

La stima dell'impatto economico delle piattaforme della tecno-scienza

Convegno Fondazione Edison- Accademia dei Lincei

LE PIATTAFORME DELLA TECNO-SCIENZA

Accademia dei Lincei, 12 ottobre 2022

Andrea Bonaccorsi
Università di Pisa

Premessa

Le piattaforme sono forme organizzative

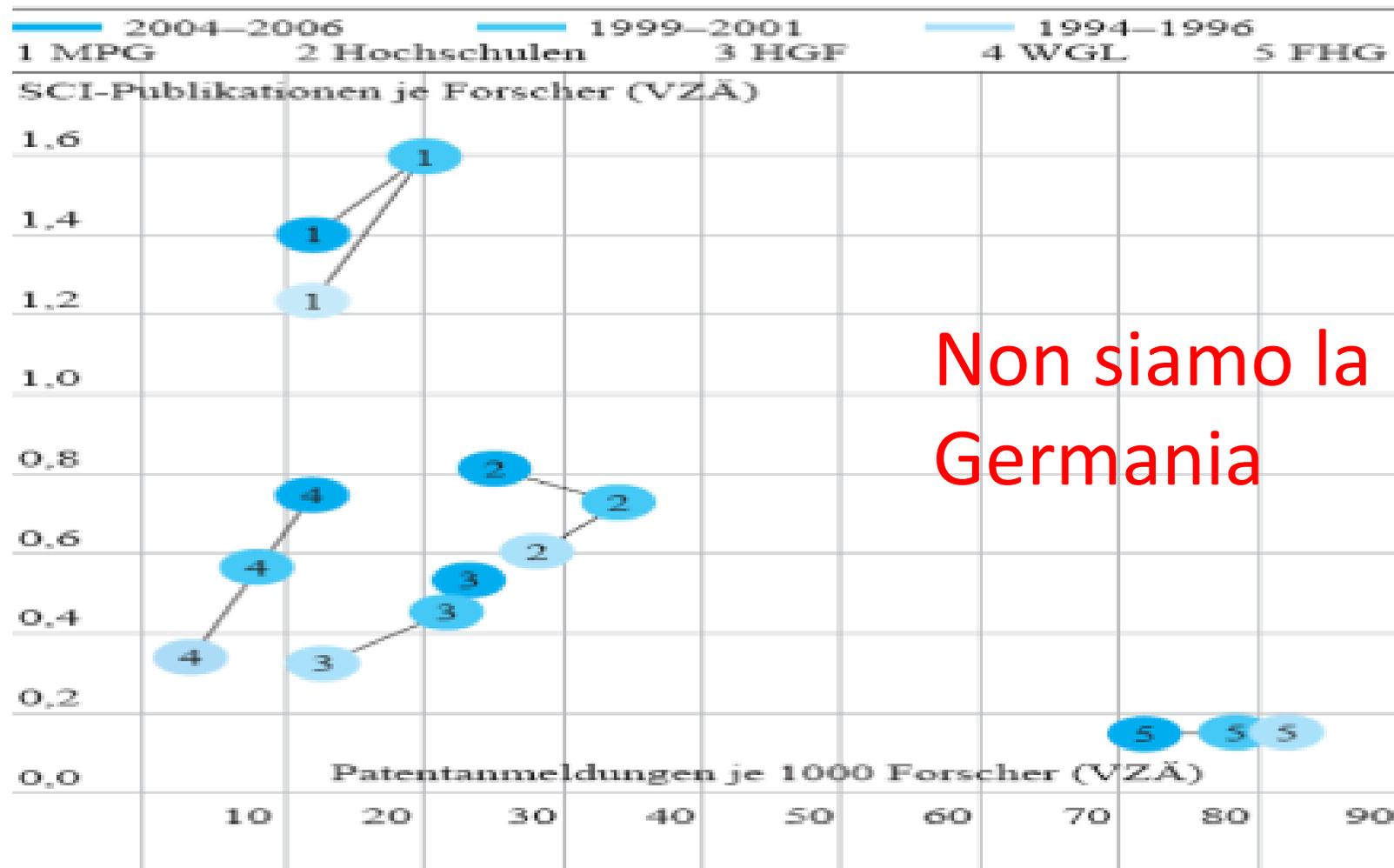
- permanenti
- aperte
- ibride

Per questa ragione l'impatto delle piattaforme tende ad avere caratteristiche peculiari

- impatto a breve ma anche a lungo (o lunghissimo) termine
- effetti non interamente appropriabili (creazione di esternalità positive)
- varietà di forme di impatto
 - stakeholder multipli
 - indicatori differenziati
 - necessità di chiarezza su output e outcome

Veränderung der Publikations- und Patentintensität der großen Wissenschaftsorganisationen und Hochschulen⁶¹

ABB 03



Non siamo la
Germania

MPG, HGF, WGL, FhG: siehe Abkürzungsverzeichnis.

