



Vincenzo Guardabasso

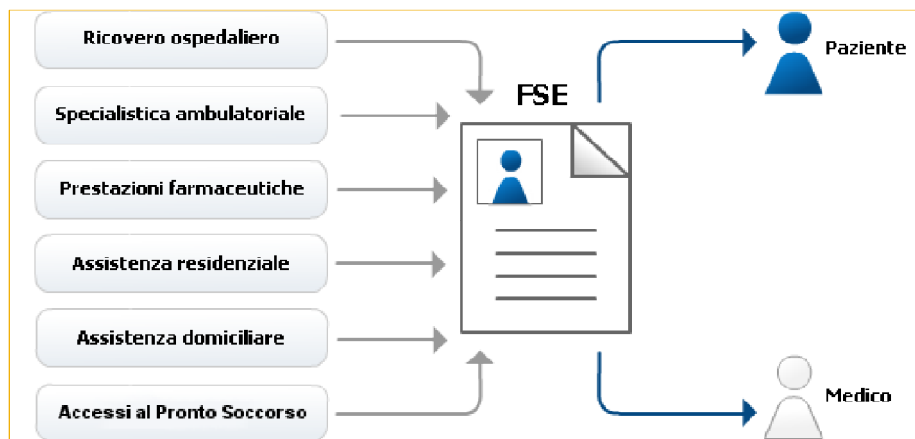
Prospettive di collegamento dei database regionali: ostacoli e facilitazioni

Workshop "Algoritmi per l'integrazione di dati sanitari: database, record-linkage, anonimizzazione"
Milano, 14 Aprile 2016

Progetto ARCHES

- Progetto di ricerca "*Electronic health databases as a source of reliable information for effective health policy*"
- finanziato da **Regione Puglia** e **Ministero della Salute**
(Ricerca Finalizzata 2010, RF-2010-2315604)
- Obiettivo generale: valutare gli **ARCHivi Elettronici Sanitari** come fonte di informazione a fini epidemiologici e di programmazione sanitaria

Fascicolo Sanitario Elettronico



DPCM 29/9/2015 , n. 178 (Regolamento FSE)

da: <http://www.nsis.salute.gov.it>

Codice: nel FSE troviamo...

- DPCM 29 settembre 2015 , n. 178
Regolamento in materia di fascicolo sanitario elettronico
(G.U. 263/ 11 novembre 2015)

ART. 1 (*Definizioni*)

1. Ai fini del presente decreto si intende per:

...

m) "**codice univoco**", il codice assegnato, attraverso una procedura automatica, ad ogni assistito [a partire dal codice fiscale](#), tale da non consentire la identificazione diretta dell'assistito durante il trattamento dei dati personali;

Il Censimento

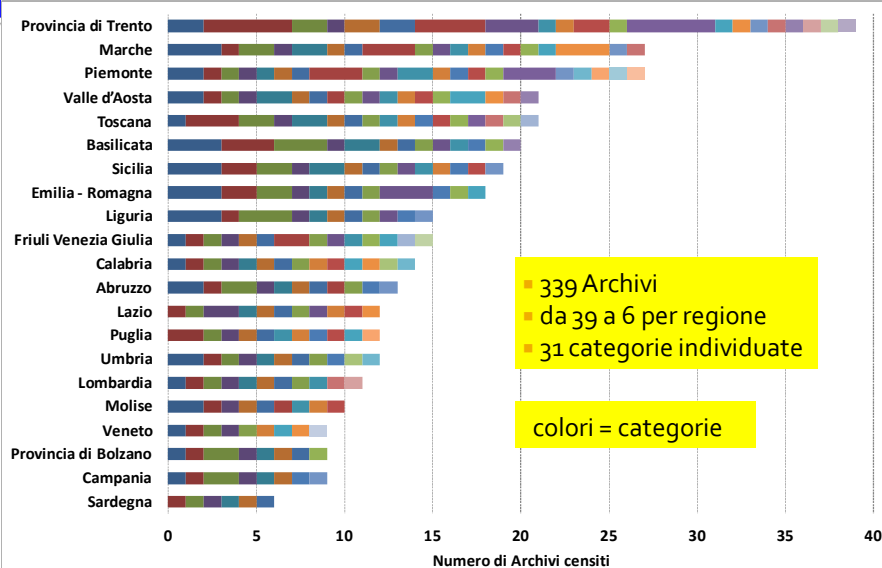
da: Gesuita et al., I
Database Elettronici
Amministrativi di natura
sanitaria in Italia...
Atti VIII Congresso
Nazionale SISMEC, Torino
16-19 Settembre 2015,
p. 396.

