POLITECNICO DI TORINO
Repository ISTITUZIONALE

Innovare la valutazione d’impatto sociale di iniziative culturali. Il caso delle Capitali Europee della Cultura

Original
Innovare la valutazione d’impatto sociale di iniziative culturali. Il caso delle Capitali Europee della Cultura / Mereghetti, C.. - In: @ DIGITCULT. - ISSN 2531-5994. - ELETTRONICO. - 7:2(2022), pp. 53-69. [10.36158/97888929562234]

Availability:
This version is available at: 11583/2974296 since: 2023-01-02T09:51:13Z

Publisher:
Tab Edizioni

Published
DOI:10.36158/97888929562234

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

25 November 2026

Innovare la valutazione d'impatto sociale di iniziative culturali. Il caso delle Capitali Europee della Cultura

Innovating the social impact evaluation of cultural initiatives. The case of the European Capitals of Culture

Cecilia Mereghetti
Politecnico di Torino
cecilia.mereghetti@polito.it

| abstract

Questo articolo esplora la possibilità di introdurre delle innovazioni nel campo della valutazione degli impatti sociali, proponendo di partire da fonti di dati alternative. In particolare, vengono prese in considerazione le analisi degli impatti nel campo delle politiche e delle iniziative culturali, usando come caso esemplificativo il programma delle Capitali Europee della Cultura (ECoC). Il testo si sviluppa in due parti. La prima parte, teorica, affronta un'analisi della letteratura esistente nel campo delle valutazioni d'impatto delle ECoC, con particolare riferimento agli impatti sociali. L'analisi della letteratura evidenzia come, nonostante la numerosità degli studi a riguardo, vi siano alcune tematiche – e impatti – meno indagate di altri, e come l'approccio metodologico sia rimasto pressoché invariato negli anni, con le metodologie qualitative come strumento principale. Una volta evidenziati i limiti – contenutistici e metodologici – degli studi d'impatto delle ECoC, l'articolo prosegue con una parte empirica, basata su una ricerca attualmente in corso. La seconda parte dell'articolo introduce un metodo alternativo per la valutazione degli impatti sociali delle ECoC, con un'attenzione particolare al tema del benessere soggettivo. La letteratura recente sottolinea un legame tra il benessere soggettivo e i contenuti pubblicati sui social media. Il metodo proposto in questo articolo si basa proprio sui contenuti testuali spontanei pubblicati sui social media che, insieme ad altre fonti scritte, vengono raccolti e analizzati con algoritmi per l'analisi testuale automatizzata. L'obiettivo è duplice: valutare l'impatto delle ECoC sul benessere individuale (soggettivo) degli utenti e valutare il successo dell'iniziativa nel suo insieme, comparando obiettivi dichiarati e risultati. L'articolo illustra sia il metodo di analisi sia la costruzione del database, soffermandosi sul processo di implementazione e sottolineando criticità e soluzioni, ma anche limiti tuttora in essere. Infine, vengono presentate e introdotte alcune potenzialità future del metodo, che in questo momento è ancora in fase sperimentale.

This article explores the possibility of introducing innovations in the field of social impact evaluation, proposing to start from alternative data sources. In particular, it regards impact analyses in the field of cultural policies and initiatives, using the European Capitals of Culture (ECoC) program as an illustrative case. The text is comprised of two parts. The first part explores theoretical issues, dealing with review of the existing literature in the field of social impact evaluations of ECoCs. The literature review shows how, despite the large number of studies on the subject, there are some issues – and impacts – less investigated than others and how the methodological approach has remained virtually unchanged over the years – with qualitative methods taking the lead. Once the limits of ECoC-related impact studies have been highlighted, the article continues with an empirical part, based on research currently underway. The second part of the article introduces an alternative method for assessing the social impacts of ECoCs, with a special focus on the issue of subjective wellbeing. Recent literature emphasizes a link between subjective wellbeing and content posted on social media. The method proposed in this article is based precisely on spontaneous textual content posted on social media, which, together with other written sources, is collected and analyzed with algorithms for automated textual analysis. The objective is twofold: to assess the impact of ECoCs on the individual (subjective) wellbeing of users and to evaluate the success of the initiative as a whole by comparing stated goals and results. The article illustrates both the method of analysis and the construction of the database, dwelling on the building process and pointing out critical issues and solutions, as well as limitations that are still in place. Finally, some future potential developments of the method, which at this time is still in an experimental stage, are presented and introduced.

doi 10.36158/97888929562234

Introduzione

Gli studi sull'impatto sociale di attività antropiche si sono evoluti negli anni, e hanno abbracciato una varietà di ambiti sempre maggiore. L'interesse verso le tematiche sociali – inclusione, partecipazione, gentrificazione, per citarne alcune – si è visto anche nel settore culturale (Biondi et al., 2020; Nagy, 2018). A livello globale, l'UNESCO (2019) ha presentato alcuni indicatori che fanno riferimento alla misurazione delle attività culturali e ai loro impatti sui territori, con particolare riferimento proprio alla sfera sociale – a conferma della presa di coscienza collettiva della rilevanza degli impatti non solo economici delle attività del settore.

Uno dei programmi culturali a livello europeo che più ha abbracciato la nuova rilevanza di queste tematiche è quello delle Capitali Europee della Cultura (da qui in avanti, ECoC). Questo Programma, come verrà esposto nel corso del presente articolo, è stato uno dei più studiati dalla letteratura degli ultimi trent'anni, da diverse angolazioni.

L'obiettivo del presente articolo è quello di inserirsi nel discorso sulle ECoC e sui loro effetti sulla sfera sociale, contribuendo al dibattito con l'introduzione di un metodo alternativo per la valutazione degli impatti sociali che parta da fonti di dati "nuove" rispetto al passato – come auspicato dai promotori del programma stesso (European Commission, 2018, p. 8) – e usi strumenti innovativi per indagare una tematica già ritenuta rilevante.

Il metodo proposto si fonda sull'uso di fonti scritte di diversa natura (documenti ufficiali, rassegna stampa, contenuti dei social media), utilizzando tecniche di Machine Learning per l'analisi testuale volte a ricercare l'impatto sociale del programma. L'analisi testuale mira nello specifico a individuare la presenza e l'utilizzo nei dati di termini e tematiche legate alla sfera sociale (con particolare riferimento al benessere soggettivo). Essendo la sfera sociale ormai considerata parte integrante degli obiettivi del Programma, la presenza o meno di queste tematiche e di termini ad esse legate – così come l'analisi del punto di vista di diversi attori coinvolti – è anche il punto di partenza per una verifica dell'ottenimento dei risultati dichiarati rispetto agli obiettivi iniziali.

La necessità di trovare nuovi indicatori, nuove forme di analisi e nuove fonti di dati è presente – implicitamente o esplicitamente – nella letteratura di settore da almeno una decade.

Per quanto riguarda le ECoC, già dai tempi del cosiddetto "modello Liverpool" (Boland, 2010; Connolly, 2013; Cox & O'Brien, 2012; O'Brien, 2011) ci si interroga sulla valutazione degli impatti e sul "successo" dell'iniziativa nel raggiungere i propri obiettivi. Molti autori si sono espressi approcciando il problema da diverse angolazioni, dall'analisi dei sistemi di governance (Lähdesmäki, 2013; Németh, 2016), all'analisi di impatto sociale (Demartini et al., 2018; Fišer & Kožuh, 2019; Ftjar et al., 2013; Nagy, 2018), all'analisi del settore culturale della città ospitante (Campbell, 2011).