Quellen: Fraunhofer ISI, Auswertung aus SCISEARCH (STN) und PATSTAT (EPO). BMBF (2008). StaBA: Fachserie 14, Reihe 3.6. Fachserie 11, Reihe 4.4. Fachserie 22, Reihe 4.5. OECD (2009a). Berechnungen des ZEW.

Principali approcci alla misurazione dell'impatto della ricerca pubblica

Stima impatto monetario

- Stima della spesa e metodo del moltiplicatore
- Cost-benefit analysis
- Impatto degli acquisti ad alta tecnologia
- Stima degli spillover

Stima impatto multidimensionale

- Stima dell'impatto settoriale
- Metodi basati su casi
- Counterfactual impact evaluation (CIE)

Metodi basati su Intelligenza Artificiale

- Tracking dei flussi di conoscenza
- Mappatura dei beneficiari

Stima impatto monetario

- Stima della spesa e metodo del moltiplicatore
- Cost-benefit analysis
- Impatto degli acquisti ad alta tecnologia
- Stima degli spillover

Stima della spesa e metodo del moltiplicatore

Stima dell'impatto complessivo = **spesa diretta totale * moltiplicatore «keynesiano»**

Metodo largamente adottato nelle università USA, soprattutto nel settore pubblico (State University) per rendicontare al governo e alla opinione pubblica (Elliott, Levin and Meisel, 1988; Kott, 1987-88; Beeson and Montgomery, 1993; Blackwell, Cobb and Weinberg, 2002).

Prime applicazioni per università italiane (IRPET)

Direct expenditures of students and faculty

- expenditure of students and faculty who move into the area from other regions;
- expenditures by the institution;
- funds for research, salaries, and other expenditures (equipment, infrastructure, supplies).

Long term direct and indirect effects

- enhancement of workers' skills;
- relationship between research and local industry;
- positive effects on business location (attraction of foreign investment);
- business creation (creation of new firms and technologies).

Stima della spesa e metodo del moltiplicatore

Metodo adottato anche per la stima di impatto della spesa per beni culturali.

Problemi

- **eterogeneità** nella definizione delle spese da inserire nel calcolo
- scelta di **localizzazione** di difficile interpretazione (mobilità degli studenti, pendolarismo, localizzazione personale)
- **allocazione delle spese alla singola istituzione** spesso problematica (es. costi fissi di infrastrutture, accesso condiviso a infrastrutture e attività)
- definizione convenzionale o arbitraria del **moltiplicatore**
- ignorati gli effetti di **equilibrio economico generale**

Cost-benefit analysis

Metodo adottato per la stima di impatto del CERN (Florio, Forte e Sirtori, 2015; 2016; Florio, 2019)

$$NPV = \sum_i \frac{(SC_{t_i} + TE_{t_i} + HC_{t_i} + CU_{t_i}) - (K_{t_i} + LS_{t_i} + LO_{t_i} + O_{t_i})}{(1+r)^{t_i}} + EXV_0.$$

Benefits

- Knowledge output (scientific publications)
- Technological externalities
- Human capital formation
- Cultural effects

Stima di valore economico basata su stato dell'arte dei metodi di economia pubblica (*culture economics, economics of education, environmental economics*)

Problemi metodologici:

- **robustezza** rispetto alle assunzioni e alla definizione delle proxy.

Impatto degli acquisti ad alta tecnologia

Primi esperimenti nel contesto Science of Science (USA) e al CERN

Approccio USA (Goldschlag, Lane et al. 2018)

- alleanza di università USA per la condivisione di dati
- liste nominative di fornitori di tecnologie avanzate delle università (apparecchiature scientifiche, materiali, componenti, software)
- fornitori delle università confrontati con un campione di controllo di non-fornitori

Approccio CERN ((Bianchi-Streit et al. 1984; Autio et al. 2003)

- questionario rivolto a un sottoinsieme selezionato di fornitori

Ipotesi

- impatto dell'università misurato in termini di aumento della produttività delle imprese
- **effetti confermati** dalla analisi : (a) sofisticazione della domanda dei ricercatori vs. domanda media di mercato; (b) lead users e sviluppo di nuovi prodotti e/o nuove linee di business; (c) attivazione o aumento della spesa in R&D; (d) effetti reputazionali.

Problemi

- **possibile errore nel campione di controllo** (= non fornitori di altre università in altri Stati USA o paesi esteri).

