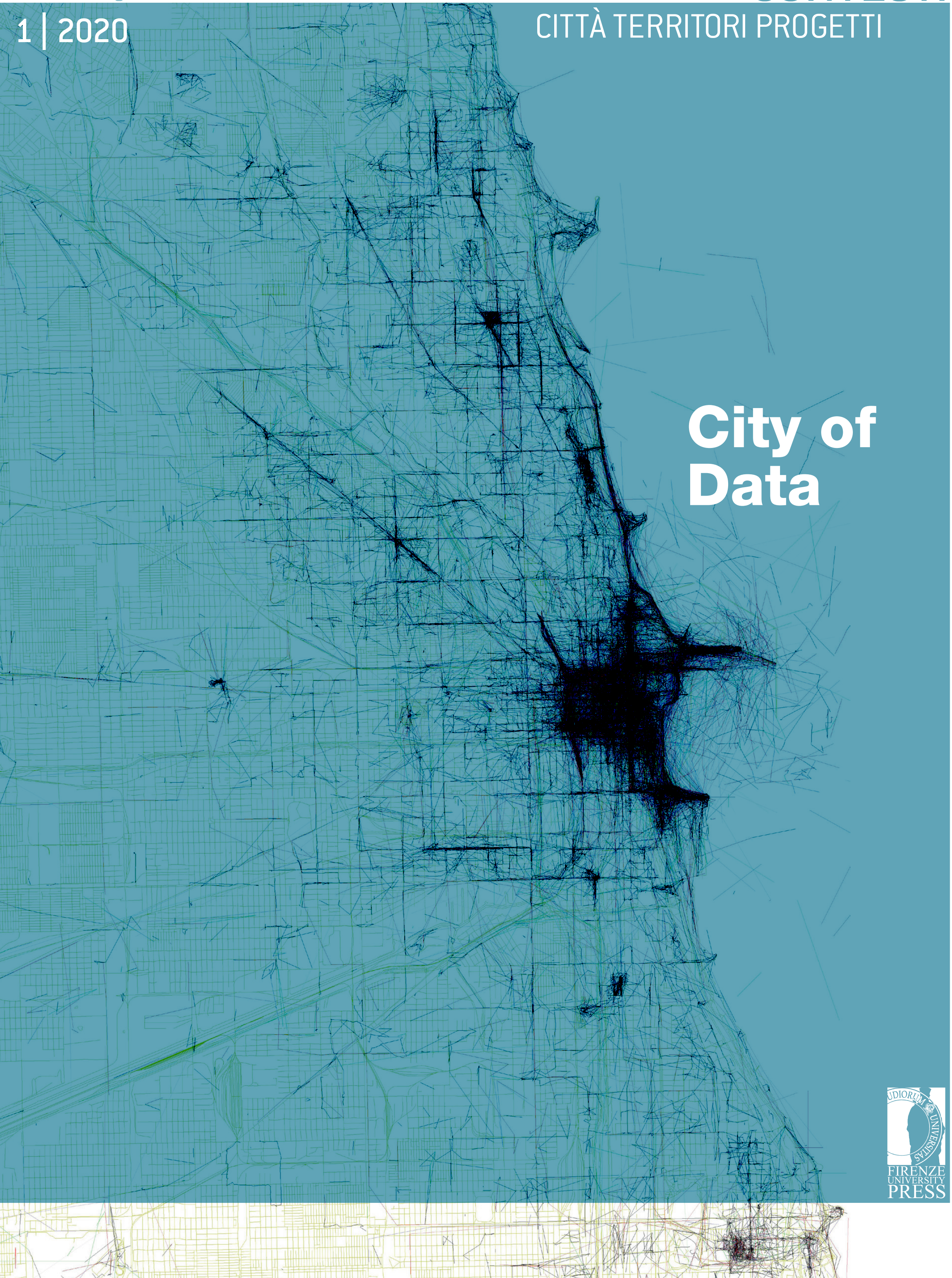
Rivista di Urbanistica e
Pianificazione del Territorio
Università degli Studi di Firenze

CONTESTI

CITTÀ TERRITORI PROGETTI

1 | 2020



City of Data

FIRENZE UNIVERSITY PRESS

CONTESTI
CITTÀ TERRITORI PROGETTI

Rivista di Urbanistica e
Pianificazione del Territorio
Università degli Studi di Firenze

UNIVERSITÀ
DEGLI STUDI
FIRENZE

UNIVERSITÀ DEGLI STUDI FIRENZE
DIDA DIPARTIMENTO DI ARCHITETTURA

CONTESTI
CITTÀ TERRITORI PROGETTI

1 | 2020
Firenze University Press
ISSN 2035-5300

Contatti
Contesti. Dipartimento di Architettura
Via della Mattonaia 8, 50121, Firenze, Italy
contesti@dida.unifi.it

progetto grafico
didacommunicationlab
Dipartimento di Architettura
Università degli Studi di Firenze

CITY OF DATA

SOMMARIO

City of Data

Carlo Pisano
Fabio Lucchesi
DIDA, Università degli Studi di Firenze
carlo.pisano@unifi.it
fabio.lucchesi@unifi.it

Firenze University Press.
DOI: 10.13128/contest-12101
www.fupress.net/index.php/contesti/

The amount of data generated by our daily activities and interactions is constantly increasing, as is the number of digital devices continuously monitored by increasingly complex and accurate satellite instruments. This condition provides information useful for making high-quality choices on multiple aspects of contemporary life.
This amount of information has driven numerous researchers and institutions to examine this revolution, starting from the hypothesis that, with the increasing complexity of the metropolis and the contemporary lifestyle, today Data appears to be the most reliable tool for providing an accurate view of both social interactions and the environment that staged those interactions.
The various research trajectories, proposed in this issue of Contesti, pursue this hypothesis, trying to provide operational tools and trace future lines of research.

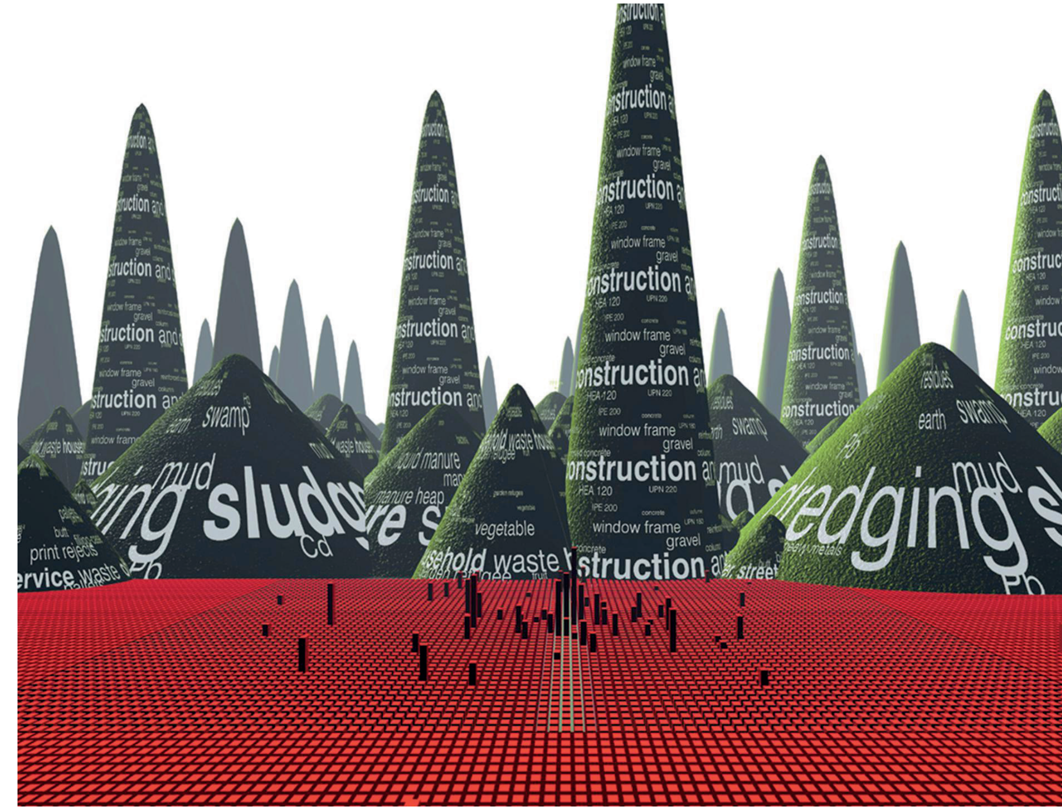
Sul finire dello scorso millennio Winy Maas, capo dello studio MVRDV, pubblicava un breve testo sul progetto Metacity/Datatown. In quel contesto, l'architetto olandese descriveva la futuristica Datatown come una città fondata esclusivamente sui dati; una città descrivibile solo attraverso il flusso di informazioni; una città che non conosce topografia, nessuna ideologia, nessuna rappresentazione e nessun contesto. Solo enormi e puri dati (Maas, 1999). A partire da queste provocazioni, Maas si domandava quali fossero le implicazioni prodotte da questa città e quali ipotesi progettuali e agende strategiche potessero scaturire da questo approccio numerico.
Nonostante il carattere provocatorio del testo e delle immagini proposte, certamente il lavoro di Maas ha rappresentato uno dei primi tentativi di far penetrare, nei regni dell'architettura e dell'urbanistica, il tema emergente dei Big Data e di manifestare il loro impatto sul mondo contemporaneo.
La Datatown trova, forse, in *Notes on the Synthesis of Form* uno dei suoi antenati più illustri. Verso la metà degli anni Sessanta Christopher Alexander (1964), infatti, stimolò il panorama delle teorie architettoniche e urbanistiche attraverso la pubblicazione di un libro in cui proponeva di mettere al centro del progetto non più l'intuizione individuale, l'atto creativo, il momento di forza o fermento di rottura dello *status quo*, ma la riflessione deduttiva resa possibile dall'uso del calcolatore elettronico, strumento a quel tempo ancora molto lontano dal dibattito quotidiano.
Situandosi nella diretta continuazione del neopositivismo e facendo riferimento ai lavori sul ruolo del linguaggio di Carnap (1947) e Wittgenstein (1953), Alexander propone di inventare un nuovo modo di descrizione e progettazione analitica utilizzando gli strumenti, nuovi per l'architetto, della logica, della matematica e dell'elaborazione delle informazioni formalizzate tramite il calcolatore, inserite nel paradigma cibernetico-sistemista e più generalmente delle epistemologie funzionaliste (Gerosa, 2002).
Alexander propose la scomposizione del problema progettuale in problemi più coerenti dal punto di vista analitico, trasferendo parte del processo di progettazione alle macchine, prefigurando scenari di lavoro ormai entrati nella nostra contemporaneità e che hanno permesso, con molta probabilità, l'inizio di una nuova epoca urbana.

In poco più di un ventennio dal lavoro di Maas e di mezzo secolo da quello di Alexander, la quantità di dati generati dalle nostre attività e interazioni quotidiane è enormemente aumentata, così come il numero di dispositivi digitali costantemente monitorati da strumenti satellitari sempre più complessi e accurati.
All'alba di quello che è stato definito il quarto paradigma della scienza, la società contemporanea sta vivendo una rivoluzione segnata dalla crescente importanza dei Big Data, della *Virtual e Augmented Reality* e alla diffusione di dispositivi computazionali via via più potenti, con capacità ormai prossime a quelle umane, ovvero la cosiddetta *Artificial Intelligence e Machine Learning* ed i relativi algoritmi predittivi (Bibri, 2019). Questa condizione fornisce un potenziale patrimonio di informazioni utili per operare scelte informate e di alta qualità su molteplici aspetti della vita contemporanea,

permettendo di apportare potenzialmente notevoli benefici agli utenti legati all'ottimizzazione dei processi produttivi, al monitoraggio e miglioramento della sicurezza pubblica, della salute dei cittadini, dell'efficienza energetica degli edifici, del traffico e dei modelli di utilizzo degli spazi pubblici, delle risorse ecosistemiche e della biodiversità urbana.
Questa rivoluzione, che a diritto si può affiancare a quelle della macchina a vapore, dell'elettricità e dell'elettronica, ha spinto numerosi ricercatori e istituzioni a confrontarsi sulle interazioni possibili tra il mondo fisico e quello virtuale, a partire dalla consapevolezza che, con il crescere della complessità delle metropoli e dello stile di vita contemporaneo, oggi i dati appaiono lo strumento più attendibile per fornire una visione accurata delle interazioni sociali, ma anche dell'ambiente che ha messo in scena tali interazioni.

A fronte di questo rapido cambiamento, i modelli di pianificazione, progettazione e gestione della città e dei territori sono ancora fortemente ancorati ad approcci statici, in grado di processare un numero ancora limitato di informazioni, capaci di coinvolgere solo parzialmente cittadini e parti interessate nei processi decisionali.
Sin dai primi anni del XVII – quando Maximilien de Béthune, duca di Sully, introducendo il ruolo dello stato francese come stazione appaltante, istituì i provveditori alle opere pubbliche, dando avvio al processo amministrativo di affidamento lavori ancora oggi vigente – il progetto è inteso come l'allegato di un contratto, un capitolato d'appalto fisso, rassicurante e da rispettare in tutte le sue parti.
Allo stesso modo il piano urbanistico, nella sua componente operativa, ma anche in quella strategico/strutturale, prevede spesso uno spostamento verso approcci fortemente burocratizzati che privilegiano il carattere stabile e la dimensione normativa delle modalità di trasformazione urbana. La lettura dei territori e la costruzione del quadro conoscitivo prevedono un periodo prestabilito di realizzazione, portando alla definizione di un documento di piano che, una volta approvato, ha l'ambizione di rimanere inalterato per periodi spesso molto lunghi.
Questi processi garantiscono risultati ottimali in sistemi a tassi di innovazione molto bassi, in cui informazioni, materiali, usi hanno la tendenza a variare in modo lento e fortemente prevedibile, entrando invece in crisi di fronte a cambiamenti radicali, repentini e continui.

L'ipotesi alla base di questo numero di Contesti è che l'innovazione apportata dai Big Data nelle discipline del progetto non costituisca solo un inspessimento della cassetta degli attrezzi a disposizione di un progettista, ma possa rappresentare l'occasione per un generale ripensamento dei processi cognitivi e progettuali, delle modalità di confezionamento di piani e progetti e della loro stessa forma, creando le condizioni per un diretto miglioramento della vivibilità e della resilienza delle nostre città e territori.



Metacity/Datatown
Fig. 1
progetto dello studio olandese MVRDV: Winy Maas, Jacob van Rijs, Nathalie de Vries, 1999.
In Maas, 1999.

A partire da questa ipotesi, i contributi presenti in questo numero articolano diverse traiettorie di ricerca, alcune delle quali corrono parallele, altre divergono andando a sondare campi e punti di vista specifici. Alcune traiettorie partono dall'alto, da proposizioni e descrizioni generali di teorie o contesti, cercando, il più delle volte, un punto di applicazione e di un'eventuale validazione. Altre traiettorie partono invece dal basso, dal quotidiano, dall'esperienza dei luoghi sviluppata attraverso ricerche o esplorazioni specifiche che cercano di ricondurre verso proposizioni più generali e astratte.
Questo doppio registro, formato da saggi descrittivi e ricerche esplorative, ci è apparso il criterio più significativo per organizzare i contributi di questo numero: se i saggi propongono descrizioni più generali dello stato dell'arte, indagando sia campi di applicazione specifici che i risultati istituzionali delle trasformazioni legate ai Big Data, le ricerche rintracciano nei singoli casi studio l'emergere di temi di carattere generale.

Saggi
Il primo gruppo di contributi è accomunato dall'obiettivo di capire come i Big Data, o gli strumenti ad essi riconducibili, possono aprire nuove prospettive nella gestione urbana e territoriale, rendendola più efficace, efficiente e affidabile. Alcuni saggi segnalano come gli strumenti legati alla *Earth Observation*, siano essenziali per implementare la nostra capacità valutativa e predittiva. Il programma europeo *Copernicus* è descritto in dettaglio nei contributi di Chirici e Taramelli et al. come strumento capace di offrire un numero di servizi operativi unici al mondo, gestibili solo con metodologie di *Information and Communications Technology* (*ICT*) avanzate. Chirici descrive, in particolare, i numerosi servizi offerti da *Copernicus* ai settori dell'urbanistica, dell'agricoltura e della gestione sostenibile delle risorse forestali. Taramelli, Castellani e De Bernardinis descrivono il programma *Copernicus* evidenziando, invece, i suoi limiti di impiego. Attualmente limitato prevalentemente dalla Comunità accademica e per finalità di ricerca, *Copernicus* non è ancora riuscito a trovare un diretto utilizzo da parte dell'utenza finale, al servizio tanto delle amministrazioni pubbliche (PPAA) quanto delle imprese. Per ovviare a questo limite, i tre autori illustrano le azioni intraprese, in particolare descrivendo il *Copernicus User Uptake* (UU) e l'istituzione della *Copernicus Academy (CA)*.

Pennacchia e Cinquepalmi affrontano invece lo studio dei *Digital twin models*, sistemi di modellazione urbana che grazie all'acquisizione e all'analisi dei Big Data ottenuti attraverso sensori, GPS, social media, reti intelligenti, dati istituzionali o registrazioni di clienti e transazioni, sono in grado di offrire un supporto per affrontare le sfide complesse che coinvolgono l'ambiente costruito.
Furgaci si interroga infine sui modi in cui i Big Data possano assistere progettisti e pianificatori nel migliorare la resilienza delle aree urbane, valutando, in particolare, il ruolo che le proprietà della ridondanza, diversità, modularità e auto-organizzazione possono ricoprire nella gestione dei sistemi urbani.

Ricerche
Se ormai l'impatto dei Big Data nella nostra vita risulta evidente, e i saggi appena menzionati ne tracciano un interessante spaccato, gli ultimi quattro contributi si propongono di chiarire quali interrelazioni siano rilevabili, grazie all'utilizzo dei dati, nei processi decisionali delle nostre città e in che modo essi possano alterare il tradizionale apparato pianificatorio e progettuale.
L'estrazione di conoscenza dai Big Data, attraverso ad esempio l'uso dell'intelligenza artificiale, e l'impiego di tale conoscenza per il miglioramento delle attività decisionali, sono infatti fortemente subordinati alla definizione di processi che consentano di visualizzare, gestire e trasformare in modo efficiente i dati. Furlan et al. e Zamperlin indagano su questo aspetto evidenziando la necessità di un vaglio qualitativo e fortemente progettuale per definire quali siano in effetti i dati che realmente possono essere utili e quale forma essi debbano assumere. Furlan, Wandl, Geldermans e Sileryte analizzano in particolare le possibili relazioni esistenti tra le tecnologie del GIS e quelle legate ai Big Data nello studio dei flussi dei rifiuti e degli scarti dell'Area Metropolitana di Amsterdam (AMA), sviluppate attraverso la metodologia 'Activity-based Spatial Material Flow Analysis'. Zamplerlin descrive invece la Snap4City, una Big Data Smart City Platform, utilizzata in numerose città e regioni europee, come ad esempio Anversa, Firenze, Pisa, Santiago etc. La piattaforma ha lo scopo di fornire un supporto per gli urbanisti, combinando potenti strumenti per l'integrazione dei dati, la loro analisi, previsione e visualizzazione con la possibilità di impostare Living Labs per migliorare la collaborazione tra gli operatori.
Gli ultimi due contributi descrivono, infine, i risultati fisici che la rivoluzione dei Big Data sta lasciando sul territorio. Lopez e Diguet osservano in particolare le architetture dei *data center* e le razionalità che hanno guidato la loro distribuzione nel territorio francese e statunitense, così come i modi con cui questi oggetti modificano paesaggio e rapporti sociali ed economici. Trombadore, Calcagno e Pierucci condividono l'esperienza del progetto internazionale Med-EcoSuRe (Mediterranean University as Catalyst for Eco-Sustainable Renovation) e in particolare la metodologia Digital Twin sperimentata nel Progetto Pilota di Santa Verdiana, che unisce le prestazioni dell'edificio con i dati sull'ambiente interno e le caratteristiche di benessere.

CITY of/with DATA
Come evidenziato in precedenza, dopo circa un ventennio dal lavoro di Maas e Metacity/Datatown, le tecnologie, così come la quantità di dati generati dalle nostre attività e interazioni quotidiane, sono enormemente aumentate, aprendo prospettive nuove che solo in parte al momento riusciamo a comprendere. Questo fenomeno, che prospetta una vera e propria rivoluzione epistemologica (Masiero, 2014), sta lasciando numerose tracce nella organizzazione sociale contemporanea, nelle sue gerarchie e apparati, nel modo di comunicare e di muoversi, negli scambi e nei consumi. È stata necessaria una crisi pandemica, come quella del Covid-19, per fare in modo che la società globale prendesse coscienza, almeno parzialmente, di quanti e quali opportunità questa condizione potesse offrire.

