

# I numeri (non) parlano da soli

*Il ruolo dei grandi network che producono dati in campo educativo*

**Orazio Giancola e Assunta Viteritti**

*Dipartimento di Scienze Economiche e Sociali "Sapienza"*

## 1. Obiettivi, metodi e storia del paper

Il paper intende analizzare come si sono formati i due macro network che nel mondo si occupano di produrre dati in campo educativo (IEA e OECD). Sarà analizzato come le diverse configurazioni, articolazioni, estensioni, artefatti, conoscenze dei due grandi network contribuiscono a creare dati che divengono patrimonio di conoscenza sul piano pubblico, visibili e utilizzabili da media, governi, lobby groups e istituzioni. Dal punto di vista metodologico il paper utilizza l'analisi della letteratura, l'analisi documentale e l'analisi dei siti dei due grandi network.

Il paper prosegue un filone di studi che gli autori stanno conducendo da alcuni anni (Giancola e Viteritti 2014; Giancola e Viteritti 2015). Contributi su questo tema sono stati presentati dagli autori a recenti convegni internazionali<sup>1</sup>. In questo paper lo studio si estende all'analisi e alla comparazione dei due grandi network che producono dati in educazione sul piano internazionale.

## 2. Ruolo dei big data nella vita sociale

Sul piano più generale i numeri/dati/big-data diventano fenomeni politici sotto altra forma, configurandosi come fenomeni collettivi strutturanti della stessa vita sociale. I dati si vanno affermando come "fatti naturali" una *seconda natura del sociale* (Savage 2013) una nuova *base sociale* (Giddens 1990; Ciborra 2008).

Le nostre vite sono in continua relazione con i dati e tramite i dati si possono produrre visioni dei fenomeni sociali sotto forma di rappresentazioni e flussi informativi.

---

<sup>1</sup> (Orazio Giancola e Assunta Viteritti) "Data for what? Converging and diverging trajectories of Large Scale Assessment in Education: OECD versus IEA. Paper presentato a ECER 2016 "Leading Education: The Distinct Contributions of Educational Research and Researchers" will take place at University College Dublin from 22-26 August, 2016. (Orazio Giancola e Assunta Viteritti) "The Life and Death of Big data in Education". Paper presentato a 4S|AESST|Conferenze|Barcellona 2016 - SCIENCE AND TECHNOLOGY BY OTHER MEANS - Exploring collectives, spaces and futures - August 31 - September 3.

I dati costruiscono relazioni, reti e effetti sociali e fanno parte di un più ampio quadro di mutamento che riguarda ciò che in letteratura viene definito come *datificazione* dei fenomeni sociali (Davenport, Barth, Bean 2012). Necessitano peraltro di essere conservati, spostati, comunicati, classificati, manipolati e interpretati.

I dati sono divenuti elementi relazionali presenti in molti ambiti di vita. La rilevanza e significatività dei grandi dati è oggi esteso a tutti i fenomeni sociali, pensiamo ad es. alla proliferazione dei dati sui tanti fenomeni quali la povertà, l'andamento dell'economia, il ruolo della finanza, la distribuzione della ricchezza, la disoccupazione, la salute, le malattie, l'immigrazione, la scienza e la tecnologia, ecc.

I grandi dati – soprattutto con l'ausilio delle risorse del web, dei grandi server e degli algoritmi, complessi sistemi di calcolo che trasformano i dati da stock a flussi – sono divenuti una componente centrale viva del modo di rappresentare, visualizzare, valutare e costruire il sociale causando una radicale trasformazione nella produzione della conoscenza e del suo ruolo nelle società contemporanee (Mazzotti 2015; Giancola e Viteritti 2015).

### **3. Big data in Education**

Nella ricerca educativa per tutto il secolo passato l'educazione è stata dominata dalle "*big theories*" (funzionalismo, le stratificazioni sociali, il capitale umano, ecc.) mentre nella parte finale del secolo e nell'inizio del nuovo (a partire dagli anni novanta) la riflessione si sposta sul ruolo dei "*big numbers*", sulle quantità, sui numeri come forma di analisi e di spiegazione sociale.

I "big data" in educazione paiono prendere il posto delle teorie e parlare al loro posto, spesso parlano al posto delle politiche e oggi pare vincere il nuovo mito dell'"evidence based education".