Stima degli effetti di spillover

Effetto di miglioramento delle performance delle imprese **localizzate nelle vicinanze** delle università grazie a spillover da attività di ricerca

Principali effetti:

- creazione di capitale umano (educazione terziaria nelle università)
- produzione di nuova conoscenza (ricerca)

Variabili indipendenti

- numero di laureati per disciplina
- numero di pubblicazioni/ citazioni per disciplina
- qualità delle pubblicazioni (top 10 o 25% riviste scientifiche, top 10 o 25% citazioni)

Variabili dipendenti

- **natalità delle imprese in Italia** (Bonaccorsi et al. 2013; 2014; Guerini et al. 2022)
- **performance delle imprese in Europa** (crescita di fatturato, occupati, totale attivo, totale attivo immateriale) (Bonaccorsi et al. 2019; 2022)

Fonte dei dati

- ETER per microdati sulle università
- Scopus per pubblicazioni
- ORBIS per le imprese

Stima di spillover

Georeferenziazione delle unità

- imprese
- università (dati ETER georeferenziati)

Costruzione di «anelli» geografici di distanza crescente dal centroide della impresa

- 0-10 km
- 11-25
- 26-50
- 51-100
- Oltre 100 km

Sommatoria di tutti gli input (= università) in ciascuno degli anelli come misura degli spillover complessivi
Stima di impatto sulla crescita delle variabili di performance per le imprese in funzione dello spillover totale delle università localizzate a diverse distanze.

Problemi

- Analisi cross-section (= no causalità)
- Endogeneità

Table 1. Regression results for models of impact of higher education on firm's growth. Total variation 2011-2015. Separate estimation of education and research impacts.

	(1)	(2)	(3)	(4)
	Turnover	Employment	Total assets	Intangible assets
Panel A. Number of graduates (ISCED 5-7)				
Graduates within 10km	0.011*** (0.002)	-0.008*** (0.001)	0.003* (0.001)	0.020*** (0.004)
Graduates between 10km and 20km	0.011*** (0.001)	0.002* (0.001)	-0.002* (0.001)	-0.001 (0.003)
Graduates between 20km and 50km	0.012*** (0.002)	0.001 (0.001)	-0.000 (0.001)	-0.007 (0.004)
Graduates between 50km and 100km	0.006*** (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.002 (0.003)
N	1,385,535	1,385,535	1,385,535	1,385,535
R ²	0.2122	0.1696	0.1698	0.1549
Panel B. Total number of citations received				
Citations within 10km	0.019*** (0.003)	-0.009*** (0.002)	0.008*** (0.002)	0.044*** (0.007)
Citations between 10km and 20km	0.009*** (0.001)	0.001 (0.001)	-0.001 (0.001)	0.004 (0.004)
Citations between 20km and 50km	0.008*** (0.001)	0.001 (0.001)	0.000 (0.001)	0.002 (0.005)
Citations between 50km and 100km	0.006*** (0.001)	-0.000 (0.001)	-0.001 (0.001)	0.002 (0.004)
N	1,334,040	1,334,040	1,334,040	1,334,040
R ²	0.2124	0.1703	0.1701	0.1567

Note: Standard errors in parentheses. Dependent variables are measured as the log growth of each variable between 2011-2015. Independent variables are standardized. Controls included (size, sector, legal form, region, initial values of dependent variables, R&D intensity, quality of institutions).

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Stima impatto multidimensionale

- Stima dell'impatto settoriale
- Metodi basati su casi
- Counterfactual impact evaluation (CIE)

Metodi di stima dell'impatto su outcome settoriali

Metodi largamente adottati in ricerca biomedica (Lomas, 1997; Lavis et al. 2002; Kothari, Birch and Charles, 2005; Kuruvilla et al. 2007; Newson et al. 2015; Cohen et al. 2015).

- Clinical Guidelines
 - Documenti autorevoli e di adozione generale
 - Citazioni selezionate alla letteratura scientifica



Tracking the impact of research on policy and practice: investigating the feasibility of using citations in clinical guidelines for research evaluation

David Kryl,¹ Liz Allen,² Kevin Dolby,² Beverley Sherbon,³ Ian Viney³

Recenti sviluppi con dataset basati su social media

- Overton (archivio di documenti di policy)

A large-scale validation of the relationship between cross-disciplinary research and its uptake in policy-related documents, using the novel Overton altmetrics database

