



L'importanza dell'interoperabilità per l'accountability pubblica e il supporto alle politiche: l'esperienza dell'ontologia nazionale su Alta Formazione, Ricerca e Innovazione

Dall'integrazione alla fruizione dei dati: Gli strumenti di data visualisation e business intelligence

Luca Cipriani (Regione Toscana) e Annamaria Donnarumma (SIRIS Academic)



Regione Toscana



Il valore aggiunto della visualizzazione dei dati

Visualizzare i dati è fondamentale per **comprendere** meglio le caratteristiche dei dati e della loro distribuzione

Ad esempio: il **quartetto di Anscombe**

Ref: Anscombe, F. J. (1973). "Graphs in Statistical Analysis".
doi:10.1080/00031305.1973.10478966

Quattro **serie di dati** che hanno
una **statistica descrittiva
semplice quasi identica...**



	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10	8.04	10	9.14	10	7.46	8	6.58
	8	6.95	8	8.14	8	6.77	8	5.76
	13	7.58	13	8.74	13	12.74	8	7.71
	9	8.81	9	8.77	9	7.11	8	8.84
	11	8.33	11	9.26	11	7.81	8	8.47
	14	9.96	14	8.1	14	8.84	8	7.04
	6	7.24	6	6.13	6	6.08	8	5.25
	4	4.26	4	3.1	4	5.39	19	12.5
	12	10.84	12	9.13	12	8.15	8	5.56
	7	4.82	7	7.26	7	6.42	8	7.91
	5	5.68	5	4.74	5	5.73	8	6.89
Media	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Varianza	11.0	4.1	11.0	4.1	11.0	4.1	11.0	4.1

Fonte dei dati:
Wikipedia

Il valore aggiunto della visualizzazione dei dati

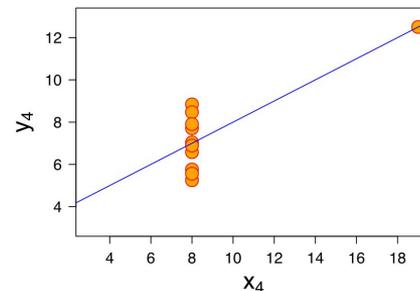
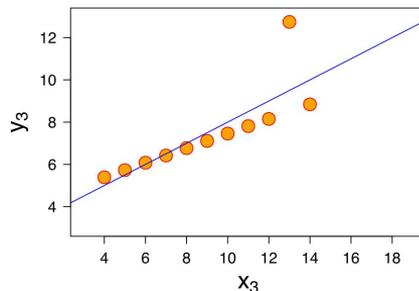
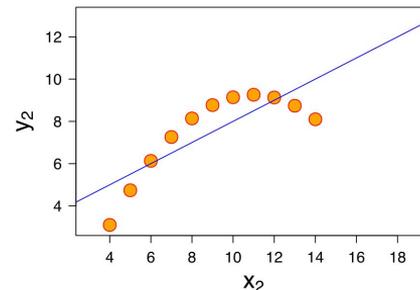
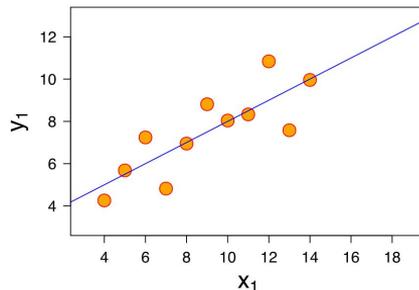
Visualizzare i dati è fondamentale per **comprendere** meglio le caratteristiche dei dati e della loro distribuzione

Ad esempio: **il quartetto di Anscombe**

Ref: Anscombe, F. J. (1973). "Graphs in Statistical Analysis".

doi:10.1080/00031305.1973.10478966

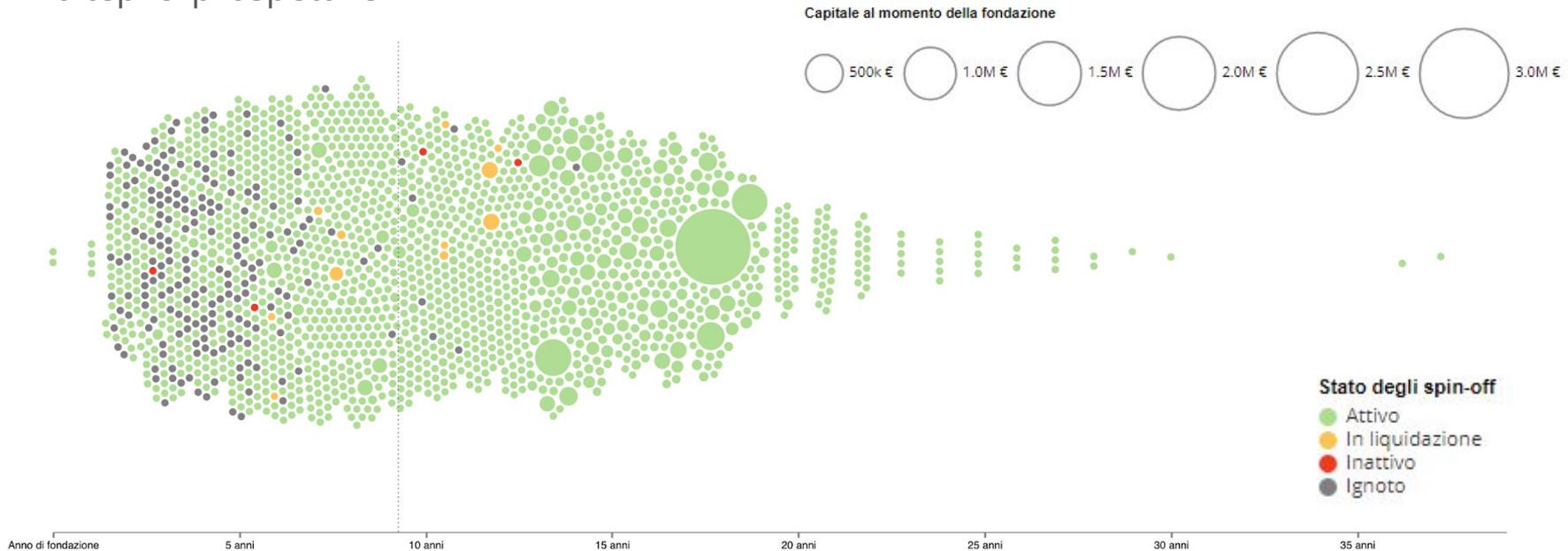
Quattro **serie di dati** che hanno una **statistica descrittiva semplice quasi identica...**
... Ma che risultano molto **differenti** quando **visualizzate**