Analizzando le trasformazioni che la rivoluzione digitale e i Big Data possono apportare nel mondo dell'architettura e dell'urbanistica, nel 2012 Ratti proponeva questa considerazione: sino a poco tempo fa per vincere un gran premio automobilistico ci voleva una buona macchina, quindi una buona meccanica e un buon pilota. Oggi per vincere un gran premio bisogna essere dotati di un sistema di telemetria: migliaia e migliaia di sensori installati sulla vettura, che raccolgono informazioni in tempo reale e le trasmettono ai computer situati nei box dove si trova il team; qui le informazioni vengono analizzate e si prendono le decisioni che permettono di vincere la gara (Masiero, 2014).

Con questo semplice aneddoto, Ratti evidenzia quanto il modello digitale si discosti in modo radicale da quello utilizzato nelle scienze moderne, che hanno impiegato, in nome del rigore disciplinare, della logica e dell'efficienza, saperi fortemente settoriali e stabili nel tempo. Il modello digitale, restituito dai Big Data, supera cioè lo *standard*, la verità archetipale, basica e omologante, per offrire sistemi di controllo più flessibili, *fuzzy* e capaci di auto-organizzarsi che, non a caso, fanno spesso riferimento al mondo delle scienze biologiche, cioè a quel complesso di discipline che studiano la vita in tutte le sue forme e manifestazioni.

Se guardate in una prospettiva comune, le varie traiettorie di ricerca, proposte in questo numero di Contesti, perseguono proprio questo cambio di paradigma, cercando di fornire strumenti operativi e tracciare future linee di ricerca. Gli strumenti legati alla *Earth Observation* come il programma Copernicus, le potenzialità legate alle interazioni tra il mondo fisico e quello virtuale ricostruite attraverso i *Digital twin models* sia urbani che edilizi, ma anche il controllo, in tempo reale, dei flussi di energia, merci, persone e rifiuti, sono tutte opportunità oggi a disposizione dei ricercatori, in attesa di trovare una loro collocazione sinergica che oltrepassi gli usi sporadici e settoriali che spesso fanno riferimento al concetto di *smart city*.

Ciò che forse ancora è assente dal dibattito contemporaneo sono proprio i lavori in grado di aprire scenari futuri radicalmente nuovi, immaginativi e utopici. È proprio la spinta utopica e radicale, che ha sempre preceduto e guidato i grandi cambiamenti dell'umanità, la grande assente dal dibattito, quasi a testimoniare che l'evoluzione tecnologica abbia preceduto e superato l'immaginazione umana degli architetti e degli urbanisti, lasciando forse alla sola fantascienza, letterarie e cinematografica, il compito di prefigurare il mondo nuovo che è in marcia.

Bibliografia

Alexander C. 1964, *Notes on the Synthesis of Form*, Harvard University Press, Cambridge.

Bibri S.E. 2019, *Big Data Science and Analytics for Smart Sustainable Urbanism*, Springer International Publishing, Switzerland.

Carnap R. 1947, *Meaning and Necessity*, University of Chicago Press, Chicago.

Gerosa P.G. 2002, *Le ipotesi metodologiche dell'ultimo razionalismo funzionale*, in P. Di Biagi (a cura di), *I classici dell'urbanistica moderna*, Donzelli, Roma, pp. 251-268.

Maas W. 1999, *Metacity / Datatown*, Uitgeverij 010, Rotterdam.

Masiero R. 2014, *Smart*, in A. Bonomi e R. Masiero, *Dalla smart city alla smart land*, Marsilio, Venezia, pp. 89-107.

Wittgenstein L. 1953, *Philosophische Untersuchungen*, in G.E.M. Anscombe e R. Rhees (a cura di), *Philosophical investigations,* Blackwell, Oxford.

saggi
essays

Una nuova era nell'uso del telerilevamento a supporto della pianificazione sostenibile del territorio

Big Data e intelligenza artificiale a portata di mano

Lavoro svolto nell'ambito delle attività dell'Unità di Ricerca "Copernicus - Earth Observation and Spatial Analysis" dell'Università degli Studi di Firenze.

Gherardo Chirici
DAGRI, Università degli Studi di Firenze
gherardo.chirici@unifi.it

DOI: 10.13128/contest-11988
www.fupress.net/index.php/contesti/

keywords
remote sensing
gig data
artificial intelligence
copernicus

Since the seventies of the last century, remote sensing technologies for Earth observation have generated a constant flow of data from different platforms, in different formats and with different purposes. From these, through successive steps, spatial information useful to support the monitoring and planning of territorial resources is generated. The remotely sensed information is recognized as a classic example of big data ante litteram. Today, new cloud computing technologies (such as Google Earth Engine) make it possible to tackle the complex problem of managing and processing big data from remote sensing with new strategies that have revolutionized the way of understanding the use of these data sources. Also thanks to the advent of artificial intelligence technologies. From experiments on small study areas, today we work globally thanks to the possibility of operationally processing vast multidimensional and multitemporal datasets. The greater availability of information from space is exemplified by the numerous services offered by the European Copernicus program that allow the generation of important information useful to support urban planning, agriculture and sustainable management of forest resources.

Introduzione

Il telerilevamento

Il telerilevamento è definibile come la tecnologia atta a misurare a distanza le caratteristiche di un oggetto o di una superficie (Sabins, 1978).

L'osservazione della Terra (*Earth Observation* - EO) si riferisce all'uso di tecnologie di telerilevamento per monitorare la terra, le acque (mari, fiumi, laghi) e l'atmosfera. L'EO è basata su satelliti equipaggiati con sensori per raccogliere dati della superficie terrestre. Questi vengono quindi elaborati e analizzati al fine di estrarre diversi tipi di informazioni che possono servire per una vasta gamma di applicazioni commerciali e di ricerca.

Le tecnologie di EO utilizzano diversi tipi di sensori:

- Telerilevamento passivo: sensori ottici o termici che rilevano l'energia ricevuta dalla Terra a causa della riflessione e riemissione dell'energia del Sole dalla superficie o dall'atmosfera terrestre. Operano nelle lunghezze d'onda comprese tra il visibile e l'infrarosso.
- Telerilevamento attivo: i sensori radar (*Radio Detection and Ranging*) operano nella parte inferiore dello spettro (microonde) inviando energia alla Terra e monitorando l'energia ricevuta di ritorno dalla superficie terrestre o dall'atmosfera, consentendo il monitoraggio anche notturno e in tutte le condizioni meteorologiche; i sensori lidar (*Light Detection and Ranging*) inviano invece impulsi laser nel campo dell'infrarosso e calcolano la posizione nello spazio dei bersagli colpiti, permettendo la ricostruzione tridimensionale degli oggetti presenti sulla superficie terrestre. .

Sulla base della risoluzione spaziale (o geometrica) del sensore, ovvero della dimensione dei pixel analizzati dai sensori, i dati da telerilevamento si classificano in:

- Bassa e media risoluzione: più di 30 metri per pixel;
- Alta risoluzione (HR): tra 1 e 30 metri per pixel;
- Risoluzione molto alta (VHR): meno di 1 metro per pixel, chiamate anche immagini sub-metriche.

La risoluzione temporale definisce la frequenza con cui i dati vengono acquisiti, le orbite della maggior parte dei satelliti per l'osservazione terrestre a medio-alta risoluzione (pixel di 10-30 metri) sono dell'ordine delle poche settimane, per aumentare la frequenza di acquisizione è però sufficiente porre satelliti gemelli sulla stessa

orbita. Con questa tecnica è possibile acquisire un'immagine completa di tutta la Terra una o più volte al giorno.
Nel caso dei sensori ottici un altro parametro importante è la risoluzione spettrale. Questa è definita dall'ampiezza delle bande dello spettro che possono essere distinte dal sensore. In genere maggiore è la risoluzione spaziale e minore sarà la risoluzione spettrale.
Le applicazioni e le tecnologie di osservazione della terra sono utilizzate in un'ampia gamma di applicazioni che riguardano il monitoraggio terrestre, marino e atmosferico.
Secondo le stime fornite dal *Copernicus Market Report* (2017) nel 2017, l'economia globale dell'EO è stimata tra 9,6 e 9,8 miliardi di EUR, diviso tra le vendite di satelliti (la sezione più a monte della catena di fornitura), l'acquisizione, l'elaborazione e la trasformazione dei dati in prodotti informativi per gli utenti finali (la sezione più a valle della catena). Il mercato globale è guidato dalla sezione a monte, che costituisce circa il 70% dei ricavi totali. Il mercato a valle dell'EO globale è stimato essere compreso tra 2,6 e 2,8 miliardi di Euro, principalmente guidato dalle applicazioni in campo pubblico, che rappresentano tra il 50% e il 60% dei ricavi. Mentre i ricavi a monte tendono ad oscillare nel corso degli anni a seconda delle fluttuazioni per la necessità di mettere in orbita grandi satelliti, il mercato a valle mostra una crescita costante, a un CAGR (Compound Annual Growth Rate) previsto del 7% fino al 2022.
Il mercato dei prodotti da EO ha subito un forte impatto derivante dal processo di democratizzazione dei dati, ovvero l'apertura verso tutti dei dati da telerilevamento, prima riservati a una strettissima fascia di utilizzatori. Oggi recuperare i dati da telerilevamento non è più un problema e per una vasta gamma di questi l'accesso è gratuito.
A seguito di questa politica in favore dei dati open access, anche i dati forniti da ditte private a pagamento stanno subendo un graduale abbassamento del prezzo di mercato. Compensata però dall'espansione della base dei clienti, che è in costante aumento.
La disponibilità a pagare è diversa a seconda del mercato, con attori della difesa e dell'intelligence disposti a pagare di più per immagini ad alta risoluzione, rispetto ad altri settori (come agricoltura e foreste) dove la diffusione dei dati da EO, seppur in crescita, è frenata dalla forte sensibilità al prezzo dei prodotti.

Big data

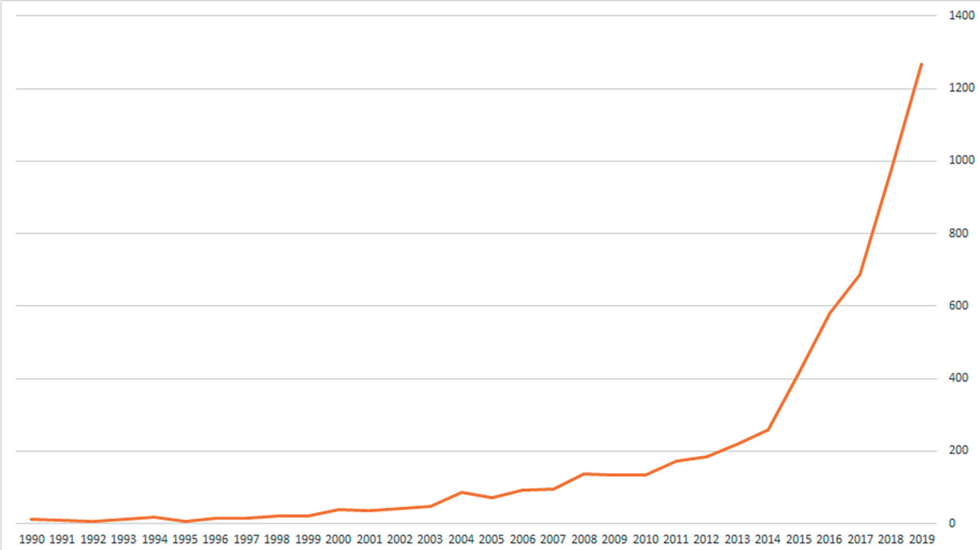
Una delle molte definizioni che viene data ai big data è di un insieme di dati che per dimensione o tipo vanno oltre la capacità dei database relazionali tradizionali di acquisirli, gestirli ed elaborarli con bassa latenza.
Kitchin (2013) dettaglia meglio le caratteristiche dei Big Data:

- enormi in volume, costituiti almeno da terabyte o petabyte di dati;
- la cui generazione avviene ad alta velocità, creati in tempo reale o quasi;
- diversi per varietà, essendo strutturati e non strutturati in natura;
- di portata esaustiva, catturano intere popolazioni o sistemi (n = tutti);
- a grana fine nella risoluzione ma univoci nell'identificazione;
- di natura relazionale, contenente campi comuni che consentono l'unione di diversi set di dati;
- flessibile, mantenendo i tratti di estensionalità (si possono aggiungere facilmente nuovi campi) e scalabilità (possono espandersi rapidamente di dimensione).

Secondo Laney (2020) non importa se si tratti di dati generati da progetti governativi, applicazioni commerciali o per ricerca accademica, i big data devono essere caratterizzati dalle 3V: ovvero un elevato Volume, Velocità di generazione e Varietà nelle caratteristiche del dato.
La rappresentazione spaziale in forma digitale della superficie terrestre utilizzata dal telerilevamento è un ottimo esempio *ante litteram* di big data, big data ancora prima che il termine venisse coniato. Nonostante il recente sviluppo di alcune tecnologie di telerilevamento laser (lidar) che producono dati grezzi in formato vettoriale, la vasta maggioranza dei prodotti da telerilevamento è in formato raster, ovvero di matrice.
Facciamo un esempio. Se vogliamo rappresentare tutte le terre emerse della Terra con una risoluzione di 10 m generiamo un'immagine contenente circa 1,5 bilioni di pixel. Considerando anche soltanto l'ormai desueta profondità di 8 bit ogni immagine 'peserebbe' 1,53 terabyte, considerando le consuete frequenze di acquisizione, un solo satellite in un'anno genererebbe approssimativamente tra i 30 e i 50 terabyte per ognuno dei canali di acquisizione. La creazione di un semplice indice di attività fotosintetica globale con 32 bit di profondità peserebbe poco meno di 6 terabyte (almeno in un formato di dati non compresso).
Sebbene dunque la produzione di big data sia esistita in alcuni domini, come il telerilevamento, le previsioni meteorologiche e i mercati finanziari, ormai da svariati decenni, solo molto recentemente una serie di sviluppi tecnologici, come l'*ubiquitous computing*, il funzionamento diffuso e ad alta velocità di Internet e i nuovi design usati nella gestione delle banche dati permettono non solo l'archiviazione ma anche un processamento dei big data in tempi ragionevoli. Questo passaggio secondo Kitchin (2014) costituisce un vero punto di svolta per la generazione di routine di analisi specificatamente progettate per far fronte a una tale abbondanza di dati.
Queste nuove tecniche di processamento sono radicate nella ricerca sull'intelligenza artificiale e sui sistemi esperti che hanno cercato di produrre sistemi di apprendimento automatico in grado rilevare pattern significativi nei big data attraverso modelli computazionali atti alla generazione di modelli predittivi, basati, il più delle volte,

Generazione	Quando	Approccio scientifico	Metodo
Prima	pre-rinascimento	Scienza sperimentale	Empirico, osservazione dei fenomeni naturali
Seconda	pre-computer	Scienza teorica	Modellistico, generalizzazione
Terza	pre-big data	Scienza computazionale	Simulazione di fenomeni complessi
Quarta	oggi	Scienza esploratoria	Data mining, data intensive, intelligenza artificiale

Le quattro fasi o paradigmi della scienza, modificato
Tabella 1
da Kitchin (2014)



Il trend di pubblicazioni su telerilevamento e big data.
Fig. 1
Fonte: SCOPUS.

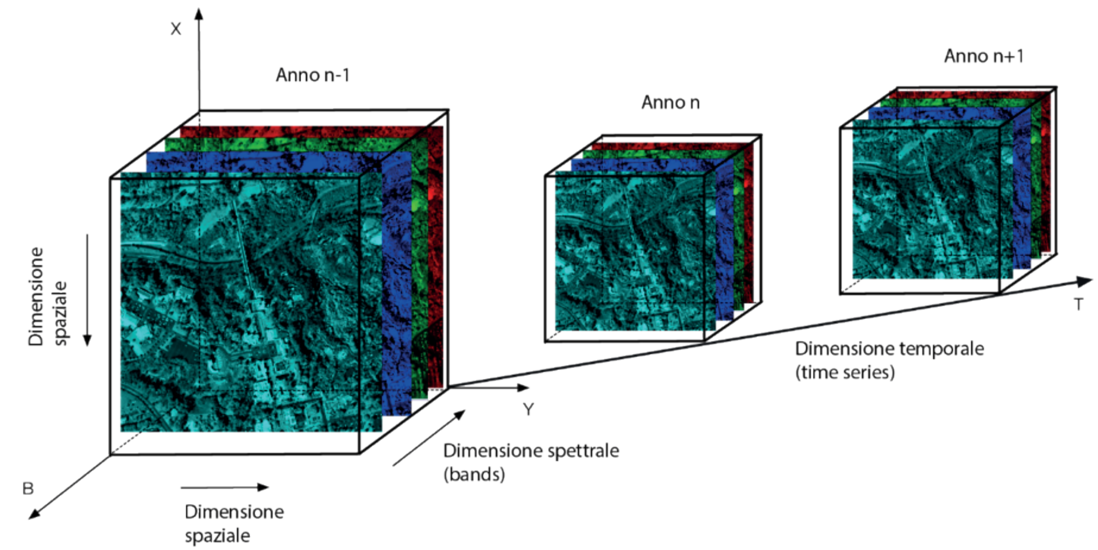
sulle architetture delle reti neurali (Han et al., 2011; Hastie et al., 2009).
Come stabilito da Kuhn (1962), un paradigma costituisce un modo accettato di interrogare il mondo e sintetizzare la conoscenza comune a una proporzione sostanziale di ricercatori in una disciplina, in un determinato momento nel tempo. Periodicamente, sostiene Kuhn (1962), emerge un nuovo modo di pensare che sfida teorie e approcci precedentemente accettati.
È ipotizzabile che l'avanzamento tecnologico che ha reso maggiormente disponibili in forma organizzata e accessibile al processamento i big data possa aver generato il passaggio a una nuova generazione della ricerca scientifica basata su approcci esplorativi ai dati derivati dall'applicazione dell'approccio all'intelligenza artificiale e non più meramente orientato alla simulazione dei fenomeni (Tabella 1)

È quanto molto probabilmente è accaduto negli ultimi anni anche nel campo del telerilevamento. Facciamo un esempio. Con i metodi scientifici di terza generazione tipicamente una ricerca scientifica basata su telerilevamento veniva realizzata su un certo ambito spaziale di studio (l'area di studio) e in un certo ambito temporale di studio (il periodo indagato). Se da un lato la ricerca poteva essere realizzata in un periodo relativamente breve e con risorse limitate, i risultati che se ne ottenevano soffrivano inevitabilmente di un limite implicito: la mancanza di generalizzabilità. I risultati sarebbero stati veri anche fuori dall'area di studio e in un altro periodo?
Oggi, con un approccio di quarta generazione, la possibilità di gestire e manipolare i big data permette di affrontare il problema alla radice: l'area di studio sarà quindi tutta la Terra, il periodo di studio è da oggi indietro fino alla data di acquisizione della prima immagine telerilevata disponibile.

Big data e telerilevamento
Il volume di dati prodotto dalle principali agenzie come NASA (National Aeronautics and Space Administration USA), JAXA (Japan Aerospace Exploration Agency), ESA (European Space Agency) (in particolare grazie al rinnovato vigore offerto dal programma Copernicus) e a tutti gli altri gestori di piattaforme di EO, ha introdotto una sfida per lo sviluppo di piattaforme on-line atte a gestire, elaborare e diffondere questi dati.
Si è in tal modo creato un nuovo mercato e quindi un'opportunità per i grandi attori del settore Information & Comunication, come Google e Amazon, che avevano tecnologie già sviluppate e quindi si sono trovati meglio posizionati per affrontare la questione dei big data.
Il nuovo mercato dei *Big Data Analytics* (BDA) nel campo EO si basa sull'elaborazione, l'analisi e la fusione di più immagini con altre fonti di dati al fine di creare informazioni non precedentemente disponibili. Si sono quindi affermati modelli di business per l'EO orientati a offrire servizi digitali basati su una forte potenza di calcolo, piattaforme basate su *cloud* e la fusione di un numero sempre maggiore di fonti di dati.
Anche nel campo della ricerca è evidente l'impegno della comunità scientifica a esplorare le opportunità offerte dall'approccio di quarta generazione. Lo testimoniano i numeri speciali dedicati a questo argomento da molte delle principali riviste di telerilevamento.