Henrique Pinheiro , Etienne Vignola-Gagné , and David Campbell 

Science-Metrix, an Elsevier company, Montreal, Quebec, Canada

Metodi basati su casi

Premesse di metodo

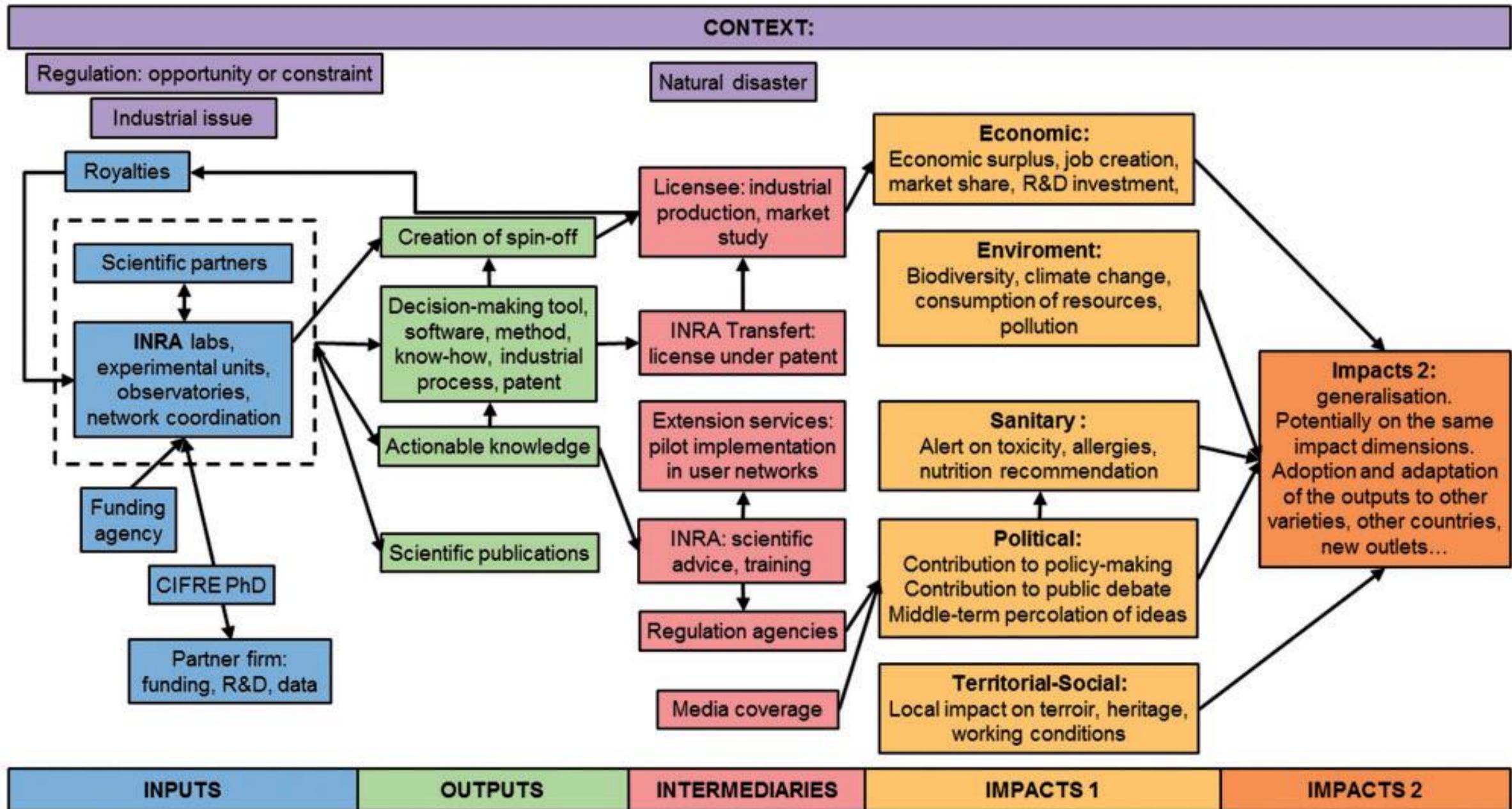
- Eterogeneità delle attività di ricerca
- Impatto multidimensionale
- Diversità delle scale temporali
- Diversi stakeholder possono descrivere l'impatto con categorie non commensurabili
 - Economico
 - Politico
 - Sociale-territoriale
 - Ambientale
 - Salute

Esperienza INRA (modello ASIRPA) (Joly et al.)

- Effetti diretti vs indiretti
- Ruolo di organizzazioni intermedie
- Standardizzazione dei casi di studio
- Rappresentazione multidimensionale (radar)

