

DATI PUBBLICI COME MOTORE DELL'ECONOMIA DIGITALE

*OSCAR SOVANI
PRESIDENZA – REGIONE LOMBARDIA*

16 maggio 2019



Regione
Lombardia

APOFTEGMA

«Data! Data! Data! I can't make bricks without clay»

Frase memorabile pronunciata da Sherlock Holmes nel racconto di Sir Arthur Conan Doyle «Le avventure dei faggi rossi»

GOVERNMENT BY DATA

Governance regionale del patrimonio informativo pubblico e ruolo di Regione Lombardia quale **Digital Information Hub** per la condivisione dei dati, l'accrescimento e arricchimento dei dati e il coinvolgimento del mercato e degli stakeholder per la generazione e lo scambio di nuovi dati.

Esperienze già avviate di data sharing

- Servizi Data as a service per il patrimonio informativo sanitario e socio sanitario
- Progetto DataLab
- Progetto Open Data

7 ANNI DI OPEN DATA LOMBARDIA

- **Tutti** i sensori **qualità dell'aria**, qualità delle acque e meteo dal 1968, oltre 650 milioni di record, aggiornati ogni 6 ore
- Decine di dataset della **Sanità**, inclusi dati di performance
- Sezione dedicata all'**Osservatorio Epidemiologico**
- Parco Autoveicoli (**8,5 mln**)
- **Catasto Impianti Termici** (caldaie) – oltre 3 mln di record
- **Catasto Certificazioni Energetiche Edifici** – oltre 1 mln di record

Statistiche (da marzo 2012)

+3.100 Datasets pubblicati
(Regione, Sistema Regionale
ed Enti Locali)

+14 mln pagine viste

+3 mln visualizzazioni dataset

+2 mln downloads dataset

2 «*Ossessioni*»

Automazione

156 dataset automatizzati
308 dataset semi-automatizzati



1.453 aggiornamenti/mese

Qualità dei dati

- ✓ Miglioramento continuo
- ✓ Correzione di ogni segnalazione
- ✓ Tool di verifica qualità
- ✓ Progetto con Univ. Milano Bicocca

OLTRE LA REGIONE, IL TERRITORIO

Delibera Open Data
per EELL

Dic
2012

18 EELL e 4
micrositi dedicati

2014-2017

136 EELL aderenti con
> 1.500 datasets

Dic
2017

Lug
2014

Primo elenco
di dataset
consigliati

Ott
2017

«Paniere» di dataset e
Co-finanziamento per
EELL

2018

Formazione e
supporto

2 CONDIZIONI

10 o 25 dataset
scelti nel **paniere** definito
da Regione Lombardia



Processi di pubblicazione
automatizzati

COLLABORAZIONI E RIUSI

Università e Ricerca : didattica, laboratori, tesi, ricerca



UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI INFORMATICA



UNIVERSITÀ
DEGLI STUDI
DI MILANO



UNIVERSITÀ
DEGLI STUDI
DI MILANO
BICOCCA



POLITECNICO DI MILANO
DIPARTIMENTO DI
ARCHITETTURA E
STUDI URBANI



UNIVERSITÀ
DEGLI STUDI
DI MILANO

SEDE DI CREMA
DIPARTIMENTO DI INFORMATICA
www.di.unimi.it



Università Commerciale
Luigi Bocconi
ASK
Art, Science
and Knowledge



UNIVERSITÀ
DEGLI STUDI
DI BRESCIA



UNIVERSITÀ DEGLI STUDI
DI BERGAMO
Dipartimento di Ingegneria gestionale,
dell'informazione e della produzione



POLITECNICO
DI TORINO
DENERG
Dipartimento Energia



LUISS
CMCS
Centre for Media and Communication
Studies Massimo Baldini

Civic hackers: diffusione, confronto



HeartInData



DataJournalism: diffusione, confronto, riuso

L'ECO
DI BERGAMO

VN
VareseNews
www.varesenews.it

la Provincia
PAVESE

MONZATODAY

netweek

LECCONOTIZIE.com

COLLABORAZIONI E RIUSI (in primis la Regione stessa)



- Legge Regionale
- Legge Regionale Statutaria
- Regolamento Interno
- Regolamento Regionale

Homepage Ricerca Ricerca avanzata Torna al sito del consiglio

Banca dati delle leggi regionali

Benvenuti in **OpenLeggiLombardia**

In questa banca dati sono presenti tutte le leggi e i regolamenti della Regione Lombardia, corredati da informazioni e documenti riguardanti l'iniziativa, la formazione e l'attuazione della legge e la valutazione delle politiche.