Questo articolo si posiziona all'interno del dibattito sulla necessità di innovare la valutazione degli impatti sociali, proponendo un nuovo metodo partendo dal caso delle ECoC – sulla base di una ricerca attualmente in corso sul caso studio di Matera 2019. Nella prima sezione verrà presentato in breve il contesto di riferimento, quello degli studi sugli impatti delle ECoC. Nella seconda sezione verrà presentato il metodo proposto, con un approfondimento sul tipo di fonti utilizzate e sulle tecniche di analisi previste. Nella sezione successiva verranno invece esposti gli obiettivi principali dell'uso di questo metodo, ma anche alcune delle criticità emerse nella sua costruzione, possibili soluzioni adottabili e limiti ancora in essere. Infine, l'articolo si conclude con alcune delle motivazioni che rendono interessante l'introduzione del metodo proposto, e alcuni potenziali usi futuri.

DOI 10.36158/97888929562234

Gli impatti delle Capitali della Cultura: stato dell'arte e nuove direzioni

Negli ultimi trent'anni anni si è visto un progressivo aumento dell'interesse verso gli impatti delle politiche e iniziative culturali, visti come parte integrante dell'ambìto obiettivo dello sviluppo sostenibile (Rayman-Bacchus & Radavoi, 2020). Particolare interesse è stato suscitato da alcune branche di ricerca, come gli studi sulla percezione degli effetti del turismo nelle comunità ospitanti (Ap, 1990), sugli impatti di eventi e festival (Bracalente et al., 2011; Langen & Garcia, 2009), sugli effetti della partecipazione culturale in diverse aree urbane (Biondi et al., 2018, 2020; Piber et al., 2017), e sull'impatto dell'arte e della cultura sull'inclusione sociale (Belfiore, 2002), sulla qualità della vita (Blessi et al., 2016; Grossi et al., 2011) e sulla salute delle persone (Cicerchia & Bologna, 2017; Grossi et al., 2012).

A livello europeo, nello stesso periodo, si è visto un aumento dell'interesse verso iniziative e programmi culturali che dimostrassero un certo "potere trasformativo" nei territori di implementazione. L'esempio più prestigioso e immediato è quello delle Capitali Europee della Cultura (ECoC), il programma forse più famoso e tra i più ambìti a livello comunitario.

Le ECoC risultano quindi tra i casi più studiati in letteratura, quando si parla di impatti di iniziative culturali. Il Programma, nato nel 1985, è stato studiato da diverse prospettive, soprattutto per analizzare il suo impatto a livello urbano (Bianchini, 2013; Ooi et al., 2014). Sono stati studiati diversi aspetti relativi agli impatti del programma, sia da una prospettiva sociale che economica, e in generale, è emersa una visione critica del programma negli ultimi anni (Ooi et al., 2014). Diverse città sono state analizzate come casi studio, rivelando tre macroaree di ricerca in cui vanno a inserirsi le valutazioni di impatto sul territorio. Una prima area riguarda il progressivo uso del Programma ECoC come strumento di city-branding, per aumentare il turismo (Campbell, 2011; Liu, 2014; Richards, 2000; Richards & Wilson, 2004; Quinn, 2009). Un secondo tipo di impatti molto studiati invece riguarda il potere "rigenerativo" del Programma (Bianchini, 2013), partendo spesso dall'analisi di strutture di governance in grado di prendere decisioni strategiche e creare reti strette e interconnesse di attori locali che lavorano insieme per fornire nuove opportunità alla città (Demartini et al., 2018; Piber et al., 2017; Németh, 2016; O'Brien, 2011).

Infine, alcuni autori hanno cercato di individuare gli impatti dell'ECoC sulla sfera sociale delle città. In particolare, negli ultimi anni, è stata prodotta sempre più letteratura legata all'analisi delle diverse sfumature degli impatti sociali, dall'inclusione alla coesione sociale, partecipazione e problematiche legate alla gentrificazione (Demartini et al., 2018; Fitjar et al., 2013; Nagy, 2018). Anche l'analisi della sfera "emotiva" è presente in letteratura. Alcuni studi hanno affrontato questo tema rispetto al Programma ECoC, specialmente in termini di sviluppo di un senso di coesione e orgoglio civico (Collins, 2016; Fišer & Kožuh 2019). Gli stessi autori sottolineano tuttavia come la dimensione emotiva sia complessa da approcciare e da analizzare.

La sensibilità accademica è andata di pari passo con l'attenzione sempre maggiore della stessa Commissione Europea verso le tematiche sociali. (UNESCO, 2019; European Commission, 2018). Gli impatti sociali nel loro insieme rientrano nello specifico tra gli obiettivi dichiarati del Programma ECoC (European Commission, 2018); l'adozione di tali impatti come punti fermi di una politica così rilevante rende evidente, quindi, quanto l'esplicitazione e l'aumento di questo interesse siano tutt'altro che momentanei.

Gli studi sugli impatti sociali sono basati su ricerche di tipo prevalentemente qualitativo, e utilizzano diverse tecniche e strumenti come interviste, focus group e questionari

DOI 10.36158/97888929562234

(Woosnam, 2014), che permettono di ottenere informazioni qualitativamente rilevanti (in termini di profondità e specificità) interagendo direttamente con gli attori interessati (Garbarino & Holland, 2009; Rao & Woolcock, 2003). In alcuni casi vengono utilizzate anche fonti secondarie (Demartini et al., 2018), per completare il quadro di informazioni e analizzare il contesto di riferimento. Un'analisi sistematica di questi studi ha evidenziato alcuni limiti – riconducibili in parte alla natura incerta degli impatti che vogliono catturare e in parte all'uso prevalente del sondaggio come strumento di raccolta dati e analisi (Jaidka et al., 2020). In primo luogo, la mancanza di definizioni standard per l'impatto sociale (o le diverse componenti dell'impatto sociale) e di misurazioni condivise (Norman & MacDonald, 2004), portano a risultati poco solidi, che spesso necessiterebbero di ulteriori approfondimenti. Nel caso delle ECoC, la natura stessa del Programma porta a una sorta di polarizzazione dei risultati, con studi dalle visioni decisamente positive che si scontrano con visioni al contrario molto critiche (Boland, 2010; O'Brien, 2011).

Un ulteriore elemento che si evidenzia riguarda i temi indagati. Nonostante la pluralità di declinazioni del termine "impatto sociale", si notano una serie di aspetti più studiati – come, ad esempio, la partecipazione – rispetto ad altri. Una delle tematiche meno indagate è quella legata al benessere soggettivo e alla sua percezione nelle città che sono diventate ECoC (Steiner et al., 2015). Questo gap contenutistico è in parte dovuto alla mancanza, anche in questo caso, di una definizione univoca di cosa sia il benessere soggettivo e quali siano le sue caratteristiche (Ballas, 2013; Goodman et al., 2017; Oishi, 2018; Seligman, 2018; Wills-Herrera et al., 2009), che porta a una dispersione dei tentativi di analisi. Questa tematica è tuttavia parte fondamentale degli impatti sociali, così come sono stati concepiti in termini obiettivi all'interno del Programma.

Infine, il questionario come metodo di indagine preferenziale per quanto riguarda gli impatti sociali – e il benessere soggettivo, nello specifico – presenta alcuni limiti: in primo luogo, i questionari richiedono una elevata quantità di tempo e risorse (ad esempio, personale con le adeguate competenze), che lo rendono uno strumento assolutamente utile, ma pian piano sempre meno efficiente (Jaidka et al., 2020). In secondo luogo, i questionari hanno bisogno di un numero di rispondenti molto elevato per fornire informazioni rilevanti, e le risposte ottenute non sono sempre complete o possono risentire di una componente aleatoria data dalla potenziale diversa interpretazione della scala da parte dei soggetti cui è sottoposto (data da distrazione, fretta o a volte da una diversa attribuzione di valore agli estremi della scala rispetto alle intenzioni del ricercatore) (Jaidka et al., 2020; Voukelatou et al., 2021).

Emerge dunque una duplice necessità: l'esplorazione di nuove tematiche legate agli impatti sociali, e l'introduzione di nuovi metodi di analisi per la valutazione dei risultati.

