

**Tim Phillips [00:00:00]:**

Oggi su VoxTalks Economics, un nuovo ambizioso progetto per scavare nei microdati e renderli accessibili ai ricercatori. Benvenuti a un altro VoxTalks, registrato dal vivo presso il Center for Economic Policy Research, Paris Symposium 2023. Mi chiamo Tim Phillips. I microdati raccolti a livello di nucleo familiare, di azienda o di individuo sono straordinariamente preziosi per i ricercatori interessati al funzionamento delle politiche. Ma anche se a volte sembra che stiamo affogando nei dati, non ci sono abbastanza dati di questo tipo, accessibili e di alta qualità, per soddisfare i ricercatori e i politici. L'anno scorso, quando VoxTalks ha visitato Parigi, abbiamo parlato con Filippo di Mauro di Compnet e Ugo Panizza del Graduate Institute di Ginevra, in un bar dell'hotel, mi sembra di ricordare, delle loro ambizioni di migliorare questa situazione. Li ho raggiunti ancora una volta per scoprire quali progressi sono stati fatti. Bentornato, Filippo.

**Filippo di Mauro [00:01:16]:**

Grazie.

**Tim Phillips [00:01:17]:**

E anche Ugo.

**Ugo Panizza [00:01:18]:**

Grazie Tim.

**Tim Phillips [00:01:20]:**

Lei definisce i microdati un bene pubblico. Può farmi un esempio del perché di questa affermazione, una domanda politica nel 2024 per la quale i microdati saranno importanti?

**Ugo Panizza [00:01:32]:**

Ok, Tim, fammi usare il Next generation UE. Si tratta di un programma enorme. L'Unione Europea ha in programma di investire più di 800 miliardi in questo programma. Ci sono diversi obiettivi. Renderla verde, renderla digitale, renderla sana, renderla forte, renderla uguale. Ne scelgo uno. Renderla forte. Il pilastro per renderla forte è incoraggiare i giovani a studiare la scienza e la tecnologia, che apre le porte ai lavori verdi del futuro, sostenendo l'istruzione e offrendo prestiti e sovvenzioni ai giovani imprenditori. È un'ottima cosa, ma come facciamo a sapere cosa funziona, cosa non funziona, come facciamo a sapere come fare nella pratica? È qui che i microdati possono aiutare. Avere informazioni dettagliate sulle imprese, sulle persone, sull'istruzione ci aiuterà a valutare quali sono i modi più efficienti per raggiungere questi obiettivi. E per farlo, abbiamo bisogno di microdati.

**Filippo di Mauro [00:02:24]:**

Vorrei aggiungere una cosa: la competitività è la parola d'ordine del 2024. A quanto pare il presidente della Commissione ha stabilito che si tratta di una priorità. Mario Draghi è stato incaricato di redigere una relazione per giugno. Come si farà con la mancanza di dati a livello microeconomico? Immaginiamo di voler verificare come le imprese siano state colpite dal COVID e dalla guerra ucraina e di capire quanto sia stato differenziato l'impatto tra grandi imprese e le imprese più produttive o le aziende che sono state coinvolte nella catena del

valore globale o meno. Ebbene, non è possibile farlo. Forse si può fare per la Francia, perché ha dati fantastici che collegano la produttività e i risultati dell'impresa con dati personalizzati, ma difficilmente si può fare con altri Paesi. Quindi buona fortuna a Mario Draghi, purtroppo.

**Tim Phillips [00:03:18]:**

Quasi tutto ciò che facciamo sembra creare dati da qualche parte. Il problema è che i dati non sono disponibili o che non esistono?

**Ugo Panizza [00:03:30]:**

Nella maggior parte dei casi, è solo che non sono disponibili. Naturalmente possiamo pensare a qualche problema per il quale non abbiamo dati, ma per quasi tutte le domande che mi vengono in mente, penso che ci siano dati in giro. Il problema è: come si accede a questi dati, quanto costa accedervi? Forse ci sono dei dati che vorrei avere e che non esistono. Ma se potessi accedere a quelli che esistono, credo che potrei rispondere al 99% delle domande che ho in testa.

**Tim Phillips [00:03:52]:**

Mi faccia un esempio di quali microdati potrebbero essere utilizzati e che al momento esistono, ma sono difficilmente accessibili.

