

LA SFIDA DELLA COMPLESSITÀ NELL'ERA DEI BIG DATA: STRATEGIA E TATTICA PER UN APPROCCIO MULTIDISCIPLINARE ALL'ISTITUTO SUPERIORE DI SANITÀ



Cecilia Bossa, Roberta Valentina Gagliardi, Alessandro Giuliani, Olga Tcheremenskaia, Igor Branchi, Barbara Caccia, Andrea Galluzzi, Guido Gigante, Maurizio Mattia, Antonio Pazienti, Evaristo Cisbani, Carla Daniele, Giuseppe D'Avenio, Giuseppe Esposito, Daniele Giansanti, Sandra Morelli, Alessandra Palma, Cecilia Poli, Francesco Facchiano, Gianluca Frustagli, Loredana Le Pera, Elisabetta Pizzi, Irene Ruspantini, Ornella Moro, Orietta Picconi, Sabrina Tait
per il Gruppo Sistemi Complessi e *Data Science* dell'ISS
Istituto Superiore di Sanità

RIASSUNTO - Per superare la mancanza di avanzamenti nella scienza fondamentale della ricerca odierna, che sembra produrre solo avanzamenti tecnologici, si ha bisogno di integrare diversi punti di vista scientifici. Tale processo d'integrazione sta diventando una realtà in molte parti del mondo scientifico. L'approccio multidisciplinare, rispetto alla soluzione di problemi complessi, è alla radice dell'esistenza dell'Istituto Superiore di Sanità (ISS), ente di ricerca in cui lavorano insieme, ad esempio, medici e biologi impegnati nell'epidemiologia statistica e nell'analisi multidimensionale dei dati, nonché fisici che partecipano a progetti di neuroscienze. Da questa realtà, multidisciplinare per nascita e sviluppo, trae origine il Gruppo di lavoro trasversale e multidisciplinare "Gruppo Sistemi Complessi e *Data Science*" dell'ISS.

Parole chiave: sistemi complessi; *Data Science*; multidisciplinarietà

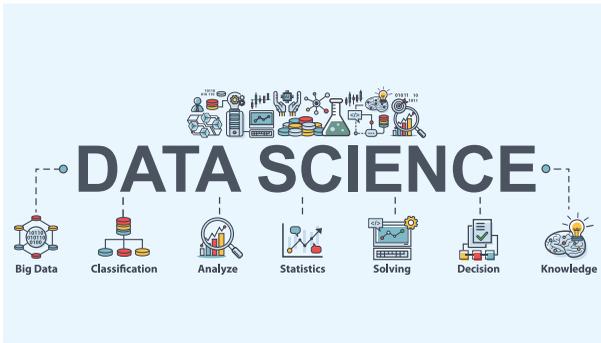
SUMMARY (*The challenge of complexity in the big data era: strategy and tactic for a multidisciplinary approach to the Istituto Superiore di Sanità - the National Institute of Health in Italy*) - We need to integrate different scientific viewpoints to overcome the lack of true innovation as for basic science, being today's research only focused on technological progress. This integration process is becoming a reality in several scientific fields. The multidisciplinary approach to solving complex problems is at the base of the very existence of the Istituto Superiore di Sanità, where physicians and biologists work together on statistical epidemiology and multidimensional data analysis, as well as physicists participating in neuroscience projects. From this multidisciplinary reality in both birth and development, comes this transversal and multidisciplinary working group: ISS Complex Systems and Data Science Group.

Key words: complex systems; *Data Science*; multidisciplinarity

barbara.caccia@iss.it - alessandro.giuliani@iss.it

Nel corso del 2022 sugli *Annali dell'Istituto Superiore di Sanità* (ISS) è stato pubblicato un Editoriale dal titolo "The challenge of complexity in the Big Data era: how to ride the wave of high-dimensional data revolution" (1). La pubblicazione è stata curata da un gruppo di ricercatori e ricercatrici dell'ISS che operano in diversi settori di ricerca dell'ISS, ma che hanno una comune visione del lavoro scientifico che si colloca sotto la dizione "Sistemi complessi e *Data Science*".

Affrontare un problema scientifico complesso restando ancorati al punto di vista della propria disciplina di appartenenza può compromettere il raggiungimento di una soluzione soddisfacente. Per superare la mancanza di avanzamento nella scienza fondamentale (o di base) del lavoro di ricerca odierno (2), che sembra produrre solo avanzamenti tecnologici, è necessaria l'integrazione di diverse scienze. Questo processo di integrazione è in atto (3) e i ricercatori/ricercatrici dell'ISS ne fanno parte. In questo ►



contesto, la rivitalizzazione dell'antica tradizione di interdisciplinarietà dell'ISS è una delle ragioni che ha favorito la creazione di questo gruppo dove è possibile incontrare medici, biologi e bioinformatici (impegnati nell'epidemiologia statistica e nell'analisi multidimensionale dei dati), fisici che partecipano a progetti di neuroscienze o diagnostica per immagini e ingegneri coinvolti negli aspetti regolatori dei dispositivi medici.

Per tale motivo, negli ultimi anni, in altri contesti scientifici internazionali si è avuta la stessa necessità che ha portato alla nascita di gruppi interdisciplinari. Basti citare, ad esempio, quello del "Theoretical and Scientific Data Science Group (TSDS)" della Scuola Internazionale Superiore di Studi Avanzati (SISSA) di Trieste. Istituito nel 2020 con un importante investimento da parte del governo italiano, il Gruppo TSDS mette a disposizione competenze interdisciplinari per affrontare i problemi dell'apprendimento automatico e le loro applicazioni alle scienze naturali.