La consultazione avviene nel testo vigente; sono forniti anche i documenti allegati (tabelle, grafici, ecc.) se non particolarmente complessi. È inoltre possibile consultare i testi storici e confrontarli tra loro grazie alle apposite evidenziazioni.



Il modello Lombardo per la presa in carico



La delibera «Governo della domanda: avvio della presa in carico di pazienti cronici e fragili» (DGR n. X/6164 del 30.1.2017) e la successiva «Riordino della rete di offerta e modalità di presa in carico dei pazienti cronici e/o fragili» (in attuazione dell'art. 9 della L.R. 33/2009 DGR n. X/6551 del XX.05.2017) definiscono il nuovo assetto del sistema sanitario a partire dal tema della cronicità.



Medici di Medicina Generale e Pediatri di Libera Scelta (Ricerca puntuale)

DATASET ORIGINALE MMG-PLS

Esporta API

Mostrando all riga Riscala gli assi dopo il Filtraggio

| ATS | COGNOME | NOME | IN COOPERATIVA | NOME COOPERATIVA | CO-GESTORE | INDIRIZZO |
|---------|-------------|------------|----------------|-------------------------------|------------|-------------|
| MILANO | ABBATANGELO | ANGELA | NO | | SI | VIALE UMB |
| BRESCIA | ABBATE | LUIGI | SI | IN SALUTE SOC. COOP. | NO | VIA SAN BE |
| BERGAMO | ABBATI | GIUSEPPINA | SI | ME-DI.CO COOPERATIVA S C R.L. | NO | VIA IV NOVI |

COLLABORAZIONI E RIUSI



Sagre Lombarde

LuMa23 Development Viaggi e info locali
PEGI 3

Contiene annunci
L'app è compatibile con il tuo dispositivo.



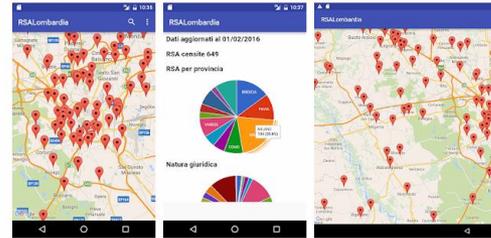
RSA Lombardia

Valerio Finazzo Salute e fitness
PEGI 3

Contiene annunci
L'app è compatibile con il tuo dispositivo.

Aggiungi alla lista desideri

Installa



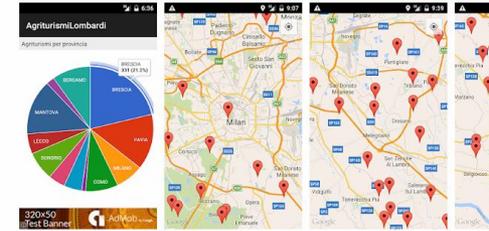
Agriturismi Lombardi

Valerio Finazzo Viaggi e info locali
PEGI 3

Contiene annunci
L'app è compatibile con il tuo dispositivo.

Aggiungi alla lista desideri

Installa



Coop Sociali Lombardia

Valerio Finazzo Viaggi e info locali
PEGI 3

Contiene annunci
L'app è compatibile con il tuo dispositivo.

Aggiungi alla lista desideri

Installa



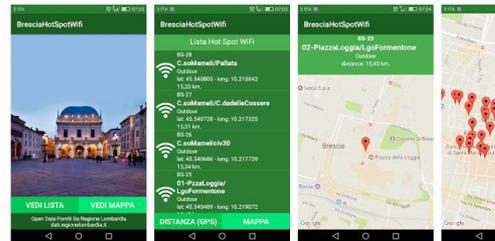
Brescia Free Wi-Fi

pierangelo orizio Viaggi e info locali
PEGI 3

L'app è compatibile con il tuo dispositivo.