**Ugo Panizza [00:04:00]:**

Vi faccio un esempio di un Paese, e non si tratta nemmeno di un Paese ricco, ma di un Paese a medio reddito, che dispone dei dati più sorprendenti che un ricercatore possa immaginare. Si tratta di un Paese che dispone di dati su ogni singola azienda del Paese, collegati con ogni singola banca del Paese. Conosciamo i prestiti concessi da una banca a una determinata azienda, e in più le aziende sono collegate tra loro attraverso il registro dell'IVA, e le banche sono collegate alle banche internazionali attraverso i prestiti internazionali, e le aziende sono collegate ai lavoratori attraverso le buste paga. Immaginate quindi questo insieme di dati. Si può seguire un lavoratore attraverso un'azienda, una banca, altre aziende, si possono studiare cose incredibili. E questi dati esistono già. Il problema è che è molto difficile accedere a questi dati in questo Paese, a causa, diciamo, di una situazione politica complicata. Ho iniziato un progetto, poi il direttore dell'istituzione che controllava questi dati è cambiato e il progetto è stato interrotto. Quindi, prima di tutto, questo modello potrebbe essere replicato da altri Paesi. E poi c'è il problema di garantire ai ricercatori l'accesso a questo tipo di dati.

**Filippo di Mauro [00:05:01]:**

Permettetemi di aggiungere qualcosa. È vero, i dati ci sono, e in effetti abbiamo buone notizie in alcuni Paesi, come la Francia e i Paesi Bassi in particolare, dove l'accesso ai dati è del tutto trasparente, è effettivamente soggetto a tariffe per remunerare, ovviamente, gli statistici che lo fanno di mestiere e l'accesso è garantito ai ricercatori che hanno un buon progetto da eseguire. Quindi è fantastico. Il problema è che questo non è affatto generalizzato in Europa. In particolare, c'è un aspetto dell'interpretazione legale della protezione dei dati, che è fondamentalmente estremamente eterogeneo tra i vari Paesi. E questo è davvero inaccettabile. Se si pensa che l'Europa è un mercato comune. Si tratta quindi di un aspetto che dobbiamo correggere. Ci sono Paesi in cui i dati sono

estremamente buoni, ma fondamentalmente hanno una protezione dei dati, l'accesso ai dati, che è inferiore a quella degli altri Paesi. In particolare, in alcuni Paesi l'accesso è consentito solo ai ricercatori nazionali. Quindi deve esserci un ricercatore di quello stesso Paese che ha accesso ai dati. Questo non è accettabile perché ci sono altri Paesi in cui queste procedure sono molto diverse e confidenziali. Ogni buon ricercatore europeo può accedere ai dati, purché vengano mantenute determinate condizioni di riservatezza.

**Tim Phillips [00:06:22]:**

Voi vi siete dati da fare da quando abbiamo parlato l'anno scorso e una delle cose che avete fatto è stata quella di chiedere ai ricercatori quali sono i loro problemi nell'accedere ai microdati. Cosa vi hanno risposto?

**Ugo Panizza [00:06:33]:**

Ci sono problemi di ogni tipo, che variano a seconda del campo in cui lavora il singolo ricercatore. Un problema comune è il costo. Alcuni database sono forniti su base commerciale e sono incredibilmente costosi. E questo, tra gli altri problemi, crea disuguaglianze nella ricerca. Giusto? Quindi, se si appartiene a istituzioni ricche e si dispone di un grande budget, si possono fare certe cose. Se si appartiene a un'istituzione più piccola e povera, non è possibile. Il costo è un problema. Alcuni set di dati sono di libero accesso, ma solo alcune persone possono accedervi. Filippo ha detto che se si è cittadini di un certo Paese, si ha accesso a determinati dati. Questo è un problema. A volte si tratta solo di chi si conosce. Se si conosce la persona giusta, si ottiene l'accesso, mentre se non si conosce la persona giusta, non si ottiene l'accesso. Ci sono quindi diversi problemi, ognuno dei quali deve essere affrontato a modo suo.

**Filippo di Mauro [00:07:28]:**

In realtà, ancora una volta, il punto di vista di Ugo è che questi dati sono scarsamente disponibili. E quindi, quando si vuole veramente fare uno studio internazionale in Europa, si devono utilizzare dati commerciali molto costosi che, tra l'altro, e mi dispiace dirlo, non sono necessariamente della migliore qualità perché si tratta di dati disponibili, pubblicamente disponibili. Ma naturalmente, dato che abbiamo detto che c'è un problema di accesso agli istituti nazionali di statistica, ovviamente non sono i migliori, perché l'istituto nazionale di statistica di alcuni Paesi avrebbe dati molto migliori.