Nel caso specifico delle ECoC, la necessità di nuove metodologie e nuove fonti di dati è esplicitata su più fronti, ma in particolare la Commissione Europea (2018) nelle sue linee guida per la valutazione dei risultati delle ECoC esprime chiaramente la necessità di ampliare l'indagine sugli impatti culturali, esplicitando alcune possibilità circa l'uso di nuove fonti di dati.

La ricerca alla base del presente articolo indaga gli impatti del programma ECoC sul benessere soggettivo delle persone coinvolte (utenti, cittadini, altri stakeholders): ma quali sono questi impatti? Come possiamo misurarli? Queste sono le domande a cui lo studio cerca di trovare una risposta. In questo contesto si inserisce questo articolo, introducendo un metodo innovativo che verrà spiegato nella sezione successiva

DOI 10.36158/97888929562234

Un nuovo metodo e nuove fonti di dati

Costruzione del database

Partendo dai presupposti spiegati nella sezione precedente e sulla base delle linee guida a livello europeo (European Commission, 2018, p. 8), sono stati individuati come fonte alternativa di dati primaria i contenuti sui social media.

Un filone di letteratura molto recente ha indagato come i contenuti postati sui social media – Twitter (Jaidka et al., 2020) e Facebook (Liu et al., 2015) in particolare – possano essere analizzati a partire dal linguaggio usato per ottenere una predizione del benessere individuale degli autori (Schwartz et al., 2016). In base al linguaggio usato nei post, si possono utilizzare tecniche e algoritmi non solo per catturare la percezione immediata (il "sentiment"), data dall'uso di parole con accezioni positive o negative (Basile & Nissim, 2013; Liu, 2012), ma anche per tentare di ricostruire pattern che – integrati con altri indicatori – possano predire il livello di benessere soggettivo degli utenti (Jaidka et al., 2020; Schwartz et al., 2016). I primi risultati ottenuti da queste analisi con l'utilizzo di tecniche data-driven, in particolare di Machine Learning, hanno a che fare con il benessere soggettivo degli utenti a livello regionale e nazionale in diverse zone del mondo (Mencarini et al., 2019; Schwartz et al., 2016), e si sono rivelati particolarmente incoraggianti (Jaidka et al., 2020). Esistono tuttavia ancora margini di sperimentazione e campi di applicazione inesplorati – come quello della cultura e delle ECoC presentato in questo articolo.

Un altro motivo per cui i contenuti postati sui social media possono essere una fonte di dati importante ha a che fare con la presenza – e il tentativo di minimizzare – i bias iniziali del ricercatore. Come accennato, l'utilizzo di tecniche qualitative di ricerca è molto vasto in campo culturale, ancora di più se si indaga la percezione degli impatti di iniziative culturali ed eventi, o la sfera personale degli stessi. Queste tecniche hanno alla base la scelta del ricercatore di porre le domande in un determinato modo per ottenere una risposta interessante, o di comporre un questionario in modo da poter ottenere tutte le informazioni che servono per la ricerca. Questa scelta implica che le risposte – all'intervista o al questionario – contengano un bias. La presenza di bias non è di per sé una cosa negativa, e, anzi, questi metodi sono molto diffusi e permettono di ottenere informazioni preziose. Tuttavia, in questo modo si rischia di non ottenere la percezione spontanea che l'intervistato o il rispondente avrebbe espresso a riguardo senza l'input del ricercatore. L'idea di usare i contenuti spontanei espressi sui social media – post o commenti – nasce anche dalla necessità di provare a minimizzare questi bias.

L'analisi dei contenuti presenti sui social media, per quanto rilevante nel misurare il benessere soggettivo, non basta.

In primo luogo, l'uso esclusivo di questi dati implica una visione parziale del fenomeno – racconta un solo punto di vista. L'iniziativa delle ECoC è invece molto sfaccettata, e la sua ideazione prima, e implementazione poi, è frutto dell'intreccio con altri punti di vista, che possono arricchire l'analisi e aumentare la rilevanza dei risultati.

In secondo luogo, nell'ottica più ampia di valutare una corrispondenza o divergenza tra risultati (in termini di benessere soggettivo) e obiettivi del programma, è necessario integrare i dati presi dai social media con altre fonti che esprimano con chiarezza le aspettative e gli effetti della Capitale oggetto dello studio.

Per tentare di costruire un metodo solido e un database ampio per l'analisi, che possa fornire informazioni su diversi ambiti del Programma ECoC, si è optato quindi per un allargamento delle fonti di dati. Il primo passo in questo senso è stata l'individuazione

DOI 10.36158/97888929562234

degli attori principali interessati dal Programma e alla riduzione della complessità fino a ottenere tre soggetti distinti, legate ognuna a un tipo di fonte scritta: gli utenti finali (con i già citati contenuti sui social media legati alle ECoC, in particolare post su Twitter e post e commenti su Facebook), le istituzioni (con la documentazione ufficiale della ECoC, da parte della città e delle istituzioni europee) e la stampa (con articoli di giornale, nazionali e locali, riguardanti le ECoC).

Nello specifico, la documentazione ufficiale prodotta dall'ente organizzatore restituisce informazioni utili su quelli che le istituzioni percepiscono essere gli impatti sul benessere soggettivo degli utenti delle ECoC, mentre gli articoli di giornale permettono di ottenere informazioni da una prospettiva mediata: da un lato, raccontano l'opinione della stampa su quanto percepito dagli utenti, dall'altra recepiscono anche il punto di vista delle istituzioni.

Alla base del metodo proposto quindi ci sono tre diversi tipi di fonti scritte, ognuna con caratteristiche specifiche e portatrici di contenuti diversi, analizzate con lo stesso approccio.

Ogni fonte diventa un corpus distinto, omogeneo internamente per caratteristiche dei testi contenuti, e una volta messi insieme creano un database eterogeneo di fonti scritte, con testi di formati e lunghezze diverse; i tre distinti corpora vengono analizzati con lo stesso approccio. La struttura del database è sintetizzata dalla tabella 1.

Tabella 1. Composizione del database per l'analisi testuale: corpora suddivisi per fonte.

Corpus A	Corpus B	Corpus C
Fonti istituzionali	**Fonti giornalistiche**	**Social media**
Documenti di candidatura Documenti di monitoraggio Documenti di valutazione	Articoli su stampa nazionale Articoli su stampa locale Articoli su stampa di settore	Post su pagine Facebook ufficiali Commenti a post su pagine Facebook ufficiali Tweet

Il Corpus A è costituito dalle fonti istituzionali, Per questo tipo di programma, sono previsti diversi tipi di documenti: preparatori, di monitoraggio in itinere, di valutazione ex-post. Queste tre fasi della ECoC sono cronologicamente conseguenti l'una all'altra, e spaziano in un arco temporale di circa 5 anni (dal bando ai report finali). In particolare, per le ECoC i documenti ufficiali sono sia quelli richiesti e prodotti su sollecitazione dalla Commissione Europea, sia quelli prodotti poi dall'ente organizzatore e dai suoi partner.

Nello specifico, si individuano in generale questi documenti:

- Bando: dove sono espressi i criteri di selezione delle ECoC e gli obiettivi generali da perseguire.
- Dossier di candidatura: documento in cui sono racchiusi gli obiettivi, la struttura di governance, le attività e in generale il progetto e la strategia della città per essere selezionata.
- Documento di selezione: In questo documento viene motivate la scelta di una città rispetto alle altre
- Documento di assegnazione: documento del ministero della cultura di assegnazione del titolo.
- Report di monitoraggio in itinere: Documenti intermedi prodotti nel periodo tra la selezione e l'implementazione per monitorare il completamento degli obiettivi preliminari esplicitati in fase di candidatura.

DOI 10.36158/97888929562234

- Dossier finali di valutazione: le città che diventano ECoC devono produrre una serie di documenti per la valutazione dei propri risultati e dei propri impatti sul territorio, in linea con gli obiettivi dichiarati.