Aggiungi alla lista desideri

Installa

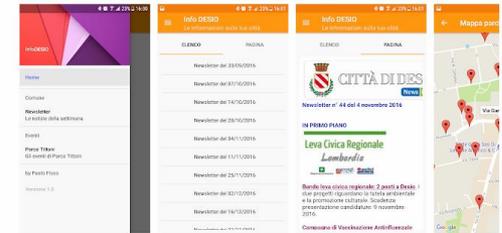


InfoDesio

Paolo Fisco Viaggi e info locali
PEGI 3

L'app è compatibile con il tuo dispositivo.

Installata



Home > openbarbetta > lombardia

Dove dormire in Lombardia

Elenco alloggi e strutture ricettive della Lombardia

provincia

classificazione struttura

AVVA

Ricerca una struttura ricettiva.

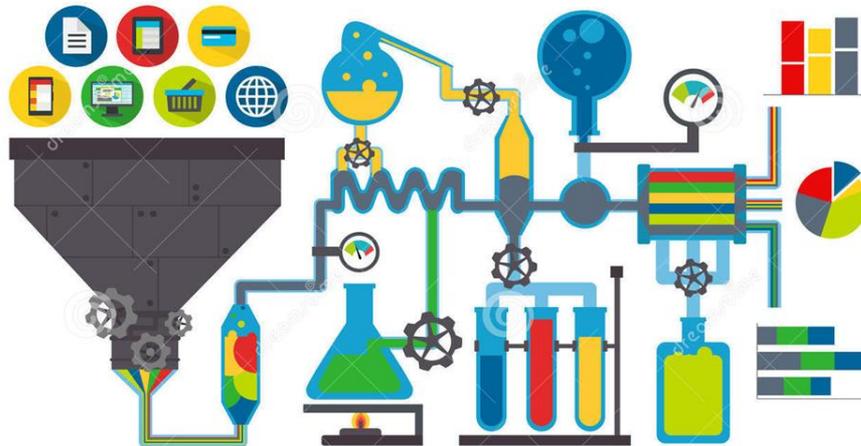
Attualmente in Regione Lombardia ci sono 3238 strutture ricettive (sono conteggiate solamente quelle attive nel database di fabiodiconzi.com).

| | | |
|----------------------|---------------|---------------|
| BERGAMO (283) | BRESCIA (825) | COMO (372) |
| CREMONA (49) | LECCO (129) | LODI (57) |
| MONZA E BRIANZA (72) | MILANO (208) | MANTOVA (178) |
| PAVIA (128) | SONDRIO (80) | VARESE (78) |



**Il laboratorio innovativo di analisi dei dati a supporto
della progettazione e del monitoraggio delle politiche regionali**

Cos'è il DATALab



DATA Lab è un progetto che vede **Polis Lombardia e Lombardia Informatica** lavorare insieme, mettendo a fattor comune le rispettive **competenze sui dati e sui processi regionali**, con l'obiettivo di costruire un **innovativo laboratorio di analisi dei dati** che supporti la **progettazione e il monitoraggio delle politiche regionali**, attraverso l'utilizzo di modelli, tecniche e strumenti di **statistica e di intelligenza artificiale** (machine learning, deep learning) sfruttando le più moderne tecnologie di **elaborazione dei dati**



DATA Lab: ambiti di analisi iniziali

PRELIMINARE

Il DATA Lab ha l'obiettivo di fornire supporto per la progettazione e il monitoraggio delle politiche regionali



| Ambiti | Lavoro  | Sicurezza  | Turismo  | ...altro?  |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| Casi d'uso | Occupabilità | Incidenti stradali | <i>In fase di definizione</i> | Ambiente Energia Famiglia Welfare... |
| Dati | Comunicazioni Obbligatorie (COB), Dichiarazioni immediata disponibilità (DID), Dote Unica Lavoro (DUL), Garanzia Giovani (GG) | Dati Istat: incidenti stradali, Archivio Statistico delle Imprese Attive (ASIA), Censimento della PA; dati Osservatorio Contratti Pubblici | Dati Istat presenze turistiche; dati Osservatorio del Turismo | <i>Da definire / analizzare</i> |



DATA Lab: Analisi di occupabilità



OBIETTIVO

Regione Lombardia fornisce sostegno alle persone in cerca di lavoro tramite interventi di politica attiva.