La rilevanza dello *spazio internazionale dell'educazione* è cresciuto nel corso degli anni grazie anche alle molte infrastrutture di dati che producono grandi indagini internazionali (Large Surveys) e che sono diventati i "nuovi attori socio-tecnici" in Italia, in Europa e a livello internazionale in grado di fornire dati a governi, politiche, opinioni pubbliche e attori sociali.

Sul piano internazionale partire dagli anni sessanta del secolo scorso (prima IEA e poi OCDE) le grandi indagini producono dati molto ampi e articolati che influenzano e condizionano visioni, culture e scelte in campo educativo.

I dati in educazione (e non solo) diventano, come affermava Latour, “oggetti di valore”, co-costruiscono mondi sociali e interpretazioni (Latour 2007): anche in educazione siamo nel mondo della *datificazione*.

I due più importanti network internazionali, IEA e OCDE, diventano negli anni i più importanti *actor-net* capaci di produrre, collezionare, disseminare, assemblare e analizzare i dati educativi a livello internazionale con ripercussioni e capacità di influenzamento anche a livello nazionale e locale.

Il valore dei dati in educazione nasce sul piano internazionale a partire dal dopoguerra, fase in cui l'educazione inizia a diventare per i paesi industrializzati (in Europa ma non solo) “capitale umano”: l'educazione diventa un fattore per movimentare la crescita sociale e economica, un valore per le economie dei paesi in fase di rilancio e di sviluppo dopo la guerra.

Dalla fine del ventesimo secolo i dati (attraverso l'OCSE, la Commissione europea, la IEA e le agenzie nazionali) sono diventati quindi strumenti di governo in materia di istruzione: esprimono mondi, valori, suggeriscono scelte politiche, incorporano conoscenza tecnica interdisciplinare; diventano indicatori di efficienza dei sistemi scolastici; vengono assunti come *dati di fatto* piuttosto che come *elementi di interesse da discutere e dibattere* (Latour 2004).

Le grandi indagini mobilitano considerevoli risorse economiche dei paesi che partecipano, complessi network di esperti e fanno crescere set di investigazione sociale. Peraltro, in varie forme, nelle grandi indagini sono interessati centinaia di migliaia di persone (insegnanti, studenti, dirigenti scolastici e famiglie) che sono coinvolti nell'indagine attraverso centinaia di migliaia di questionari. Allo stesso tempo sono coinvolti un enorme numero di esperti e ricercatori di differenti discipline (includendo l'education, la pedagogia, l'information technology, l'economia, la sociologia).

Questi ampi volumi di dati sono processati, immagazzinati, analizzati e usati per classificare individui e paesi, per mostrare trend, similarità e per sottolineare differenze tra e dentro i paesi.

Le Large Scale Survey mobilitano:

- ✓ considerevoli risorse economiche dei paesi partecipanti
- ✓ complessi network di esperti (in espansione e differenziazione)
- ✓ un crescente set di investigazioni e metodi
- ✓ una pluralità di utilizzatori esperti dei dati (ricercatori, accademici, analisti politici media, giornalisti e opinione pubblica)

- ✓ la moltiplicazione di interessi economici e politici (a livello internazionale, nazionale e locale).

#### **4. I due più importanti macro-network internazionali analizzati**

In educazione i due più importanti network nel mondo sono IEA and OECD. Questi due macro network producono LSA (Large Scale Assessment). I LSA producono valori, etichette, concezioni, alleati e nemici. I due grandi network internazionali OCDE e IEA rendono disponibili a governi, scuole, esperti, accademia, dati, misure e comparazioni sul piano nazionale e internazionale in campo educativo. Questi due macroattori si formano dal dopoguerra, con la fine dei conflitti mondiali, con l'affermarsi della forza degli stati nazionali e di organismi internazionali (quali l'Unesco e l'ONU). Nasce in questa fase l'esigenza della comparazione, della modernizzazione dei sistemi educativi dopo il piano Marshall, il piano politico-economici statunitense per la ricostruzione dell'Europa dopo la seconda guerra mondiale.

Il network IEA sviluppa le importanti LSA come PIRLS (che si occupa di reading literacy), TIMSS (che si occupa di matematica, scienze) e CIVED (che si occupa di civicness).

Il network OCDE è emerso a partire degli anni novanta dello scorso secolo e ora produce varie grandi survey tra cui: PISA (che si occupa di matematica, scienze e reading) e PIACC (che si occupa di adult literacy e numeracy).