Il Corpus B contiene le fonti giornalistiche, e include una sorta di rassegna stampa a più livelli. Nello specifico, si identificano articoli sulla ECoC in oggetto su:

- Giornali nazionali, specialmente con inserto culturale: notizie di rilievo nazionale, come ad esempio articoli sulla selezione, o sugli eventi di apertura o chiusura dell'anno di candidatura.
- Giornali locali: articoli sugli avvenimenti del territorio, anche se meno rilevanti dal punto di vista nazionale.
- Magazines di settore e riviste online: articoli specializzati sul settore culturale, con ampio rilievo dato ad eventi e politiche culturali.

Infine, il Corpus C è legato ai contenuti sui social media, e deve sottostare a una limitazione di tipo tecnico. In generale, l'accesso ai dati dei social media è stato oggetto di restrizioni dopo il 2018, con lo scandalo di Cambridge Analytica (Mancosu and Vegetti 2020). Dopo un iniziale periodo di chiusura totale, negli ultimi anni si è visto un ammorbidimento della posizione dei provider, a favore soprattutto della ricerca. Ogni piattaforma, tuttavia, ha le sue regole e i suoi limiti. Nel caso della ricerca alla base di questo articolo, si è scelto di utilizzare contenuti su Twitter (molto accessibile) e Facebook (solo post e commenti su pagine pubbliche). Dunque, il corpus in questo caso risulta formato da:

- Tweet di utenti diversi, liberamente postati sulla ECoC in oggetto;
- Post su Facebook della pagina/gruppo pubblico della ECoC in oggetto;
- Commenti (pubblici) ai post della/e pagina/e individuata/e.

Metodo di analisi

Il metodo che viene qui proposto per analizzare queste fonti prevede l'uso di tecniche di analisi testuale automatizzata, nello specifico basate su algoritmi di Machine Learning, per analizzare fonti documentarie riferite a una città individuata come caso studio. L'analisi automatizzata di testi più o meno lunghi è stata applicata con successo in diversi ambiti di ricerca, dalla letteratura alla storia (Di Maggio et al., 2013; Ferri et al., 2018), al management (Hannigan et al., 2019), alla finanza (Nguyen & Shirai, 2015). Si tratta di un campo di analisi molto ampio, in cui ricadono tecniche diverse che catturano informazioni e danno risultati molto diversi tra loro.

In questo caso, le due principali tecniche proposte sono topic modeling (Blei et al. 2013) e sentiment analysis (Basile & Nissim, 2013; Liu, 2012), con la possibilità di implementare alcune ulteriori analisi integrative e modelli predittivi (Jaidka et al., 2020; Schwartz et al., 2016) sulla base dei risultati delle prime analisi.

Il topic modeling è una tecnica analitica basata sulla statistica bayesiana (Ferri et al., 2018; Hannigan et al., 2019). Consente l'analisi di grandi volumi di testi analizzando iterativamente i testi e raggruppando le parole in base alle loro co-occorrenze (Ferri et al., 2018), cioè creando "topics", argomenti. Il più famoso algoritmo su cui si basa il topic modeling è chiamato Latent Dirichlet Allocation (LDA; Blei et al., 2003; Ferri et al., 2018; Di Maggio et al., 2013). L'algoritmo LDA si basa sulla statistica bayesiana e permette di inquadrare e

DOI 10.36158/97888929562234

annotare gli argomenti presenti nel testo in modo completamente automatizzato attraverso un modello probabilistico (Di Maggio et al., 2013; Ferri et al., 2018).

Le principali caratteristiche del topic modeling lo rendono uno strumento prezioso per la ricerca e l'analisi del testo (Di Maggio et al. 2013). In primo luogo, permette al ricercatore di analizzare grandi quantità di testi che sarebbero altrimenti impossibili da gestire per un essere umano. In secondo luogo, gli argomenti sono prodotti automaticamente senza bisogno di una categorizzazione a priori (il che significa meno pregiudizi nel processo di estrazione). In terzo luogo, "il topic modeling categorizza le parole, non i documenti" (Ferri et al., 2018, p. 176), cioè il modello riconosce i diversi significati della stessa parola in base al loro contesto (le altre parole che la circondano). Infine, sia gli argomenti che il processo sono espliciti, il che significa che la trasparenza del processo è estremamente elevata, e altri ricercatori possono riprodurre l'analisi o utilizzare il processo per adattarlo alle proprie domande di ricerca (Ferri et al., 2018), con un miglioramento anche dell'affidabilità della tecnica (DiMaggio et al., 2013).

Ciò che rende il topic modeling particolarmente interessante in questo caso è la possibilità di scoprire temi emergenti (non espliciti) rilevanti in diversi tipi di testo – scritti e concepiti con scopi diversi – rendendoli comparabili tra di loro.

La sentiment analysis è un tipo di analisi testuale ampiamente utilizzata per conoscere lo stato d'animo di chi scrive, e la sua opinione su un dato argomento (Liu, 2012; Mencarini et al., 2019). Con questo tipo di analisi è possibile identificare l'opinione dell'autore su un argomento dalle parole che usa, anche se il testo in questione è puramente descrittivo o impersonale. Questo è anche il motivo per cui viene chiamato "opinion mining" (Liu, 2012).

Nella sentiment analysis, le parole vengono identificate e categorizzate in base al "sentiment" che le sottende, cioè vengono viste come positive o negative in base al vocabolario di riferimento (Liu, 2012; Basile & Nissim, 2013). In breve, la sentiment analysis automatizzata si basa generalmente su un algoritmo che – dato un vocabolario predefinito di parole, costrutti e ricorrenze riconosciute dal computer – scompone il testo in pezzi più piccoli (token) e assegna a ciascuno di questi una valenza positiva, negativa o neutra (Basile & Nissim, 2013). Secondo la percentuale di termini positivi, negativi o neutri, il testo (o parti di esso) sarà considerato come positivamente, negativamente o neutralmente connotato. Questo tipo di analisi è molto importante per catturare le percezioni di un soggetto (istituzione, gruppo o individuo) su un dato argomento (Liu, 2012). Il linguaggio utilizzato è fondamentale per determinare la percezione (opinione) sottostante. Quindi, dato un testo, la sentiment analysis restituisce la percezione dell'autore, espressa – più o meno esplicitamente – in quel testo (Mencarini et al., 2019). Combinata con il topic modeling e ulteriori tecniche, completa i risultati e fornisce approfondimenti più precisi su come i diversi argomenti vengono trattati dai diversi attori.

Esistono anche tecniche non basate direttamente su vocabolari annotati di parole, ma su costrutti derivati dal "training" di un modello predittivo su altri testi (in genere un *corpus* molto grande di testi non correlati con la ricerca in corso), poi testato sul *corpus* dello specifico caso studiato (Jaidka et al., 2020) per trovare una corrispondenza. Si possono utilizzare raccolte di testi già preparati, oppure altri tipi di materiale scritto, come i risultati di sondaggi o interviste alla base di altre ricerche (a patto però che indaghino lo stesso argomento della ricerca in corso). Tramite l'uso di algoritmi specifici, si calcola la corrispondenza tra quanto emerge dall'analisi testuale automatizzata e risultati già "consolidati" (Schwartz et al., 2016), di fatto ottenendo una predizione del livello di benessere soggettivo nel *corpus* oggetto nella nuova ricerca.

DOI 10.36158/97888929562234

Questo metodo non ha la pretesa di risolvere tutti i problemi legati alla valutazione di impatto sociale in campo culturale, ma si pone come metodo alternativo o, meglio, complementare a quelle già in essere. Unita a una profonda analisi del contesto di riferimento, questo metodo ha il potenziale per svelare alcune tematiche non esplicite e restituire informazioni utili per l'analisi degli impatti del Programma sulla vita delle persone coinvolte.

Interpretazione dei risultati

Le tecniche brevemente presentate in questo articolo sono spesso usate separatamente, ma in questo caso sono pensate per essere complementari, con un duplice obiettivo.