Si inizia nel 2015 con *IEEE JSTARS* che pubblica il numero speciale su "*Big Data in Remote Sensing*" e il *Journal of Applied Remote Sensing* con "*Management and Analytics of Remotely Sensed Big Data*". Seguono nel 2016 *IEEE Geoscience and Remote Sensing Magazine* con lo special issue su "*Big Data from Space and GeoInformatics*" e di Springer in *Geoinformatica* con "*Big Spatial and Spatiotemporal Data Management and Analytics*". Nel 2017 è la volta di *Remote Sensing of Environment* con "*Big Remotely Sensed Data: tools, applications and experiences*". In casa MDPI a partire dal 2018 sulla rivista *Remote Sensing* si aprono diversi special issue "*Advanced Machine Learning and Big Data Analytics in Remote Sensing for Natural Hazards Management*", "*SAR in the Big Data Era*" e "*Analysis of Big Data in Remote Sensing*". Dello stesso anno lo speciale "*Social Sensing and Big Data Computing for Disaster Management*" nella rivista *International Journal of Digital Earth*.
In SCOPUS usando una query del tipo "*remote* AND *sensing* AND *big* AND *data*" al 2019 risultano 5852 documenti, con un evidente trend di crescita (Figura 1).

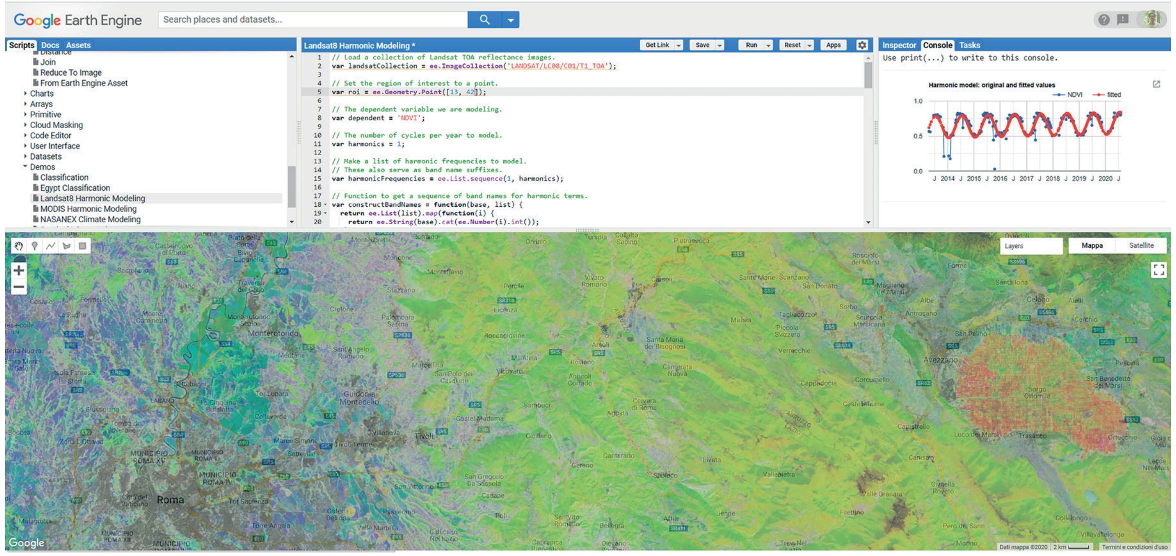
L'incremento della mole di dati da telerilevamento è dovuto alla loro multidimensionalità
Nel telerilevamento il termine big data non si riferisce solo al volume e alla velocità di produzione dei dati (lo *streaming*), ci sono infatti anche altri elementi che



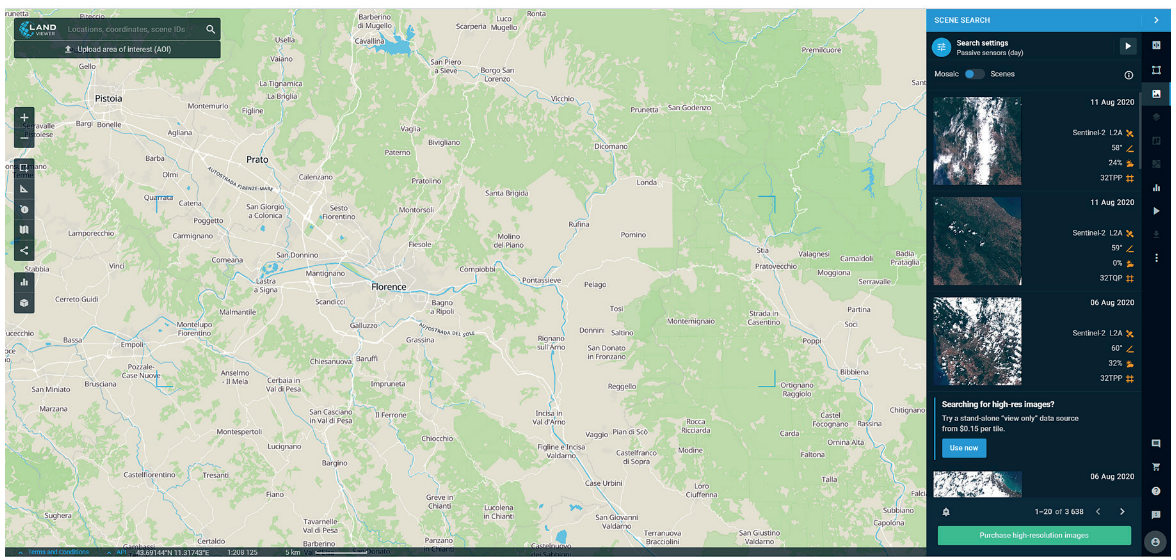
La multidimensionalità dei dati telerilevati
Fig. 2
Da Ma et al,. 2015, modificato

mettono in crisi le capacità di archiviazione e di elaborazione di questi dati quali la loro varietà e complessità. In altre parole la loro multidimensionalità, che porta inevitabilmente a una maggiore complessità.
I dati da telerilevamento normalmente servono molte discipline, dalle scienze della terra, al monitoraggio ambientale, i processi terrestri, l'atmosfera, l'idrologia e l'oceanografia, dallo studio delle aree urbane, delle coste, delle aree agricole, alle foreste e agli altri biomi. L'ampia gamma delle discipline coinvolte dà luogo anche a diverse forme di pre-trattamento dei dati grezzi che crea spesso la duplicazione degli archivi. Per esempio per i dati ottici si troveranno le immagini grezze, quelle normalizzate radiometricamente, quelle con diversi livelli di ortocorrezione, quelle con valori *Bottom of Atmosphere* (BOA) e quelle *Top of Atmosphere* (TOA). Senza contare i mosaici creati con diverse tecniche per la eliminazione delle coperture nuvolose.
Ma et al. (2015) riporta per gli archivi della NASA circa 7000 tipi di set di dati. Nella maggior parte dei casi, i set sono memorizzati in file strutturati utilizzando vari formati standard, inclusi HDF, netCDF, GeoTIFF, FAST, ASCII, JPG2000 e così via.
Oggi il progetto ESDIS (*Earth Science Data and Information System*) della NASA raccoglie metriche da tutto il sistema di dati e informazioni del sistema di osservazione della terra (EOSDIS) la cui dimensione è ad oggi di 33.6 petabyte, in crescita giornaliera di circa 20 terabyte, i dati distribuiti on-line agli utenti sono circa 103 terabyte al giorno (https://earthdata.nasa.gov/eosdis/system-performance). Il programma COPERNICUS riporta invece una crescita annua di circa 8 terabyte al giorno, con una consistenza complessiva stimata in circa 130 petabyte, che lo rende il data provider geografico più grande del mondo (https://www.copernicus.eu/en/what-can-you-do-130-petabytes-data).
Le ragioni per il costante trend di aumento (cioè l'aumento della velocità di crescita dei dati da EO) è da ricercarsi in diverse ragioni:

1. Per il proliferare del numero di missioni e del numero di satelliti per missione. Attualmente sulla base della banca dati OSCAR (*Observing Systems Capability Analysis and Review Tool*) disponibile open access all'indirizzo https://www.wmo-sat.info/oscar/satellites, il numero di piattaforme di osservazione della terra operative sono 240 (ognuna talvolta con più satelliti come i Sentinel, i MODIS, Landsat, NOAA AVHRR, Planet, ecc.), includendo anche le missioni meteorologiche il totale tra missioni passate, presenti e previste per il prossimo futuro è pari a 779. Come conseguenza del diffondersi delle missioni con satelliti gemelli o con micro-satelliti si è assistito alla nascita di un telerilevamento in near-real-time con immagini ad alta risoluzione giornaliere. Vale la pena ricordare il caso forse più clamoroso, ovvero la nuova costellazione di nanosatelliti PlanetScope. La missione PlanetScope (https://www.planet.com/) è stata avviata da Planet Labs Inc., una società di EO con sede a San Francisco (USA), che ha messo in orbita una costellazione di nano-satelliti denominati CubeSat 3U (Leach et al., 2019). Questi nanosatelliti chiamati *Doves* misurano circa 10 cm x 10 cm x 30 cm e sono dotati di una fotocamera multispettrale relativamente semplice che acquisisce i dati in quattro bande nei canali del visibile e NIR (tra 455 e 860 nm). A partire dal primo *flock* schierato il 22 giugno 2016 (i microsatelliti sono portati ed immessi in orbita a gruppi), il numero di satelliti è costantemente aumentato fino a 149 satelliti in orbita da settembre 2019 (Francini et al., 2020).
2. Per l'aumento della risoluzione geometrica. I satelliti con risoluzione metrica o sub-metrica sono in costante proliferazione ma anche la risoluzione dei normali satelliti ottici come i Sentinel-2 è nove volte superiore alla precedente generazione di satelliti NASA Landsat.
3. L'aumento della profondità, ovvero della memoria occupata da ogni pixel di ogni immagine. Fino a dieci anni fa lo standard era di 8 bit (2^8 ovvero un range di 256 valori o tonalità di grigio). Oggi ci si spinge frequentemente a 11 o 12 bit.
4. Per l'aumento del numero di bande dei sensori ottici. Si moltiplicano le missioni da satellite iperspettrali come Hyperion da 220 bande spettrali con una risoluzione di 10 nm, lo strumento WIS con 812 bande, o le 128 bande a bordo del HJ-1°, fino alle 250 bande dei sensori a bordo della missione Italiana PRISMA.
5. L'avvento di nuove tecnologie. Alle tradizionali tecnologie di acquisizione su base ottica passiva e radar a microonde si aggiunge da qualche anno la tecnologia lidar. Con alcune missioni per il monitoraggio della consistenza dei ghiacci (ICESAT) ma oggi in corso di sperimentazione in forma di batterie di sensori con acquisizioni in *fullwave form* dal programma GEDI (*Global Ecosystem Dynamics Investigation*) a bordo della Stazione Spaziale Internazionale.



L'API on-line di GEE
Fig. 3



LandViewer un servizio on-line per la ricerca di immagini telerilevate basato sulla piattaforma AWS
Fig. 4

Big Data Analytics: dai dati alle informazioni
Questa proliferazione di dati senza precedenti ha posto sfide significative nella gestione, elaborazione e interpretazione dei big data da EO. Sono stati compiuti grandi sforzi per incorporare il paradigma del calcolo ad alte prestazioni (*High Performance Computing* - HPC) nelle applicazioni di telerilevamento (Plaa et al., 2007). Gli approcci HPC stanno diventando dominanti per poter affrontare gli enormi requisiti computazionali introdotti dalla necessità di processare grandi moli di dati da EO.
L'approccio HPC all'analisi dei big data nel campo del telerilevamento si basa sulla scelta di due strategie alternative: 1) avere a disposizione un centro di calcolo ad alte prestazioni; 2) avvalersi di servizi di cloud computing sviluppati da terze parti.
Il primo approccio, dominante fino al decennio passato, ha però il notevole fattore limitante, oltre al costo di implementazione e manutenzione, di dover creare una copia dei dati da telerilevamento in locale. Visto l'aumentato flusso dello streaming appare oggi molto più vantaggioso il secondo approccio.
Infatti nell'ultimo decennio, il cloud computing e gli altri sistemi HPC sono diventati disponibili per il ricercatore e il consumatore medio con schemi pay-per-use (Barr, 2006). Il concetto alla base del cloud computing è l'idea di offrire il computing come servizio: i clienti hanno una serie sempre più diversificata di esigenze e spesso desiderano poter personalizzare i propri sistemi computazionali a un livello piuttosto basso. Prima del *cloud computing*, i prodotti software generalmente presupponevano (1) la proprietà dell'hardware da parte dell'utente finale o (2) specificità di dominio piuttosto rigide. Con il cloud computing, l'idea è di fornire risorse computazionali gestite come un prodotto commerciale, consentendo all'utente finale di scegliere e distribuire il software sopra i livelli sottostanti, senza soluzione di continuità, hardware / rete / sistema operativo (Hegeman et al., 2014).
Vediamo ora i principali sistemi di *cloud computing* partendo dai primi due che offrono servizi specifici per l'EO.
Google Earth Engine (GEE) è una piattaforma basata su cloud per l'analisi geospaziale su scala planetaria che consente di accedere alle enormi capacità di calcolo di Google. GEE permette di sostenere una varietà di programmi/servizi di monitoraggio su diverse questioni sociali ad alto impatto tra cui deforestazione, siccità, disastri, malattie, sicurezza alimentare, gestione dell'acqua, monitoraggio del clima e protezione ambientale. È unico nel campo come piattaforma integrata progettata per potenziare non solo gli scienziati del telerilevamento tradizionale, ma anche un pubblico molto più ampio che non ha la capacità tecnica necessaria per utilizzare i supercomputer tradizionali o le risorse di cloud computing su larga scala (Gorelick et al., 2017)
GEE è costituito da un catalogo di dati pronto per l'analisi multi-petabyte insieme a un servizio di calcolo intrinsecamente parallelo ad alte prestazioni. Vi si accede e si controlla tramite un'API (*Application Programming Interface*) accessibile da Internet e un ambiente di sviluppo interattivo (IDE) associato basato sul web che consente la prototipazione rapida e la visualizzazione dei risultati.
L'uso di GEE è gratuito per gli utenti Google purchè non si richiedano risorse particolarmente elevate (necessarie per esempio per analisi globali multitemporali con dati a media o alta risoluzione. In questi casi è necessario stipulare accordi specifici

con Google. Per lo storage dei risultati sono necessari account *Google Drive* di dimensioni adeguate.
Tra le poche alternative a GEE *l'Amazon Web Services* (AWS), una sussidiaria di Amazon che fornisce API e piattaforme di cloud computing su richiesta a privati, aziende e governi, con pagamento in base al consumo. Questi servizi web di cloud computing forniscono una varietà di infrastrutture tecniche astratte di base e blocchi e strumenti informatici distribuiti. Uno di questi servizi è *Amazon Elastic Compute Cloud* (EC2), che consente agli utenti di avere a disposizione un cluster virtuale di computer, sempre disponibile, tramite Internet. La versione dei computer virtuali di AWS emula la maggior parte degli attributi di un computer reale, comprese le unità di elaborazione centrale hardware (CPU) e le unità di elaborazione grafica (GPU); memoria locale / RAM; archiviazione su disco rigido / SSD; una scelta di sistemi operativi; networking; e software applicativo precaricato come server web. La tecnologia AWS è implementata nelle server farm di tutto il mondo e gestita dalla filiale di Amazon.

Simile nella filosofia ad AWS l'IBM Compute è un server cloud ad alte prestazioni che permette la virtualizzazione di sistemi HPC ottimizzando le configurazioni di base per soddisfare le esigenze di carico di lavoro in termini di RAM, SSD, GPU e altro. E' organizzato in 60 data center. Altri servizi simili di cloud computing sono offerti da importanti multinazionali quali Oracle e Microsoft.
In questo quadro emerge chiaramente il ruolo di leader del programma Europeo Copernicus, attualmente il principale provider di informazioni da EO e quindi uno dei principali motori del mercato Big Data Analytics.

Il programma Europeo Copernicus

Copernicus è il programma dell'Unione Europea per l'osservazione della Terra e i servizi associati.
Dal 2008 al 2020, si prevede che gli investimenti totali nel programma Copernicus raggiungeranno gli 8,2 miliardi di euro.
Nello stesso periodo, questo investimento genererà benefici economici compresi tra 16,2 e 21,3 miliardi di euro (esclusi i benefici non monetari). Questo valore economico è generato attraverso il valore aggiunto creato nell'industria spaziale a monte, e le vendite di applicazioni basate su Copernicus a valle.
Il programma Copernicus è cofinanziato dalla Commissione Europea e dall'ESA.
L'ESA, insieme a EUMETSAT (*European Organisation for the Exploitation of Meteorological Satellites*) su aspetti specifici, è prevalentemente responsabile dello sviluppo delle componenti spaziali, denominate anche infrastrutture in Copernicus, mentre la Commissione europea è responsabile dell'implementazione e della gestione di tutte e tre le componenti: spaziale, in-situ, servizi. Copernicus riceve, elaborare e offre set di dati da varie fonti - i satelliti di EO (le proprie Sentinelle, e le altre missioni partner) e una moltitudine di sensori in situ a terra, in mare, o in volo.
Copernicus dispone di sei servizi tematici (Terra, Mare, Atmosfera, Clima, Emergenza e Sicurezza) supportando lo sviluppo di molte applicazioni. I servizi Copernicus elaborano e analizzano i dati, li integrano con altre fonti, offrendo *Geographic Information Systems* (GIS) ai propri utenti e al servizio delle autorità pubbliche e delle imprese commerciali. A partire dal 2017 la Commissione europea ha lanciato l'iniziativa di servizi di accesso on-line (DIAS - *Data and Information Access Services*) per sviluppare meglio la diffusione dei dati e delle informazioni Copernicus.
La maggior parte di questi consente solo il download e talvolta lo sfruttamento delle informazioni a cui si accede tramite un servizio di mappatura web (WMS). È da notare che attraverso appositi accordi di collaborazione molte delle più importanti banche dati Copernicus sono oggi accessibili attraverso i due più comuni ambienti di sviluppo per il cloud computing: ovvero i già citati GEE e, in misura più ridotta, AWS.
Sul piano nazionale è attivo il Forum Nazionale degli Utenti del Programma Copernicus, coordinato dall'Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA). Il Forum Nazionale consiste in uno strumento preposto alla raccolta dei requisiti degli utenti finali ed intermedi con l'obiettivo di giungere ad una posizione nazionale concordata ed armonizzata sia per gli sviluppi nazionali in seno alla Space Economy e sia verso il Forum e Comitato Europeo del Copernicus, coerentemente con le esigenze istituzionali, della ricerca, dell'impresa e dell'industria del Paese. Le attività del progetto Copernicus e del Forum sono state recentemente presentate in una giornata di studi svoltasi il 19/06/2020 presso l'Università degli Studi di Firenze (https://www.isprambiente.gov.it/it/news/copernicus-il-programma-europeo-di-osservazione-della-terra-e-le-sue-applicazioni).

Evoluzione temporale dei lanci delle Sentinelle Copernicus
Fig. 5

Esempio dei dati disponibili sul TEP

Fig. 6a-6b
https://urban-tep.eu/puma/tool/
Per l'area di Zhengzhou nella Provincia di Henan in Cina l'area urbanizzata nel 1975 (in alto) e quella al 2014 (in basso) secondo il dataset GHSL.

La componente spaziale

La componente spaziale include lo sviluppo, il lancio e la gestione dei satelliti Sentinel e di una serie di missioni operate da partner.
In primo luogo, i Sentinel sono una famiglia di satelliti di proprietà dell'UE, che sono stati specificatamente progettati per soddisfare le esigenze dei servizi Copernicus e dei loro utenti. Attualmente, Sentinel-1, 2, 3 e 5p sono già stati lanciati (Figura 5). Sentinel-4, 5 e 6 sono attesi nei prossimi anni. Sentinel-1, 2, 3, 5p e 6 sono dedicati all'osservazione della Terra attraverso vari sensori (ad esempio radar, ottico, altimetrico), mentre Sentinel 4 e 5 sono satelliti meteorologici operati da EUMETSAT.
Le Sentinelle sono integrate da altre missioni, che esistono indipendentemente da Copernicus. Queste missioni sono gestite da ESA (ad esempio ERS), EUMETSAT (ad esempio METEOSAT), gli Stati membri dell'UE, altri paesi terzi (es. Landsat o MODIS negli USA) o fornitori commerciali (es. Deimos, Pleiadi). Queste missioni sono importanti perchè forniscono a Copernicus dati ad altissima risoluzione che non possono essere ottenuto tramite le Sentinelle o compensare le Sentinelle che non sono state ancora lanciate.