Vediamo ora velocemente le caratteristiche del network IEA. Nato negli anni 50 dello scorso secolo il network IEA ha realizzato diverse indagini *test based*. IEA include molte grandi indagini tra cui: PIRL e TIMSS. Il network IEA include molti accademici e ricercatori di molti paesi e le indagini realizzate sono orientate soprattutto alle questioni curriculari. IEA è diventata un ente legale nel 1967 ma la sua origine data nel 1958 quando un gruppo di studiosi, soprattutto psicologi, pedagogisti dell'educazione, psicometrismi e sociologi si incontrano all'UNESCO Institute for Education ad Hamburg, in Germania per discutere i problemi della scuola e della valutazione degli studenti dopo la guerra. IEA (International Association for the Evaluation of Educational Achievement) realizza diverse indagini internazionali tra cui TIMSS (Trends in Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study) e ICCS (International Civic and Citizenship Education Study, indagine incentrata sulle conoscenze civiche) sono invece condotte dalla IEA (International Association for the Evaluation of Educational Achievement). Attualmente aderiscono alla IEA 53 paesi in tutto il mondo. Il gruppo di

ricerca comprende psicologi, psicometristi, sociologi ed esperti dell'educazione provenienti per lo più da università tedesche che, americane e del nord Europa. Lo studio PIRLS (Progress in International Reading Literacy Study) è realizzato in quasi 60 paesi ha come principale obiettivo la valutazione comparativa dell'abilità di lettura dei bambini al quarto anno di scolarità. L'indagine iniziata nel 2001 viene realizzata con cadenza quinquennale. L'indagine TIMSS (Trends in International Mathematics and Science Study) analizza il rendimento degli studenti in Matematica e Scienze in oltre 60 Paesi. L'indagine misura la performance degli studenti alla IV classe della scuola primaria e alla III secondaria di I grado.

L'OCSE è un'organizzazione sovranazionale, costituita a Parigi nel 1961. Supporta i governi a far fronte alle sfide economiche, sociali e ambientali poste dall'economia mondiale. Le indagini OCDE in campo educativo sono l'esito di "reti lunghe" a partire dagli anni sessanta sotto la spinta dell'UNESCO. E' oggi sostenuto da più di 70 paesi che partecipano e finanziano le ricerche in campo educativo insieme a organismi internazionali quali la Banca Mondiale.

Vediamo ora in sintesi il network OCDE. Il network si è molto esteso negli anni e include un ampio raggio di discipline e di esperti che con le loro indagini riescono oggi a influenzare governi, politiche scolastiche (a vari livelli). OCDE produce in campo educativo tre grandi survey: PISA (math, science e reading a cui ha incluso nelle ultime edizioni anche "financial literacy"); PIACC (che si occupa di adult literacy numeracy) e TALIS (che si occupa di teachers head-teachers practices). In particolare PISA (Programme for International Student Assessment) è la principale indagine sui temi educativi internazionale dell'OCSE. Ha lo scopo di valutare, con periodicità triennale, il livello di istruzione degli adolescenti dei principali paesi industrializzati negli ambiti di alfabetizzazione letteraria, matematica e scientifica. L'indagine è effettuata in più di 60 paesi ogni 3 anni. Le prove hanno come destinatari fra 4.500 e 10.000 allievi in ogni paese. PIAAC (*Programme for the International Assessment of Adult Competencies*) è il Programma OCSE, a cui hanno partecipato 24 paesi nel mondo. L'indagine ha lo scopo di conoscere attraverso un questionario e dei test cognitivi specifici le abilità fondamentali della popolazione adulta compresa tra i 16 e i 65 anni, ovvero quelle competenze ritenute indispensabili per partecipare attivamente alla vita sociale ed economica nel 21esimo secolo.

## **5. Le infrastrutture digitali su cui viaggiano i dati, differenze della forma dei dati tra i due macro-network**

I dati prodotti dai due macro-network sono disponibili sulle infrastrutture digitali dedicate, sul sito di IEA e OCDE. I database su cui viaggiano i dati dei due grandi network sono infrastrutture informative complesse e in continuo allestimento che: richiedono l'uso e l'assemblaggio di conoscenze complesse e eterogenee; necessitano dell'alleanza di molteplici discipline e esperti professionali (informatici, economisti, ingegneri, sociologi, pedagogisti, ecc.).