Il primo obiettivo è quello, già menzionato, di individuare gli impatti sociali delle ECoC, in particolare in termini di benessere soggettivo. Una prima fase di analisi delle tematiche ricorrenti (con il topic modeling) restituisce la situazione in termini di argomento, evidenziando in maniera chiara di che cosa si stia parlando e in quale testo. La seconda fase è quella dell'analisi del linguaggio e della percezione (sentiment analysis), che evidenzia invece il modo in cui gli autori dei testi parlano degli argomenti individuati. Combinare le due tecniche serve come punto di partenza per individuare gli impatti sulle persone, gli utenti finali (Jaidka et al., 2020; Schwartz et al., 2016). Come anticipato nella sezione precedente, in base ai risultati preliminari delle due fasi è possibile costruire un modello predittivo del benessere soggettivo nel contesto della ECoC. Che sia attraverso l'uso di un training set basato su testi annotati riguardanti il benessere, o attraverso la correlazione dei risultati del topic modeling con i risultati di altre ricerche (questionari o interviste) (Schwartz et al., 2016), è possibile inferire il livello di benessere soggettivo espresso nei corpora che formano il database analizzato – e quindi avere un proxy per il benessere soggettivo degli individui coinvolti nella ECoC.

Il secondo obiettivo dell'analisi è quello di effettuare un confronto tra gli obiettivi dichiarati, i risultati attesi e i risultati effettivi del Programma, attraverso la comparazione dei contenuti prodotti dai diversi soggetti individuati, durante le diverse fasi del Programma. Il risultato dell'analisi di ogni corpus può essere confrontato con il risultato delle altre due, dando come prima informazione la diversa struttura tematica dei soggetti coinvolti, cioè quale soggetto si occupi di quali tematiche – e, successivamente, come. Divergenze tra i corpora possono significare una diversa visione delle tematiche affrontate, oltre che una differenza nel linguaggio adottato e nel sentiment sottostante. Per poter però arrivare a un confronto tra obiettivi e risultati è necessario prendere in considerazione anche la variabile temporale dell'evento. Come visto nella sezione dedicata alla descrizione delle fonti e nella tabella 1, le ECoC sono un evento a più fasi, dalla candidatura si passa alla fase di progettazione e poi all'implementazione degli eventi. Una distribuzione temporale delle fonti secondo queste fasi permette di avere un'analisi più precisa dei contenuti. Dalle fonti istituzionali possiamo avere un'idea degli obiettivi dichiarati in fase di candidatura, ma anche dei risultati della fase di progettazione, e dei risultati finali così come sono raccontati nei documenti ufficiali. Dalle fonti giornalistiche è possibile ottenere il racconto sia degli obiettivi che dei risultati, grazie alla datazione degli articoli, "mediata" dalla percezione della stampa. Infine, dai social media è possibile ottenere quelli che vengono percepiti come i risultati del Programma dagli utenti.

Una volta ottenuti i risultati delle diverse analisi, è necessario interpretarli. Per poter interpretare i dati correttamente – o meglio, il più correttamente possibile – è necessario

DOI 10.36158/97888929562234

integrarli con altri tipi di informazioni. Le informazioni integrative sono fornite innanzitutto dall'analisi approfondita del contesto dell'ECoC di riferimento, per esempio lo studio di fonti secondarie, come letteratura accademica, interventi in eventi pubblici, ma anche interviste con attori del territorio. Altre informazioni importanti possono arrivare dall'analisi di indicatori socioeconomici e demografici sul benessere relativi al territorio studiato – facendo attenzione, però, alla diversa concezione di benessere alla base di ogni indicatore. Infine, molto utili possono essere le informazioni contenute nei metadati dei contenuti pubblicati sui social media (geografici, per esempio, ma anche relativi all'engagement degli utenti e al loro "consenso").

Questa integrazione è dunque fondamentale per l'interpretazione dei risultati e il loro inserimento nel contesto della ECoC di riferimento.

Criticità e possibili soluzioni

Quello qui proposto è un metodo ancora di fatto sperimentale, attualmente in fase di applicazione ai casi studio di una ECoC, Matera 2019. Nella messa a punto sono state individuate alcune criticità a cui, con il tempo, è stato possibile trovare alcune possibili soluzioni.

Una prima criticità, parlando di ECoC, è quella della lingua. Essendo un Programma di livello europeo, ogni città ospitante ha la sua lingua; pertanto, molti dei contenuti prodotti nel contesto della ECoC saranno scritti in quella lingua; ma non solo: gran parte dei documenti della Commissione Europea, tra cui il Bando per la candidatura e i report di monitoraggio, sono redatti in lingua inglese. In questo caso, dunque, vi sarebbe una duplice necessità. Da un lato, la necessità di dividere, all'interno del corpus delle fonti istituzionali, i documenti in base alla lingua e analizzarli separatamente. Dall'altra, sebbene siano stati sviluppati tecniche e programmi che permettono l'analisi di quasi tutte le lingue, la conoscenza del ricercatore della lingua di riferimento, ma anche del contesto e delle regole sintattiche è fondamentale. Si tratta di una criticità di cui tenere conto, per esempio nella scelta del caso da studiare.

Una seconda criticità, di tipo interpretativo, è quella dell'individuazione precisa dei punti di vista, che potrebbe non coincidere totalmente con la divisione delle fonti per tipologia ipotizzata in questo articolo. Un esempio di questo è la possibilità che il punto di vista delle istituzioni sia espresso non solo attraverso i documenti ufficiali, ma anche tramite comunicati stampa (che sono raccolti insieme alle fonti giornalistiche) o sulle pagine ufficiali sui social media (raccolti in un corpus a parte insieme ai contenuti degli altri utenti non istituzionali). Questa mancata corrispondenza potrebbe creare dei problemi nella fase di interpretazione dei dati. Una possibile soluzione a questo problema è una ripartizione diversa dei testi nei corpora, ovvero non per tipologia di fonte (istituzionale, giornalistica, social media), ma per autore. Si tratterebbe quindi di non intendere i corpora come raggruppamento di fonti dello stesso tipo, ma come espressione dello stesso punto di vista. In pratica, nel corpus A delle fonti istituzionali verrebbero raccolti i documenti ufficiali, i comunicati stampa e i post delle pagine Facebook ufficiali delle ECoC; nel corpus B delle fonti giornalistiche si avrebbero solo gli articoli di giornale (un cambiamento minimo rispetto alla struttura iniziale); infine, nel corpus C dei social media verrebbero mantenuti solamente i commenti degli utenti ai post delle pagine ufficiali su Facebook (ora spostati nel corpus A) e i tweet. Questa nuova divisione è sintetizzata nella tabella 2.

DOI 10.36158/97888929562234

Tabella 2. Composizione del database per l'analisi testuale: corpora suddivisi per punto di vista.

Corpus A	Corpus B	Corpus C
Punto di vista istituzionale	**Punto di vista della stampa**	**Punto di vista dei fruitori**
Documenti ufficiali: Documenti di candidatura Documenti di monitoraggio Documenti di valutazione Comunicati stampa istituzionali Post su pagine Facebook ufficiali	Articoli su stampa nazionale Articoli su stampa locale Articoli su stampa di settore	Commenti a post su pagine Facebook ufficiali Tweet

Procedendo in questo modo si otterrebbe un'assegnazione più corretta dal punto di vista contenutistico dei testi al soggetto/punto di vista, ma si aumenterebbe la complessità e l'eterogeneità interna ai singoli corpora. In questo caso potrebbe essere più semplice anche procedere con l'analisi "per fasi" dell'ECoC, dalla candidatura, all'implementazione, all'anno dell'evento. Si tratta di un trade-off tra la correttezza contenutistica e la praticabilità dell'analisi, ancora tutto da verificare.