Fonte dell'immagine: Wikipedia

Il valore aggiunto della visualizzazione dei dati

Visualizzare i dati permette di **comunicare** l'informazione più efficacemente e secondo molteplici prospettive



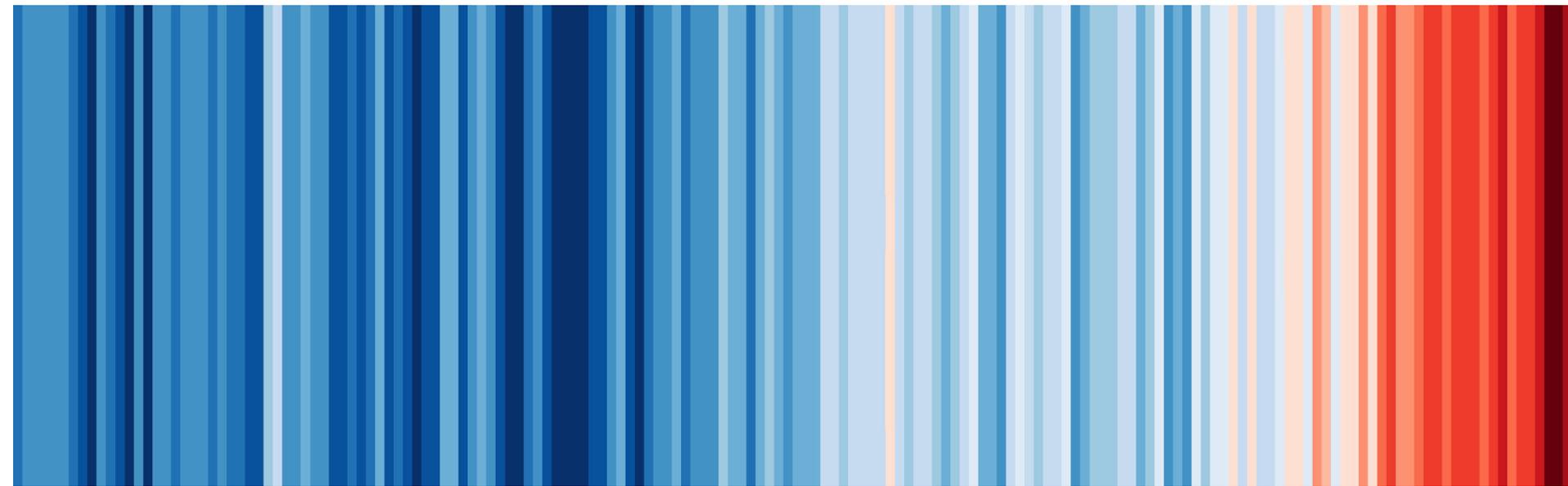
Distribuzione delle spin-off per durata di vita dell'azienda a partire dall'inizio della sua attività

Credits: Fondazione Sistema Toscana - Regione Toscana - SIRIS Academic

Fonte dell'immagine: http://www.toscanaopenresearch.it/esplora-i-dati/#/ricerca_e_innovazione?toscana_spinoff

Il valore aggiunto della visualizzazione dei dati

Visualizzare i dati permette di **coinvolgere** l'utente ed aumentare l'impatto dell'informazione