Con riferimento ai due siti analizzati quello IEA si presenta esclusivamente come un repository, una vetrina di dati, mentre il sito OCDE (con riferimento a PISA come caso esemplare) si presenta come una infrastruttura relazionale. Con riferimento alla forma dei dati i due network presentano sia dati aperti e accessibili (per ricercatori e scuole come il caso di PISA OCDE) che dati meno aperti e a disposizione soprattutto degli esperti (come nel caso delle indagini IEA). Se infatti si volessero i dati IEA degli anni novanta si avrebbero grosse difficoltà. I database IEA sono infatti legati solo sul piano tematico ma non sul piano dei dati, hanno una tematica comune ma non hanno una metrica (modelli di misurazione) comune. Le diverse indagini IEA utilizzano framework e indicatori differenti di conseguenza una vecchia indagine non è collegabile alle precedenti e successive: l'aggancio intertemporale tra le varie indagini è quindi molto problematica. Per usare i dati IEA bisogna scaricare un software per uso esperto che poi si collega a SPSS che poi si aggancia a IEA (siamo davanti alla logica del software proprietario e chiuso). I dati OCDE si possono invece avere in formato OPEN e si possono aprire con qualsiasi software. Per es. PISA e TIMS ti danno i database che si possono aprire con qualunque software e i database producono delle misurazioni "che hanno la pretesa di essere confrontabili nel tempo". In sintesi si può dire che i dati IEA sono dati chiusi, dati proprietari e non sono confrontabili nel tempo. Mentre i dati OCDE sono dati relazionali, i dati IEA sono dati isolati e statici, ogni indagine è chiusa e risolta in se stessa e non si collega alle altre. Nel mondo OCDE infatti anche se il dato è vecchio consente confronti nel tempo. Nel mondo IEA i dati diventano presto obsoleti, nessuno (ad es.) cita i dati delle indagini TIMS del 90, il dato vecchio diventa obsoleto e isolato, non è più utilizzabile (IEA) e diventa inerte. C'è una chiusura autoreferenziale dei dati, la comunità degli utenti potenziali è limitata. I dati OCDE (es. PISA) possono essere sempre attuali e vivi perché confrontabili, consentono la narrazione di una storia longitudinale, i dati hanno un ciclo di vita lungo. Il dato "resta in

vita” fin quando è usato o usabili attraverso pubblicazioni, conferenze, l’alleanza con gli utenti (ricercatori-scholars, data journalist, media, policy makers ecc). Gli indicatori OCDE vivono una seconda vita anche perché sono messi dentro il report Educational at Glance, un rapporto sintetico e divulgativo che semplifica i dati prodotti dalle indagini e li mette in circolazione tra i meno esperti. Si tratta di un grande oggetto di frontiera e di mediazione che da un lato si lega ai dati prodotti dalle indagini e dall’altro raggiunge lo spazio pubblico creando dibattito culturale.

I report IEA sono solo rapporti tecnici, disponibili solo per esperti e confinati allo spazio accademico. Non sono scaricabili on line e anche se sono gratuiti necessitano, per essere scaricati di una serie di passaggi sulla piattaforma che sono solo per i più esperti.

## **6. La differente forza dei due network**

Le diverse configurazioni, articolazioni, estensioni, saperi, dei due macro-attori contribuiscono a co-costruire:

- ✓ dati più *vivi* (dati aperti e pubblici, visibili e utilizzabili da governments, lobby groups and institutions around the world) come il caso dei dati OCDE
- ✓ dati più *inerti e obsoleti* (dati prodotti in reti chiuse accademiche che non divengono oggetti pubblici) come il caso dei dati IEA.

Facendo una analisi delle citazioni dei dati dei diversi network si vede che le ricerche OCDE hanno più citazioni delle indagini IEA. Le indagini OCDE hanno più lettori, raggiungono più utilizzatori, sono più visibili nei media e sono una componente del dibattito pubblico. In relazione ai tipi di dati che viaggiano nelle infrastrutture IEA E OCDE emergono due tipi di uso dei dati. Nel caso delle indagini IEA i dati sono utilizzabili solo come *stock* fissi, frutto di procedure e standardizzazioni progressive (come il caso delle LSA IEA). In questo caso I dati sono inerti, immobili, possono essere solo consultati. Nel caso delle indagini OCDE (per esempio l’indagine PISA) i dati disponibili sul web possono essere usati come flussi: dati comparativi ed intertemporali/in serie storica. I dati possono essere assemblati, adattati, sono versatili, sono disponibili sotto varie forme.

