione Toscana

*[info-bigdata@regione.toscana.it](mailto:info-bigdata@regione.toscana.it)*

[www.open.toscana.it](http://www.open.toscana.it)



# Prime sperimentazioni

- Nel corso del 2016, Regione Toscana ha avviato lo studio e la sperimentazione di alcune tecnologie nell'ambito dei big data, con particolare riguardo agli aspetti e alle potenzialità connesse con il mondo della business intelligence e del data warehouse.
- Il caso d'uso da cui siamo partiti è l'analisi statistica dei log applicativi, inizialmente con riferimento ai servizi del fascicolo sanitario elettronico, in seguito estendendo l'ambito di interesse a tutti i servizi digitali qualificati messi a disposizione dei cittadini e dei professionisti.



# Integrazione con Hadoop

- A tale riguardo Regione Toscana si è interessata ad esplorare l'integrazione della propria piattaforma di monitoraggio eventi di tipo generico (una implementazione dello stack tecnologico Elasticsearch/Logstash/Kibana) con gli strumenti di classe big data messi a disposizione dalla distribuzione Apache Bigtop.



# Nuovi casi d'uso

- In aggiunta all'esigenza di analizzare i log, in questo percorso si sono innestati, come era auspicabile, ulteriori casi d'uso ai quali è stato possibile applicare questi strumenti, affinandone la conoscenza ed impostando un percorso per la loro evoluzione, sia in termini di corretto dimensionamento dell'infrastruttura complessiva, che di efficace scelta dei singoli moduli di cui è composta.

# Modernizzazione data warehouse

La strategia è quella di utilizzare le tecnologie big data non in sostituzione, ma come ausilio delle tecnologie tradizionali, mettendo a disposizione un ambiente:

- di prima raccolta di tutti i dati;
- di persistenza dei dati storici in un archivio interrogabile e più economico;
- di efficientamento dei processi batch di elaborazione e trasformazione dei dati su larga scala prima del loro inserimento nel data warehouse;
- di esecuzione di processi di data discovery anche attraverso l'applicazione di algoritmi e metodi di machine learning.



# Federazione e virtualizzazione

In ottica di modernizzazione, le sperimentazioni condotte hanno altresì portato all'adozione di tecnologie orientate alla data virtualization, che vengono impiegate per la predisposizione di un “data warehouse logico” e che sono ottimali per:

- ridurre livelli intermedi di storage ed elaborazione dei dati, riducendo il time-to-market e i costi di sviluppo e manutenzione del data warehouse;
- mettere a disposizione un accesso ampio e trasversale a dati aggiornati real-time.



# Contesto organizzativo

- Da gennaio 2017 è operativa la modifica e l'integrazione dell'attuale gruppo di lavoro interdirezionale "Open data" nel nuovo gruppo di lavoro interdirezionale a carattere permanente "Open Data & Big Data"
- Affidamento 2016-2020 per lo sviluppo e la manutenzione del data warehouse e dei sistemi informativi direzionali
- Utilizzo delle infrastrutture per i big data e la virtualizzazione messi a disposizione presso il Tuscany Internet eXchange (TIX)
- Partecipazione al processo di condivisione della propria esperienza con altre pubbliche amministrazioni (Agenda Digitale, gruppo interregionale IT Sanità, network OT11/OT2, ...)



## Attività in corso

- Ottimizzazione e potenziamento dell'attuale piattaforma HDP®;
- Adozione diffusa e a regime delle componenti già sperimentate con successo su tematiche fiscali, amministrativo-contabili e relative al ciclo acquisti;
- Realizzazione di nuovi proof of concept sugli strumenti di machine learning messi a disposizione dalla piattaforma (log analysis mediante clustering non supervisionato, classificazione e indirizzamento dei protocolli in arrivo mediante classification supervisionato, sentiment analysis sui tweet connessi alle politiche amministrative, sociali e sanitarie della Regione);
- Estensione degli ambiti di interesse e di sperimentazione agli strumenti e ai metodi per la predisposizione e l'utilizzo di linked data.



Regione Toscana

 **PENTOSCANA**



**GRAZIE!**

[www.open.toscana.it](http://www.open.toscana.it)