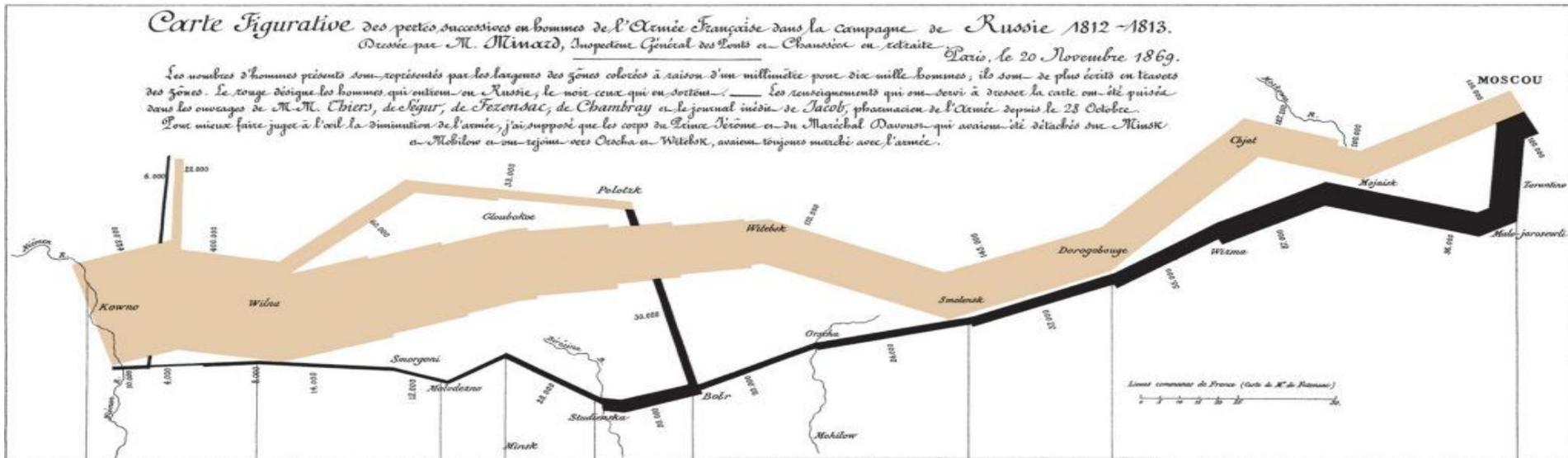
“Warming Stripes”, sugli effetti del riscaldamento globale.

Credits: Ed Hawkins (University of Reading)

Fonte dell'immagine: <https://showyourstripes.info/>

Il valore aggiunto della visualizzazione dei dati

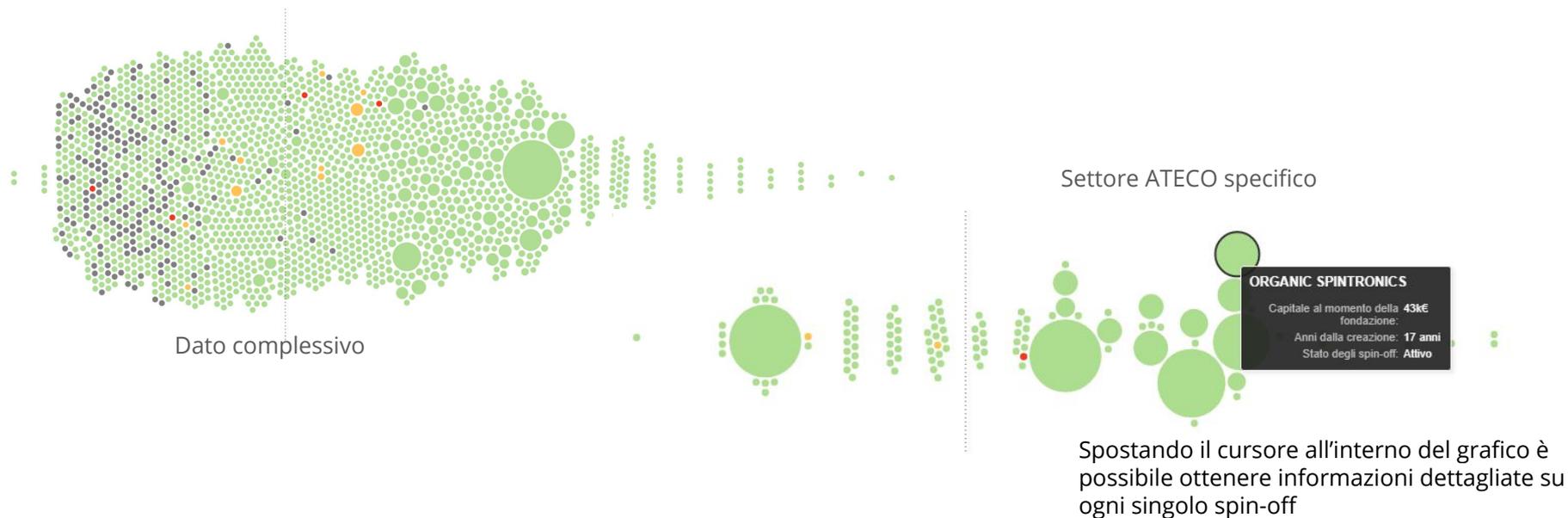
Visualizzare i dati permette di **raccontare** in modo sintetico e dinamico un fenomeno



La campagna di Russia secondo Charles Joseph Minard
Fonte dell'immagine: https://en.wikipedia.org/wiki/Charles_Joseph_Minard

Il valore aggiunto dell'interattività

Le rappresentazioni non sono statiche: l'utente può **interagire** con esse secondo i propri interessi di analisi.