Il nostro obiettivo dovrebbe quindi essere quello di tirare fuori i dati che si trovano nei caveau degli istituti nazionali di statistica, riducendo così i costi per tutti, piuttosto che prendere una scorciatoia, che è: "Ok, lasciatemi usare i dati che sono disponibili". Sono un ricercatore ricco, quindi posso permetterlo. Questo non va bene.

**Tim Phillips [00:08:19]:**

Ora, negli ultimi anni molte organizzazioni hanno reso l'accesso ai loro dati, i dati che raccolgono, come un sottoprodotto di ciò che fanno quotidianamente, parte della loro strategia commerciale, monetizzando quei dati. Se diciamo che questi dati sono un bene pubblico, l'accesso ai dati per la ricerca dovrebbe essere libero o limitato a qualche livello?

**Ugo Panizza [00:08:42]:**

Le istituzioni pubbliche, a mio avviso, dovrebbero avere il dovere di fornire questi dati gratuitamente o senza alcun profitto, perché questo fa parte del loro mandato di produrre beni pubblici. Ora, se si tratta di istituzioni private a scopo di lucro, è difficile dire loro cosa fare con qualcosa che è di loro proprietà. Ma credo che sarebbe ideale e probabilmente anche conveniente per loro avere una chiara strategia di prezzo per i ricercatori, perché ora alcune di queste istituzioni hanno a che fare con grandi aziende del settore privato e dicono che accedere a questi dati costa mezzo milione di dollari. Se sei una grande banca, va bene. Se sei un ricercatore, non va bene, e poi devi chiamarli e ognuno ha un prezzo diverso. Quindi, se questa istituzione avesse un prezzo accademico chiaro o un accordo con il consorzio sarebbe sicuramente un buon passo.

**Filippo di Mauro [00:09:29]:**

Credo, Tim, che il finanziamento sia una questione fondamentale. Gli istituti di statistica sono strutturalmente sotto organico, sottovalutati, quasi considerati non all'altezza degli economisti. Giusto. Questo è un aspetto che deve essere cambiato, perché la qualità media del loro lavoro è semplicemente incredibile. La ricchezza dei dati che hanno e la qualità del materiale che hanno, anche se sono completamente sottovalutati, è fantastica. Ho un'esperienza incredibile in diversi Paesi. Per me, quindi, la questione del finanziamento è fondamentale, perché dobbiamo assicurarci che non ci sia più una carenza cronica di personale in grado di fornire i servizi quando servono. Ho citato il caso della Francia e dei Paesi Bassi, un ottimo caso, perché in realtà sono stati monetizzati. I loro servizi, che non sono terribilmente costosi rispetto ai set di dati commerciali, sono solo una frazione di quel prezzo. Per questo motivo, io opterei per un metodo di finanziamento sistematico, e quindi per la partecipazione dei ricercatori. Utilizzare i fondi normalmente disponibili per il finanziamento dei set di dati, eccetera, credo sia la strada giusta da percorrere.

**Tim Phillips [00:10:30]:**

Naturalmente, nel settore privato non si può obbligare la gente a fare qualcosa con i propri dati. Nel settore pubblico, invece, c'è l'opportunità di avere un processo standard per l'accesso, costi standard.

**Filippo di Mauro [00:10:40]:**

Esattamente. Vogliamo che sia trasparente come lo è in Francia e nei Paesi Bassi. Anche in questo caso, si tratta di un modello che l'Europa dovrebbe seguire.

**Ugo Panizza [00:10:49]:**

Così, ad esempio, l'Amministrazione italiana della previdenza sociale nel 2015 ha creato il programma Visit INPS. Non tutti possono avere accesso ai dati perché le capacità sono limitate. Ma c'è un processo chiaro. Vado sul loro sito web, c'è una procedura di richiesta. Così tutti hanno le stesse informazioni. Questo dovrebbe essere il gold standard ovunque, avere una procedura chiara di cui tutti sono a conoscenza. Poi dipende dall'istituzione, ovviamente, e dal tipo di dati. Ma il fatto che tutti debbano avere lo stesso tipo di informazioni e, in linea di principio, la stessa possibilità di accedere ai dati, credo sia molto importante.