Obiettivo specifico:

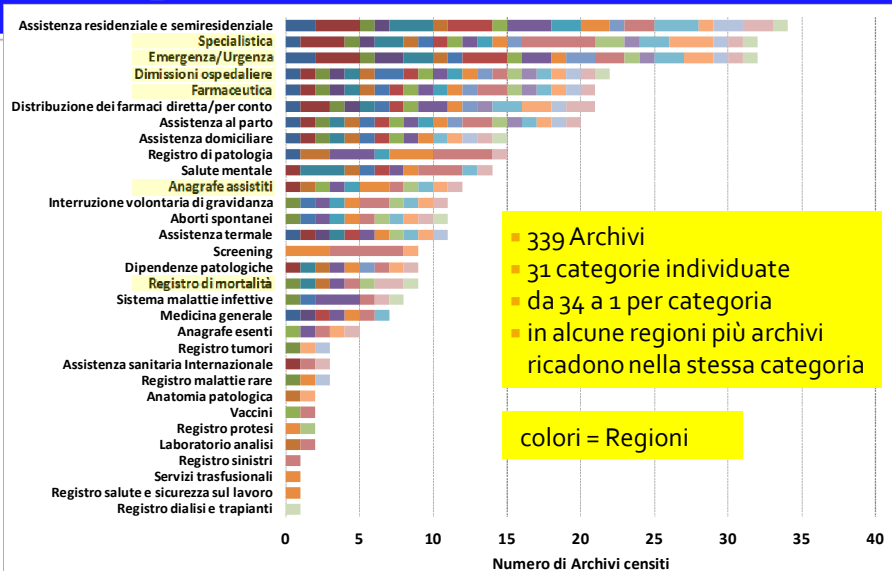
- censire le banche dati
sanitarie automatizzate
esistenti sul territorio
nazionale [al livello regionale]
- ... e verificarne le potenzialità
di utilizzo per la realizzazione
di studi osservazionali



Archivi censiti nelle Regioni



Categorie di Archivi censiti



Altri Obiettivi specifici del Progetto

- [Individuare e analizzare] le procedure di *record-linkage*...
 - utili per l'integrazione degli Archivi Sanitari Elettronici
- Verificare la migliore modalità di anonimizzazione dei dati...
 - per garantire la *privacy*...
 - [nella] trasmissione dei dati sanitari fra strutture intra- e inter-regionali.

Aspetti rilevanti per l'integrazione

- Presenza di una "chiave di *linkage*"?

- è possibile indicare procedure di *record-linkage* efficaci?



SA-NT DataLink, AU

- Anonimizzazione e "Privacy"

- è possibile indicare procedure che garantiscano adeguatamente la Privacy in accordo con le normative vigenti?