L'obiettivo dell'analisi è stato quello di valutare un sistema di profilazione, basato sulle **probabilità di ritrovare lavoro, per i cittadini in cerca di occupazione** (stima dell'occupabilità).



I DATI

Il dataset contiene i contratti dei cittadini in cerca di occupazione che hanno rilasciato almeno una Dichiarazione di Immediata Disponibilità (DID);

Periodo: Dicembre 2015 – Aprile 2018;

Contiene informazioni:

- Sul lavoratore;
- Storia contrattuale;
- Riguardanti la DID.



VARIABLE TARGET

Problema di classificazione:

- Cittadini in cerca di occupazione che **non trovano lavoro entro 180 giorni** dalla dichiarazione di disponibilità (*Target 0*);
- Cittadini in cerca di occupazione che **trovano lavoro entro 180 giorni** dalla dichiarazione di disponibilità (*Target 1*).

Occupabilità Fase 1

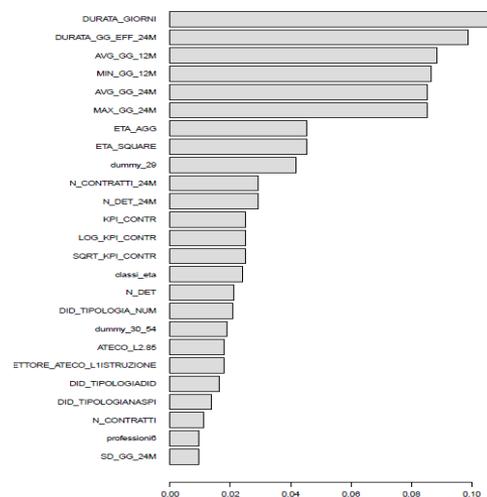
Baseline Model... Decision tree

CONFUSION MATRIX

| | | Actual | |
|-----------|------------------|------------------|--------------|
| | | Non trova lavoro | Trova lavoro |
| Predicted | Non trova lavoro | 33390 | 12964 |
| | Trova lavoro | 11016 | 23584 |

DETAILS

| | | | | |
|-----------------------------|-----------------------------|---------------------------|------------------------|--------------------|
| Sensitivity 0.645 | Specificity 0.752 | Precision 0.682 | Recall 0.645 | F1 0.663 |
| Accuracy 0.704 | | Kappa 0.399 | | |



- Threshold = 0.42
- CP: 0.0001
- Discreta accuratezza (0.704)
- F1 score basso pari a 0.663

- **Variabili più rilevanti:** feature estratte dagli ultimi 2 anni di contratti.
- **Caratteristiche del lavoratore:** la variabile età.
- **Tipologia della DID**
- **L'istruzione** sembrerebbe **NON influenzare** l'occupabilità

Occupabilità Fase 2

Analisi della «sopravvivenza» con il modello di Cox

Modello a rischi non proporzionali

- **Eventi prima del rilascio della DID, non sono ancora stati considerati gli eventi di politica attiva**
- Discreta performance del modello: **Concordanza 0.74**

Effetti Positivi

- Gli uomini hanno più probabilità di ritrovare lavoro rispetto alle donne (**il «rischio» di ritrovare lavoro è il 64% in più rispetto alle donne**)
- **Titoli di studio più elevati aumentano la probabilità di ritrovare lavoro entro 180 giorni dal rilascio della DID**
- Avere una **cittadinanza italiana** rispetto ad una straniera
- L'effetto congiunto tra **titolo di studio e la cittadinanza italiana**
- **Effetto congiunto tra il numero di contratti e l'età** (Esperienza?)
- Effetto congiunto **tra titoli di studio alti e l'età** (Anni di esperienza di persone altamente qualificate)