Distribuzione delle spin-off per durata di vita dell'azienda a partire dall'inizio della sua attività

Credits: Fondazione Sistema Toscana - Regione Toscana - SIRIS Academic

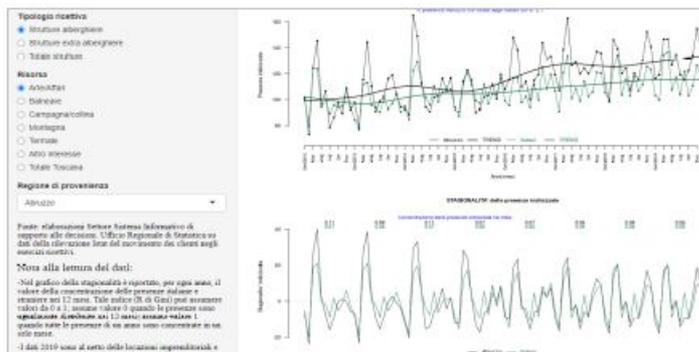
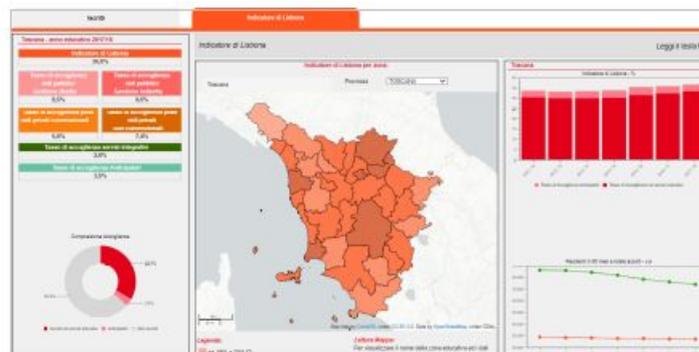
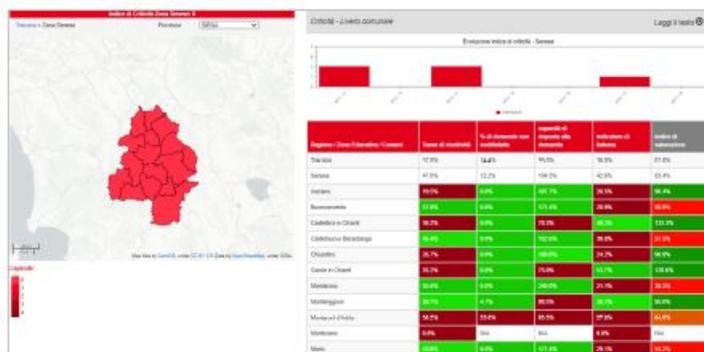
Fonte dell'immagine: http://www.toscanaopenresearch.it/esplora-i-dati/#/ricerca_e_innovazione?toscana_spinoff

Alcuni obiettivi della data visualization per la PA

- Identificare i processi di trasformazione in atto, in modo da adattare/informare le (re)azioni politiche;
- Rafforzare la fiducia e la cooperazione con e tra gli stakeholder e i cittadini;
- Garantire la responsabilità del processo di elaborazione delle politiche;
- Informare e diffondere conoscenza con modalità intuitive e facilmente accessibili;
- Evidenziare il valore potenziale e contribuire alla diffusione dei dati *open*.

Diversi strumenti per utenti e scopi informativi differenti

Policy-makers, cittadini e utenti della Regione: immediatezza, capacità di sintesi, ergonomia

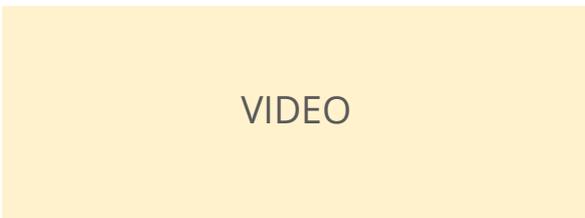
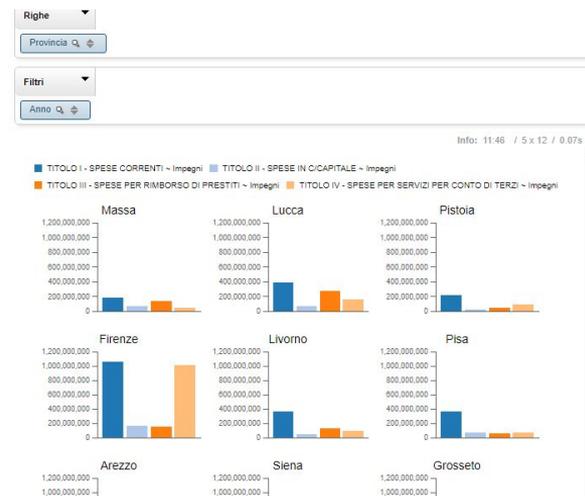
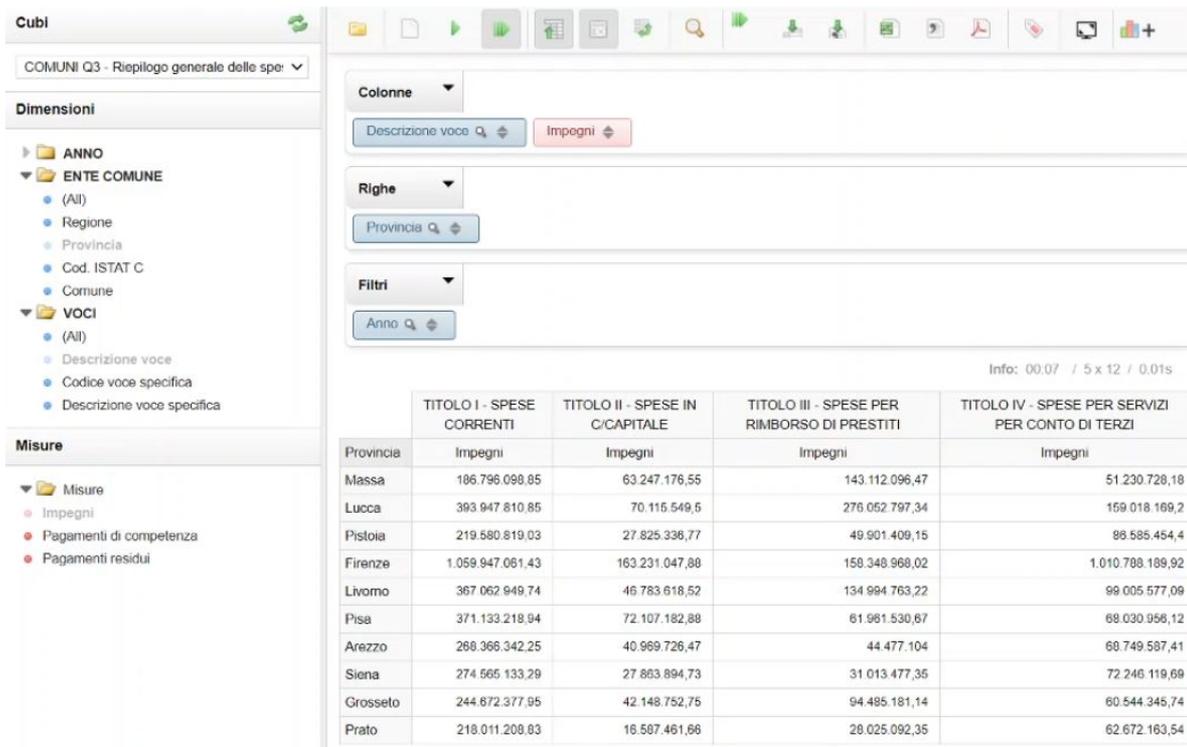