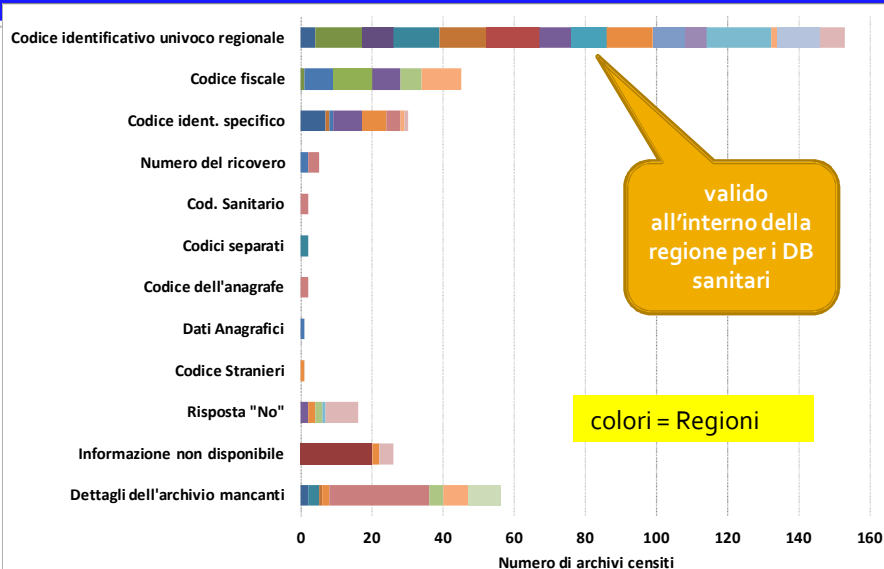
iq oncept/123RF Archivio Fotografico

Tre domande nel questionario

17	Codice identificativo dell'individuo	Indicare il codice con cui ciascuna unità di osservazione è associata all'individuo che la genera. Questa voce rappresenta la chiave di link fra i database. Essa permette di utilizzare a fini epidemiologici i database amministrativi, potendo passare dalla unità di osservazione (non necessariamente il singolo soggetto) del database all'individuo, generalmente l'unità di osservazione rilevante nella ricerca epidemiologica. Codice identificativo dell'assistito costruito dalla Regione, unico per tutti i DB di interesse sanitario, costruito a partire dall'anagrafica dei residenti/assistiti
18	Codice identificativo è stato anonimizzato	Indicare se il codice identificativo dell'individuo è stato anonimizzato Il codice utilizzato è generato da un programma in gestione presso un ufficio regionale, unico deputato alla sua creazione, e garantisce l'impossibilità per chiunque di risalire all'individuo, fatto salvo lo stesso ufficio regionale che l'ha generato.
19	Se sì, con quale procedura	Descrivere brevemente la procedura di anonimizzazione Tramite la secretazione e, quindi, la protezione dell'abbinamento univoco tra codice e anagrafica della popolazione residente /assistiti

- In nero, la domanda
- In blu, la spiegazione
- In arancione, un esempio di risposta