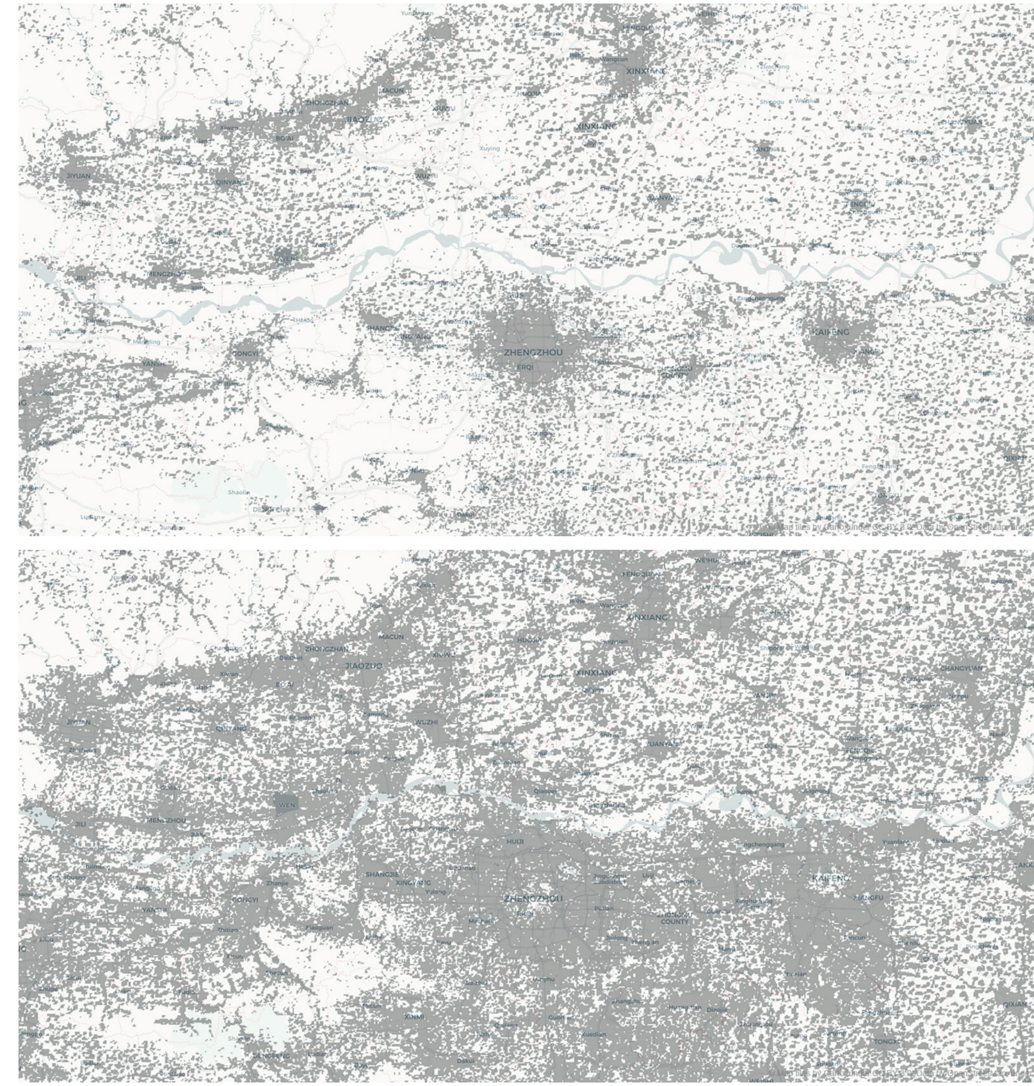
La componente in situ

Copernicus si basa anche su dati in situ acquisiti a terra, per lo più appartenenti agli Stati membri dell'UE o a infrastrutture di ricerca internazionali, e che vengono messe a disposizione di Copernicus tramite specifici accordi.
Sono incluse le osservazioni a terra, dal mare e dall'aria, così come i riferimenti geospaziali e i dati ausiliari concessi in licenza o forniti per l'uso nel programma Copernicus. Più recentemente, i dati in situ hanno trovato nuove fonti nei sensori e nelle immagini raccolte da veicoli aerei senza pilota (UAV) o da crowdsourcing (ad esempio OpenStreetMap). I dati in situ hanno due ruoli:

- Calibrare, integrare e convalidare i dati ottenuti dalle missioni satellitari per garantirne l'affidabilità nel tempo.
- Integrare le informazioni Copernicus dallo spazio per produrre prodotti e fornire servizi più vicini alle esigenze degli utenti finali.

La componente dei servizi

I servizi sfruttano i dati Sentinel, delle missioni partner e dei dati in situ, per fornire prodotti e informazioni a supporto commerciale, applicazioni istituzionali e di ricerca.
Il progetto Copernicus ha permesso di ottenere notevoli benefici, qui di seguito se ne analizzano i principali nel settore dell'urbanistica, e per i comparti agricoli e forestali. I dati sono ottenuti dall'ultimo Copernicus Market Report 2019 disponibile on-line https://www.copernicus.eu/sites/default/files/2019-02/PwC_Copernicus_Market_Report_2019_PDF_version.pdf



Urbanistica

Gli utenti intermedi dei dati forniti da Copernicus sono principalmente società di servizi a valore aggiunto, startup e PMI, che preparano poi prodotti e servizi per gli utenti finali, che sono prevalentemente autorità locali, regionali e nazionali. Tra i dati forniti da Copernicus quelli più utilizzati in campo urbanistico sono i dati radar Sentinel-1, utilizzati per il rilevamento delle modifiche all'edificato e per lo sviluppo di modelli 3D. I dati ottici Sentinel-2 forniscono dati rilevanti per il monitoraggio della crescita urbana grazie alla elevata risoluzione spaziale. A partire da queste fonti di dati sono stati sviluppati importanti dataset globali come il *Global Human Settlement Layer* (GHSL) coordinato dal *Joint Research Centre* della Commissione Europea (https://ghsl.jrc.ec.europa.eu/) e il *World Settlement Footprint* 2015 coordinato dall'European Space Agency (https://www.esa.int/

Applications/Observing_the_Earth/Mapping_our_global_human_footprint). Entrambi i progetti hanno portato alla produzione di un vasto spettro di strati informativi inerenti lo studio dell'evoluzione temporale delle aree urbanizzate, tutti consultabili, insieme ad altre banche dati tematiche inerenti, attraverso la piattaforma *Urban Thematic Exploitation Platform* (TEP) https://urban-tep.eu/puma/tool/.

Non mancano d'altra parte gli strati informativi sviluppati direttamente dal *Copernicus Land Monitoring Service* (CLMS) quali le immagini telerilevate che possono essere utilizzate per la creazione di prodotti o i dati cartografici già elaborati come il *Corine Land Cover*, gli *High Resolution Layers*, o il *Land Surface Temperature*, tutti disponibili on-line https://land.copernicus.eu/.

I risultati delle ricerche inerenti lo studio globale dei trend di popolazione dipingono un quadro estremamente grave. Secondo il *World Urbanization Prospects* 2018 (https://population.un.org/wup/) con grandi concentrazioni di persone, proprietà e ricchezza, le città si stanno espandendo come mai prima d'ora, con la metà della popolazione globale che vive nelle aree urbane. Il mondo si sta urbanizzando con un tasso di crescita mai così alto: oltre il 50% della popolazione mondiale vive nelle città, che crescerà al 66% entro il 2050, su un totale di 9,8 miliardi di persone. A fianco della moltitudine di opportunità in rapida espansione, le città devono affrontare anche un certo numero di sfide. Disagi economici, conflitti sociali e disastri ambientali si verificano sempre più frequentemente proprio nelle città che si stanno espandendo più velocemente. La popolazione urbana mondiale sta crescendo quattro volte più velocemente della popolazione rurale. Il 90% di questa crescita si sta verificando in regioni in via di sviluppo e l'Africa attualmente ha il tasso di crescita urbana più veloce, a più del 5% all'anno.

Evidentemente l'espansione urbana deve essere monitorata per poter garantire che proceda in forma sostenibile, senza compromettere le risorse ambientali e territoriali, né peggiorare la qualità della vita e la sicurezza della popolazione urbana.

L'osservazione della Terra, in questo senso, rappresenta uno strumento fondamentale per la generazione di informazioni utili a supportare le più opportune scelte in ambito urbanistico. Oltre ai già citati servizi finalizzati alla mappatura multitemporale delle aree urbane si annoverano il monitoraggio delle temperature e delle isole di calore, il monitoraggio del verde urbano, analisi dell'*urban sprawl*, il supporto alla gestione e alla pianificazione delle aree di trasporto, alla ricostruzione tridimensionale dell'edificato, e le deformazioni derivanti da fenomeni come terremoti, frane e alluvioni.

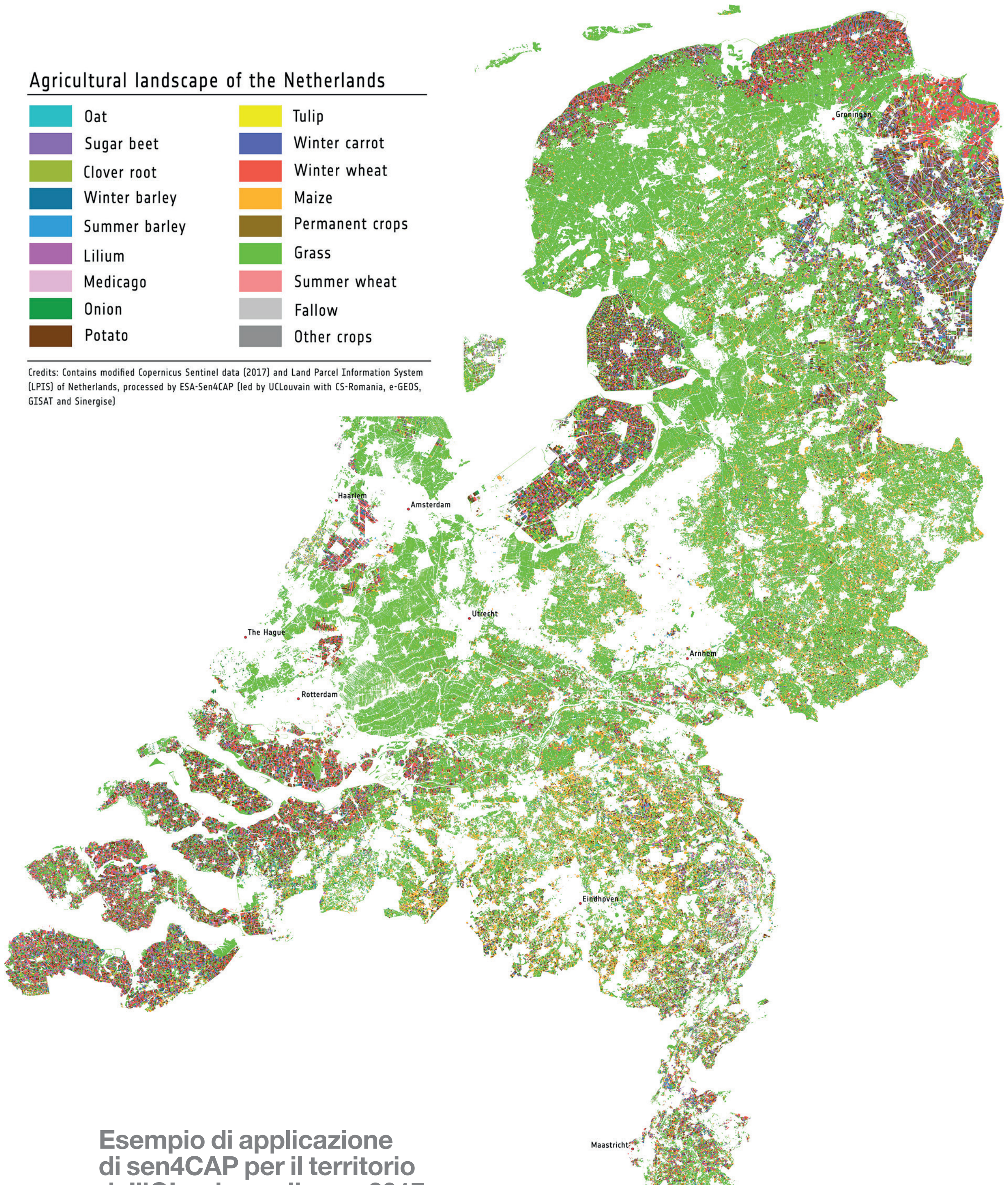
Agricoltura

L'agricoltura è stato uno dei primi mercati del telerilevamento e continua a esserci un crescente interesse per il modo in cui i prodotti da telerilevamento possono supportare le sfide affrontate dal settore. Le immagini da satellite possono dare un contributo significativo all'efficienza di moderne pratiche agricole come la *precision agriculture*. Spesso i dati da EO vengono combinati con i dati in situ (Di Biase et al., 2018) al fine di produrre risultati più mirati e localizzati per l'agricoltore, mentre i dati aerei o da droni possono fungere da dati complementari.

Nonostante la lunga tradizione di uso, il mercato del telerilevamento soffre di una forte frammentazione con utenti intermedi che includono una varietà di attori: start-up, piccole e medie imprese (PMI), grandi società e attori puramente scientifici come organismi di ricerca e università. Gli utenti finali sono abbastanza bilanciati tra attori pubblici (come le autorità governative che necessitano di precise informazioni per politiche e regolamenti specifici, o per monitorare questioni specifiche) e attori privati (cooperative agricole, aziende alimentari, consulenti agronomici, ecc.). Gli agricoltori, che dovrebbero essere i principali utenti finali dei prodotti agricoli basati sui dati EO, incontrano invece una certa difficoltà di accesso ai servizi Copernicus a causa della mancanza di infrastrutture (ad esempio una buona connessione a banda larga) o di conoscenza digitale.

Il settore agricolo gioca un ruolo cruciale nell'economia globale, ed è in rapida evoluzione a causa dell'emergere e lo sviluppo di numerose innovazioni tecniche, in un quadro di rilevanti dinamiche ambientali in rapida evoluzione. Negli anni futuri, l'agricoltura dovrà affrontare diverse pressioni significative. La popolazione mondiale dovrebbe crescere dagli attuali 7,4 miliardi di 2018 a 9,8 miliardi entro il 2050, causando la necessità di accrescere la produzione di cibo dagli attuali 8,4 miliardi di tonnellate a circa 13,5 miliardi di tonnellate all'anno. I cambiamenti climatici e l'intensificazione nella concorrenza per terra, acqua ed energia imporrano scelte strategiche per assicurare una corretta ma sostenibile produzione agricola.

La *Food and Agriculture Organization* delle Nazioni Unite (FAO) riporta che per il 2018 il numero di persone con alimentaizone insufficiente è aumentata dai circa 804 millioni del 2016 agli 821 milioni del 2017. In Europa il settore è sostenuto dalla Politica Agricola Comunitaria (un investimento di 60 miliardi di euro all'anno) che sostiene gli agricoltori Europei nel difficile lavoro di mantenere l'attività assicurando una vantaggiosa produzione di alimenti sicuri, riducendo l'impatto sull'ambiente e sul clima, e proteggendo la biodiversità.



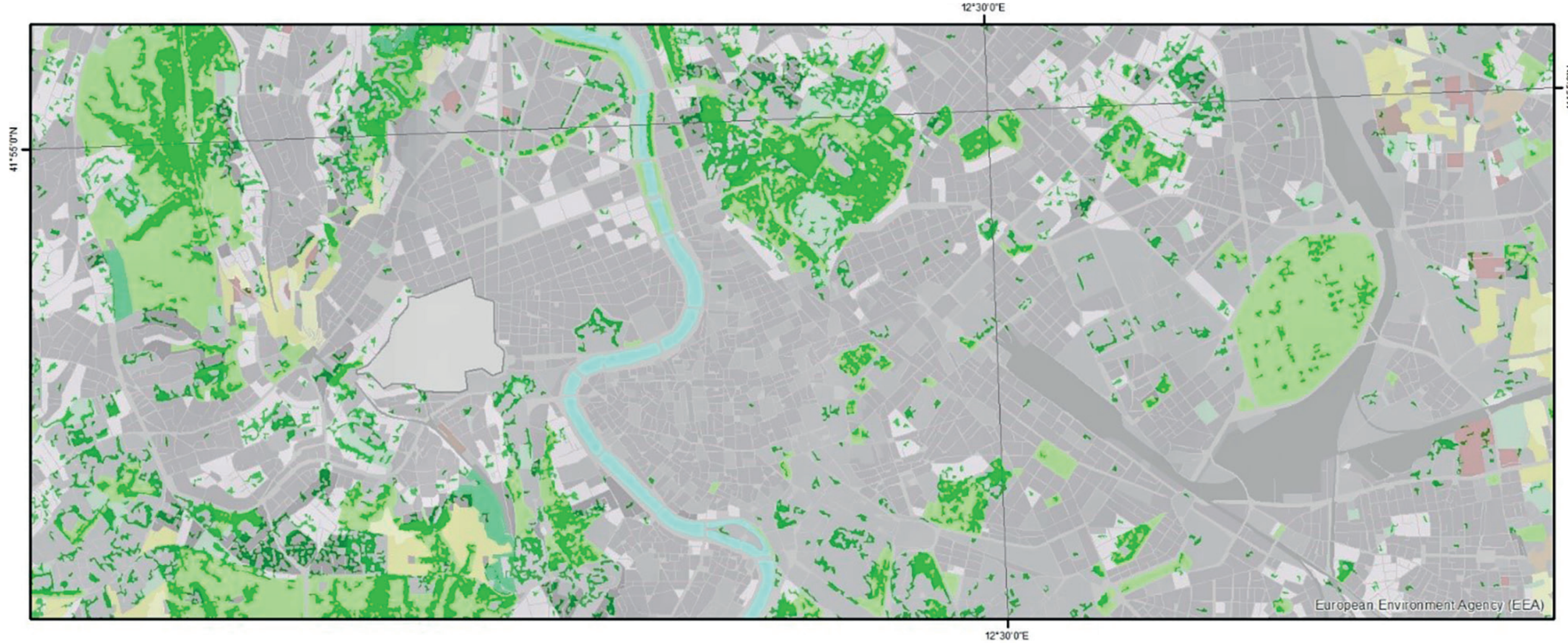
Esempio di applicazione di sen4CAP per il territorio dell'Olanda per l'anno 2017
Fig. 7
Da https://www.esa.int/ESA_Multimedia/Images/2018/05/Crop_map.

Le applicazioni basate su EO possono aggiungere valore a una serie di sottosettori agricoli, comprese le attività relative alle colture (ad es. monitoraggio della salute delle colture, inventari delle colture, monitoraggio del tipo di colture), gestione delle risorse idriche (es. umidità del suolo, irrigazione delle colture, monitoraggio della siccità) e ottimizzazione e gestione dell'uso del suolo (per esempio per il monitoraggio della biodiversità).
Alcuni esempi significativi dell'utilizzo di big data è il riconoscimento e la mappatura automatica delle principali tipologie di culture agricole basate sull'analisi della firma fenologica di scene Sentinel2. L'applicazione, sviluppata da ESA è disponibile all'indirizzo http://esa-sen4cap.org e il suo utilizzo operativo è previsto per il supporto ai controlli nei pagamenti dei contributi di supporto alla PAC. In Figura 7 l'applicazione di sen4CAP su tutto il territorio dell'Olanda.

I prodotti derivanti da sen4CAP dovrebbero essere a breve disponibili sulla piattaforma Copernicus sia in termini di serie multitemporali di osservazioni della fenologia ad alta risoluzione (10 m) (*High Resolution Vegetation Phenology and Productivity*) sia in termini di mappe annuali dei principali tipi di coltivazione (*High Resolution Layer Crop Types*).

Foreste

Le foreste in Europa coprono una superficie di circa 215 milioni di ettari (circa il 33% delle terre emerse), questa superficie è in costante aumento. Lo stato di salute delle foreste è considerato in modo positivo, le fonti di stress che producono danni (incendi, danni da vento, attacchi di patogeni) riguardano solo circa il 3% delle foreste Europee. Preleviamo legno e legname dalle nostre foreste ma a un ritmo che, se pur in crescita, è ancora ampliamente sostenibile.
Le foreste in Europa assimilano ogni anno 719 milioni di tonnellate di CO_2, circa il 9% delle nostre emissioni. Le nuove tecnologie impiegate nella trasformazione del legno permettono oggi la creazione di prodotti che possono sostituire materiali tradizionali (come cemento e acciaio) la cui produzione determina enormi emissioni di CO_2 (Forest Europe, 2015)
Anche il settore forestale, come quello agricolo, è riconosciuto come uno degli utilizzatori storici delle informazioni derivanti dal telerilevamento. Iniziative globali,



Per il centro di Roma, sullo sfondo dello Urban Atlas, in verde le informazioni dello Street Tree Layer.
Fig. 8
Entrambi i prodotti sono vettoriali e distribuiti dal servizio Copernicus Land.

come il programma delle Nazioni Unite sulla riduzione delle emissioni dovute alla deforestazione e al degrado delle foreste (*Reducing Emissions from Deforestation and Forest Degradation* - REDD), i programmi permanenti di monitoraggio delle risorse forestali a livello globale (FAO – *Global Forest Resource Monitoring*) ed Europeo (Forest Europe – *Interministerial Conference on the Protection of Forests in Europe*) richiedono ai paesi partecipanti la disponibilità di dati estremamente accurati e precisi sulle foreste.
Le foreste producono una vasta gamma di servizi ecosistemici: produzione di prodotti come legno, legname e cellulosa, protezione dal rischio idrogeologico (alluvioni, frane e valanghe), protezione della biodiversità, capacità di contrastare il cambiamento climatico grazie all'assimilazione della CO_2 atmosferica, socio culturale (sostegno alle popolazioni che vivono grazie al reddito derivante dalla filiera foresta-legno ma anche per la fruizione ricreativa dei boschi e al loro valore culturale, estetico e paesaggistico).
Una gestione sostenibile delle risorse forestali ha per obiettivo il mantenimento e la conservazione per le future generazioni della capacità di produrre servizi ecosistemici. Questo evidentemente si può raggiungere solo attraverso la quantificazione dei servizi ecosistemici attraverso indicatori spazializzati. Per molti di questi indicatori i dati da telerilevamento possono essere utilizzati con successo.