Fonte: www.regione.toscana.it

Osservatorio Regionale Educazione e Istruzione, Osservatorio Regionale Lavoro e Apprendistato, Banca dati Turismo in Toscana.

Diversi strumenti per utenti e scopi informativi differenti

Utenti esperti, ricercatori, utenti della Regione: **analisi interattive per bisogni emergenti**



Fonte: www.regione.toscana.it

Osservatorio Finanza Autonomie Locali, Banca dati Imprese e Occupazione, Demografia in Toscana, Turismo in Toscana.

Dati da diverse fonti e con diversa struttura

Archivi amministrativi

Quotidiana operatività, numerosi, dinamici e disomogenei per formati, accuratezza e completezza.

Monitoraggio e rendicontazione

La loro funzione deriva maggiormente da esigenze di programmare e valutare le attività, se non anche i loro impatti.

Fonti statistiche

Elevata qualità e accuratezza, adeguatezza a supportare analisi di contesto e a valutazioni di medio/lungo periodo.

Archivi interni

Basamenti informativi che fanno capo all'attività operativa giornaliera e variano in base al processo di volta in volta gestito.

Archivi esterni

Disponibili presso altre organizzazioni pubbliche o private, acquisiti tramite accordi o rilasciati come dati aperti.

Dati con una struttura

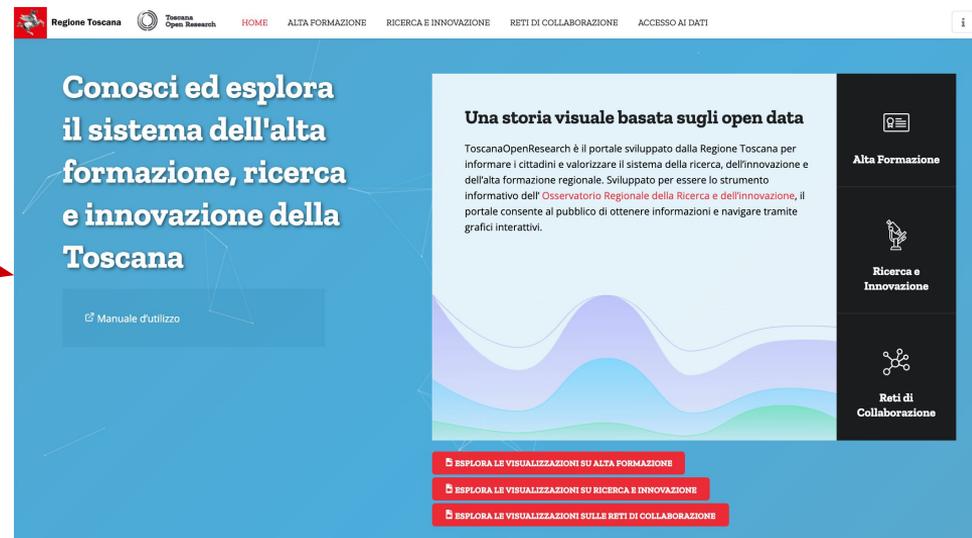
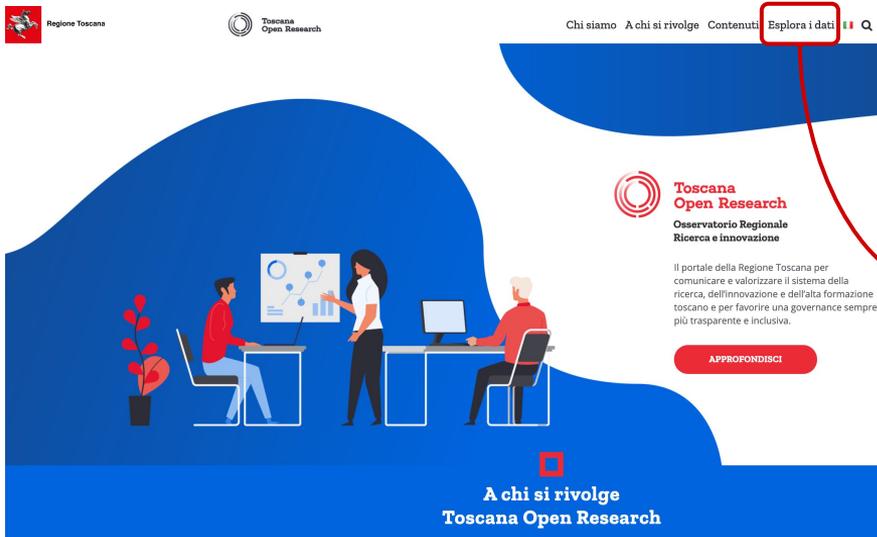
REC_NK_COMUNE	REC_CD_COMUNE_CATAST	REC_DENOMINAZIONE	REC_AMBITO_OTTIMALE	REC_AMBITO_TURISTICO	123 SUPERFICIE_STAT	123 ALTITUDINE
045008	D735	Fosdinovo	28 - Lunigiana	Massa Carrara	48,63	500
046031	L533	Vagli Sotto	27 - Garfagnana	Lucca	41,22	575
046016	E059	Giuncugnano	27 - Garfagnana	Lucca	19,1	876
046016	E059	Giuncugnano	27 - Garfagnana	Lucca	19,1	876
046018	F035	Massarosa	47 - Versilia	Versilia	68,27	10
046018	F035	Massarosa	47 - Versilia	Versilia	68,27	10
047003	B251	Buggiano	44 - Val di Nievole Ovest	M.tecatini T. - Val Nievole	16,04	41
046020	F283	Molazzana	27 - Garfagnana	Lucca	31,33	474
046020	F283	Molazzana	27 - Garfagnana	Lucca	31,33	474
046020	F283	Molazzana	27 - Garfagnana	Lucca	31,33	474
046025	G648	Pieve Fosciana	27 - Garfagnana	Lucca	28,76	369
046026	G882	Porcari	11 - Area Lucchese Est	Lucca	18,05	32