**Tim Phillips [00:11:27]:**

Quando abbiamo parlato l'anno scorso, voi avete avuto l'idea di creare una sorta di micro hub di dati per l'Europa.

Come stanno procedendo questi piani ?

**Filippo di Mauro [00:11:36]:**

Tim, fortunatamente stanno procedendo molto bene. Devo dire che abbiamo ottenuto un finanziamento dalla Commissione europea, la DG Finanze per le riforme, un cosiddetto meccanismo di intervento di supporto tecnico. Questo ci ha fornito le risorse per assumere un po' di personale, giovani economisti che ci aiutano a fare da intermediari tra gli istituti di statistica, i comitati nazionali per la produttività, che dovranno contribuire alla formazione e all'assistenza per utilizzare in modo appropriato questo set di dati. Stiamo quindi facendo enormi progressi. Sono già cinque i Paesi che hanno creato l'infrastruttura di microdati, l'MDI, e ci aspettiamo che altri pochi Paesi si aggiungano. Ieri ho parlato con l'ISTAT italiano. È stato garantito loro che entro pochi giorni saranno a bordo. Ho avuto un incontro a Malta. Saranno a bordo tra qualche settimana. Abbiamo già la Germania e la Francia e i Paesi Bassi, come già detto, sono già campioni per noi. Quindi stiamo procedendo bene. A proposito, ringrazio molto la comunità del CEPR per averci aiutato in questa impresa.

**Tim Phillips [00:12:38]:**

Posso immaginare che non sia un problema facile da risolvere. Come funziona l'MDI?

**Filippo di Mauro [00:12:42]:**

Sì, non è affatto facile. Stabiliamo l'accesso ai dati paese per paese. Le procedure di accesso sono molto diverse. Per questo motivo, continuiamo ad avere contatti bilaterali con gli istituti nazionali di statistica, assicurandoci di firmare un memorandum d'intesa su il tipo di dati che vogliamo, il tipo di procedura che dobbiamo seguire per avere questi dati. A quel punto, creiamo dei codici che fondamentalmente armonizzano la raccolta dei dati, in modo che quando eseguiamo un codice per, ad esempio, la Francia, sia lo stesso codice per i Paesi Bassi. In linea di principio, in una situazione di stato stazionario, un ricercatore ci proporrà una serie di codici, ad esempio sulla dimensione della produttività delle esportazioni, e noi forniremo i risultati di quel codice in poche settimane per il numero di contee che partecipano.

**Tim Phillips [00:13:36]:**

Qual è l'obiettivo finale, Filippo? L'MDI è progettato per essere un archivio di dati aperto a tutti, o è destinato a ricercatori e progetti di ricerca specifici?

**Filippo di Mauro [00:13:46]:**

Dobbiamo distinguere due cose, giusto? Da un lato abbiamo la serie di dati dei componenti, come ricorderete. Ok, ho detto al pubblico di cosa si tratta. Si tratta fondamentalmente di un insieme di indicatori che costruiamo ogni anno e che sono fissati, e sono circa 300 o 400, che studiano la relazione tra la produttività e i fattori che la determinano. Si tratta quindi di una serie di dati fissi. Questo sarà pubblico. L'MDI, invece, sarà un tipo di strumento molto più modulare, perché a quel punto, pur avendo un set di dati di base, dipenderà dal tipo di domande che l'utente ci invierà, giusto? In questo caso, quindi, sarà molto più modulare e personalizzato. Quindi, ogni due o tre mesi, eseguiremo una serie di codici, ma questa sarà diversa.

da quelli che si candidano solo qualche mese prima. Questo è il concetto di differenza.

**Tim Phillips [00:14:35]:**

Ogni giorno nei telegiornali si sente parlare di IA e di big data, e questo ha reso le persone molto più consapevoli di quanti dati ci siano, di quanti dati ci siano su di loro. Come pensa che questo influirà sul suo progetto? Pensa che renderà le persone diffidenti perché sono consapevoli della quantità di dati che vengono acquisiti e condivisi su di loro? Oppure pensate che le persone saranno entusiaste perché si renderanno conto del potenziale di quei dati per aiutare le loro vite?