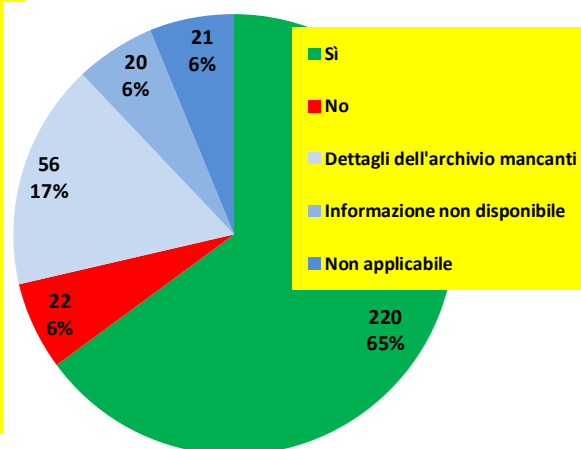
Domanda 17: Codice identificativo



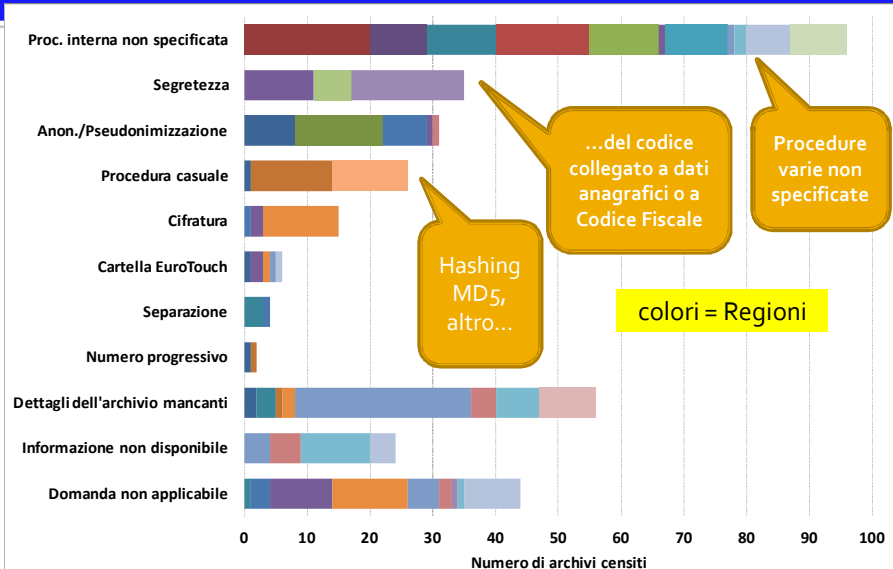
Domanda 18: Anonimizzazione

■ E' garantito l'anonimato?

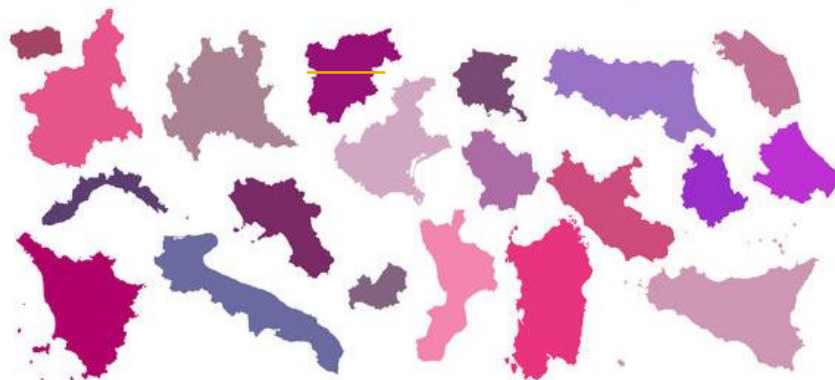
- Sì
- ... ma le informazioni sono incomplete



Domanda 19: Procedura usata



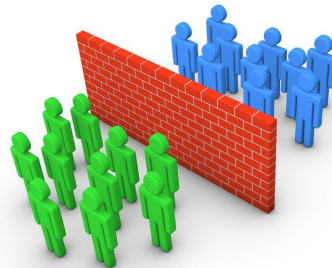
Il quadro nazionale:



Autore: ynochka / 123RF Archivio Fotografico

Ostacoli ...

- Frammentazione
 - U nella giurisdizione
 - U nei sistemi usati
- Variabilità della presenza di archivi
- Difficoltà ad individuare alcuni interlocutori