Limiti

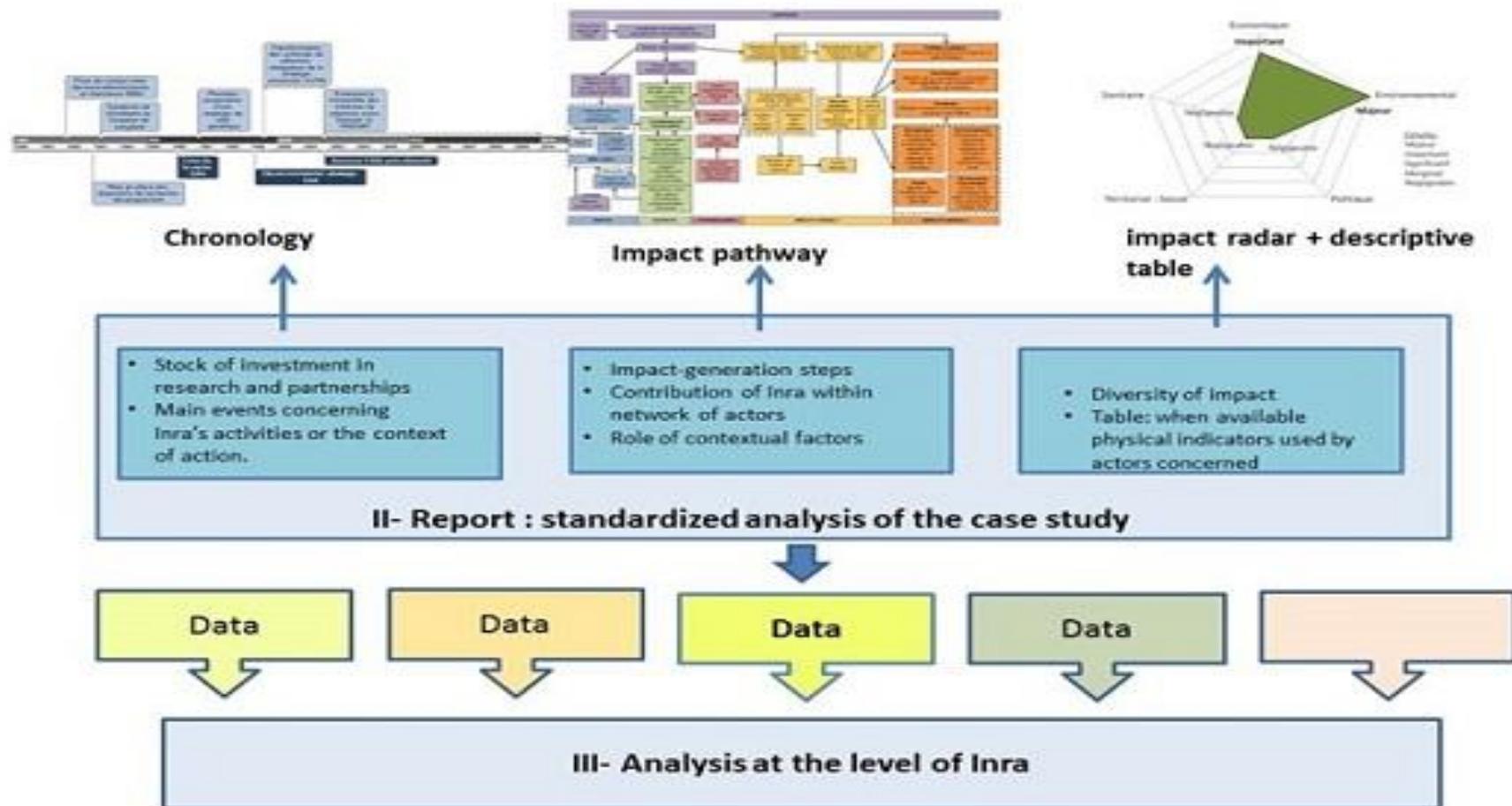
- Assenza di benchmark di riferimento
- Standardizzazione limitata dei casi di studio e analisi descrittiva



Source: INRA

I- Case selection: criteria

- Significant case- diversity – reasonably recent (publication less than 15 years)



Counterfactual Impact Evaluation (CIE)

Applicazione al tema della valutazione di impatto economico della ricerca pubblica del metodo controfattuale

- policy maker applica una variante del principio di addizionalità
- Regression Discontinuity Design (RDD): identificazione di una variabile esogena (evento storico, regola di selezione) rispetto alla quale sia possibile confrontare due campioni di unità statisticamente simili rispetto alla variabile di *outcome*
 - *Ammessi e non ammessi in una graduatoria di selezione per progetti di ricerca*
 - *Prima e dopo un cambiamento di legislazione*
 - *Prima e dopo la creazione di una nuova istituzione*
- **variabile di discontinuità= data di creazione della università o della piattaforma**

Basi teoriche

- **Teoria della causalità (controfattuale)** come esperimento attivo, ovvero confronto tra trattati e non trattati
- Esperimenti (quasi) naturali
- Tecniche di costruzione dei campioni di non trattati

Counterfactual Impact Evaluation (CIE)/2

Esempi

- In a large-scale study covering more than 15,000 universities in 78 countries Valero and Van Reenen (2019), considering the 1950-2010 time interval, estimate that **doubling the number of universities per capita generated an increase in income per capita of 4% in subsequent years**
- A recent study exploits the opportunity offered by the creation of Universities of Applied Sciences in Switzerland to carry out counterfactual analysis and finds that **the creation of a new HEI has a strong impact on the regional economy: 8.5 to 14 % increase in regional patent activity, and 2 to 3.6 % increase in citation per patent** (Pfister et al., 2017).