**Ugo Panizza [00:15:01]:**

Credo che sarà soprattutto un'opportunità. La maggior parte di noi ha letto il libro di Daron Acemoglu e Simon Johnson sull'IA e il loro punto di vista secondo cui l'IA può essere usata per il bene e per il male, e quindi spero che venga usata per il bene. Penso che in termini di accesso ai dati, possiamo fare cose incredibili, in parte con i dati che sono già lì quando parliamo con gli storici dell'economia, la loro sfida principale, quindi lì, i dati esistono, ma sono in qualche archivio, scritti in qualche strano linguaggio, difficili da leggere, e ora gli storici dell'economia, stanno trascrivendo le cose. A questo punto penso che l'IA possa davvero cambiare le carte in tavola, perché se riuscissimo a scattare una foto e a far sì che l'IA la riconosca e digitalizzi i dati, potremmo fare qualcosa di davvero incredibile.

**Filippo di Mauro [00:15:44]:**

Sono assolutamente d'accordo con il potenziale sottolineato da Ugo. Tuttavia, non siamo ancora in quella fase, perché a questo punto il nostro obiettivo è creare i dati di base per il maggior numero possibile di Paesi. Alla fine, però, quando il set di dati, l'MDI, sarà pronto, ovviamente l'intelligenza artificiale ci aiuterà a gestire le richieste, per esempio, o a fare in modo di raggruppare le richieste, in modo che il sistema sia tremendamente efficace e tremendamente più efficiente grazie all'AI. Penso quindi che questo sia un enorme vantaggio per il nostro progetto.

**Tim Phillips [00:16:16]:**

Quindi, ultima cosa, Ugo, qual è il ruolo del CEPR in tutto questo? Cosa può fare?

**Ugo Panizza [00:16:22]:**

La nostra presidente Beatrice ci dice sempre che C sta per centro ma dovrebbe essere N per network (rete), giusto? Perché siamo davvero una rete di ricercatori. E credo che questa rete possa davvero svolgere un ruolo nel migliorare l'accesso ai dati facendo molte cose. In primo luogo, come centro, il CEPR può coordinare le persone, ma come rete possiamo informare i ricercatori, informare le persone e informare i politici. E questo è davvero fondamentale perché, come ha detto Filippo, le risorse sono fondamentali. E alla fine, queste risorse, le risorse pubbliche, sono controllate dai politici. Quindi, se i responsabili politici comprendono appieno i vantaggi di disporre di dati migliori, ad esempio per la valutazione dell'UE di prossima generazione o per la valutazione della competitività, allora saremo disposti a destinare risorse agli uffici statistici e a creare incentivi affinché questi ultimi consentano ai ricercatori di accedere ai dati. E credo che in questo caso la rete CEPR, la comunità CEPR, che è molto visibile nei singoli Paesi, possa svolgere un ruolo importante.

**Filippo di Mauro [00:17:20]:**

Per me è molto importante anche cercare di convincere i nostri colleghi a cambiare mentalità. L'idea di avere un progetto che riguardi un solo Paese, magari piccolo, non è necessariamente il modo migliore di procedere. Dobbiamo davvero puntare il più possibile e il prima possibile a un meccanismo multilaterale di raccolta dei dati e a una ricerca basata su di esso. Voglio dire, negli Stati Uniti lo fanno. Non fanno tanto per uno Stato, il Missouri contro il Michigan. Smettiamo di farlo in Europa. Ma questo, di nuovo, non è il CEPR. Ma spero che in realtà molte persone condividano il mio punto di vista e si impegnino in tal senso.

**Tim Phillips [00:18:00]:**

Un'opportunità enorme, ma anche una sfida enorme. Quindi grazie, Filippo.

**Filippo di Mauro [00:18:03]:**

Grazie.

**Tim Phillips [00:18:03]:**

Grazie, Ugo.

**Ugo Panizza [00:18:04]:**

Grazie, Tim.

**Tim Phillips [00:18:13]:**

Se volete leggere le ricerche fatte finora, c'è il documento di discussione Microdata for Economic Research in Europe. Sfide e proposte. Molti gli autori, Filippo e Ugo sono due di loro. Si tratta del documento di discussione 18640 del CEPR.

**[Voce fuori campo] [00:18:34]:**

Questo è un VoxTalk registrato al Simposio di Parigi del CEPR. Quando controllate Instagram, cercate le clip degli episodi passati su VoxTalks Economics. Seguiteci ovunque riceviate i vostri podcast. Noi siamo lì. E se vi piace quello che ascoltate, perché non ci lasciate anche una recensione? La prossima settimana su VoxTalks Economics, la partecipazione agli utili migliora i dipendenti o le aziende?