## 7. Ruolo sociale dei dati e rischi della datificazione

La produzione dei big data è cambiata nel tempo: da sapere proprio di comunità chiuse di esperti (statistici, economisti, pedagogisti, sociologi) è diventato uno strumento per l'azione pubblica, un campo sociale di decisioni e oggi in letteratura si parla di governo tramite i numeri (Grek, Ozka, 2010).

Nel caso delle indagini OCDE i dati hanno una ciclicità, e rimangono vivi e disponibili nelle mani degli utenti esperti; gli indicatori delle varie indagini sono tra loro comparabili; hanno una estesa diffusione tra esperti e policy makers; sono disponibili in molti network eterogenei (politici, professionisti, esperti, utenti); si tratta di dati open, hanno una maggiore presenza dei dati nel dibattito pubblico dei vari paesi; hanno una maggiore influenza sulle politiche educative nazionali e internazionali; sono dati vivi, hanno una ampia circolazione; sono tradotti e usati in molte reti eterogenee; producono effetti di policy, sono oggetti di valore che influenzano agende politiche e scelte educative; sono il frutto di una elevata standardizzazione di metodi, strumenti e risultati; sono diventati un attore politico.

Nel caso delle indagini IEA si tratta di dati inerti, con una certa dose di obsolescenza poichè hanno ciclicità piuttosto ridotta; i dati sono prodotti e diffusi solo dentro network accademici e hanno una circolazione limitata tra esperti. I dati IEA sono prodotti e circolano solo dentro reti accademiche e reti esperte, ogni indagine introduce innovazioni di metodo, di campionamento e hanno una minore standardizzazione.

Le reti lunghe e eterogenee (politici, esperti, scuole, famiglie, studenti, ecc) consentono che i dati abbiano maggiore riconoscibilità e circolazione. Le reti fatte solo di esperti accademici producono dati che hanno una minore circolazione e visibilità.

Le dimensioni della visibilità e vitalità dei big data in educazione sono:

- ✓ maggiore standardizzazione dei metodi;
- ✓ comparabilità dei risultati e dei metodi;
- ✓ maggiore visibilità pubblica dei dati, dei report e dei dibattiti;
- ✓ ampio numero di partecipanti e users (esperti e meno esperti);
- ✓ circolazione dei dati all'interno di reti eterogenee tra loro interconnesse;
- ✓ presenza di reti lunghe (dai progettisti delle indagini agli utilizzatori finali dei dati)
- ✓ molti oggetti materiali che veicolano e traducono i dati (report, libri, media, devices, software libero, ecc.);
- ✓ disponibilità dei dati su piattaforme relazionali;



✓ i dati come oggetti di dibattito pubblico tra esperti, policy makers, operatori scolastici.

Ma quali i rischi che producono i grandi dati? Vediamoli sinteticamente.

- I macro-network (come OCSE e IEA) sono diventati (in modo diverso) agente fondamentali nel processo di europeizzazione nella formazione: essi sono un potente attori-rete, dotati di grande potere sul piano dei significati, sui piani linguistici, culturali e dei valori nonché in termini politici
- La datificazione si configura come un processo egemonico nello spazio educativo: o spazio internazionale di istruzione si sta costruendo principalmente attraverso la concatenazione di grandi banche dati, di tecnologie digitali e di strumenti di visualizzazione dei dati. Quindi con una pervasività dilagante sui sistemi di istruzione, a livello nazionale e locale
- La massiccia presenza e l'influenza dei dati depoliticizza e tecnicizza il campo educativo, con un forte accento sull'efficienza, la massimizzazione delle risorse limitate, la creazione e l'uso di standard (spesso assunti come "dati per scontato")
- I dati provenienti dal LSA nella percezione degli attori sociali diventano "entità reali", grandi infrastrutture socio-tecniche che agiscono sia a livello rappresentazioni sociali e sulle pratiche politiche.

## **8. I rischi della semplificazione**

La domanda di strumenti di misura in campo educativo è stata alimentata sia dal basso (i decisori politici) che dall'alto (OECD, UE, UNESCO ecc.), trovando diverse organizzazioni pronte a fornirli. La differenza tra successo e l'insuccesso in questo caso l'ha fatta l'apparato tecnologico e i veicoli di diffusione utilizzati. Riguardo l'apparato tecnologico, l'organizzazione OECD-PISA per prima e a differenza delle altre (tra tutte IEA) ha investito nella realizzazione di una poderosa e complessa infrastruttura informatica e si è avvalsa della collaborazione di nuove figure come data scientists. La combinazione di questi fattori ha permesso la creazione di numerosi indici (tra cui ESCS, *Index of Economic, Social and Cultural Status / Indicatore dello status socio-economico e culturale degli studenti*)<sup>2</sup>, strumenti di misura basati su moderni e complessi algoritmi, mediante i quali sono state estrapolate numerose informazioni dai dati raccolti nelle grandi indagini internazionali