Problemi

- **utilizzo di variabili iper-semplificate** (variabile binaria: presenza/assenza di università ad ogni data della serie storica)
- data constraint: **assenza/ indisponibilità di serie storiche lunghe con dati più «ricchi»** sugli input della ricerca (spesa, addetti, infrastrutture) e sugli output (pubblicazioni, brevetti, altri indicatori di impatto)
- inapplicabilità di molte delle tecniche di CIE.

Counterfactual Impact Evaluation (CIE)

Applicazione del metodo controfattuale alle infrastrutture di ricerca

- Data di inizio attività
 - raccolta documentale
 - piattaforma IRFS
 - testare più modelli in funzione della «finestra» di operatività della infrastruttura
- Definizione del perimetro geografico di impatto
 - aspetti istituzionali (regole di accesso, collaborazioni internazionali)
 - econometria spaziale
 - funzione endogena (exponential decay)
 - anelli di dimensione crescente intorno al centroide della infrastruttura
- Definizione delle variabili di outcome
 - **problema di identificazione causale** (es. impatto sui risultati di ricerca in termini di pubblicazioni, brevetti, spinoff, imprenditorialità + formazione capitale umano + altri effetti)

Metodi basati su Intelligenza Artificiale

- Tracking dei flussi di conoscenza
- Mappatura dei beneficiari

Premesse di metodo

- **L'impatto della ricerca «lascia sempre una traccia» in elementi testuali**
- Il formato dei testi potenzialmente rilevanti è tuttavia eterogeneo (non solo pubblicazioni e brevetti ma documenti ufficiali, news, social media, report e letteratura grigia)
- I dati sono in generale non strutturati
- Utilizzo di metodi avanzati di Natural Language Processing (NLP) per la collezione, il processamento e l'analisi di testi.

Tracking dei flussi di conoscenza

Creazione di **link tra dati eterogenei**

- progetti di ricerca
- pubblicazioni
- brevetti
- autorizzazioni amministrative
(es. AIC farmaci)

Link creato attraverso parole chiave (con sviluppo di queries articolate per minimizzare falsi positivi e falsi negativi) + controlli di omonimia e di affiliazione

Lunghe serie storiche di dati (minimo 10-20 anni)

Stima dei tempi di transizione tra le fasi

- «demografia» delle idee: entrata (novelty), sopravvivenza (survival), uscita (exit)- e possibile ripresa in tempi successivi
- tempistica dettata da procedure autorizzative vs. libera

Valore economico dei nuovi prodotti sul mercato

Paper-patent citation linkages as early signs for predicting delayed recognized knowledge: Macro and micro evidence

Jian Du^{a,b}, Peixin Li^b, Robin Haunschild^c, Yinan Sun^b, Xiaoli Tang^{b,*}

OPEN ACCESS Freely available online

 PLOS ONE

Tracking 20 Years of Compound-to-Target Output from Literature and Patents

Christopher Southan¹, Peter Varkonyi², Kiran Boppana³, Sarma A. R. P. Jagarlapudi³, Sorel Muresan^{2,4*}



The credibility of research impact statements: A new analysis of REF with Semantic Hypergraphs

Andrea Bonaccorsi,* Nicola Melluso, Filippo Chiarello, and Gualtiero Fantoni

Scientometrics
<https://doi.org/10.1007/s11192-020-03803-z>

Mappatura dei beneficiari (user groups)

Impact for whom? Mapping the users of public research with lexicon-based text mining

