



Centro di Competenza sul Calcolo Scientifico

Le sfide e le opportunità computazionali nella ricerca in Economia e Statistica

Rosalba IGNACCOLO
rosaria.ignaccolo@unito.it

Dipartimento



Economia e Statistica
Cognetti de Martiis

- *Big data*: sfide per *storage* e analisi MA
opportunità di rispondere a nuove domande
- Microsimulazioni, simulazione stocastica e modelli ad agenti: opportunità di analizzare scenari possibili per valutare effetti di *policy*
- Statistica computazionale: opportunità di valutare nuovi metodi statistici e di modellare dati complessi

- **Analisi di dati brevettuali** in Economia dell'Innovazione: circa 3.3 milioni di brevetti con informazioni ausiliarie sugli inventori per studiare il ruolo delle interazioni nella produzione di conoscenza tecnologica.
- **Indice di morosità Turin-Index:** offre una visione dinamica del fenomeno della povertà, utile per individuare fasce di utenza soggette al rischio di ritardare i pagamenti. Necessità di integrare dati da fonti diverse con attuali 5.560.000 bollette (circa 510.000 utenti).
- Modello per **stimare la domanda di lavoro** nel quale l'impresa sceglie tra le opzioni di potenziali candidati (occupati e disoccupati); l'insieme di scelta sarebbe l'intero mercato del lavoro. Esempio Veneto: circa un milione di proposte l'anno, per 10.000 candidati, per 8 anni = 80 miliardi di record

- **JSTOR e l'evoluzione della conoscenza in Economia:** analisi semantica dell'intero testo di 500.000 articoli in economia presenti su JSTOR a partire dal 1840, attraverso tecniche di *deep learning* e lo sviluppo di algoritmi di *dynamic topic modelling*.
- **Analisi dei consumi culturali tramite Abbonamento Musei:** dati sui comportamenti di 300.000 possessori in vari anni della tessera musei, con informazioni ausiliarie sulla loro residenza e sul probabile percorso casa-museo con Google API.
- **TOBIKE e BIKE2WORK:** analisi dei dati di esercizio – sempre in aumento - del servizio di *bike sharing* cittadino, fonte informativa che sarà integrata da una base di geo-dati della *challenge* BIKE2WORK che prevede il tracciamento e la registrazione dei percorsi di tutti i partecipanti per comprendere tendenze e criticità nella navigazione urbana dei ciclisti torinesi.

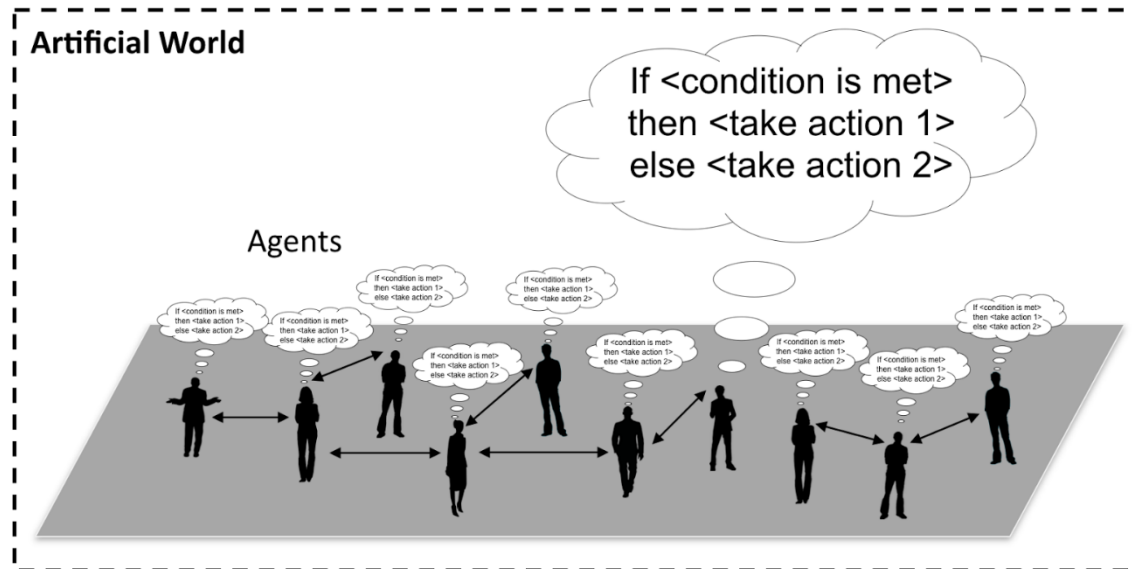
Una microsimulazione considera dati osservati su un campione grande di individui (famiglie o organizzazioni) e usa una matrice di transizione per aggiornare lo stato di ogni individuo nel campione come se scorresse il tempo. La possibilità di introdurre misure di politica economica nelle regole permette di effettuare analisi di scenario ex-ante.

- Approccio utilizzato in Economia del Lavoro per studiare implicazioni distributive di potenziali **riforme nel sistema pensionistico italiano**.
- **WHIP**: banca dati di storie lavorative individuali, costruita a partire dagli archivi gestionali dell'Inps.

Simulazioni stocastiche per i parametri di un modello deterministico i cui valori iniziali sono stimati a partire da dati reali.

- Approccio utilizzato in Economia dell'Ambiente per prevedere possibili scenari introducendo il **cambiamento climatico** nello studio di politiche di supporto ad **attività produttive che utilizzano risorse naturali** (scenari per acquacoltura di mitili in Italia fino al 2050).

Simulazione somigliante alle interazioni umane all'interno di uno spazio geografico o culturale.



Un sistema è modellizzato come un insieme di entità autonome, gli agenti, che prendono decisioni. Ogni agente valuta il suo stato e prende decisioni in base ad un insieme di regole che ne definiscono il comportamento sociale. Gli agenti interagiscono tra loro e con l'ambiente in cui si trovano.

- Modello per riprodurre le **scelte di trasporto urbano** nella città di Varese per studiare l'impatto di politiche pubbliche volte ad incentivare scelte di mobilità con emissioni inferiori di PM.
- Simulazione del **mercato dell'energia elettrica** per permettere l'analisi di politiche tenendo conto della regolamentazione, della griglia di distribuzione, dei comportamenti dei consumatori e delle strategie dei venditori.

- La **modellazione di dati complessi** richiede spesso calcolo ‘intensivo’
 - Modelli gerarchici per dati spaziotemporali via *Markov Chain Monte Carlo*
- **Esperimenti di simulazione** per esplorare le proprietà di una procedura statistica
 - Previsione spaziale di un’intera curva su un sito non monitorato e valutazione dell’incertezza associata via *bootstrap*

- Possibilità di «research data center as a service» ?
- Con C3S si potrà aumentare la velocità di calcolo e la capacità computazionale in molti progetti, e rendere possibili ricerche altrimenti impossibili
- Con C3S si favorisce la collaborazione e l'interscambio di idee ed esperienze per favorire la ricerca interdisciplinare