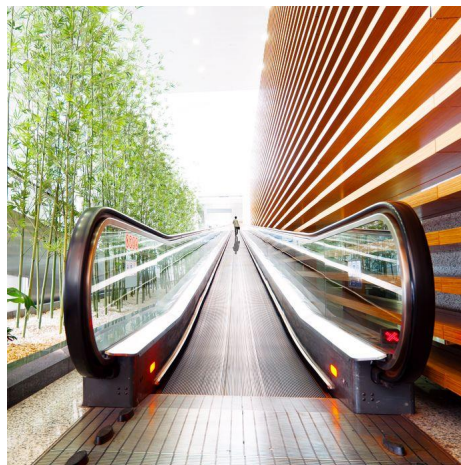


viktor88/123RF Archivio Fotografico

... e facilitazioni

- numero limitato di interlocutori
- sistema rodato: tempestività

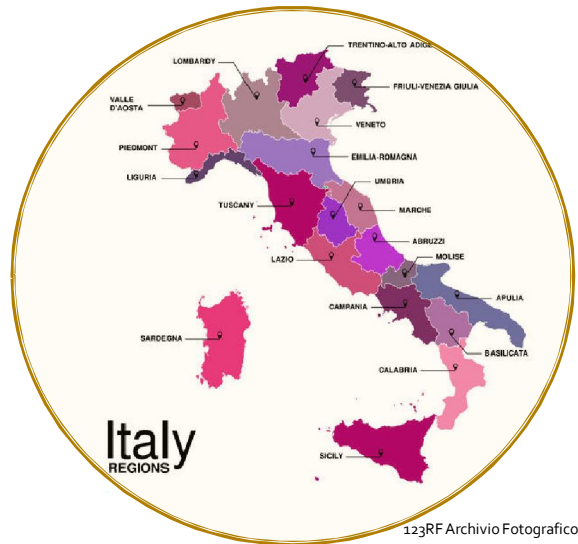
Storia retrospettiva:
i più importanti archivi
risalgono a molti anni indietro



gyn9037/123RF Archivio Fotografico

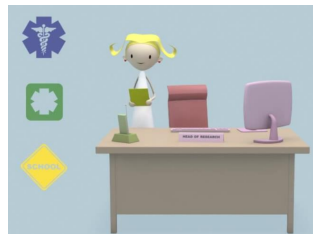
Ricomporre il quadro nazionale?

- E' possibile ricomporre il quadro?
- Sì, certo...



Si può fare ricerca sui dati degli Archivi sanitari regionali?

- Ministeri e Regioni certamente possono, con le limitazioni previste dalle leggi (ad es. FSE)
- Poniamoci dal punto di vista del ricercatore che non è il destinatario dei flussi di dati:
 - che può fare?



Big data?



<http://www.hostingtalk.it>

■ Di quanti dati stiamo parlando?

Tipo di informazione	Anno 2014
Schede di Dimissione Ospedaliera	9,5 Milioni
Accessi al Pronto Soccorso	20,5 Milioni
Ricette farmaceutiche	609 Milioni
Prestazioni diagnostiche	?
Visite medici di base/pediatri	?

Come fare ricerca? Soluzione 1

<http://95.110.213.190/PNEed15/>

- Riunire i dati nello stesso posto
 - Esempio: Programma Nazionale Esiti
 - dove: Dipartimento di Epidemiologia del Servizio Sanitario Regionale - Regione Lazio

Come fare ricerca? Altre soluzioni

- Soluzione 2: usare dati e indicatori aggregati
- Soluzione 3: formare un consorzio *ad hoc*
 - EU-ADR
 - EU₄Research
 - Healthcare Research & Pharmacoepidemiology
- Soluzione 4:
 - il "Sistema australiano": *l'agenzia di record-linkage*

L'approccio australiano:

- *The Privacy Protecting Model*
 - ... a *linkage service* to enable research and policy analysis of *de-identified* data from *multiple* databases, instead of researchers having access to individual's personal details.
- *The Separation Principle*
 - a model that restricts the use of identified data and anonymises data to allow researchers access to only de-identified data for research studies.



De-identificazione di “parte terza”

- il **ricercatore** contatta l'Agenzia e dice chi sono i vari **gestori** dei database per l'analisi.
- L'agenzia esamina i soli dati **identificativi** e fornisce a tutti i gestori un codice **anonimo** per ogni individuo.
- Il ricercatore riceve i dati **sanitari** con codici anonimi.
- I dati possono essere collegati...



SA-NT DataLink, AU



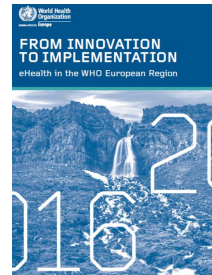
Conclusione: questioni aperte

... per ricerca epidemiologica e di governo:

- CHI può fare ricerca?
 - chi può legalmente fare studi longitudinali o trasversali su episodi di cura comprendenti più regioni?
- COME fare ricerca?
 - serve un **codice anonimo**
 - serve l'**integrazione** trasversale (nel Paese) e longitudinale (nel tempo)
 - servono una **tassonomia** e **metadati**
 - ... per avere **open data** come in altri paesi

Indicazioni da (per)seguire

- *Lack of data governance*
 - principale ostacolo all'uso dei *big data* per la salute



© World Health Organization 2016



© European Union, 2012



- *Liberate the data!*
 - Leva n. 2 per il cambiamento