Andrea Bonaccorsi¹  · Filippo Chiarello¹ · Gualtiero Fantoni¹

Research Evaluation, 2021, 1–21
doi: 10.1093/reseval/rwab008
Article

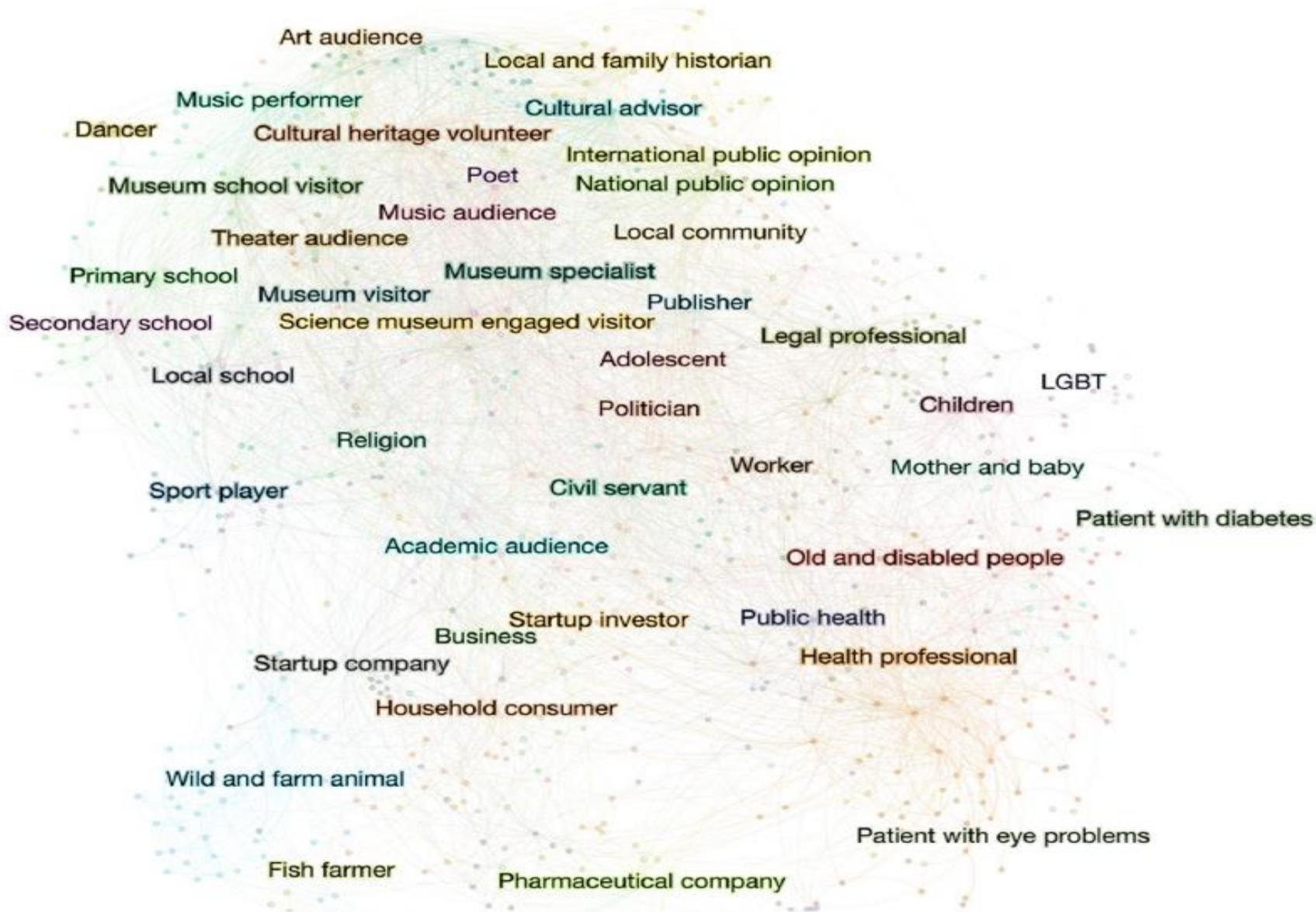


SSH researchers make an impact differently. Looking at public research from the perspective of users

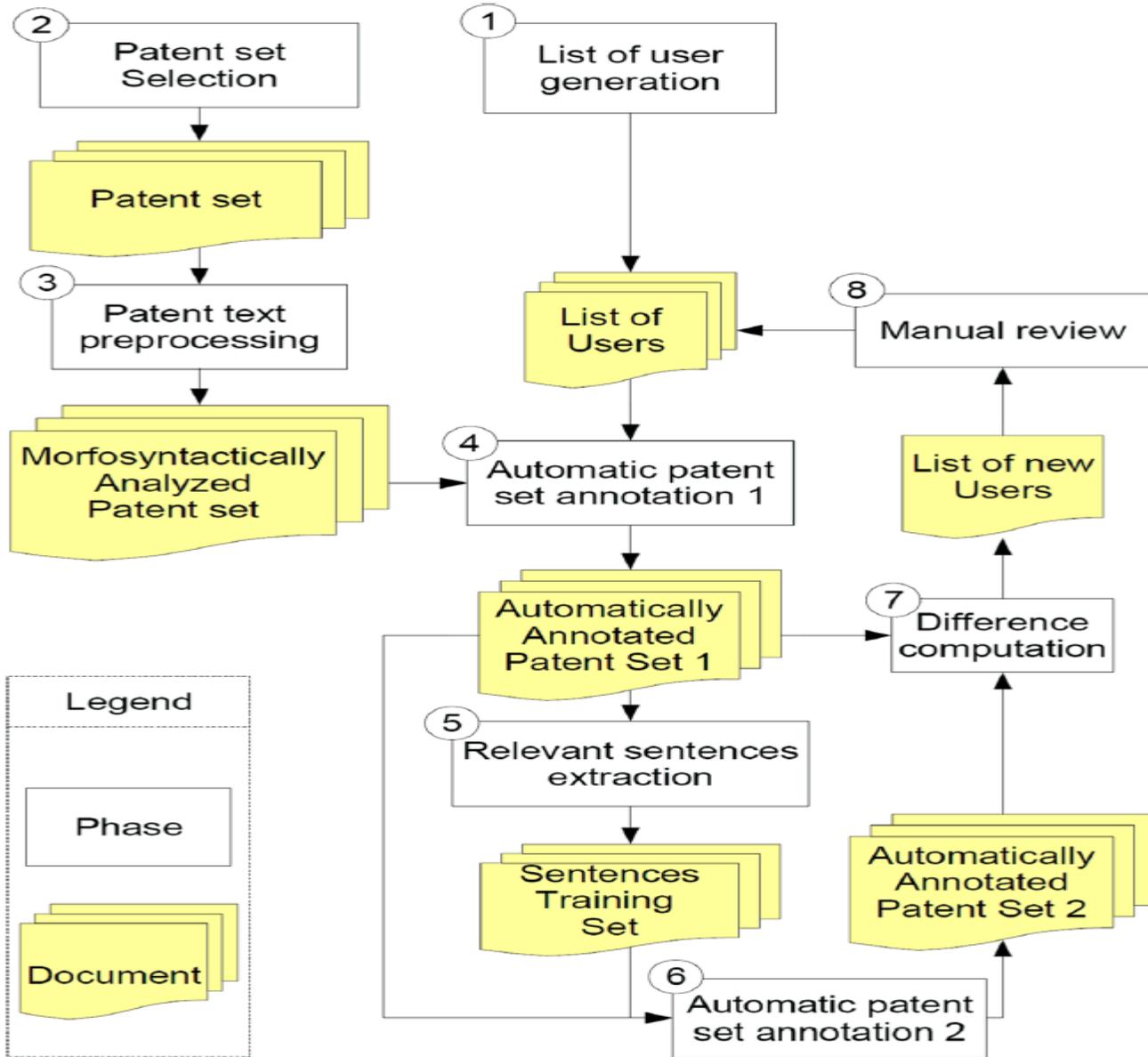
Andrea Bonaccorsi ^{1,*}, Filippo Chiarello¹ and Gualtiero Fantoni²