Dati non strutturati



Esplorazione dei dati su Alta Formazione, Ricerca e Innovazione tramite il portale ToscanaOpenResearch (TOR)

Per cittadini, esperti, stakeholders...



Fonte dell'immagine: Portale *ToscanaOpenResearch*:
www.toscanaopenresearch.it

Fonte dell'immagine: Portale *ToscanaOpenResearch* - Sezione "Esplora i dati":
www.toscanaopenresearch.it/esplora-i-dati/#/

L'approccio in ToscanaOpenResearch

Dai dati alla conoscenza: partire dalle domande

Alta Formazione

Ricerca e
Innovazione

Reti di
Collaborazione



Dove si iscrivono gli studenti toscani che frequentano l'università al di fuori della regione Toscana?

Flusso dei residenti in Toscana che frequentano le università con sede fuori dalla regione, filtrabili per settore prevalente, per ateneo e per classi di laurea



Qual è la distribuzione degli immatricolati per classe di laurea, rispetto al diploma di scuola superiore?

Flusso degli immatricolati negli atenei toscani per classe di laurea, rispetto al tipo di diploma di scuola secondaria



Quali sono e che caratteristiche hanno gli spin-off creati in Toscana?

Dettagli sugli spin-off accademici creati in Toscana



Qual è la distribuzione di docenti e ricercatori per età, ruolo e area CUN?

Età media del personale docente in Toscana, per ruolo e per area CUN



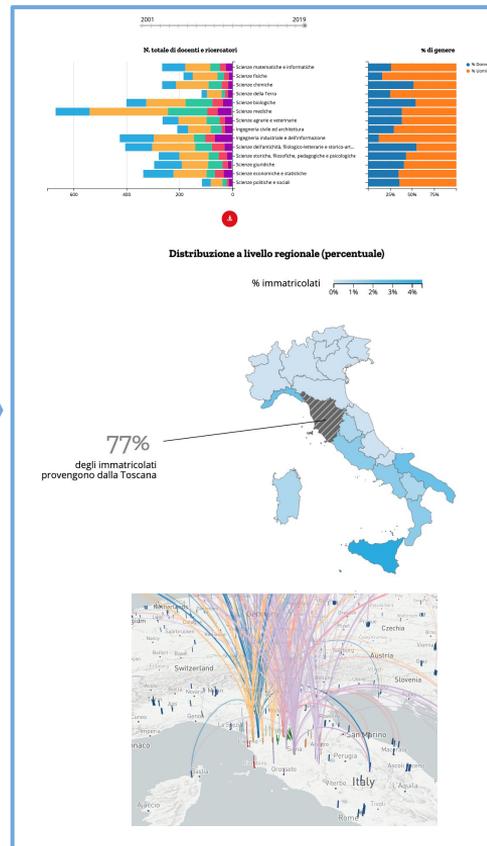
Reti di Collaborazione CTN

Collaborazioni nazionali sui progetti dei Cluster Tecnologici Nazionali



Reti di Collaborazione sui progetti europei

Le collaborazioni degli enti toscani ai progetti di ricerca e innovazione a finanziamento europeo



L'approccio in ToscanaOpenResearch

Fornire strumenti per la lettura e la comprensione dei dati



The screenshot shows the top navigation bar with the ToscanaOpenResearch logo and 'HOME' link. The main content area has a blue background with the text 'Accesso ai Dati' in large white font. Below this, there are three links: 'SPARQL endpoint', 'Documentazione dell'ontologia', and 'Manuale d'utilizzo versione estesa'. At the bottom, a red-bordered box highlights the 'Manuale d'utilizzo' link.