Una terza criticità è data dalla presenza limitazioni nella raccolta dei dati, che potrebbero rendere difficile ottenere un campione significativo. Questo è il caso della raccolta di dati dai social media. Come accennato nella sezione precedente, l'accesso ai dati dei social media è stato oggetto di restrizioni piuttosto severe dopo lo scandalo di Cambridge Analytica nel 2018 (Mancosu & Vegetti, 2020), soprattutto per quanto riguarda Facebook. Su insistenza della comunità accademica la posizione delle parti lese si è andata ammorbidendo, fino a una parziale riapertura all'uso dei dati della piattaforma per scopi di ricerca. Tuttavia, rimangono delle restrizioni a ciò che è consentito scaricare, e i dati ottenuti devono essere anonimizzati a norma di GDPR per evitare la profilazione degli utenti e garantirne la privacy. Ogni piattaforma, comunque, conserva le proprie particolarità in termini di regole e limitazioni. Nel caso specifico qui proposto, Facebook e Twitter hanno politiche diverse per l'accesso e l'ottenimento di dati su post e commenti.

Nel caso di Facebook, per ottenere dati testuali oggi è necessario avvalersi di Facepager[9] (Jünger & Keyling, 2019), un programma che permette di scaricare post e commenti pubblici di una determinata pagina, in un range temporale stabilito dal ricercatore tra i parametri di ricerca. Per l'analisi dei metadati, invece, è possibile richiedere l'accesso a CrowdTangle[10], un'estensione creata da Facebook (oggi Meta) per aiutare i ricercatori a fruire dei dati presenti sulla piattaforma, senza venir meno alle cautele imposte dagli eventi del passato.

Twitter, invece, è più accessibile. Ci sono diversi canali che permettono ai ricercatori di ottenere il testo di tweet pubblici, a partire proprio dall'API (Application Programming Interface) di Twitter che permette, in sintesi, ai ricercatori registrati con un account "Developer"[11] di fare ricerche per parole chiave e di scaricare facilmente i contenuti per le proprie analisi. Un'alternativa è usare anche in questo caso Facepager, o avvalersi di database già esistenti (Mencarini et al., 2019; Basile & Nissim, 2013).

9. Si veda: https://github.com/strohne/Facepager/releases.
10. Si veda: https://www.crowdtangle.com.
11. Si veda: https://developer.twitter.com/en/products/twitter-api/academic-research.

DOI 10.36158/97888929562234

Limiti del metodo

Nella sezione precedente sono state elencate alcune criticità riscontrate nella costruzione del database, e le soluzioni adottate nell'ambito della ricerca in corso. Restano tuttavia aperti alcuni punti sui quali è necessario fare ulteriori riflessioni.

Una cosa che, almeno per ora, non è possibile fare è una distinzione, all'interno dello stesso corpus, tra "voci" diverse. Ad esempio, all'interno del corpus legato alla percezione degli utenti (Corpus C) non è possibile al momento distinguere tra la percezione dei cittadini e degli abitanti della città diventata ECoC, e gli altri – in questo caso fruitori occasionali e turisti (Ap, 1990; Balduck et al., 2011). Una distinzione di questo tipo non è strettamente necessaria perché l'analisi funzioni e produca risultati interessanti ai fini della ricerca. Secondo la Commissione Europea (2017), i destinatari degli impatti previsti dal Programma sono infatti sia le comunità locali (in particolare tramite la partecipazione e il coinvolgimento attivo nell'organizzazione e fruizione degli eventi) sia i visitatori e in generale la comunità europea, in una logica di condivisione, interazione e scambio reciproci. Perciò non è sbagliato analizzare gli impatti sugli utenti senza distinzioni, perché questi ricadono sia sui cittadini che sui fruitori in generale. Tuttavia, una distinzione di questo tipo potrebbe rivelarsi utile per studi più specifici e per comprendere meglio gli impatti degli eventi – o mega-eventi (Németh, 2016) – sulle comunità ospitanti (Ap, 1990) e fornire spunti per lo sviluppo locale. Una potenziale soluzione a questo problema è l'utilizzo dei metadati. I metadati completi di solito includono sia la geolocalizzazione del contenuto che le informazioni sull'utente. Da questi dati potrebbe essere possibile filtrare i risultati per includere solo i membri della comunità locale (ad esempio, selezionare i record/utenti che hanno un id di localizzazione permanente nella città o che si siano registrati nella stessa località per un determinato periodo di tempo) Questo tipo di soluzione, tuttavia, può essere complicata e si basa sulla qualità dei dati raccolti. Non tutti i dati scaricati da Facebook o Twitter hanno un geotag, ad esempio, e non tutti i dati registrati dagli utenti sono verificabili (Acker, 2018). Ciò significa che individuare i membri della comunità ospitante attraverso i metadati potrebbe non essere tecnicamente possibile – e non per ragioni imputabili al ricercatore. Si tratta di un problema complesso, che rimane tuttora aperto.

Un secondo limite è legato alla natura stessa dei social media. Come accennato, parte del motivo per cui si è scelto di utilizzare i social media come fonte primaria di dati per questo metodo è il tentativo di superare alcuni bias tipici delle analisi qualitative. Tuttavia, i social media stessi non sono esenti da altri bias. Un primo problema è che i social media – Twitter, in particolare – non sono totalmente rappresentativi della popolazione (Iacus et al., 2020). Questo significa che la composizione demografica degli utenti che usano Twitter non rispecchia necessariamente tutte le caratteristiche presenti nel contesto studiato. In secondo luogo, bisogna tenere in considerazione la selection bias (Iacus et al., 2020), cioè in questo caso il fatto che non tutti gli utenti dei social media li utilizzino commentando o creando contenuti nello stesso modo, con utenti più attivi e altri invece più schivi, che rischiano di autoescludersi dal campione. In terzo luogo, i dati raccolti dai social media – di nuovo, in particolare da Twitter – vengono selezionati attraverso una ricerca per parole chiave (Basile & Nissim, 2013). Il rischio, in questo caso, è quello di escludere a priori alcuni contenuti sulla base di una selezione troppo stringente dei termini più significativi.

Oltre a questi limiti tecnici, ve ne è uno procedurale. Affinché i risultati dell'analisi abbiano senso, è necessario bisogna procedere per fasi ben ordinate. Prima di tutto, si

DOI 10.36158/97888929562234

procede con l'analisi del Corpus C, che restituisce la linea base della percezione del benessere soggettivo da parte degli utenti. Dopodiché si passa all'analisi e al confronto dei Corpora A e B. A questo punto, si ha la risposta al primo obiettivo, ovvero quello di valutare l'impatto della ECoC sul benessere soggettivo. Al termine di questo primo "pacchetto" di analisi, si può passare al secondo obiettivo, quello cioè di comparare gli obiettivi e i risultati. Questa precisa sequenza di passaggi può essere vista come un limite, in quanto non è consigliabile saltare un passaggio per arrivare a un risultato prima che a un altro che può rivelarsi più interessante dal punto di vista della ricerca. È inoltre opportuno ricordare che i risultati dei singoli passaggi – così come quelli finali – devono sempre essere interpretati; quindi, può rivelarsi necessario utilizzare altre fonti per integrare la propria conoscenza del contesto e giungere alle corrette conclusioni.

Conclusioni

Nel corso di questo articolo si è visto come, a partire dalla letteratura degli ultimi decenni, sia emersa la necessità di portare innovazione nella pratica della valutazione degli impatti sociali delle iniziative culturali, portando l'esempio delle ECoC.