Entrambi i dati Sentinel-1 e 2 sono fonti riconosciute di preziose informazioni per la mappatura della superficie forestale e della copertura arborea (i così detti *Trees Outside Forests* o *Small Woody Features*), la stima della biomassa e del carbonio atmosferico assimilato (Puletti et al., 2018).
Le immagini satellitari permettono anche l'individuazione delle principali forme di disturbo derivanti da fonti biotiche (insetti e patogeni) e abiotiche (utilizzazioni, incendi, danni da vento).
Gli utenti intermedi dei dati da EO nel dominio forestale includono diversi attori privati (dalle microimprese ai più grandi attori), istituti di ricerca pubblici e organizzazioni di gestione forestale.
I prodotti EO finalizzati alla silvicoltura sono attualmente utilizzati principalmente da utenti finali pubblici (90%), piuttosto che da utenti privati (10%).
Tra i servizi Copernicus più utilizzati in campo forestale vi sono gli *High Resolution Layers* per la mappatura della densità della copertura arborea (*Tree cover density*), dei *Dominant leaf type* (latifoglie vs. conifere), e del *Forest type product* specificatamente sviluppato per coincidere il più precisamente possibile con la definizione internazionale della FAO utilizzata in campo inventariale.
A queste informazioni si aggiungono quelle relative alla copertura di alberi fuori foresta: gli *Small Woody Features* e lo *Street Tree Layer* in ambiente urbano (Figura 8).

Conclusioni

Le tecnologie di osservazione della Terra costituiscono un chiaro caso di big data *ante litteram*. In questo contributo abbiamo cercato di presentare le principali motivazioni che hanno portato all'aumento della velocità di produzione di questi dati: più missioni, più satelliti per missione, maggiore risoluzione geometrica, maggiore risoluzione spettrale e radiometrica, internet più veloce, democratizzazione nell'accesso ai dati (oggi spesso gratuito).

In questo scenario spicca il programma Copernicus che, tramite uno sforzo economico senza pari della Commissione Europea e dell'Agenzia Spaziale Europea (ESA), è ad oggi il principale distribuitore globale di dati e informazioni derivanti, o comunque collegate, alle tecnologie di EO.

La proliferazione dei dati dà luogo a una ulteriore crescente complessità dei dati da EO intrinsecamente caratterizzati già da una elevata multidimensionalità.

L'elaborazione di questi dati per la produzione di informazioni (*Big Data Analytics*) non può basarsi sugli approcci modellistici tipici di un periodo storico nel quale la mole di dati era infinitamente inferiore. Ma deve invece poter esplorare più in profondità la complessità tipica dei big data, con l'uso di approcci di data mining basati sull'intelligenza artificiale.

Tali approcci sono però ad alta intensità computazionale e per essere operati in tempi ragionevoli richiedono sistemi HPC che devono includere: (1) cluster di supercomputer e piattaforme di cloud computing ottimizzati per carichi ad alta intensità di dati; (2) file system e database paralleli che si basino sulla disponibilità di accesso veloce ai dati; (3) strumenti di gestione dei dati per il controllo del posizionamento dei dati in memoria locale dei dati multilivello; (4) pianificazione delle attività di elaborazione in parallelo.

Operare con tali requisiti attraverso infrastrutture di calcolo locali appare oggi poco efficace.

In questo scenario si sono sviluppati diversi servizi commerciali che permettono di operare attraverso sistemi HPC virtuali residenti su servizi di cloud esterni. Nel panorama dei competitor più importanti dell'IC quali Amazon, Oracle, IBM e Microsoft spicca Google Earth Engine, ad oggi la piattaforma più utilizzata, specie nel settore della ricerca.

È probabile che nel prossimo futuro possano emergere alcuni nuovi problemi con l'ulteriore aumento della quantità e dell'applicazione diffusa dei dati da telerilevamento. Tra queste la crescente domanda di capacità di elaborazione in tempo reale o in tempo quasi reale da parte di molte applicazioni EO per fornire informazioni ritenute critiche dal punto di vista temporale (per motivi di sicurezza o per emergenze ambientali). In questi casi il trattamento dei dati dovrà essere fatto in tempi brevissimi e quindi saranno necessarie infrastrutture di calcolo parallelo specificatamente destinate.

Bibliografia

Barr J. 2006, *Amazon EC2 Beta*, Amazon Web Services Blog, <http://aws.typepad.com/aws/2006/08/amazon_ec2_beta.html> (05/20)

Boyd D., Crawford K. 2012, *Critical questions for big data*, «Information, Communication and Society», vol. 15, n. 5, pp. 662–679.

Di Biase R.M., Fattorini L., Marchi M. 2018, *Statistical inferential techniques for approaching forest mapping. A review of methods*, «Annals of Silvicultural Research», vol. 42, n. 2, pp. 46-58.

FOREST EUROPE 2015, *State of Europe's Forests*, Ministerial Conference on the Protection of Forests in Europe, <https://www.foresteurope.org/docs/fullsoef2015.pdf> (07/20)

Francini S., McRoberts R.E., Giannetti F., Mencucci M., Marchetti M., Scarascia Mugnozza G., Chirici G. 2020, *Near-real time forest change detection using PlanetScope imagery*, «European Journal of Remote Sensing», in press.

Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R. 2017, *Google Earth Engine: Planetary-scale geospatial analysis for everyone*, «Remote Sensing of Environment», vol. 202, pp. 18-27.

Hegeman J.W., Sardeshmukh V.B., Sugumaran R., Armstrong M.P. 2014, *Distributed LiDAR data processing in a high-memory cloud-computing environment*, «Annals of GIS», vol. 20, pp. 255-264.

Kitchin, R. 2013, *Big data and human geography: Opportunities, challenges and risks*, «Dialogues in Human Geography», vol. 3, n. 3, pp. 262–267.

Laney, D. 2012, *3D Data Management: Controlling Data Volume, Velocity and Variety*, <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-DataVolume-Velocity-and-Variety.pdf> (08/20).

Leach N., Coops N.C., Obrknezev N. 2019, *Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies*, «Computer and Electronics in Agriculture», vol. 164 <https://doi.org/10.1016/j.compag.2019.104893> (08/20).

Liu P., Di L., Du Q., Wang L. 2018, *Remote Sensing Big Data: Theory, Methods and Applications*, «Remote Sensing», vol. 10, pp. 711.

Ma Y., Wu H., Wang L., Huang B., Ranjan R., Zomaya A., Jie W. 2015, «*Remote sensing big data computing: Challenges and opportunities*, Future Generation Computer Systems», vol. 51, pp. 47-60.

Mayer-Schonberger V., Cukier K. 2013, *Big Data: A Revolution that Will Change How We Live, Work and Think*, John Murray, London.

Mashey J.R. 1998, *Big Data … and the Next Wave of InfraStress*, Slides from invited talk, Usenix. <https://static.usenix.org/event/usenix99/invited_talks/mashey.pdf> (08/20).

Antonio J., Plaa Chein I. 2007, *Chang High Performance Computing in Remote Sensing*, Chapman & Hall/CRC.

Puletti N., Chianucci F., Castaldi C. 2018, *Use of Sentinel-2 for forest classification in Mediterranean environments*, «Annals of Silvicultural Research», vol. 42, n. 1, pp. 32-39.

Rathore M.M., Ahmad A., Paul A., Rho S. 2016, *Urban planning and building smart cities based on the Internet of Things using Big Data analytics*, «Computer Networks», vol. 101, pp. 63-80.

Skytland N. 2012, *Big data: What is NASA doing with big data today*, Open. Gov open access article < https://open.nasa.gov/blog/what-is-nasa-doing-with-big-data-today/> (09/20).

Jin J., Gubbi J., Marusic S., Palaniswami M. 2014, *An information framework for creating a smart city through Internet of things*, «IEEE Internet Things J.», vol. 1, n. 2, pp. 112-121.

Kitchin A. 2014, *The real-time city? Big data and smart urbanism*, «GeoJournal», vol. 79, pp. 1–14.

Kuhn T. 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Han J., Kamber M.P. 2011, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, Waltham.

Hastie T., Tibshirani R., Friedman. J. 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, New York.

Sabins F.F. 1978, *Remote Sensing: Principles and Interpretation*, Freeman.

Il programma Copernicus e il ruolo della sua Academy all’epoca dei Big Data

Andrea Taramelli
IUSS Pavia, Delegato Nazionale al Committee e allo User Forum Europeo Copernicus
Bernardo De Bernardinis
Coordinatore della Rete nazionale Copernicus Academy
Maria Vittoria Castellani
Ispra, coordinamento Rete nazionale Copernicus Academy
andrea.taramelli@isprambiente.it
bdb.posta@gmail.com
maria.castellani@isprambiente.it

Received: May 2020
Accepted: September 2020

DOI: 10.13128/contest-11868
www.fupress.net/index.php/contesti/

keywords
earth observation
geodata
geoinformation
ict
higher education & training

Copernicus is a very successful European Earth Observation Program, producing a unique volume of data and information and number of operational services in the world. However, for a long time, satellite data, more than the information produced through its operational services, has been mainly used for research purposes. This is not the primary aim of the Programme, that has to be addressed to several others user communities, such as those gathering public administrators, professionals and enterprises, which primarily make use of the produced and free of charge available information. To overcome this limit, since 2016, specific actions have been carried out at the European and National level. The most important of them is the Copernicus User Uptake, aimed at informing and making the end-users aware of the usefulness of what Copernicus produces and provides free of charge. Within this action, the activities designed and implemented through the Copernicus Academy Networks have a particular and great relevance. In order to take advantage of the opportunities offered by Copernicus, knowledge of and skill with advanced Information and Communications Technology (ICT) – such as those related to big data and hypercomputing facilities management, along with the most innovative and available cloud approach – are needed. This paper traces the history and illustrates the role of Copernicus and of its Academy, relating them to the most recent innovative ICT evolution.

Copernicus e i suoi numeri[1]

Il Programma europeo di osservazione della terra Copernicus è il primo al mondo per impegno finanziario, organizzativo e operativo nel monitoraggio ambientale e dei territori. Con un volume di nel 2019 di 158 Pebibyte è il 3° fornitore di dati e informazioni, accessibili gratuitamente, con oltre 300.000 utenti registrati ai numerosi servizi offerti agli abitanti europei.

L’obiettivo primario di Copernicus sono i servizi di comune interesse della Commissione Europea (CE) e degli Stati Membri (SM) dell’Unione Europea, detti ‘*Core Services*’ (*CSs*), destinati a “fornire informazioni sullo stato dell’atmosfera, anche a livello locale, nazionale, europeo e mondiale; informazioni sullo stato degli oceani, anche mediante l’istituzione di un raggruppamento europeo specifico per il monitoraggio marino; informazioni per il monitoraggio del territorio a sostegno dell’attuazione di politiche locali, nazionali ed europee; informazioni a sostegno delle politiche di adattamento e mitigazione dei cambiamenti climatici; informazioni geospaziali a sostegno della gestione delle emergenze, anche attraverso attività di prevenzione, e della sicurezza civile compreso il sostegno all’azione esterna dell’Unione.” [2]

I *CSs* trasformano i dati, raccolti sia da satelliti che in situ, in informazioni a valore aggiunto, detti prodotti (Fig.1).

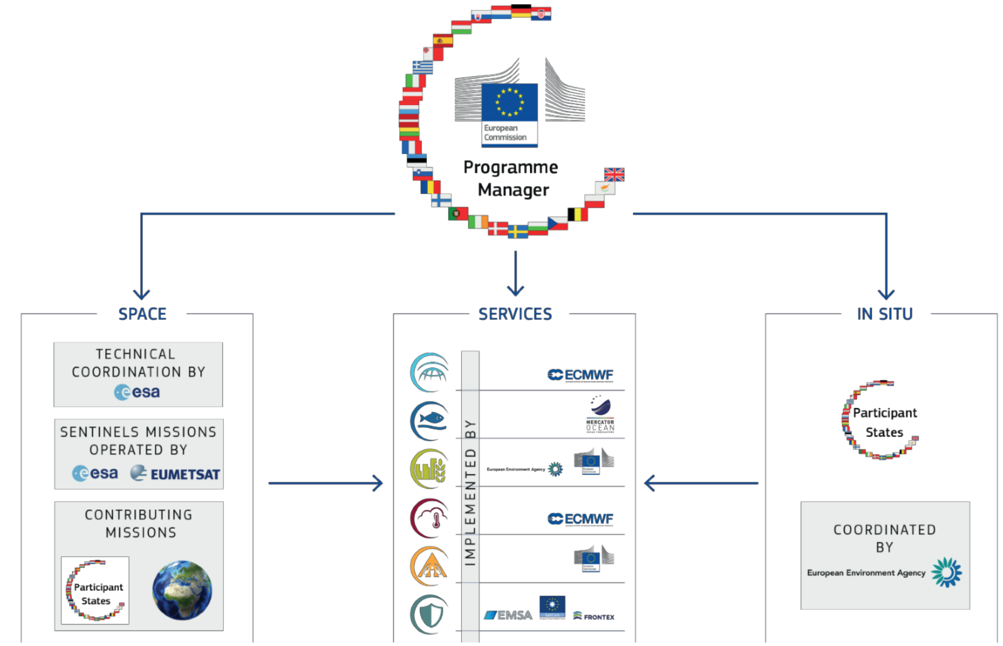
Infatti, i dati dopo essere stati adeguatamente processati, alimentano modellazioni interpretative e simulazioni dei fenomeni monitorati di origine naturale e antropica, e, integrando ulteriori fonti informative, producono informazioni validate ed utili agli utenti finali del Programma (Fig.2). Vengono così prodotte mappe spaziali multi-temporali, identificate caratteristiche e anomalie di eventi e processi biotici ed abiotici, attesi o in atto e rese disponibili elaborazioni statistiche. Il Programma, quindi, presenta tre Componenti strettamente connesse tra loro: quella spaziale, quella relativa ai dati *in situ* e quella relativa ai servizi applicativi che producono le informazioni. Esso rappresenta non solo una grande sfida per l’*Earth Observation* (*EO*), per la *Geoinformation & Geomatic* (*GI*), e per

Dai dati ai prodotti di Copernicus
Fig. 1
da S. La Terra Bella, DG Grow, CE

Il ciclo di Copernicus

Fig.2
da S. La Terra Bella, DG Grow, CE

le *ICT* e le *High Computing Facilities* (*HCF*) attraverso, rispettivamente, la gestione dei *Big Data* e del relativo *Number Crunching*, ma anche una grande sfida culturale. Infatti, il Programma, i suoi obiettivi e contenuti, le sue regole organizzative, funzionali ed operative, i suoi sviluppi e gli investimenti a tal fine necessari, sono primariamente ed ufficialmente guidati dai fabbisogni e dai requisiti espressi dagli utenti finali, organizzati in Comunità, e non determinati dai possessori delle tecnologie per ottenerli. Il Programma Copernicus è governato dalla Commissione con il concorso degli SM, le cui rappresentanze portano le istanze, le opinioni e le volontà degli utenti finali e dei rispettivi Governi nazionali, nell'ambito dello *User Forum* (CEUF) e al *Committee* (COM). Le attività del Programma e i relativi investimenti sono proposti e discussi nell'ambito del CEUF e del COM e, quindi, portati all'attenzione ed approvazione delle diverse Istituzioni dell'UE. A partire dal 2014, il finanziamento complessivamente riconosciuto al Programma dal Parlamento europeo (PE) è pari ad oltre 4.3 Bilioni di Euro. L'implementazione e la gestione delle attività sono state affidate alle *Entrusted Entities* (EE) di riferimento europeo, come *ESA, EEUMETSAT, EEA, JRC, ECMWF, MERCATOR OCEAN, EMSA, SATCEN e FRONTEX*, ciascuno con i propri compiti (Fig.3).

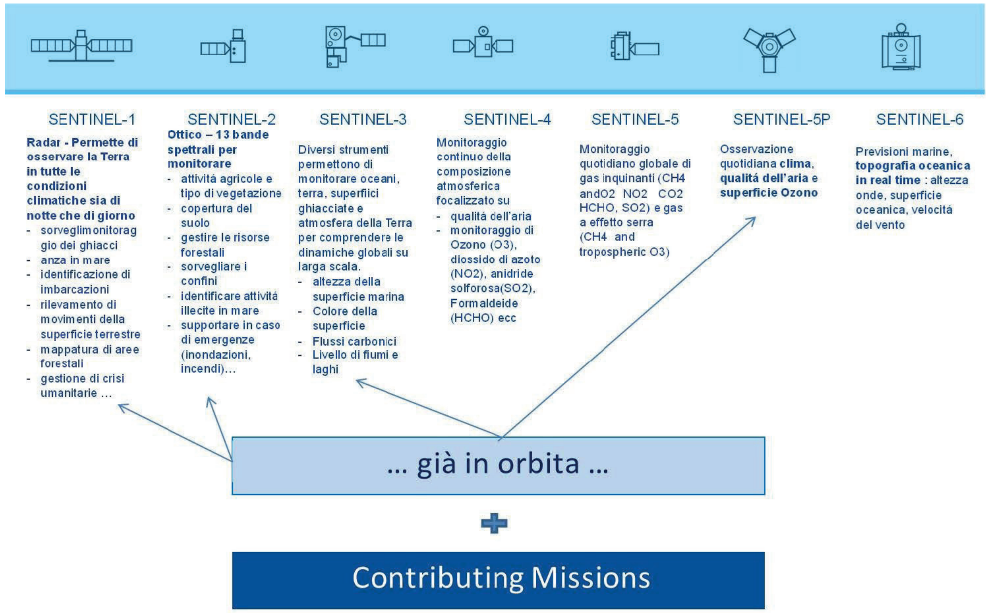


La struttura, le Componenti e le Entrusted Entities di Copernicus

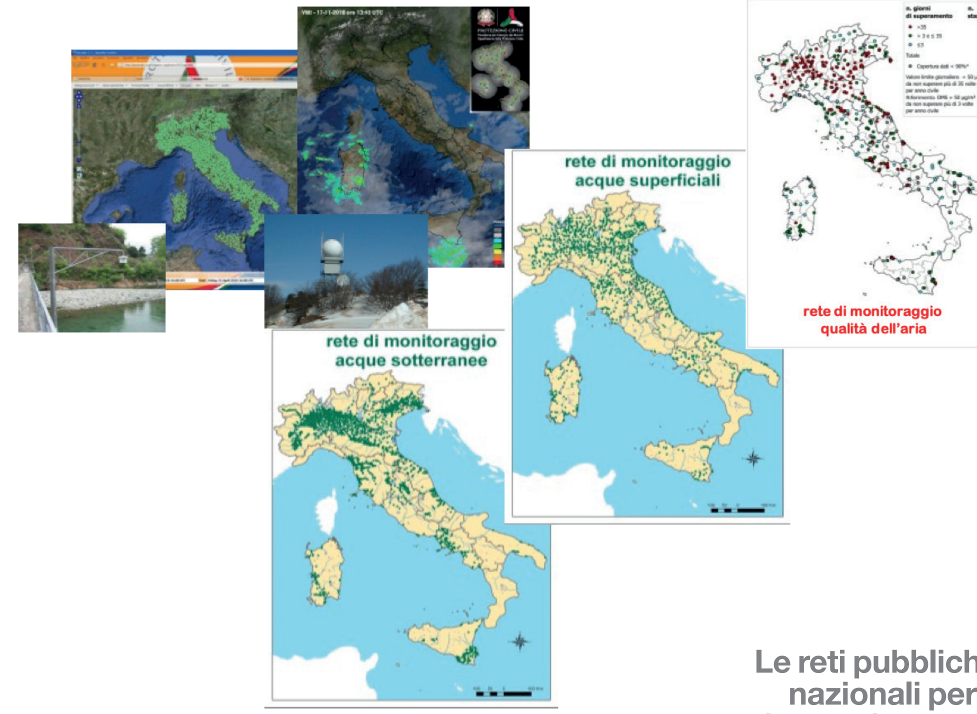
Fig.3
da S. La Terra Bella, DG Grow, CE

La Componente Spazio di Copernicus
L'UE, così come nel caso del Programma Galileo relativamente alla tecnologia e ai servizi di geolocalizzazione, ha deciso di assicurarsi una propria capacità osservativa, autonoma rispetto a quella dei Paesi non appartenenti all'UE, realizzando una costellazione di sistemi, o sub costellazioni, di satelliti di varia natura e capacità, denominate *Sentinel*, e avvalendosi di quelle degli stessi SM, denominate *Contributing Mission*. Le *Sentinel* ad oggi programmate, e finanziate, sono 6, di cui 4 in orbita e già attive. La fig.4 illustra le caratteristiche e le finalità di ciascuna delle sub costellazioni *Sentinel* fornendo un quadro delle *Contributing Mission* oggi disponibili anche per Copernicus. Tra queste ultime è presente anche la costellazione nazionale COSMO *Skymed*, gestita dall'Agenzia Spaziale Italiana.