Effetti negativi

- **L'aumento dell'età e del numero di contratti (non congiunto ma come effetti principali)** diminuisce la probabilità di ritrovare il lavoro
- L'effetto congiunto tra il **genere maschile e la cittadinanza italiana** riduce la probabilità di ritrovare lavoro rispetto alle donne con cittadinanza straniera.

```
coef      exp(coef)    se(coef)    robust
ANAG_GENEREM      4.979e-01  1.645e+00  1.804e-02  1.931e-
ISCED_ISTRUZIONE1 -2.741e-01  7.602e-01  1.655e-01  1.850e-01
ISCED_ISTRUZIONE2  2.664e-01  1.305e+00  3.872e-02  4.614e-02
ISCED_ISTRUZIONE3  7.503e-01  2.118e+00  4.317e-02  5.168e-02
ISCED_ISTRUZIONE5  1.674e+00  5.335e+00  7.930e-02  1.126e-
ISCED_ISTRUZIONE6  1.554e+00  4.731e+00  7.681e-01  1.098e+
N_CONTRATTI      -5.064e-02  9.506e-01  3.720e-03  1.264e-
ETA_LAVORATORE   -5.622e-02  9.453e-01  7.432e-04  1.480e-
CITT_ITAITALIANA  -4.863e-01  1.626e+00  1.936e-02  2.080e-
ANAG_GENEREM:CITT_ITAITALIANA -2.269e-01  7.970e-01  1.998e-02  2.135e-
ISCED_ISTRUZIONE1:ETA_LAVORATORE  4.154e-03  1.004e+00  4.130e-03  4.576e-
ISCED_ISTRUZIONE2:ETA_LAVORATORE -2.537e-03  9.975e-01  9.176e-04  1.097e-
ISCED_ISTRUZIONE3:ETA_LAVORATORE -1.817e-02  9.820e-01  9.380e-04  1.158e-
ISCED_ISTRUZIONE5:ETA_LAVORATORE -3.752e-02  9.632e-01  1.575e-03  2.574e-
ISCED_ISTRUZIONE6:ETA_LAVORATORE -9.199e-02  9.121e-01  1.876e-02  2.927e-
ISCED_ISTRUZIONE1:CITT_ITAITALIANA -4.750e-01  6.219e-01  1.021e-01  1.079e-
ISCED_ISTRUZIONE2:CITT_ITAITALIANA -1.957e-01  8.223e-01  2.487e-02  2.758e-
ISCED_ISTRUZIONE3:CITT_ITAITALIANA -1.118e-01  8.942e-01  3.165e-02  3.572e-
ISCED_ISTRUZIONE5:CITT_ITAITALIANA  2.488e-01  1.282e+00  6.248e-02  7.877e-
ISCED_ISTRUZIONE6:CITT_ITAITALIANA  1.026e+00  2.789e+00  5.142e-01  5.359e-
N_CONTRATTI:ETA_LAVORATORE      2.417e-03  1.002e+00  7.549e-05  4.243e-
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

exp(coef)  exp(-coef)  lower .95  upper .95
ANAG_GENEREM      1.6453      0.6078      1.5842      1.7087
ISCED_ISTRUZIONE1  0.7602      1.3154      0.5290      1.0925
ISCED_ISTRUZIONE2  1.3052      0.7662      1.1923      1.4287
ISCED_ISTRUZIONE3  2.1176      0.4722      1.9136      2.3433
ISCED_ISTRUZIONE5  5.3352      0.1874      4.2785      6.6528
ISCED_ISTRUZIONE6  4.7312      0.2114      0.5500      40.7013
N_CONTRATTI      0.9506      1.0519      0.9274      0.9745
ETA_LAVORATORE    0.9453      1.0578      0.9426      0.9481
CITT_ITAITALIANA  1.6263      0.6149      1.5613      1.6939
ANAG_GENEREM:CITT_ITAITALIANA  0.7970      1.2547      0.7643      0.8311
ISCED_ISTRUZIONE1:ETA_LAVORATORE  1.0042      0.9959      0.9952      1.0132
ISCED_ISTRUZIONE2:ETA_LAVORATORE  0.9975      1.0025      0.9953      0.9996
ISCED_ISTRUZIONE3:ETA_LAVORATORE  0.9820      1.0183      0.9798      0.9842
ISCED_ISTRUZIONE5:ETA_LAVORATORE  0.9632      1.0382      0.9583      0.9680
ISCED_ISTRUZIONE6:ETA_LAVORATORE  0.9121      1.0964      0.8613      0.9660
ISCED_ISTRUZIONE1:CITT_ITAITALIANA  0.6219      1.6079      0.5033      0.7684
ISCED_ISTRUZIONE2:CITT_ITAITALIANA  0.8223      1.2161      0.7790      0.8680
ISCED_ISTRUZIONE3:CITT_ITAITALIANA  0.8942      1.1183      0.8337      0.9591
ISCED_ISTRUZIONE5:CITT_ITAITALIANA  1.2825      0.7798      1.0990      1.4966
ISCED_ISTRUZIONE6:CITT_ITAITALIANA  2.7889      0.3586      0.9756      7.9729
N_CONTRATTI:ETA_LAVORATORE      1.0024      0.9976      1.0016      1.0033

Concordance= 0.74 (se = 0.001 )
Rsquare= 0.035 (max possible= 0.756 )
Likelihood ratio test= 41158 on 21 df, p=<2e-16
Wald test = 25652 on 21 df, p=<2e-16
Score (logrank) test = 45241 on 21 df, p=<2e-16, Robust = 30871 p=<2e-16
```