Attraverso l'analisi delle diverse tematiche affrontate dagli studiosi e dei loro metodi di analisi, è emerso come le metodologie attualmente applicate in questo campo presentino alcune limitazioni, o siano usate principalmente per studiare alcune tematiche a discapito di altre. Uno degli impatti meno studiati è quello sul benessere soggettivo nel contesto delle ECoC (Steiner et al., 2015), che risulta dal canto suo un ambito particolarmente adatto per sfruttare nuove fonti di dati e nuove metodologie di ricerca. Partendo dalla letteratura recente, si è poi mostrato come un metodo promettente per investigare questa tematica sia partire dai contenuti espressi sui social media – Twitter, Facebook – per comprendere, attraverso testi prodotti spontaneamente, la percezione degli impatti da parte degli utenti e poter successivamente sviluppare modelli predittivi del benessere soggettivo (Basile & Nissim, 2013; Mencarini et al., 2019; Jaidka et al., 2020; Schwartz et al., 2016) in un contesto specifico, delimitato chiaramente dalla scelta di un determinato evento. Poter indagare la tematica del benessere soggettivo attraverso l'uso dei social media renderebbe la fase di raccolta e di analisi dei dati altamente efficiente dal punto di vista del consumo di tempo e risorse, superando uno dei limiti principali delle ricerche qualitative fatte finora in questo campo (Jaidka et al., 2020).

Nel contesto delle ECoC è tuttavia importante considerare anche altri punti di vista, e di conseguenza, in questo caso, altre fonti. Il metodo alternativo proposto è dunque basato sull'analisi di fonti testuali di diversa natura – documenti ufficiali della ECoC, articoli di giornale e contenuti pubblici sui social media – attraverso tecniche automatizzate derivanti dal Machine Learning e, in generale, dall'analisi del linguaggio (NLP) (Blei et al., 2013; Liu, 2012). In particolare, l'uso combinato di topic modeling, sentiment analysis e modelli predittivi (Ferri et al., 2018; Jaidka et al., 2020; Liu, 2012), porta a un duplice risultato. Da un lato, investigare l'impatto della ECoC sul benessere soggettivo nel suo contesto di riferimento; dall'altro, operare un confronto tematico e linguistico tra le fonti per determinare una convergenza o una divergenza tra obiettivi dichiarati, risultati attesi e risultati ottenuti.

Questo articolo nello specifico ha tentato di spiegare come si possa partire da un problema noto (l'analisi di impatto sociale di una ECoC) e usare fonti e tecniche alternative digitali per ampliare le sue potenzialità. Questo tentativo propone, di fatto, una tec-

DOI 10.36158/97888929562234

nica sperimentale, che parte da una commistione di fonti per ora inedita – al contempo in linea con le Linee Guida per l'autovalutazione delle ECoC espresse dalla Commissione Europea (2018) per il periodo 2020-2033. Si tratta di una prima esposizione di un progetto di ricerca tutt'ora in corso.

Un'ulteriore peculiarità del metodo qui introdotto è l'applicazione di tecniche di Machine Learning in un contesto nuovo rispetto a quelli in cui sono solitamente applicate, contribuendo al dibattito e alla letteratura sugli impatti sociali delle ECoC e delle politiche culturali.

Infine, nelle intenzioni dello studio di cui è parte questo articolo, questo metodo e il suo utilizzo sono il primo passo verso la creazione di un modello di analisi che potrebbe idealmente essere applicato a qualsiasi evento con un impianto simile alle ECoC, normalmente studiato con tecniche più tradizionali. Possono essere eventi di natura culturale – come festival, concerti, grandi mostre – o di altra natura – come Olimpiadi, Mondiali di calcio, grandi eventi.

Seppur con alcune differenze (ad esempio, il tipo di documentazione ufficiale prodotta, la scala, la lingua, le caratteristiche del contesto) e limiti, il metodo qui proposto permetterebbe di esplorare nuove possibilità di analisi e contribuire a innovare il dibattito sull'impatto di questi cosiddetti "mega-eventi" (Langen & Garcia, 2009; Németh, 2016), aprendo le porte a potenziali nuovi usi di fonti già note e nuove idee per ricerche future nell'ambito degli effetti delle politiche pubbliche.

Bibliografia

Acker, A. (2018). Data craft: The manipulation of social media metadata. *Data & Society Research Institute*, *13*.

Ap, J. (1990). Residents' perceptions research on the social impacts of tourism. *Annals of tourism research*, *17*(4), 610-616.

Dimitris, B. (2013). What makes a'happy city'? *Cities-The International Journal of Urban Policy and Planning*, *32*, S39-S50. https://doi.org/10.1016/j.cities.2013.04.009.

Basile, V., & Nissim, M. (2013). Sentiment analysis on Italian tweets. In *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*.

Belfiore, E. (2002). Art as a means of alleviating social exclusion: Does it really work? A critique of instrumental cultural policies and social impact studies in the UK. *International journal of cultural policy*, *8*(1), 91-106. https://doi.org/10.1080/102866302900324658

Bianchini, F. (2013). The regenerative impacts of the European City/Capital of Culture events. In *The Routledge companion to urban regeneration* (pp. 535-545). Routledge.

Biondi, L., Demartini, P., Marchegiani, L., Marchiori, M., & Piber, M. (2018). The outreach of participatory cultural initiatives: The importance of creating and exchanging knowledge. *IFKAD Book of Proceedings*.

Biondi, L., Demartini, P., Marchegiani, L., Marchiori, M., & Piber, M. (2020). Understanding orchestrated participatory cultural initiatives: Mapping the dynamics of governance and participation. *Cities*, *96*, 102459. https://doi.org/10.1016/j.cities.2019.102459

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022. https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf?ref=https://githubhelp.com (accessed on March 14, 2022).

Blessi, G. T., Grossi, E., Sacco, P. L., Pieretti, G., & Ferilli, G. (2016). The contribution of

DOI 10.36158/97888929562234

cultural participation to urban well-being. A comparative study in Bolzano/Bozen and Siracusa, Italy. *Cities*, *50*, 216-226. https://doi.org/10.1016/j.cities.2015.10.009.

Boland, P. (2010). 'Capital of Culture – you must be having a laugh!' Challenging the official rhetoric of Liverpool as the 2008 European cultural capital. *Social & Cultural Geography*, *11*(7), 627-645. https://doi.org/10.1080/14649365.2010.508562.

Bracalente, B., Chirieleison, C., Cossignani, M., Ferrucci, L., Gigliotti, M., & Ranalli, M. G. (2011). The economic impact of cultural events: The Umbria Jazz music festival. *Tourism Economics, 17*(6), 1235-1255. https://doi.org/10.5367/te.2011.0096.

Campbell, P. (2011). Creative industries in a European Capital of Culture. *International Journal of Cultural Policy*, *17*(5), 510-522. https://doi.org/10.1080/10286632.2010.543461.

Cicerchia, A., & Bologna, E. (2017). Salute, benessere e partecipazione culturale: tra narrazioni e indicatori. *Economia della cultura*, *27*(2), 313-322.

Collins, T. (2016). Urban civic pride and the new localism. *Transactions of the Institute of British Geographers*, *41*(2), 175-186. https://doi.org/10.1111/tran.12113.

Cox, T., & O'Brien, D. (2012). The "scouse wedding" and other myths: reflections on the evolution of a "Liverpool model" for culture-led urban regeneration. *Cultural Trends*, *21*(2), 93-101. https://doi.org/10.1080/09548963.2012.674749.

Demartini, P., Marchiori, M., & Marchegiani, L. (2018). Citizen engagement as a criterion to select the European capital of culture: A critical appraisal. In *13th international forum on knowledge asset dynamics (IFKAD) proceedings 2018*.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics*, *41*(6), 570-606. https://doi.org/10.1016/j.poetic.2013.08.004.

European Commission (2017). *European Capitals of Culture (ECoC) 2020-2033: A guide for cities preparing to bid*.

European Commission (2018). *European Capitals of Culture (ECoC) 2020-2033: Guidelines for the cities' own evaluations of the results of their ECOC*. https://culture.ec.europa.eu/sites/default/files/2021-04/ecoc-guidelines-for-cities-own-evaluations-2020-2033.pdf (accessed on March 14, 2022).