The screenshot shows a document titled 'Guida all'utilizzo di ToscanaOpenResearch'. It contains sections for 'Qual è lo scopo di questa guida?' and 'Cos'è il progetto ToscanaOpenResearch?'. The 'Qual è lo scopo di questa guida?' section includes text about the manual's purpose and contact information. The 'Cos'è il progetto ToscanaOpenResearch?' section describes the project's context and goals. A table of contents on the right lists various topics covered in the guide.

Qual è lo scopo di questa guida?

Questo manuale di utilizzo offre una panoramica delle informazioni contenute nel portale [ToscanaOpenResearch](#) e si propone di presentare ad un pubblico senza competenze specifiche le modalità di consultazione e accesso ai dati sulla Ricerca, Innovazione e Alta Formazione in Toscana in esso reperibili.

Con il fine quindi di accrescere l'utilità e di estendere l'utilizzo di ToscanaOpenResearch, questo documento descrive il progetto che è alla base di questo portale, sottolineandone la ricchezza e varietà dei dati contenuti e illustrando il modo in cui è possibile interrogare il sistema per ottenere le informazioni desiderate.

Per qualsiasi commento o domanda riguardo ToscanaOpenResearch, si può mandare una mail al seguente indirizzo: staff@toscanaopenresearch.it.

Cos'è il progetto ToscanaOpenResearch?

Nel contesto del proprio Osservatorio Regionale della Ricerca e Innovazione, costituito con L.R. 20/2009 al fine di disporre di una base di conoscenza qualificata e di uno strumento informativo di supporto alle politiche, la Regione Toscana ha realizzato ToscanaOpenResearch (TOR) con una duplice finalità:

- Comunicare e valorizzare l'ecosistema dell'alta formazione e ricerca regionale, favorendo una governance sempre più trasparente e inclusiva;
- Analizzare l'ecosistema dell'alta formazione e ricerca regionale per avviare un processo di orientamento delle politiche regionali in un'ottica data-driven.

TOR rappresenta quindi lo strumento informativo dell'Osservatorio, ed è costituito da differenti componenti:

- vari set di dati sulla ricerca, l'innovazione e l'alta formazione: i dati sono presentati in dettaglio nelle sezioni "Quali dati sono disponibili su ToscanaOpenResearch?" e "Che dati possono essere esplorati su TOR?";
- un'ontologia di dominio, presentata nella sezione "L'ontologia TOR";
- un endpoint SPARQL per esplorare e scaricare i dati, presentato nella sezione "Cos'è lo SPARQL Endpoint e come funziona?";
- delle visualizzazioni interattive che rispondono a domande sull'alta formazione, la ricerca e l'innovazione e le collaborazioni di ricerca in Toscana. Le visualizzazioni permettono di esplorare facilmente i dati e di scaricare direttamente i dati visualizzati in formato CSV, e sono presentate nella sezione "Qual è l'origine delle visualizzazioni TOR?".

Qual è lo scopo di questa guida?
Cos'è il progetto ToscanaOpenResearch?
Qual è l'origine delle visualizzazioni TOR?
Quali dati sono disponibili su ToscanaOpenResearch?
Che dati possono essere esplorati su TOR?
Una visione semplificata dei dati tramite mind map
Come si interrogano i dati?
Cos'è lo SPARQL Endpoint e come funziona?
L'ontologia TOR
Come leggere il diagramma dell'ontologia?
Approfondimento - I vantaggi di un accesso ai dati basato su ontologie di dominio
Come scrivere le query per esplorare e scaricare dati?

Immagini tratte dal sito ToscanaOpenResearch: Sezione *Esplora i dati* > *Accesso ai dati*:

www.toscanaopenresearch.it/esplora-i-dati/#/accesso_ai_dati

L'approccio in ToscanaOpenResearch

Fornire strumenti per la lettura e la comprensione dei dati



Regione Toscana Toscana Open Research HOME

Accesso ai Dati

- SPARQL endpoint
- Documentazione dell'ontologia
- Manuale d'utilizzo versione estesa
- Manuale d'utilizzo

Cos'è il progetto ToscanaOpenResearch?

Nel contesto del proprio Osservatorio Regionale della Ricerca e Innovazione, costituito con L.R. 20/2009 al fine di disporre di una base di conoscenza qualificata e di uno strumento informativo di supporto alle politiche, la Regione Toscana ha realizzato ToscanaOpenResearch (TOR) con una duplice finalità:

- Comunicare e valorizzare l'ecosistema dell'alta formazione e ricerca regionale, favorendo una governance sempre più trasparente e inclusiva;
- Analizzare l'ecosistema dell'alta formazione e ricerca regionale per avviare un processo di orientamento delle politiche regionali in un'ottica data-driven.