---

<sup>2</sup> Dal punto di vista tecnico-statistico, l'aggregazione degli indicatori semplici per arrivare alla formulazione di indicatori compositi, si fa utilizzando tecniche multivariate di riduzione dei dati. Nel caso del calcolo dell'indice ESCS viene utilizzata la tecnica ACP (analisi in componenti principali).

condotte. Questo comunque non sarebbe stato sufficiente, se non fosse stato affiancato da un sistema di divulgazione altrettanto esteso ed efficace.

Gli strumenti prodotti dalle indagini OECD-PISA sono adottati come riferimento non soltanto in tema di politiche pubbliche, ma questa loro capacità di occupare sempre più spazi di *mercato* ha fatto sì che anche le altre organizzazioni, dopo qualche tempo, si adeguassero adottando simili strumenti di misura e indici. Questa è la vera forza che caratterizza l'organizzazione OECD-PISA.

La spinta all'isomorfismo ha portato alla produzione di indici basati su artifici (divulgativi, teorici, algoritmici) che hanno portato ad una semplificazione da molti ritenuta come inopportuna, eccessiva, non adeguata (Bottani, 2016) a poter condurre delle valutazioni tenendo in adeguata considerazione altri elementi del contesto geografico, politico, economico e sociale dei paesi oggetto di indagine (Giancola, 2016).

L'utilizzo di tali indicatori in aspetti riguardanti l'amministrazione pubblica sta portando avanti il modello di quella che è stata definita con il concetto di cultura algoritmica (Striphas, 2015), ovvero *il ricorso a processi computazionali per ordinare, classificare, e gerarchizzare persone, luoghi, oggetti e idee*. Pratiche di significazione della realtà mediante linguaggio computazionale, espresso sotto forma di complessi algoritmi matematici.

La complessità di questi algoritmi, di questi *oggetti sociomateriali* come sono gli indici prodotti da OECD-PISA, deve essere esplicitata ed indagata dagli scienziati sociali. Prendiamo ad esempio l'indice ESCS

ebbene questa tipologia di strumenti non deve essere utilizzata come se avesse a che vedere solo con algoritmi e automazione. Quello che non vediamo è il ruolo essenziale che hanno le persone che progettano questi strumenti (Mayer-Schönberger, et al., 2013). L'opinione più diffusa è che le metodologie su larga scala possono essere considerate come neutre, questo può essere in parte vero, a patto però di riconoscere il lato umano dei dati, la loro arbitrarietà, la loro costruzione sociale ed il loro essere "carichi di teoria" (Giancola, Viteritti 2015).

C'è un ulteriore aspetto che solitamente non è considerato e riguarda la presenza dei cosiddetti metadati<sup>3</sup>. Il database di PISA in realtà fornisce molte più informazioni rispetto a quelle *mostrate* nei report ufficiali e divulgativi. Tutti i dati che servono alla costruzione degli indici delle indagini come OECD-PISA in origine contengono queste *meta informazioni*, ma una volta prodotti gli indici finali queste non sono più accessibili. La

---

<sup>3</sup> Dati che forniscono informazioni su altri dati.

possibilità di avere accesso a informazioni di dettaglio come queste costituirebbe un prezioso aiuto al fine di avere una comprensione maggiore dei fenomeni indagati e rendere più raffinati i modelli di studio.

Una delle possibilità è che, i dati raccolti dagli algoritmi statistico-matematici come nel caso dell'indice ESCS, producano una rappresentazione della realtà molto semplificata. Questo ad esempio può essere dovuto alla scelta teorica o al costo computazionale che richiedono le elaborazioni. Un rischio, quello di semplificare il processo di rappresentazione di una realtà estremamente più complessa e variegata che, in altri termini, porta a trascurare quegli elementi (variabili e particolari) caratterizzati da proprietà di tipo qualitativo e per questa ragione non facilmente adattabili ad un modello matematico di analisi. Un rischio ulteriore riguarda coloro che fanno riferimento ai dati prodotti da OECD-PISA con l'indice ESCS, nel momento in cui andranno ad utilizzare questi dati, prodotti mediante l'uso di elaborazioni algoritmiche e con tecniche statistiche (come ACP), è come se andassero ad aderire a quelle che sono le forme di categorizzazione determinate dagli scienziati OECD in sede di progettazione teorica. In questo modo, non ci si limita alla riproduzione di questi modelli di indagine, ma si accoglie la loro visione, si trasforma il proprio approccio da neutrale ad arbitrario, perché le scelte prodotte conterranno al loro interno le categorie di classificazione stabilite da OECD.