La Componente dei Dati in-situ di Copernicus
La disponibilità dei dati in situ necessari a Copernicus è affidata alla volontà e responsabilità degli SM, che, evidentemente, è data per acquisita. In molti casi è questa scelta, assieme al processo di validazione, che limita la frequenza temporale della disponibilità e disseminazione delle informazioni prodotte dai *CSs* di Copernicus e non il processo elaborativo in sé. Il processo di validazione è assolutamente necessario per



La costellazione europea delle Sentinel
Fig. 4



Le reti pubbliche nazionali per il monitoraggio meteo idrologico, della qualità delle acque e dell'aria
Fig. 5

rispondere soprattutto alle finalità applicative e di monitoraggio dell'attuazione delle diverse Direttive e Regolamenti europei presso gli SM. Al fine di dare una misura della dimensione ed estensione delle molteplici e variegate tipologie di tali dati, per il nostro Paese e facendo riferimento solo a servizi istituzionali, possiamo contare oltre 11.000 stazioni di misura delle reti e dei sistemi di monitoraggio meteoidropluviometrici e circa 4000 relative alla qualità delle acque e dell'aria con più di 360 sostanze inquinanti complessivamente monitorate (Fig.5). Larga parte di tali dati sono veicolati verso Copernicus e quindi verso i *CSs* attraverso la rete europea EIOnet gestita dall'*EEA*.

La Componente Servizi di Copernicus

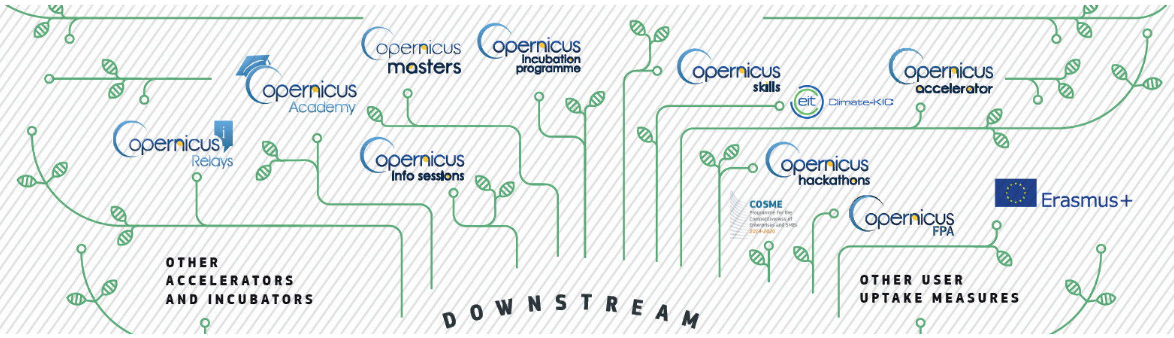
Come è stato già evidenziato, l'obiettivo primario di Copernicus sono i *CSs* e le massive informazioni da essi prodotti. Ciò è all'origine già di GMES e si avvia, sulla base delle conoscenze e dei dati allora esistenti, con tre servizi preoperativi: il *Land*, il *Marine* e l'*Emergency*, ed è quest'ultimo che per primo diventa pienamente operativo nel 2010. È a tale data che anche il loro finanziamento inizia ad essere riconosciuto nell'ambito del Piano Finanziario Pluriennale (MFF) dell'EU. Tranne il *Copernicus Security Service* (CSS), i *CSs* sono accessibili a tutti e ciascuno di essi offre gratuitamente un ricco portafoglio di prodotti con requisiti spaziali e temporali, anche metrici e giornalieri, in funzione del servizio e del prodotto stessi. Elencandoli nell'ordine di avvio della loro operatività, l'*Emergency Management* (CEMS), il *Land Monitoring* (CAMS), il *Marine Environment* (CMEMS), l'*Atmosphere* (CAMS) e il *Climate Change* (C3S), complessivamente, rendono disponibili oltre 500 prodotti di diversa complessità e livello informativo, operativamente e gratuitamente accessibili, assieme ai dati processati a tal fine, attraverso i rispettivi portali *web* dedicati[3], ciascuno dei quali con le proprie regole e procedure d'accesso, ma comunque aperti a tutti. I dati processati delle *Sentinel*, delle *Contributing Mission* o di altre missioni europee sono resi disponibili anche da *ESA*[4] e da *EUMETSAT*[5] attraverso propri portali *web* (Fig.6). Il portale *web* generale del Programma[6], oltre ad assicurare con continuità un'informativa su ogni aspetto, azione e attività relativa a Copernicus, anche presso gli SM, consente l'accesso a tutti gli altri portali *web* in essere e qui rassegnati. L'articolazione e differenziazione delle modalità di accesso ai dati e alle informazioni e la difficoltà di una gestione degli stessi da parte di utenti non specializzati, né formati allo scopo, rappresenta uno tra gli ostacoli più significativi ad un'agevole comprensione ed utilizzo del Programma stesso.
A questo aspetto si aggiunge frequentemente la critica che la non sempre elevata frequenza temporale della disponibilità e disseminazione delle informazioni prodotte dai *CSs* di

Le home page dei siti web dei diversi Copernicus Core Services
Fig. 6

Copernicus, non è adeguata ad un uso operativo quotidiano delle stesse da parte di un utente finale, in particolare se individuale e generico, che quindi le rigetta come inutili per i suoi fini. Tuttavia, ciò non è vero nel caso di moltissimi prodotti, mentre l'ulteriore, forse principale, ma raramente evidenziato, motivo di difficoltà risiede nel fatto che, per un uso operativo ed individuale di tali dati e informazioni è ineludibile la conoscenza seppur minima dei principi, dei metodi e degli strumenti di Telerilevamento, di Geomatica e di Geoinformazione, condizione non così diffusa presso gli utenti finali stessi. Quanto sopra giustifica il fatto che per lungo tempo, quantomeno sino al 2016, la Comunità della Ricerca e dell'Università ha continuato ad essere il principale e prevalente utente di Copernicus.

Il Copernicus User Uptake e la Copernicus Academy

Certamente già nel 2016 Copernicus era un successo mondiale. Tuttavia, ciò era soprattutto vero per quanto riguardava l'uso dei dati e non delle informazioni prodotte dal Programma e, in particolare, per finalità di ricerca, così com'era all'avvio di *GMES*. Infatti, se nel 2016, con in orbita le *Sentinel* 1 e 2, il solo *Sentinels Data Access* di *ESA* registrava oltre 50.000 utenti, saliti ad oltre 136.000 nel 2018 per un volume complessivo di oltre 55 PBytes, di cui solo il 43% distribuito in Europa, non un pari successo era ascrivibile alle informazioni prodotte dai Copernicus *CSs*. Ad esempio, il *CLMS*, uno dei più popolari tra i *CSs*, nato prima delle *Sentinel,* nel 2016 contava non più di 1.700 utenti registrati, superando i 2.500 nel 2018 per un volume di informazioni scaricate pari a circa 165.000 GBytes. Sino ad allora, assumendo tale risultato più o meno analogo nel caso degli altri *CSs*, le informazioni avevano ricevuto un'attenzione, da parte degli utenti finali, 10 volte inferiore a quella ricevuta dai dati prodotti dalle *Sentinel* e, in entrambi i casi, oltre il 50% degli utenti finali e oltre l'80% dei volumi scaricati erano attribuibili alla Comunità della Ricerca e dell'Università. Non era questo il risultato atteso da un investimento di svariati miliardi di euro resi disponibili prima a *GMES* e poi a Copernicus dai diversi *MFF* approvati dal Parlamento europeo. Infatti, la valutazione intermedia della CE di Copernicus, pubblicata nel giugno 2017, se da un lato evidenziava risultati molto positivi, dall'altro illustrava altresì ancora un Programma 'sulla buona strada per il successo', sottolineando nelle conclusioni finali che comunque le attività per la sensibilizzazione, familiarizzazione e l'uso da parte degli utenti avrebbero dovuto essere ulteriormente potenziate ed estese. Tuttavia, apparve subito chiaro che le ragioni del parziale insuccesso erano variegate ed in qualche modo tra loro connesse, anche se raggruppabili in alcune grandi categorie e quindi fronteggiabili e risolvibili con altrettante specifiche azioni.



L'Ecositema delle azioni di User Uptake di Copernicus
Fig. 7

Certamente tra le ragioni principali persiste la non conoscenza delle possibilità e potenzialità offerte, ma anche la complessità di accesso e di acquisizione di adeguate capacità di utilizzo, alle quali sono da aggiungerne almeno altre due:

- la dimensione e l'alta frammentazione del mercato, in difficoltà nel collegare e far incontrare in modo efficiente domanda e offerta di prodotti e servizi offerti da Copernicus, ed in generale di *EO, GI* ed *ICT*, quando si tratta di utenti non tecnici;
- la difficoltà e l'esigenza di identificare i fabbisogni degli utenti anche non tecnici; infatti, nonostante lo sforzo rilevante della CE per identificare e raccogliere tali fabbisogni a sostegno della progettazione del sistema complessivo che fa capo a Copernicus, è ancora limitata la conoscenza, l'attenzione prestata e la presa in carico dei fabbisogni degli utenti non tecnici nella progettazione di prodotti e servizi da parte di imprese industriali e commerciali.

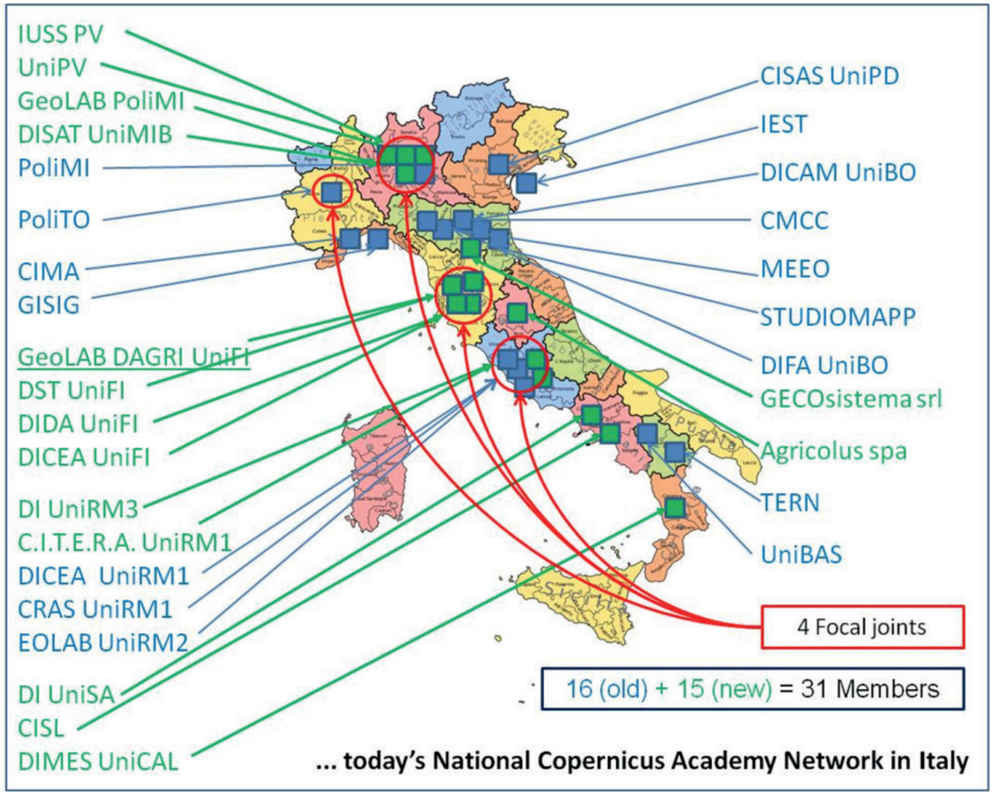
Tali ostacoli diminuiscono di molto la grande opportunità e attrattività offerta dalla gratuità e dal libero accesso a dati ed informazioni, i cui costi, in passato, erano sostenibili solo da pochi soggetti, pubblici e da grandi imprese private, determinando la nascita e la crescita di intermediari culturali e professionali, fornitori istituzionali e commerciali di strumenti e servizi di *EO, GI, ICT* e *HCF*. D'altra parte, anche la presenza e la crescita di tali soggetti sono auspicate, sostenute e trovano spazio in Copernicus, ma ciò, solo e comunque, ove avvenga sulla base di fabbisogni e requisiti espressi da utenti educati ed addestrati allo scopo e non disponibili a subire passivamente ed acriticamente un'offerta non qualificata, né utile, ancor più se onerosa.

Per fronteggiare tali ostacoli e conseguire gli obiettivi prioritari ed irrinunciabili del Programma Copernicus, a partire dal 2016, la CE e gli SM decisero di varare una serie di azioni, come:

- misure di accompagnamento a livello nazionale e regionale per massimizzarne l'uso;
- iniziative volte a migliorare l'accesso e lo sfruttamento di Copernicus, abilitando altresì processi *cross-fertilisation* con altre fonti di dati e informazioni e ciò attraverso piattaforme abilitanti servizi avanzati di *ICT*, e non solo, destinati ad accrescere la capacità e le possibilità di diffusione dei prodotti;
- iniziative volte a promuoverne ulteriormente l'adozione nelle politiche dell'UE, in particolare, per soddisfare le esigenze di autonomia e di sicurezza e per migliorare la capacità di risposta alle grandi sfide, anche culturali, quali quella del *Green New Deal*.

Tuttavia, la più ampia, complessiva e trasversale di tali azioni resta certamente quella di *User Uptake* (*UU*) che mira ad una maggiore informazione, sensibilizzazione e presa di coscienza e conoscenza da parte dei potenziali utenti finali dell'utilità dell'uso di quanto prodotto e reso loro gratuitamente disponibile da Copernicus, per le loro specifiche finalità ed attività, anche quotidiane, e per aumentare le possibilità di successo delle imprese.
Lo *UU* è un processo complessivo (Fig.7), non lineare e particolarmente articolato che vede l'implementazione di nove misure principali, tra loro connesse anche dall'implementazione fattane a valle e liberamente da ciascuno degli SM. Nel nostro Paese le attività di *UU* sono coordinate e assicurate attraverso lo User Forum Nazionale di Copernicus (CNUF), le Rappresentanze delle diverse Comunità degli utenti, i Tavoli e le Reti di coordinamento presenti in esso.

L'estensione della Copernicus Academy in Italia nel 2020
Fig. 8



È importante notare che tra tali misure:

- le Copernicus *Info Sessions* hanno una finalità informativa generale;
- i Copernicus *Masters, Hackathons, Accelerator, Incubation Programme* rappresentano una filiera per la promozione ed il supporto alla nascita di Start-up e/o nuove Imprese;
- le reti europee dei *Relays* e dei *Copernicus Academy* si propongono, in modi diversi e complementari rispettivamente, di informare, ascoltare ed assistere localmente sia le PPAA che le imprese e di informare, educare, formare ed addestrare gli utenti finali, anche potenziali;
- i Copernicus *Skills*, anche in stretto legame con il mondo dei Progetti Erasmus+, si propongono l'identificazione e la definizione delle nuove e/o rinnovate professioni e professionalità connesse con l'*EO*;
- il *Framework Partnership for Copernicus User Uptake* (FPCUP), è la misura gestita da un consorzio europeo tra CE e SM, destinato progressivamente dal 2018 a sostenere finanziariamente tutte le attività di *UU*; tale misura prevede una pianificazione triennale comune, prima a livello nazionale e quindi a livello europeo, coerente con quattro assi principali di attività: di *UU* solo nazionali, di *UU* pan ed extra europee e di sviluppo di prodotti e servizi downstream accoppiati con attività di *UU*.

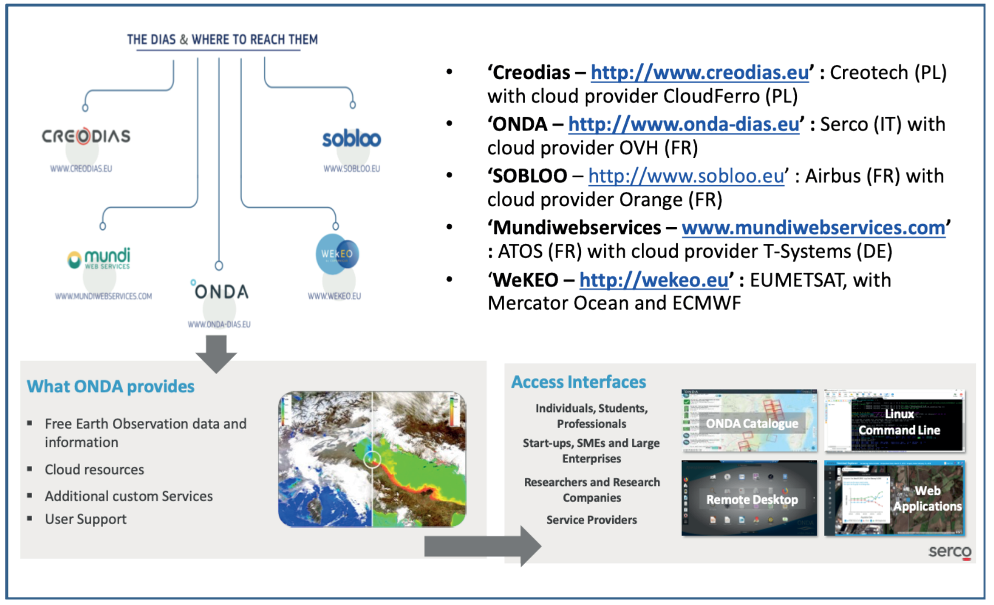
La Copernicus Academy (CA), la sua storia ed organizzazione a livello nazionale

Per capire il ruolo ed il significato della *CA* è opportuna una riflessione sul duplice ruolo che l'Accademia gioca in Copernicus. Infatti, il ruolo dell'Università, insieme ai diversi soggetti della Ricerca, sia pubblica[7] che privata, è quello di membro della grande Comunità degli attori istituzionali e delle PPAA, ma contemporaneamente in tutti gli SM è responsabile della formazione universitaria e di partecipare con un ruolo primario nell'alta formazione post universitaria dei futuri funzionari pubblici, professionisti, imprenditori, insegnanti e ricercatori, nonché della classe dirigente del Paese.

È evidente che la scelta di istituire una Rete Europea dei *CA* fa riferimento a questo secondo ruolo e nasce da almeno tre constatazioni, cioè che:

- il coinvolgere e formare, nonché addestrare i futuri ideatori, realizzatori ed utilizzatori di Copernicus circa i principi, ai metodi ed agli strumenti quanto meno di *EO, GI ed ICT*, è il modo più efficace e duraturo per raggiungere gli obiettivi di *UU* nel medio e lungo periodo;
- in quasi tutte le misure di *UU*, in un modo o nell'altro, la formazione gioca un ruolo centrale e prioritario;
- solo introducendo elementi innovativi nei percorsi educativi e nei processi formativi ed addestrativi sarà possibile promuovere nuovi profili professionali e/o modificare quelli già esistenti, per rispondere ai modelli e quindi ai processi di business e/o aziendali emergenti, connessi o contaminati dall'*EO* e dalla *GI*.

Un recentissimo rapporto su '*Business processes and occupational profiles*', prodotto nell'ambito del progetto europeo *EO4GEO*[8], ha identificato, attraverso l'applicazione della metodologia *Business Process Modelling and Notation* (*BPMN*) a svariate tipologie d'evento, almeno 30 nuovi modelli di *business*, ragione di oltre 200 nuovi obiettivi e profili professionali in domini applicativi di interesse di Copernicus, quali il Cambiamento Climatico e la Gestione delle Aree Urbane.