TOR rappresenta quindi lo strumento informativo dell'Osservatorio, ed è costituito da differenti componenti:

- vari set di dati sulla ricerca, l'innovazione e l'alta formazione; i dati sono presentati in dettaglio nelle sezioni "[Quali dati sono disponibili su ToscanaOpenResearch?](#)" e "[Che dati possono essere esplorati su TOR?](#)";
- un'ontologia di dominio, presentata nella sezione "[L'ontologia TOR](#)";
- un endpoint SPARQL per esplorare e scaricare i dati, presentato nella sezione "[Cos'è lo SPARQL Endpoint e come funziona?](#)";
- delle visualizzazioni interattive che rispondono a domande sull'[alta formazione](#), la [ricerca e l'innovazione](#) e le [collaborazioni](#) di ricerca in Toscana. Le visualizzazioni permettono di esplorare facilmente i dati e di scaricare direttamente i dati visualizzati in formato CSV, e sono presentate nella sezione "[Qual è l'origine delle visualizzazioni TOR?](#)".

La componenti di ToscanaOpenResearch sono presentate schematicamente nella figura in basso:



Schema delle componenti di ToscanaOpenResearch.

Immagini tratte dal sito ToscanaOpenResearch: Sezione *Esplora i dati* > *Accesso ai dati*:

www.toscanaopenresearch.it/esplora-i-dati/#/accesso_ai_dati

L'approccio in ToscanaOpenResearch

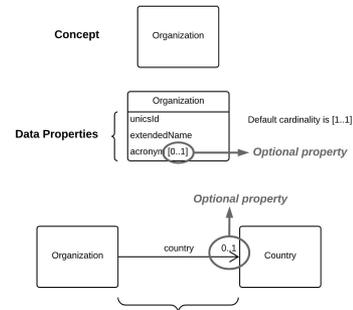
Fornire strumenti per la lettura e la comprensione dei dati

I diagrammi dell'ontologia TOR su Alta Formazione, Ricerca e Innovazione



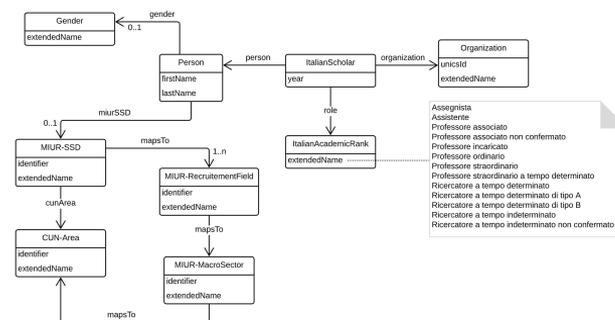
Spiegazione della notazione grafica:

Graphical notation



Esempio di diagrammi dell'ontologia:

MIUR data—Italian scholars (including Assegnisti)

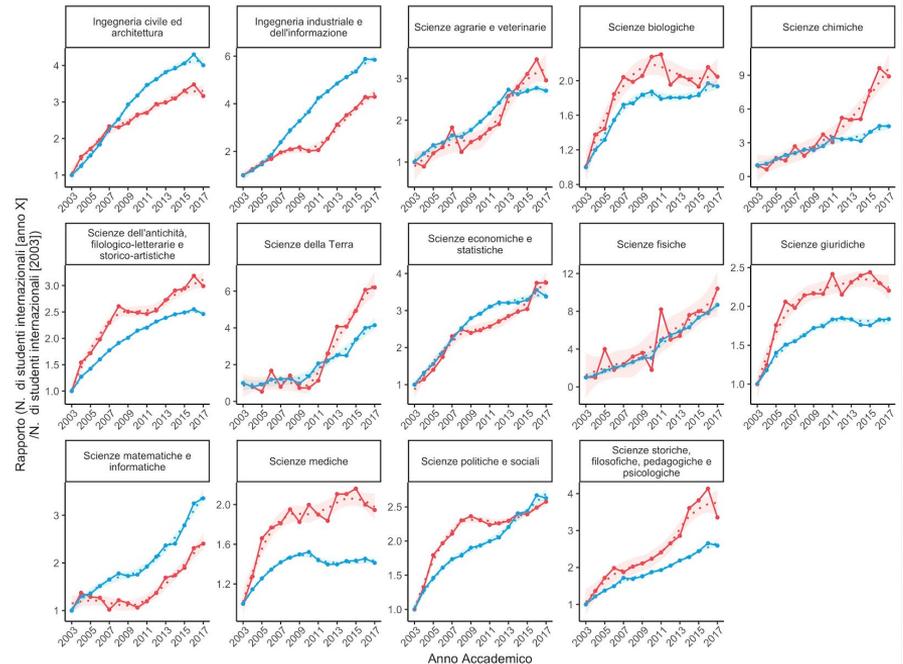


Immagini tratte dal sito ToscanaOpenResearch: Sezione *Esplora i dati* > *Accesso ai dati*:
www.toscanaopenresearch.it/esplora-i-dati/#/accesso-ai-dati

Diversi strumenti di visualizzazione per utenti e scopi informativi differenti

Policy-makers: **documenti interattivi ad-hoc per l'analisi ed il monitoraggio di indicatori consolidati**

- Grafici basati su indicatori con un livello di granularità adatto ai bisogni informativi dei policy-makers
- Combinazione di dati e grafici differenti per restituire una visione completa dei fenomeni di interesse



**“L'obiettivo è quello di trasformare
i dati in informazioni, e
le informazioni in conoscenza.”**

Carly Fiorina, ex CEO di Hewlett-Packard, 2004

“The goal is to turn data into information, and information into insight.”



Grazie per l'attenzione!



Regione Toscana