## **9. Alcune osservazioni conclusive**

La produzione dei big data è cambiata nel tempo: da sapere proprio di comunità chiuse di esperti (statistici, economisti, pedagogisti, sociologi) è diventato uno strumento per l'azione pubblica, un campo sociale di decisioni e oggi in letteratura si parla di governo tramite i numeri (Grek 2009).

Le reti lunghe e eterogenee a cui si legano i dati (politici, esperti, scuole, famiglie, studenti, ecc) consentono che i dati abbiano maggiore riconoscibilità e circolazione. Le reti fatte solo di esperti accademici producono dati che hanno una minore circolazione e visibilità.

*Ma quali i rischi che producono i grandi dati? Vediamoli sinteticamente.* I due macro-network sono diventati agenti di europeizzazione in educazione. Sono degli attori potenti in senso linguistico e culturale. La *datificazione* è diventato un processo egemonico nello spazio educativo. La pervasività dei dati investe i sistemi educativi a livello nazionale e locale. il successo globale e l'influenza che l'indagine PISA ha ottenuto sulle politiche

nazionali, fornendo un modello prototipo di politica educativa «*evidence based*» (Meyer, Benavot, 2013), ha permesso all'OECD di espandere il suo campo d'azione all'analisi globale delle abilità (*skills*) dei lavoratori attraverso l'indagine PIAAC. Con PIAAC, l'OECD si va a collocare in una posizione di spiccata egemonia in relazione alla valutazione delle abilità e dell'educazione di livello superiore, simile a quella che PISA ora detiene rispetto alla valutazione internazionale comparativa delle *performance* dei sistemi scolastici. Questo processo di valutazione sta ricreando e rafforzando il lavoro nell'educazione dell'OECD: esso ha sviluppato «macchine» impressionanti per produrre i dati internazionali comparativi e questi dati sono stati ampiamente utilizzati nelle arene politiche di molti paesi. In molti Paesi infatti i dati comparativi nazionali e internazionali vengono utilizzati come basi per le *policy high-stakes* e per riformare le decisioni. I processi di comparazione non sono disinteressati e oggettivi, come non lo sono gli strumenti e i dispositivi metodologici costruiti e usati in seno all'OECD (o allo IEA) e ciò che viene dichiarato come verità scientifica sono in effetti ideologie tecniche che producono conseguenze performanti sui sistemi sociali.

La massiva presenza dei dati sembra depoliticizzare e tecnicizzare il campo educativo. Il governo tramite i numeri enfatizza aspetti legati all'efficienza dei sistemi ponendo l'enfasi sulle misure, gli standard (Gorur, 2012; Lawn, 2011; Timmermans e Epstein 2010). Le misure, gli standard e la *evidence based view*, sono assunti come *dati per scontato*. I dati che arrivano dalle grandi survey sono percepiti come attori sociali, come entità reali. Le due grandi infrastrutture socio-tecniche agiscono sul piano delle rappresentazioni a livello politico e sociale.

La questione dell'impiego delle grandi fonti di dati («conoscitivi», ma in fin dei conti «valutativi») nell'ambito della ricerca sociale, e nello specifico nell'ambito dell'educazione, come si è visto fino ad ora, ha portato ad un approccio quasi totalmente distale della valutazione e della misurazione dei sistemi d'istruzione nel mondo. Solo a partire dal riconoscimento della non-neutralità degli strumenti di valutazione e dell'uso dei dati come strumenti “oggettivi” di analisi se ne può valorizzare il potenziale conoscitivo anche per problematizzare l'effetto “classificatorio” dei numeri e il loro uso solo apparentemente depoliticizzante.