Table 3 Clusters of user groups of research of UK universities

No.	Cluster	Main words
1	Cultural heritage volunteer	Volunteer, wider society, citizen science, heritage organization, community centre
2	Secondary school	Secondary school, secondary school student, high school, Teacher, school leader
3	Media producer	Producer, guest, media professional, academic consultant, production company
4	Church community	Leader, church leader, international community, base company, fabian society
5	Civil society	Civil society, citizen, international organization, non-governmental organization, public service
6	Public health	Public health, European Society of Cardiology, Kcl researcher, elderly, hospital
7	Publisher	Publisher, everyone, anyone, many people, ordinary people
8	Museum visitor	Unique visitor, university student, register user, unique visitor per month
9	Museum school visitor	School child, visitor per year, school engagement, local museum
10	Music audience	Listener, non academic audience, ancestor, someone, non specialist audience
11	Worker	Worker, workforce, trade union, health worker, migrant worker
12	Art audience	New audience, diverse audience, global audience, total audience, art organisation
13	Health professional	Health professional, health care professional, nurse, doctor, physician
14	National public opinion	Public audience, wide public audience, average audience
15	Academic audience	Professor, lecturer, senior lecturer
16	Civil servant	Civil servant, senior manager, senior civil servant, specialist adviser, voluntary organization
17	Theater audience	Actor, reviewer, audience feedback, theater company, assistant director
18	International public opinion	International audience, general audience, large audience, public, lay audience
19	Local/family historian	Historian, archivist, family historian, local historian, librarian
20	Wild and farm animal	Animal, farmer, healthy animal, owner, bird
21	Business	Trade, enterprise, social enterprise, international trade, entrepreneur
22	Local school	Local school, wider community, involved teacher
23	Startup company	Engineer, designer, phd student, wider community, startup company
24	Startup investor	Investor, inventor, stem cell, skilled people, private investor
25	Children	Social worker, protected child, psychologist, psychiatrist, adolescent
26	Adolescent	Friend, youth, girl, teenager, academic community
27	Politician	Politician, economist, observer, public policy, commentator



Fine-grained map
of social users of
university-based
research in UK



What does «lexicon approach» mean in NLP?

The corpus of text is processed using a comprehensive list of words that «saturate» a semantic field.

No. entries = 76,857

Named Entity Recognition (NER)

Supervised (top-down) learning approach.

Conclusioni

Punto di partenza è definizione delle **dimensioni rilevanti dell'impatto**

Sviluppo di **indicatori appropriati**

- fonti primarie/secondarie
- granularità

I metodi esistono!

- trade-off (gestibile) tra quantificazione e eterogeneità
- metodi di AI consentono di ampliare la base di dati utilizzabili

La «**rivoluzione dell'impatto**» proseguirà

Grazie dell'attenzione

andrea.bonaccorsi@unipi.it
a.bonaccorsi@gmail.com