I DIAS ed i servizi ICT del DIAS ONDA di SERCO Italia
Fig. 9

Vista l'autonomia lasciata dalla CE a ciascun SM nell'organizzare le proprie attività nazionali di *UU*, nel 2018, il *CNUF*, su proposta della Delegazione Nazionale, ha deciso di nominare un Coordinatore e di istituire un Coordinamento nazionale dei membri *CA*, chiamando questi ultimi ad animare una Rete Nazionale (RNCA), in analogia e rapporto con quella europea.
Gli obiettivi e le attività della RNCA sono focalizzati e mirati alla realizzazione, secondo format condivisi, di eventi, percorsi e processi educativi, formativi ed addestrativi, universitari e/o di alta formazione, tanto in capo ad Atenei e/o alle PPAA, come quelle regionali, quanto quelli posti in essere da soggetti privati, quali gli ordini professionali, oppure le imprese di *OT, GI* e *ICT*, che offrono assistenza a soggetti interessati a Copernicus.
Nel 2018 i membri nazionali della *CA* erano 16, soggetti singoli e in maggioranza accademici, quasi sempre non in relazione tra loro. Oggi i membri ammontano ad oltre 30, tutti partecipanti attivamente alla RNCA e quasi tutti soggetti non più individuali, ma collettivi, come Dipartimenti e Centri universitari di interfacoltà e interateneo, nonché soggetti nuovi come imprese di supporto e assistenza ad attività produttive specifiche, ma anche impegnate in attività e fornitrici di servizi di alta formazione, formazione tecnico-scientifica ed addestramento, così come mostrato in fig.8.
Tali risultati sono da attribuirsi all'attività del Coordinamento nazionale con il concorso di una RNCA in continua espansione, concretizzatasi, in particolare, in incontri e seminari presso Atenei, Dipartimenti e Centri Universitari, Enti ed Agenzie erogatrici di servizi di pubblico interesse ed imprese, in *Workshop* nazionali e *Open School* tematiche e nella coordinata partecipazione ai bandi del *FPCUP*.
Per *l'Action Plan* 2020-2023 del *FPCUP*, la RNCA ha predisposto 11 delle 18 proposte avanzate a livello europeo dal nostro Paese, contribuendo a portare l'Italia ad un 35% circa del totale, in valore, delle complessive proposte presentate da tutti gli SM.

Copernicus, i DIAS e l'Agenda digitale europea
I recenti rapidi progressi nel C*loud computing* hanno avuto un impatto significativo sul mercato dei servizi di *EO, GI, ICT e HCF* a valle di Copernicus, riducendo significativamente i costi degli spazi di memoria e della potenza di calcolo necessari a tal fine. Il *Cloud computing*, combinato con una spinta alla digitalizzazione, ha aperto le porte a molte nuove soluzioni e approcci per la gestione dei dati, ma altresì, l'approccio e le metodologie *Big data*, cioè la capacità e la possibilità di gestire e far interagire un grande volume di dati strutturati e non strutturati, hanno avuto ed avranno anche in futuro una particolare rilevanza nello sviluppo di Copernicus stesso. Per dare una dimensione al tema, quando tutte le *Sentinel* saranno operative, Copernicus sarà in grado di produrre oltre 10 PByte di soli dati satellitari ogni anno, a cui si aggiungerà l'altrettanto significativo volume di informazioni prodotte dall'ulteriore sviluppo dei servizi operativi già esistenti e da quelli nuovi che si stanno o saranno realizzati, come i servizi per il *Ground motion*, i *Cultural Heritages*, il *Costal area management*, il *CO_2 monitoring* e altri ancora che saranno decisi in futuro da CE e SM.
E' apparso quindi immediatamente chiaro che, oltre a dover gestire in modo semplice ed efficace un volume estremamente significativo e crescente di dati ed informazioni per una platea altrettanto crescente di utenti di varia natura e con obiettivi molto diversificati tra loro, era necessario garantire la possibilità di accedere ad altri servizi *ICT* e *HCF*, ancor più avanzati, per lo sviluppo di nuovi prodotti e servizi innovativi di Copernicus a valle di esso, nonché per lo sviluppo e l'efficacia di molte attività di *UU*. Ciò anche per rispondere all'esigenza di autonomia e di sicurezza dell'Europa in tale ambito, coerentemente con le politiche e gli investimenti europei per la '*Digital Transformation & Innovation*' nell'ambito dell'Agenda Industria 4.0.
Si è quindi innanzitutto ritenuto necessario semplificare gli esistenti sistemi di accesso a dati ed informazioni prodotti, tanto eterogenei tra loro quanto differentemente articolati al loro interno. Tale azione ha dato vita a cinque sistemi di *Data Information Access Services* (DIAS), ciascuno dei quali, oltre a garantire un unico punto di accesso a quanto prodotto e reso disponibile da Copernicus, omologo a quello degli altri e

sempre completamente gratuito, è capace altresì di offrire ulteriori, avanzati ed innovativi servizi a pagamento, soprattutto di *ICT* e *HCF*, utili per rispondere a domande del potenziale utente finale tanto specifiche e/o innovative da non poter essere soddisfatte da quanto già offerto da Copernicus (Fig.9).
Certamente Copernicus è un Programma europeo di successo mondiale, produttore di volumi di dati, informazioni e numero di servizi operativi unici al mondo e gestibili solo con metodologie di *ICT* avanzate, come quelle relative ai *Big Data* ed alle piattaforme *Cloud* più innovative. Tuttavia, per lungo tempo è stato utilizzato prevalentemente, anche dalla Comunità accademica, per finalità di ricerca, mancando le proprie prioritarie finalità, cioè di essere utilizzato soprattutto da un'utenza finale, pubblico funzionario, professionista o piccolo imprenditore che sia, per rispondere ai propri obiettivi, anche di mercato.
Tali criticità nascono da diversi fattori, ma soprattutto dalla complessità intrinseca del Programma e dei suoi contenuti applicativi che spaziano a tutto tondo, accresciuta ulteriormente proprio dalla significativa produzione e disponibilità di dati ed informazioni, nonchè dalla non conoscenza né delle metodologie, né degli strumenti necessari per fronteggiare tali difficoltà, a partire da quelli propri del Telerilevamento, della Geomatica e della Geoinformazione, per arrivare a quelli della gestione dei *Big Data*.

La Comunità accademica così come gli altri soggetti della ricerca non hanno tali carenze, né la complessità rappresenta una difficoltà, ma forse una sfida, mentre per l'utente finale, generico, non specificamente formato, tutto ciò rappresenta una barriera insormontabile.
I membri della *CA*, il loro Coordinamento e la RNCA, devono e possono abbattere tale barriera, nel medio e lungo periodo, promuovendo una formazione universitaria, e non solo, ad alta innovatività, trasversale e interdisciplinare, mirata anche ai nuovi profili professionali richiesti dal mercato e, nel brevissimo periodo, partecipando quale acceleratore nei diversi processi di *UU*.
È così che l'*OT*, la *GI* e l'*ICT* dei *Big Data*, del *Cloud* e dell'*Hypercomputing* e quanto prodotto da Copernicus e da altri Programmi europei come Galileo, diventano *Facilities* e *Popular,* quasi quanto gli *Smartphone* che utilizziamo quotidianamente.
È così che Copernicus consolida i suoi obiettivi di "contribuire al *soft power* dell'UE a livello globale" e di "essere uno strumento per lo sviluppo economico, chiave per l'economia digitale"[9].

Note

[1] I dati contenuti in questo articolo fanno riferimento al gennaio 2020.
[2] Regolamento (UE) N.377/2014
[3] http://land.copernicus.eu http://atmosphere.copernicus.eu http://marine.copernicus.eu http://emergency.copernicus.eu http://climate.copernicus.eu
[4] https://spacedata.copernicus.eu/ https://scihub.copernicus.eu/
[5] https://www.eumetsat.int/website/home https://www.eumetsat.int/website/home/Data/DataDelivery/CopernicusOnlineDataAccess
[6] https://www.copernicus.eu
[7] In Italia la ricerca, oltre che presso le Università è portata avanti dagli Enti Pubblici della Ricerca (EPR), sia vigilati dal MUR, come il CNR, che vigilati e indirizzati da un Ministero di riferimento, come ISPRA con il MATTM.
[8] EO4GEO è una Erasmus+ Sector Skills Alliance ormai in fase conclusiva, ma che ha saputo raccogliere ben 26 partecipanti da 13 SM, con l'obiettivo di definire una strategia sostenibile e di lungo periodo per identificare i gap tra domanda ed offerta educativa e formativa in materia Space, EO, GI e ICT alla luce dei recenti sviluppi sia tecnologici che non tecnologici, nonché proporre dei percorsi per colmarli.
[9] Stefano La Terra Bella, DG GROW, CE

Twin digital cities
la vera 'intelligenza' della città digitale nel XXI secolo

Elisa Pennacchia
CITERA - Centro Interdipartimentale Territorio Edilizia Restauro Ambiente
Federico Cinquepalmi
ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale
elisa.pennacchia@uniroma1.it
federico.cinquepalmi@miur.it

DOI: 10.13128/contest-11601
www.fupress.net/index.php/contesti/

keywords
big data
smart sustainable urbanism
urban digital twin
smart age-friendly city
energy efficiency

Cities are 'living' organisms that born, grow, change rapidly and age (Pagani, Chiesa 2016); they are increasingly complex systems linked to changing demographic, economic and environmental changes. Urban complexity therefore requires planning capable of rapidly providing suitable responses to the constantly evolving human needs.
The Big Data revolution offers an active innovative approach to building processes and urban and territorial planners, making them able to face the fluidity of contemporary urban spaces and the urgent challenges related to sustainable development. This essay is part of this debate by offering a survey on how Big Data relate to human activities and the built environment and by producing a qualitative change, supporting on the one hand planning and on the other hand a regenerative impact on the development of urban settlements.

Introduzione

All'alba di quello che è stato definito 'il quarto paradigma della scienza', la società contemporanea sta vivendo una rivoluzione scientifica che è segnata dalla crescente importanza dei *Big data* e dallo sviluppo dell'analitica predittiva, e delle relative tecnologie (Bibri, 2019)[1].
Lo scenario del XXI secolo si contraddistingue infatti sempre più per le interazioni tra il mondo fisico e quello virtuale, grazie alla progressiva creazione di uno spazio connettivo globale ad elevata intensità di flussi informativi, alle potenzialità dell'*Information Communication Technology* (ICT), così come dell'*Internet of Thinghs* (IoT), dei *Big Data*, della *Virtual e Augmented Reality* e alla diffusione di dispositivi computazionali via via più potenti, le cui elevate capacità elaborative, sia pur prive di discrezionalità, vengono definite, ovvero la cosiddetta *Artificial Intelligence e Machine Learning*, ed i relativi algoritmi predittivi.

Si tratta di una vera e propria rivoluzione che, dopo quelle legate all'introduzione della spoletta volante nel telaio meccanico, della macchina a vapore, dell'elettricità, dell'elettronica e dell'IT per l'automazione dei processi produttivi, è stata avviata grazie allo sviluppo del digitale e alla *Sharing economy*.
Al contempo il nuovo paradigma di sviluppo urbano legato alla sostenibilità rigenerativa richiede un nuovo approccio per una migliore comprensione dei processi che cambiano rapidamente il volto delle città contemporanee, approccio che permetta di individuare e progettare soluzioni più efficienti (Kamrowska-Zaluska, Obracht-Prondzynska, 2018).
L'aumento della connettività funzionale e virtuale dello spazio urbano consente di adottare soluzioni *Smart* e innovative pal fine di andare verso una maggiore sostenibilità e resilienza delle città. In questo contesto, i *Digital twin models*, grazie all'acquisizione e all'analisi dei *Big data* ottenuti attraverso sensori, GPS, *Social media*, reti intelligenti, dati istituzionali o registrazioni di clienti e transazioni, offrono un supporto fondamentale per affrontare le sfide complesse che coinvolgono l'ambiente costruito.
Il presente studio si propone di valutare le possibilità offerte da alcuni strumenti basati sull'utilizzo dei Big Data a sostegno della progettazione e della pianificazione, nel rispetto dei principi dello sviluppo sostenibile, presentando un esempio progettuale significativo per evidenziare il potenziale supporto ad un approccio rigenerativo.

Big Data e Digital twin models

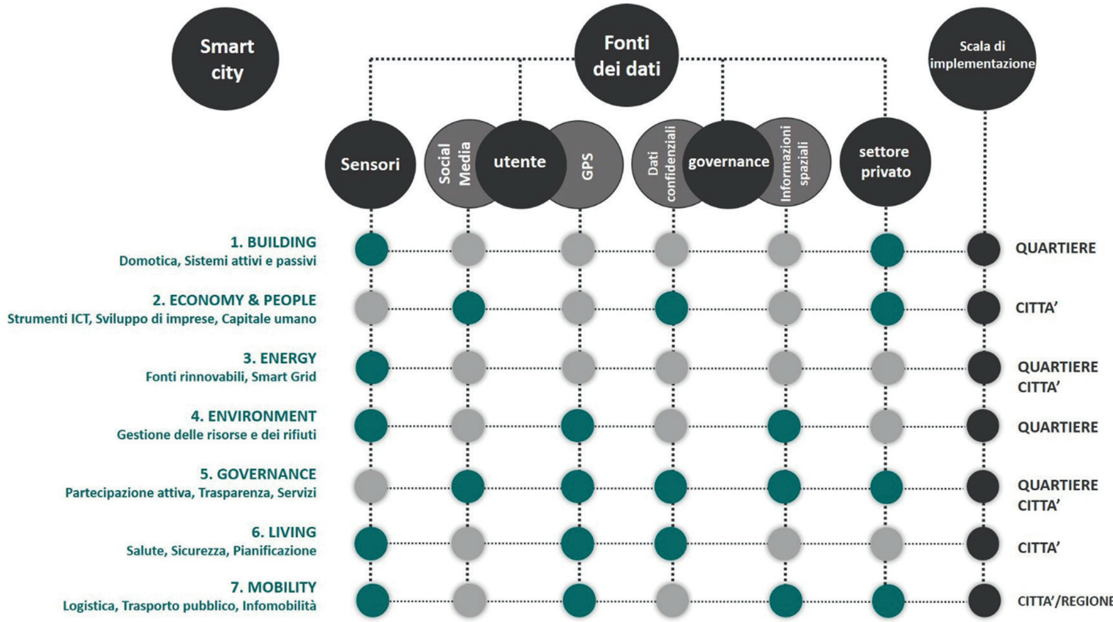
Negli ultimi anni si è assistito ad un drastico aumento della capacità di raccogliere dati da vari sensori, dispositivi cosiddetti intelligenti, dall'*Internet of Things (*IoT) e dalla proliferazione del *Cloud Computing* (Botta et al., 2016), dati che vengono raccolti in formati differenti, da applicazioni che possono essere tra loro indipendenti oppure interconnesse. Per *Big Data* si intendono gli insiemi di dati numerosi e complessi, generati, catturati ed elaborati molto rapidamente, che i tradizionali software applicativi di elaborazione dati non sono in grado di gestire (Babu Sriramoju, 2017). Il termine è stato coniato a metà degli anni '90 dall'informatico statunitense John Mashey, direttore della

Silicon Graphics Inc. (SGI) (Diebold, 2012). La necessità di confrontarsi con i *Big Data*, e la relativa opportunità di usufruirne in maniera efficace è ora un tema chiave per molti settori lavorativi e di ricerca, facendo di tale tema scientifico quello probabilmente più rilevante del nostro tempo.
Nell'ambito delle *Digital city* si fa sempre più strada la raccolta dei *Big Data* per testare e creare sofisticate simulazioni relative tanto ai processi urbani, quanto agli aspetti comportamentali dei cittadini (Dembski et al., 2020). In questo ambito i *Digital twin models* consentono di modellare e riprodurre un oggetto/sistema fisico reale all'interno di un ambito virtuale, dove poter svolgere studi, analisi costi-benefici di soluzioni alternative e simulazioni varie. Le applicazioni sono varie e coinvolgono molti settori a partire da quello aerospaziale[2], dove un *Twin model* è stato testato per la prima volta, ma anche nel settore industriale e più recentemente in quello edilizio ed urbano. Il termine *Digital Twin* risale ad una presentazione svolta da Grieves[3] nel 2003 relativa alla gestione del ciclo di vita del prodotto (*Product Lifecycle Management* - PLM) presso l'Università del Michigan (Farsi et al., 2020), ed il gemello digitale è stato impiegato inizialmente per simulare le prestazioni di un dispositivo o di un'apparecchiatura in ambienti diversi oppure con finalità predittive rispetto a guasti e malfunzionamenti.

Nell'ambito della pianificazione e della gestione urbana i *Big Data* e l'*Urban digital twin* stanno acquisendo sempre più importanza, tenuto conto della crescente complessità gestionale degli agglomerati urbani, nonché degli obiettivi di sviluppo sostenibile (SDGs) adottati dall'Assemblea generale delle Nazioni Unite il 25 settembre 2015 (la cosiddetta Agenda 2030), e dell'attuazione della Nuova Agenda Urbana dell'ONU che spetta a chiunque abbia un ruolo nella *governance* urbana, sia a livello nazionale che locale.
Le città costituiscono infatti a livello globale le fonti principali di inquinanti ambientali (Cinquepalmi, 2019) ed i luoghi dove fronteggiare la vulnerabilità ai rischi climatici ed ai relativi sconvolgimenti, affrontare le maggiori sfide socio-economiche e della salute delle persone, come ha ben dimostrato la Pandemia causata dal virus COVID-19.
Pertanto, esse sono i luoghi più importanti in cui avviare la transizione verso la sostenibilità, collegando l'agenda dello sviluppo sostenibile con quella dell'ICT della ricerca e dello sviluppo digitale pervasivo.
Sarà pertanto fondamentale, usufruendo dei dati raccolti da sorgenti diverse, come ad esempio gli archivi dell'istituto nazionale di statistica, oppure ottenibili in tempo reale grazie alla sensoristica avanzata e ai dispositivi IoT, l'adozione di sistemi integrati ed intelligenti di gestione che permettano agli amministratori del territorio di monitorare costantemente l'evoluzione dello stato dei sistemi, e di tutti i sottosistemi relativi.
Questo approccio consentirà di apportare delle variazioni sia a breve sia a lungo termine, per rispondere alle esigenze mutevoli degli utenti, pianificando al contempo azioni di riqualificazione e rigenerazione urbana.
Dalla progettazione degli attuali edifici intelligenti i possibili miglioramenti consentiti dalle tecnologie di rilevamento anche a scala urbana sono immensi, come ad esempio:

- l'uso più efficiente di risorse quali energia, acqua, materiali, e naturalmente un più efficente smaltimento dei rifiuti;
- un maggior controllo pterritoriale er disincentivare atti vandalici e criminali e garantire la sicurezza pubblica;
- la conservazione del paesaggio naturale e l'uso sostenibile del suolo;
- l'efficientamento del sistema dei trasporti pubblici;
- il miglioramento della qualità dei servizi sanitari forniti ai pazienti anche mediante telemedicina;
- l'integrazione di servizi e il miglioramento di quelli esistenti per rispondere alle esigenze delle diverse fasce d'età della popolazione.

Il gemello digitale si aggiornerà costantemente in base al notevole flusso di informazioni raccolto, implementando continuamente i possibili scenari per fornire un supporto rilevante per le più svariate operazioni, orientate all'automazione e alla gestione *smart* dell'ambiente costruito, anche attraverso l'integrazione di sistemi di Intelligenza Artificiale che possano contribuire all'ottimizzazione funzionale delle attività interconnesse al ciclo di vita degli immobili e delle infrastrutture. L'uso crescente dell'*Internet of Things* (IoT) dovuto alla diminuzione del costo della tecnologia, permette infatti l'implementazione dei futuri *Digital twin* urbani grazie al conseguente aumento della condivisione delle informazioni; il gemello digitale consente un approccio olistico in termini di ottimizzazione trasversale della progettazione, della gestione e del funzionamento delle infrastrutture urbane. La strategia apparentemente più efficace non sarebbe però quella di modellare un singolo gemello digitale per un'intera città, ma piuttosto di creare un'aggregazione e integrazione di *digital twin* specifici per sistemi definiti, come gli edifici intelligenti, le infrastrutture del traffico, le reti energetiche e idriche.
Per realizzare un *Urban digital twin* è fondamentale creare un modello 3D dell'ambiente costruito utilizzando la metodologia *Building Information Model* (BIM) che consente di definire un database strutturato da oggetti parametrici caratterizzati quantitativamente e



Esempi di fonti di dati utili per ogni settore urbano
Fig. 1 (elaborazione degli autori)

qualitativamente, in cui possono essere inseriti dati inerenti alle prestazioni termo-acustiche dei materiali e alla tecnologia HVAC[4] e i dati delle reti elettriche, idriche e fognarie possono essere così associati al gemello digitale urbano.
L'integrazione del sistema BIM con il *Geographic Information System* (GIS) che fornisce una rappresentazione del contesto territoriale/urbano in cui gli edifici sono inseriti, consente di realizzare un modello informativo digitale anche a scala territoriale.
La modellazione della città può essere integrata con dati più dettagliati ottenuti grazie a sistemi laser scanner mobile 3d a terra ed informazioni satellitari come quelle fornite dal sistema europeo Copernicus[5].
Il modello viene arricchito con i dati provenienti da una rete di sensori che rilevano costantemente misure relative ad esempio all'aria, all'acqua, al calore, al traffico ecc., permettendo la creazione di serie storiche e modelli previsionali con un'elevatissima attendibilità (Fig. 1). I dati provenienti dal GIS e dalla sensoristica permettono di acquisire una maggiore consapevolezza del contesto ambientale e territoriale, che consente una comprensione immediata degli impatti positivi o negativi legati alle scelte progettuali prima della loro realizzazione.