## Referenze

- Borer V. L., Lawn M., (2013) *Governing Education Systems by Shaping Data: from the past to the present, from national to international perspectives*, European Educational Research Journal, Volume 12, Number 1
- Ciborra C. (2008) *I labirinti dell'informazione. Sfida alla sapienza dei sistemi*. Palermo, Sellerio
- Davenport, T. H., Barth P., Bean R. (2012), "How 'Big Data' Is Different." in MIT Sloan Management Review, n.54.
- Desrosières, A. (2002) *The Politics of Large Numbers: A History of Statistical Reasoning* (translated by Camille Naish), Cambridge, Mass: Cambridge University Press.
- Giancola O., Viteritti A. (2014) "Distal and Proximal Vision: a multi-perspective research in sociology of education", in *European Educational Research Journal*, Volume 13 Number 1 2014, pp.47-57
- Giancola O., Viteritti A. (2015) "Il ruolo delle grandi survey in campo educativo. L'indagine PISA e il governo dell'educazione tramite i numeri", n.3-4, pp. 555-580, in (a cura di) Federico Neresini e Ota De Leonardis, *Il potere dei grandi numeri*, Numero Monografico Rassegna Italiana di Sociologia
- Giancola O., (2015), "Il nuovo scenario delle politiche educative: tra valutazione, quasi-mercato e l'emergere di nuovi attori" in "Neoliberismi e azione pubblica. Il caso italiano", in Moini G. (a cura di), Edizioni Ediesse, Roma
- Giancola O. (2016), "Piccoli e grandi numeri nel mondo dell'education. Una riflessione su somiglianze e differenze tra il voto scolastico e il testing su larga scala" in *Scuola Democratica* N.1/2016
- Giddens A. (1990) *The Consequences of Modernity*, Stanford University Press
- Gillespie T. (2014) "The relevance of Algorithms" in (edited by) Tarleton Gillespie, Pablo J Boczkowski, Kirsten A. Foot *Media Technologies. Essay on Communication, Materiality and Society*, MIT, Cambridge.
- Gorur R. (2014) "Towards a Sociology of Measurement in Education Policy", in *European Educational Research Journal* Volume 13 Number 1 pp. 58-72
- Gorur, R. (2011) "ANT on the PISA Trail: following the statistical pursuit of certainty", in *Educational Philosophy & Theory*, 43(5-6), 76-93.
- Gorur, R. (2012) "The Invisible Infrastructure of Standards", in *Critical Studies in Education*, 54(2), 132-142.

- Grek S., 2009, "Governing by numbers: the PISA 'effect' in Europe", in *Journal of educational policy*, Vol.24, No.1, 2009, Routledge
- Grek S., Ozga J. (2010), "Governing Education through Data: Scotland, England and the European Education Policy Space", in *British Education Research Journal* 36 (6) 937-952
- Korsnakova P., (2013) "Large-scale Comparative Studies of Educational Achievement: an  
Latour, Bruno *La fabbrica del diritto: etnografia del Consiglio di stato*. Città aperta, 2007.
- Law, J. (1992), "Notes On the Theory of the Actor Network: Ordering, Strategy and Heterogeneity", *Systems Practice* 5, pp. 379-393.
- Lawn, M., (2011) "Standardizing the European Education Policy Space", in *European Educational Research Journal*, 10(2), p.259.
- Lawn M., (2013) "Voyages of Measurement in Education in the Twentieth Century: experts, tools and centres", in *European Educational Research Journal*, Volume 12, Number 1
- Lawn M., Grek S. (2013) *Europeanizing Education. Governing a new policy space*, Symposium Books Ltd
- Massimo Mazzotti, Per una sociologia degli algoritmi, in "Rassegna Italiana di Sociologia" 3-4/2015, pp. 465-478, doi: 10.1423/81801
- Mayer-Schönberger V, Cukier K. (2012) *Big Data: A Revolution That Transforms How we Work, Live, and Think*, Houghton Mifflin Harcourt
- Meyer H.D., Benavot A., (2013), "PISA and the Globalization of Education Governance: some puzzles and problems", in Meyer H.D., Benavot A., 2013 (eds.) *PISA, Power, and Policy the emergence of global educational governance*, Oxford, Oxford Studies in Comparative Education - Symposium Books
- Savage M. (2013) "The 'Social Life of Methods': A Critical Introduction" in *Theory Culture and Society* 30: 4, pp. 3–21. Saetnan A.R.,
- Saetnam A.R., Lomell H.M, Hammer S., 2010 (eds.) *The Mutual Construction of Statistics and Society*, Routledge
- Striphas, T. (2015) "Algorithmic culture" in *European Journal of Cultural Studies*. 2015, 18(4-5) 395
- Timmermans, S. & Epstein, S., (2010) "A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization", in *Annual Review of Sociology*, 36(1), pp.69 - 89.