Regione
Lombardia



Lettura a matrice del territorio

Su alcuni temi Datalab può contribuire alla costruzione di matrici di indicatori territoriali, integrando gli indicatori PRS con la misurazione di indicatori di impatto e con analisi di natura



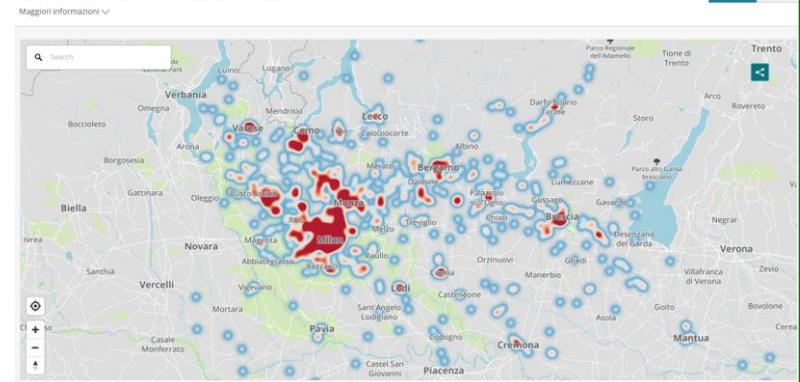
Aziende cinesi in Lombardia:

6700 dipendenti

4,7 Miliardi di fatturato

Aziende a rischio di bancarotta in Lombardia nei prossimi 18 mesi (indici di rischio 10 volte superiori alla media nazionale e con più di 5 dipendenti).