Tutti i gruppi interdisciplinari, incluso quello che si è formato presso l'ISS, condividono la necessità di integrare diversi campi di indagine per affrontare le nuove sfide che per la loro complessità, non possono essere comprese restando ancorati a un singolo punto di vista scientifico.

L'approccio interdisciplinare è stimolato da una necessità individuata già a metà del secolo scorso. Nel 1948 il matematico Warren Weaver sosteneva che in natura esistono tre diverse classi di sistemi (4): semplici, a complessità disorganizzata e a complessità organizzata.

La prima classe include i sistemi che consentono un'astrazione estrema, e possono essere descritti con poche variabili. La loro semplicità non comporta necessariamente la possibilità di risolverli in modo semplice e spesso la soluzione non è necessariamente esaustiva.

I sistemi di classe 2, sono caratterizzati da un numero estremamente elevato di variabili, ognuna delle quali può avere un comportamento individuale casuale o non noto. In questo caso i sistemi necessitano della generazione di descrittori macroscopici a grana grossa. La termodinamica è uno degli esempi più noti di questo approccio statistico: le variabili di stato emergenti, come la temperatura o la pressione, descrivono completamente il sistema senza ricorrere alla conoscenza dei dettagli microscopici (dominati dal rumore), e possono essere considerate omogenee.

I sistemi biologici, solo in pochissimi casi, soddisfano i vincoli posti dalla descrizione di sistemi di classe 1 o 2, introducendoci quindi nella terza classe dei sistemi descritti da Warren Weaver, ovvero quelli a complessità organizzata ed è questo l'ambito delle scienze biomediche, ma non solo.

I sistemi a complessità organizzata sono caratterizzati da molti (anche se non così tanti come nella classe 2) elementi non identici che interagiscono tra loro con una forza di correlazione variabile nel tempo.

La complessità organizzata presenta caratteristiche uniche ed è troppo complicata per essere trattata con le tecniche tradizionali, come ad esempio quelle puramente statistiche, che ben funzionano con i sistemi di classe 2.





Secondo Warren Weaver la sfida della scienza sarebbe stata proprio la soluzione dei sistemi complessi in fisica, in biologia, ma anche in settori apparentemente più lontani, come l'economia (5).

Affrontare lo studio di sistemi complessi tipici delle scienze biomediche richiede una profonda rielaborazione del modo di pensare sia dei "teorici" che degli "sperimentalisti". Gli scienziati devono coniugare le leggi generali con le contingenze, in una miscela di attitudini verso il modo di fare ricerca che può nascere solo dal lavorare insieme.

Alla complessità del sistema, sia disorganizzato che organizzato, si aggiunge poi la quantità enorme di dati prodotti dalle tecnologie *high throughput* (ad esempio, omiche, neuroimmagini ecc.) che hanno reso anacronistici gli strumenti di lavoro tipici della biostatistica.

La rivoluzione tecnologica ha reso necessario abbattere vecchi paradigmi in cui l'evidenza ha dimostrato che non è più sostenibile la classica separazione delle imprese scientifiche in una sequenza lineare fatta di: formulazione di ipotesi, metodi sperimentali, analisi dei dati, verifica/falsificazione delle ipotesi.

Lo studio di modi nuovi per affrontare la complessità ha portato ad affiancare l'esperimento con la simulazione numerica, ponendo la necessità di collegare l'ingresso dell'intelligenza artificiale con l'esigenza di "spiegabilità", cioè quella di coniugare la previsione di un *end-point* biomedico rilevante con un modello teorico coerente. Tutto ciò ha provocato una ripresa di interesse per alcuni pilastri storici della metodologia scientifica, come gli approcci bayesiani e di riduzione della dimensionalità.

Nel complesso, il nuovo quadro di riferimento per chi fa ricerca in ambito biomedico ha favorito la ricerca di integrazione di diversi approcci scientifici, incoraggiando la collaborazione tra chi fa attività sperimentale e chi si occupa di modelli e/o di calcolo.

Queste connessioni sono la novità epistemologica più promettente, in cui lo sforzo di tenere insieme visioni e approcci diversi è necessario per affrontare nel miglior modo possibile le sfide poste a un istituto che fa ricerca per la tutela della salute pubblica. ■

Dichiarazione sui conflitti di interesse

Gli autori dichiarano che non esiste alcun potenziale conflitto di interesse o alcuna relazione di natura finanziaria o personale con persone o con organizzazioni, che possano influenzare in modo inappropriato lo svolgimento e i risultati di questo lavoro.

Riferimenti bibliografici

1. Bossa C, Branchi I, Caccia B, et al. The challenge of complexity in the Big Data era: how to ride the wave of high dimensional data revolution. *Ann Ist Super Sanità* 2022;58(3):151-3 (<https://annali.iss.it/index.php/anna/article/view/1511>).
2. Geman D, Geman S. Opinion: science in the age of selfies. *PNAS* 2016;113(34):9384-7 (doi: 10.1073/pnas.1609793113).
3. Choi BC, Pak AW. Multidisciplinarity, interdisciplinarity, and transdisciplinarity in health research, services, education and policy: 3. Discipline, inter-discipline distance, and selection of discipline. *Clin Invest Med* 31(1):E41-8 (doi: 10.25011/cim.v31i1.3140).
4. Weaver W. Science and complexity. *Am Scientist* 1948;36(4):536-44.
5. Greco P. Alla ricerca della complessità. *Scienzainrete* 17 aprile 2013.

TAKE HOME MESSAGES

È necessario integrare diversi campi di indagine per vincere le nuove sfide che non possono essere affrontate da una singola visione scientifica.

L'Istituto Superiore di Sanità ha una struttura multidisciplinare con un grande potenziale in grado di affrontare le nuove sfide scientifiche.