Ferri, P., Lusiani, M., & Pareschi, L. (2018). Accounting for Accounting History: A topic modeling approach (1996-2015). *Accounting History*, *23*(1-2), 173-205. https://doi.org/10.1177/1032373217740707.

Fišer, Žilič S., & Kožuh, I. (2019). The impact of cultural events on community reputation and pride in Maribor, The European Capital of Culture 2012. *Social Indicators Research*, *142*(3), 1055-1073. https://doi.org/10.1007/s11205-018-1958-4.

Fitjar, R. D., Rommetvedt, H., & Berg, C. (2013). European Capitals of Culture: elitism or inclusion? The case of Stavanger2008. *International Journal of Cultural Policy*, *19*(1), 63-83. https://doi.org/10.1080/10286632.2011.600755.

Garbarino, S., & Holland, J. (2009). Quantitative and qualitative methods in impact evaluation and measuring results. http://www.gsdrc.org/docs/open/EIRS4.pdf.

Goodman, F. R., Disabato, D. J., Kashdan, T. B., & Kauffman, S. B. (2018). Measuring well-being: A comparison of subjective well-being and PERMA. *The Journal of Positive Psychology*, *13*(4), 321-332. https://doi.org/10.1080/17439760.2017.1388434.

Grossi, E., Tavano Blessi, G., Sacco, P. L., & Buscema, M. (2012). The interaction between culture, health and psychological well-being: Data mining from the Italian culture and well-being project. *Journal of Happiness Studies*, *13*(1), 129-148. https://doi.org/10.1007/s10902-011-9254-x.

Grossi, E., Sacco, P. L., Blessi, G. T., & Cerutti, R. (2011). The impact of culture on the

DOI 10.36158/97888929562234

individual subjective well-being of the Italian population: An exploratory study. *Applied research in quality of life*, *6*(4), 387-410. https://doi.org/10.1007/s11482-010-9135-1.

Hannigan, T. R., Haans, R. F., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., & Jennings, P. D. (2019). Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, *13*(2), 586-632. https://doi.org/10.5465/annals.2017.0099.

Iacus, S. M., Porro, G., Salini, S., & Siletti, E. (2020). Controlling for selection bias in social media indicators through official statistics: A proposal. http://dx.doi.org/10.2478/JOS-2020-0017.

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences*, *117*(19), 10165-10171. 10165-10171. https://doi.org/10.1073/pnas.1906364117.

Jagodzińska, K., Sanetra-Szeliga, J., Purchla, J., Van Balen, K., Thys, C., Vandesande, A., & Van der Auwera, S. (2015). Cultural Heritage Counts for Europe: Full Report.

Jünger, J., & Keyling, T. (2019). Facepager. *An application for automated data retrieval on the web. Facepager. An application for generic data retrieval through APIs. Source code and releases available*. Accessed on March 14, 2022: https://www.univie.ac.at/digitalmethods/wp-content/uploads/2013/10/Facepager_2013.pdf.

Lähdesmäki, T. (2013). Cultural activism as a counter-discourse to the European Capital of Culture programme: The case of Turku 2011. *European journal of cultural studies*, *16*(5), 598-619. https://doi.org/10.1177/1367549413491720.

Langen, F., & Garcia, B. (2009). *Measuring the impacts of large scale cultural events: a literature review*. Liverpool: Impacts 08. https://www.liverpool.ac.uk/media/livacuk/impacts08/pdf/pdf/Impacts08-FLangen_and_BGarcia_May_2009_Events_Review.pdf (accessed on March 14, 2022).

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Liu, P., Tov, W., Kosinski, M., Stillwell, D. J., & Qiu, L. (2015). Do Facebook status updates reflect subjective well-being?. *Cyberpsychology, Behavior, and Social Networking*, *18*(7), 373-379. https://doi.org/10.1089/cyber.2015.0022

Mancosu, M., & Vegetti, F. (2020). What you can scrape and what is right to scrape: A proposal for a tool to collect public Facebook data. *Social Media+ Society*, *6*(3). https://doi.org/10.1177/2056305120940703.

Mencarini, L., Hernández-Farías, D. I., Lai, M., Patti, V., Sulis, E., & Vignoli, D. (2019). Happy parents' tweets. *Demographic Research*, *40*, 693-724. https://www.jstor.org/stable/26727014.

Nagy, S. (2018). Framing culture: Participatory governance in the European capital of culture programme. *Participations*, *15*(2), 243-262.

Németh, Á. (2016). European capitals of culture-digging deeper into the governance of the mega-event. *Territory, Politics, Governance*, *4*(1), 52-74. https://doi.org/10.1080/21622671.2014.992804.

Nguyen, T. H., & Shirai, K. (2015, July). Topic modeling based sentiment analysis on social media for stock market prediction. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (volume 1: Long Papers)* (pp. 1354-1364).

Norman, W., & MacDonald, C. (2004). Getting to the bottom of "triple bottom line". *Business ethics quarterly*, *14*(2), 243-262. doi:10.5840/beq200414211.

DOI 10.36158/97888929562234

O'Brien, D. (2011). Who is in charge? Liverpool, European Capital of Culture 2008 and the governance of cultural planning. *Town Planning Review*, *82*(1), 45-60. http://www.jstor.org/stable/27975979.

Oishi, S. (2018). Culture and subjective well-being: Conceptual and measurement issues. *Handbook of well-being. Salt Lake City, UT: DEF Publishers*. doi:nobascholar.com.

Ooi, C. S., Håkanson, L., & LaCava, L. (2014). Poetics and politics of the European Capital of Culture Project. *Procedia-Social and Behavioral Sciences*, *148*, 420-427. https://doi.org/10.1016/j.sbspro.2014.07.061.

Quinn, B. (2009). The European capital culture initiative and cultural legacy: an analysis of the cultural sector in the aftermath of Cork 2005. *Event management*, *13*(4), 249-264. https://doi.org/10.3727/152599510X12621081189077.

Rayman-Bacchus, L., & Radavoi, C. N. (2020). Advancing culture's role in sustainable development: social change through cultural policy. *International Journal of Cultural Policy*, *26*(5), 649-667. https://doi.org/10.1080/10286632.2019.1624735.

Richards, G. (2000). The European cultural capital event: Strategic weapon in the cultural arms race?. *International Journal of Cultural Policy*, *6*(2), 159-181. https://doi.org/10.1080/10286630009358119.

Richards, G., & Wilson, J. (2004). The impact of cultural events on city image: Rotterdam, cultural capital of Europe 2001. *Urban studies*, *41*(10), 1931-1951. https://doi.org/10.1080/0042098042000256323.

Schwartz, H. A., Sap, M., Kern, M. L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., & Ungar, L. H. (2016). Predicting individual well-being through the language of social media. In *Biocomputing 2016: Proceedings of the Pacific Symposium* (pp. 516-527). https://doi.org/10.1142/9789814749411_0047.

Seligman, M. (2018). PERMA and the building blocks of well-being. *The journal of positive psychology*, *13*(4), 333-335. https://doi.org/10.1080/17439760.2018.1437466.

Steiner, L., Frey, B., & Hotz, S. (2015). European capitals of culture and life satisfaction. *Urban studies*, *52*(2), 374-394. https://doi.org/10.1177/0042098014524609.

UNESCO (2019). *Culture | 2030 Indicators*.

Voukelatou, V., Gabrielli, L., Miliou, I., Cresci, S., Sharma, R., Tesconi, M., & Pappalardo, L. (2021). Measuring objective and subjective well-being: dimensions and data sources. *International Journal of Data Science and Analytics*, *11*(4), 279-309. https://doi.org/10.1007/s41060-020-00224-2.

Wills-Herrera, E., Islam, G., & Hamilton, M. (2009). Subjective well-being in cities: A multidimensional concept of individual, social and cultural variables. *Applied Research in Quality of Life*, *4*(2), 201-221. https://doi.org/10.1007/s11482-009-9072-z.

DOI 10.36158/97888929562234