mappa aziende a rischio bancarotta



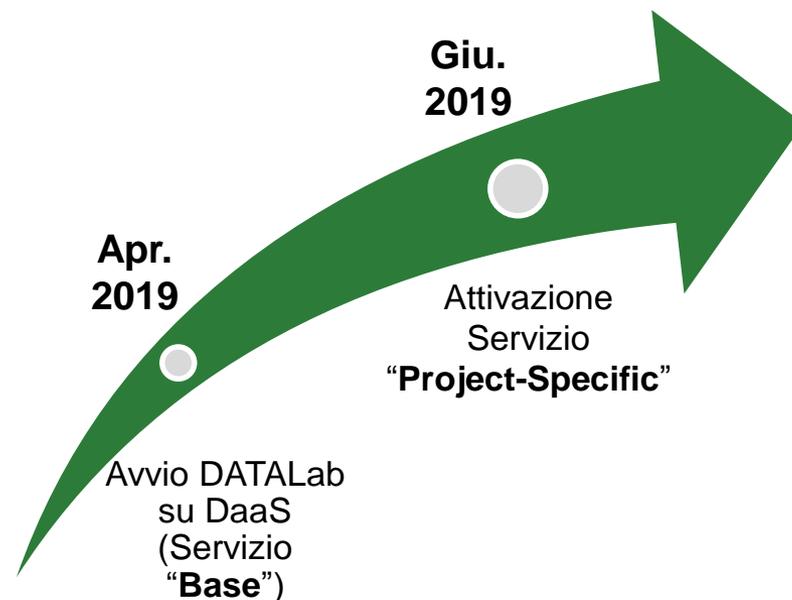
DATA Lab e scenario evolutivo Data As A Service (DAAS)



DaaS a servizio del DATA Lab

A partire da Aprile si prevede l'attivazione del servizio DaaS per DATA Lab

- Nel 2018 Llspa ha avviato il **consolidamento del modello tecnico-organizzativo del «Data as a Service» (DaaS)** in una logica di «servizio», con l'intento di condividere il Patrimonio Informativo sanitario e socio-sanitario all'interno di quanto previsto dalla DGR X/4893/2016, recentemente integrata dalla DGR XI/491/2018
- Attraverso la gestione sicura e a norma delle richieste e della messa a disposizione dei dati e delle informazioni utilizzabili è prevista **l'estensione del modello ad altri enti del Sistema Regionale**
- In particolare, sfruttando la scalabilità della tecnologia e dei processi operativi per le finalità individuate da Regione Lombardia, si vuole abilitare l'estensione del servizio al **contesto e alle specifiche attività del DATA Lab POLIS**



Modello di riferimento

L'attivazione del DATA Lab come strumento trasversale di supporto per le politiche regionali prevede la definizione di quattro elementi alla base del modello di funzionamento del DaaS

Finalità istituzionale

Definizione di finalità di tipo istituzionale

Esempio: finalità di programmazione, controllo e monitoraggio delle politiche regionali

Base giuridica

Provvedimenti normativi dotati degli elementi minimi per definire:

- Finalità e modalità del trattamento di dati personali;
- Tipologia dei dati trattati;
- Archivi e tipologia delle operazioni eseguite;
- Eventuali operazioni particolari (interconnessione, comunicazione e diffusione)

Strumenti negoziali

Accordi con le parti, inclusi gli accordi con terze parti

Esempio: contratti quadro e/o convenzioni che definiscono i rapporti economici, le caratteristiche del servizio e i processi operativi

Adempimenti specifici

Documenti atti a specificare le caratteristiche operative ed organizzative delle singole iniziative, con il coinvolgimento della/e DG di competenza

Esempio: proposte progettuali, atti di nomina a Responsabile del trattamento di dati personali e altri allegati «tecnici»



Regione
Lombardia



Privacy-by-design

Il DaaS è l'insieme di organizzazione, risorse e tecnologie per valorizzare il Patrimonio Informativo regionale coerentemente col quadro normativo vigente

Driver

Numerose richieste di poter accedere ai dati di Regione Lombardia

Vincoli

Normativa vigente per la sicurezza e la riservatezza dei dati

Fattore abilitante:

Privacy-by-design*

Misure organizzative

Misure tecniche

Misure organizzative

- Trattamento predefinito dei soli dati personali necessari per ogni specifica finalità del trattamento
- Conformità con il principio della minimizzazione dei dati

Misure tecniche

- Pseudonimizzazione e anonimizzazione dei dati personali in linea con i requisiti normativi
- Adeguati sistemi IT e opportuna progettazione della struttura e infrastruttura di rete

Accesso al DaaS

Conforme

Securizzato

Monitorato

Scalabile



Regione Lombardia