Il caso di studio: il progetto SmartMed

Una proposta di ricerca avanzata in grado di soddisfare tutti gli obiettivi di cui sopra, ma che si concentra sulle questioni energetiche, è il progetto *SmartMed - Mediterranean Flexible Energy Communities* - (SM) che riunisce due *Lighthouse Cities* (LHC), Roma e Siviglia, insieme a sette *Fellow Cities* (FC) e una grande partnership composta da istituzioni accademiche e scientifiche di ricerca e tecnologia tra cui il Centro Interdipartimentale Territorio Edilizia Restauro e Ambiente (CITERA) dell'Università Sapienza di Roma.
Il progetto mira a contribuire alla transizione energetica attraverso infrastrutture integrate intelligenti mediante la realizzazione di Blocchi Energetici Positivi (*Positive Energy Blocks* - PEB), lavorando in una prospettiva di Distretto Energetico Positivo (*Positive Energy District* - PED), idoneo per le città del Mediterraneo; il Blocco Energetico Positivo, da implementare nei Distretti *Smart* e *Low Energy* di Roma è Pietralata (Fig.2).
Il progetto fornirà una 'Piattaforma di Interoperabilità *SmartMed*', la cui progettazione seguirà il concetto di Meccanismi di Interoperabilità Minimale (*Minimal Interoperability Mechanisms* - MIM) dell'*Open and Agile Smart Cities* (OASC), con un insieme di interfacce semplici per raggiungere una maggiore sinergia.
La piattaforma urbana è un 'sistema operativo' dei servizi forniti dalle smart city, fondamentali per coordinare la crescente varietà dei dati attinenti a settori differenti e degli stakeholder. Essa promuove standard aperti, un'interfaccia di programmazione dell'applicazione aperta API (*Application Programming Interface*).
Grazie alle Open API e ai componenti dedicati forniti dalla Piattaforma di Interoperabilità *SmartMed* sarà possibile, ad esempio, creare una cabina di controllo per il monitoraggio dell'energia, per l'analisi dei dati aziendali, per il controllo continuo degli indicatori chiave di prestazione (*Key Performance Indicator* – KPI) o per la visualizzazione di dati basati su mappe per le applicazioni dell'utente finale. Questa attività è strettamente legata ai componenti GIS/BIM per la mappatura energetica del Distretto e dei PEB. Le cabine di controllo saranno progettate per essere facilmente scalabili in diversi campi di applicazione. La piattaforma per la gestione della domanda di energia termica e per la creazione della smart grid termica, precedentemente sviluppata a livello di Blocchi Energetici Positivi per definire in tempo reale i valori di domanda termica e selezionare il funzionamento delle diverse fonti di energia (geotermica, fotovoltaica, chp, accumulo termico), sarà implementata e potenziata al fine di elaborare i dati provenienti dai sensori installati sulla rete termica e di ottimizzare la produzione di energia all'interno del Distretto Energetico Positivo.
La piattaforma sarà integrata con uno strumento web GIS per consentire la partecipazione dei cittadini alla pianificazione urbana innovativa. Questo strumento fornirà mappe 2D-3D con funzionalità di geotagging collaborativo per monitorare le prestazioni delle politiche attuali e per supportare il processo decisionale nello sviluppo di interventi, politiche, nuove proposte, eventi, sessioni di co-working e storytelling.
I nuovi dispositivi elettronici intelligenti (IED) per il monitoraggio delle principali

grandezze elettriche (ad es. energia, potenza, tensione, corrente, ecc.), precedentemente installati nel perimetro del PEB, saranno estesi nella misura del PED, consentendo di aumentare l'osservabilità e l'automazione della rete in bassa e media tensione nell'area oggetto di studio. Saranno implementate nuove funzionalità nel sistema di gestione della rete con l'obiettivo di ridurre la durata e la frequenza delle interruzioni, rilevare eventuali rischi critici e definire le relative misure di mitigazione per garantire la corretta distribuzione delle risorse energetiche nel PED. In questa fase verrà implementato il sistema di comunicazione dell'architettura della piattaforma SmartMED IoT sviluppata sulla base della tecnologia LoRA[6] (acronimo di Long Range), già sviluppata e testata a livello di PEB, al fine di comunicare e sistematizzare tutte le grandezze elettriche ed ambientali relative ad esempio all'inquinamento atmosferico ed acustico, ai livelli di illuminazione, al flusso del traffico veicolare, ecc.

In particolare, la piattaforma di comunicazione IoT sarà implementata per ricevere e supportare una grande quantità di dati relativi alle misurazioni effettuate nelle aree PED più grandi. In questa fase, infatti, sarà estesa all'area del Distretto la dimostrazione di contatori smart 2G e l'illuminazione con pali intelligenti dotati di sensori per la rilevazione di alcune grandezze ambientali (in particolare NO e NO_2, umidità per il controllo degli impianti di irrigazione, rumore), inizialmente installati a livello di PEB. I dati raccolti e registrati sulla piattaforma IoT saranno analizzati e validati con due obiettivi principali: definire e aggiornare i KPI e garantire un monitoraggio costante della qualità dell'aria e delle condizioni elettriche nelle aree PED, durante e dopo la fine del progetto; i dati saranno condivisi con tutti gli *stakeholder* sulla base di un approccio orientato alla rete, ampliando la gamma del sistema di controllo intelligente dai PEB al livello PED (Fig. 3).

In quest'area sarà realizzata una pista ciclabile e pedonale intelligente e innovativa con pannelli fotovoltaici policristallini, rivestita da vetro antiscivolo e inserita in una struttura in acciaio ancorata al suolo, con una lunghezza di 1.450 m e una produzione annua di energia fotovoltaica di 400.000/kWh anno. La pista ciclabile e pedonale trasferirà l'energia in modalità SEU (Sistema Efficiente di Utenza) ai consumatori di energia situati nel PEB.

Il progetto pilota utilizzerà la *blockchain* applicata ad un'infrastruttura tecnologica, permettendo di autenticare le transazioni di dati (energetici, economici, informativi, comportamentali, commerciali, etc.) e sarà implementata e strutturata come segue:

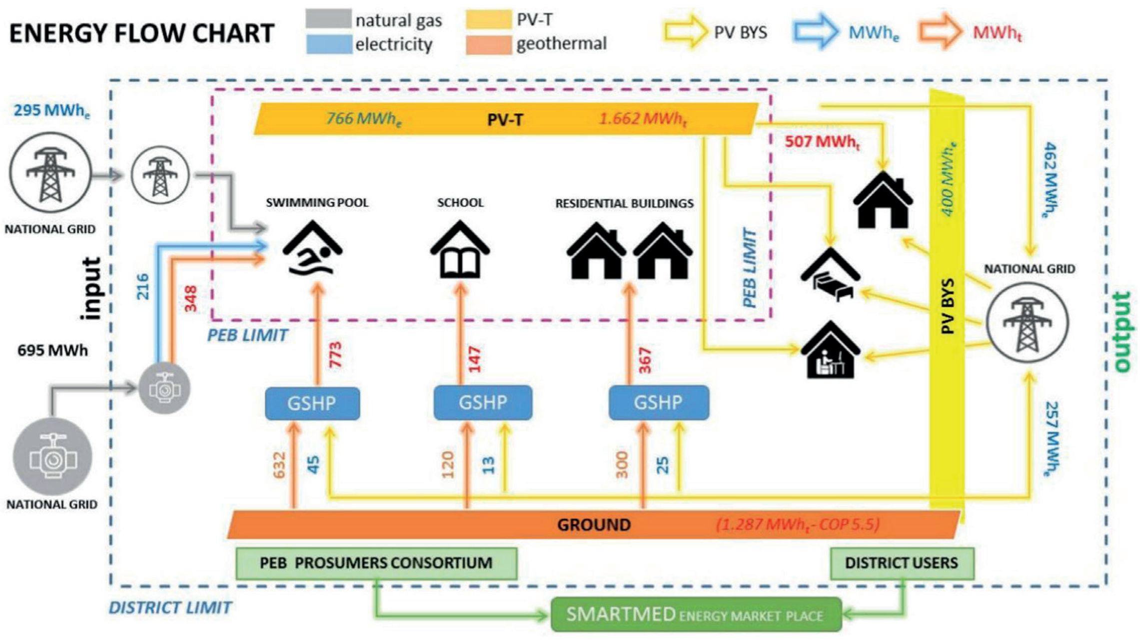
- un sistema fotovoltaico connesso alla rete elettrica (PV) che utilizza un simulatore di rete;
- un sistema di accumulo dell'energia;
- una stazione per la ricarica dei veicoli elettrici (e-car);
- un'area dedicata alle auto elettriche;
- dispositivi Internet of Things (IoT) installati presso i punti di tracciamento delle transazioni pianificate;
- App per registrare e monitorare le transazioni dell'infrastruttura-utenti, in modo tale da rendere l'energia prodotta dalle FER rintracciabile e disponibile per l'utente.

Area del Blocco Energetico Positivo di Pietralata, Roma
Fig. 2 (elaborazione degli autori)

Conclusioni

La diffusione della pandemia COVID 19 nel novembre 2019, tuttora in fase di espansione globale all'atto della pubblicazione del presente testo, pone all'attenzione la necessità di un nuovo approccio gestionale ai sistemi urbani sempre più sovrappopolati. Non è un caso che proprio nell'ambito di città affollate, dove la contiguità tra la specie umana e le specie animali serbatoi di virus è continua, si verifichino spesso quei salti di specie che generano pandemie come quella attuale, mettendo in crisi il modello stesso della città come luogo di sicurezza per i cittadini. L'idea di proporre il modello digitale della città gemella, come strumento integrato per una progettualità e una governance più aderenti alla domanda sociale e ai requisiti dello sviluppo sostenibile, e alla risposta agli eventi estremi, sembra la strada più promettente da seguire per affrontare le complesse sfide urbane e demografiche del XXI secolo.

Tuttavia, per sua natura di modello, il gemello digitale urbano non include tutte le informazioni dal mondo fisico e dalla vita reale. L'obiettivo è raggiungere somiglianze con il mondo reale a un livello di dettaglio sufficientemente accurato per affrontare problemi complessi. La ricerca futura dovrà avere come approccio iniziale l'idea che la vita in una città sia il risultato di



Flusso di energia nel Distretto che contiene il PEB, che può alimentare l'area circostante con il suo surplus di energia termica ed elettrica

Fig. 3

Fonte: Horizon 2020 Call: H2020-LC-SC3-2018-2019-2020 (BUILDING A LOW-CARBON, CLIMATE RESILIENT FUTURE: SECURE, CLEAN AND EFFICIENT ENERGY)

complesse interazioni tra molti fattori. Con la disponibilità di nuovi sensori e di ulteriori strumenti digitali per la produzione di dati di elevata qualità, la ricerca potrà essere integrata con aggiuntivi grandi set di dati che rappresentano ad esempio il flusso del traffico stradale e pedonale, la qualità dell'aria ed ulteriori aspetti di interesse specifico quali quelli sociali e di carattere sanitario. Unendo questi grandi set di dati, utilizzando computer con elevatissime prestazioni e visualizzandoli nella realtà virtuale, diventerà inoltre più facile comprendere le loro complesse interazioni, ad esempio vedere come un nuovo edificio o un cambiamento nei modelli di traffico potrebbe influenzare la qualità ambientale del contesto urbano, per offrire un prezioso supporto al processo decisionale della governance.

Il contributo offerto dalla tecnologia per la pianificazione urbana e territoriale, che solo apparentemente può sembrare disumanizzante, è frutto di quel talento tutto umano che si rispecchia negli algoritmi che supporteranno le scelte che, per quanto automatizzate, saranno sempre frutto dell'intelligenza dell'uomo.

Gli strumenti digitali per la gestione dei sistemi complessi sono a qualunque livello una grande conquista del XXI secolo, fermo restando che lo strumento più efficace dietro ad ogni *Digital device*, deve rimanere la capacità umana di discernimento.

Note

[1] E' il cosiddetto paradigma "open", definito dall'informatico Jim Gray durante un discorso tenuto al Computer Science and Telecommunications Board, l'11 gennaio 2007. Si tratta della possibilità di navigare tra una mole di dati raccolti, di usarli e di condividerli per generare nuova conoscenza. Questa rivoluzione nel modo di fare scienza è definita "quarto paradigma" perché succede al primo paradigma riguardante la descrizione dei fenomeni naturali, al secondo relativo alla scoperta delle "leggi della natura" e al terzo basato sulla simulazione (Aliprandi, 2017)

[2] Il concetto di 'twin' è stato infatti elaborato dal programma Apollo della NASA per il quale sono stati costruiti due veicoli spaziali identici; uno è stato inviato nello spazio e l'altro è stato lasciato sulla terra per permettere agli ingegneri di replicare le condizioni di quello lanciato (Tao et al., 2019, p. 5).

[3] Micheal Grieves autore di Product Lifecycle Management: Driving the Next Generation of Lean Thinking e di Virtually Perfect: Driving Innovation and Lean through Product Lifecycle Management (in pubblicazione). È un esperto riconosciuto a livello mondiale nel PLM e nei sistemi informativi e tiene conferenze in tutto il mondo sul Product Lifecycle Management.

[4] HVAC è l'acronimo inglese che sta per Heating (riscaldamento), Ventilation (Ventilazione) e Air Conditioning (Aria condizionata).

[5] Copernicus è il programma di osservazione della Terra dell'Unione Europea coordinato dalla Commissione Europea e dagli Stati Membri in collaborazione con diverse agenzie multilaterali, comunitarie e nazionali. Fornisce informazioni accurate, tempestive e facilmente accessibili per migliorare, tra l'altro, la gestione dell'ambiente, comprendere e mitigare gli effetti del cambiamento climatico e garantire la sicurezza civile.

[6] È una tecnologia wireless sviluppata per realizzare una rete a bassa potenza e ampio raggio fondamentale per l'applicazione dell'Internet of Things (IoT).

Bibliografia

Aliprandi S. (a cura di) 2017, *Fare open access. La libera diffusione del sapere scientifico nell'era digitale*, Ledizioni, Milano.

Bibri S. E. 2019, *Big Data Science and Analytics for Smart Sustainable Urbanism. Unprecedented Paradigmatic Shifts and Practical Advancements*, Springer, Cham, Svizzera.

Botta A., de Donato W., Persico V., Pescapé A. 2016, *Integration of cloud computing and internet of things: a survey*, «Future Generation Computer Systems», vol. 56, pp. 684-700.

Cinquepalmi F. 2019, *La città fragile: dalla smart alla (r)evolving city*, Didapress Università degli studi di Firenze, Firenze.

Dembski F., Wössner U., Letzgus M., Ruddat M., Yamu C. 2020, *Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany*, «Sustainability», vol. 12, pp 1-17.

Diebold F. X. 2012, *On the Origin(s) and Development of the Term "Big Data"*, Pier working paper archive 12-037, Penn Institute for Economic Research Department of Economics University of Pennsylvania.

Farsi M., Daneshkhah A., Hosseinian – Far A., Jahankhani H. 2020, *Digital Twins Technologies and Smart Cities*, Springer, Cham, Svizzera.

Kamrowska-Zaluska D., Obracht-Prondzynska H. 2018, *The Use of Big Data in Regenerative Planning*, «Sustainability», vol. 10, pp. 1-20.

Klimczuk A. 2017, *Economic Foundations for Creative Ageing Policy. Putting Theory into Practice*, Palgrave Macmillan, New York, USA.

Pagani R., Chiesa G. (a cura di) 2016, *Urban Data. Tecnologie e metodi per la città algoritmica*, FrancoAngeli, Milano.

Sriramoju S. B. 2017, *Introduction of Big Data: infrastructure and networking consideration*, Horizon Book, Gurugram, Haryana, India.

Tao F., Zhang M., Nee A.Y.C. 2019, *Digital Twin Driven Smart Manufacturing*, Academic Press of Elsevier, London.

United Nations 2019, Department of Economic and Social Affairs, Population Division, *World Population Prospects* 2019, Highlights, <https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf> (7/20)

Smart and resilient cities
How can big data inform spatial design and planning for urban resilience?

Claudiu Forgaci
Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands
AMS Amsterdam Institute for Advanced Metropolitan Solutions, The Netherlands
4TU Centre for Resilience Engineering, The Netherlands
c.forgaci@tudelft.nl

DOI: 10.13128/contest-12035
www.fupress.net/index.php/contesti/

keywords
urban design
urban planning
urban resilience
big data

Amongst other vectors of change, the development of cities as complex socio-technical-environmental systems is influenced by two notable driving forces: the accelerated development of smart city technologies enabled by the abundance pervasiveness of Big Data and the challenge of resilience to an increasing number of shocks and stresses driven by climate change and urbanisation. Recognising potential synergies between these two driving forces, the paper explores ways in which Big Data can contribute to the translation of systemic resilience targets into principles of spatial transformation and, as a result, to a better evidence base for urban design and planning decisions. To that end, the paper discusses the relationship between Big Data, urban design and planning, and urban resilience. It highlights how resilience-building is enabled by Big Data through, such as evidence-based design, spatial data visualisation and cross-scalar design. The paper also provides an overview of challenges that are brough by Big Data to urban resilience-building. The discussion ends with how Big-Data-driven urban resilience can be an approach that is complementary rather than alternative to traditional practices.

Introduction

Cities as complex socio-technical-environmental systems are facing two interlinked challenges: the abundance and pervasiveness of data enabled by accelerated developments in ICT and the challenge of resilience to an increasing number of shocks and stresses driven by climate change and urbanisation. These two challenges are at the core of two urban development paradigms that guide urban design and planning in relative disconnection: The Smart City and The Resilient City. It can be argued that these two paradigms overlap and that Big Data, mostly fuelling the Smart City movement, can provide valuable input for building urban resilience. With a shared systemic understanding of the urban environment, both signal the need for evidence-based design and planning.

This paper explores «How can Big Data help urban designers and planners improve urban resilience?» and discusses «To what extent is Big Data necessary in addition to current knowledge and practices of resilience building?» The following sections provide an overview of the impact of Big Data on urban transformations in the context of urban design and planning in general and on resilience-building in particular. The paper concludes with a discussion of challenges and opportunities of employing Big Data in resilience-building through urban design and planning.

Big Data and urban transformations

In a rapidly developing data economy, Big Data play an increasingly prominent role. Large volumes of data, mostly streamed through routine sensing and crowdsourcing, capture spatial and temporal information in great detail, progressively complementing traditional datasets underlying decisions on urban transformations (Sim & Miller, 2019). This kind of granular data is unprecedented, as it can provide insights into short-term dynamics in cities (e.g. hourly pedestrian movement patterns in cities based on GPS and real-time transportation data), potentially helping urban designers and planners transform the urban environment in a less ad-hoc and more evidence-based manner (Batty, 2013).

Combined with advances in the Internet of Things (IoT), machine learning (ML) and artificial intelligence (AI) (Allam & Dhunny, 2019), as well as developments in computational toolkits (Boeing, 2019), Big Data enable a multitude of Smart City applications with a wide range of descriptive, predictive, prescriptive and discursive functions (Data-Pop Alliance, 2015). Applications include using Big Data for designing an planning urban environments

that encourage human health and well-being (Miller & Tolle, 2016), evaluating human perception of—and emotional responses to—urban spaces through Flickr or Twitter data (Li et al., 2016), understanding and enabling self-organised urban transformations (Xu, Yan, & Huang, 2017), measuring street walkability using ML algorithms and Google Street View imagery (Yin & Wang, 2016), and agent-based simulations improved by Big Data (Scheutz & Mayer, 2016), to name a few.
In this context, the fields of urban design and planning are getting acquainted with the opportunities and challenges that Big Data bring for urban transformations. Urban design is commonly defined as "the process of making better places for people than would otherwise be produced [and it is] primarily concerned with shaping urban space as a means to make, or re-make, the 'public' places that people can use and enjoy" (Carmona, Heath, Oc, & Steve, 2010, p. 3). This definition extends beyond the scope of the Smart City paradigm, primarily concerned with optimising the performance of urban systems, to emphasise the qualitative and experiential dimension of the urban environment. With an understanding of the urban environment as shaped by the combined flows of people, goods and information (Batty, 2013), urban designers and planners are concerned with guiding the morphological (i.e. spatial, physical) transformation of cities at multiple spatial and temporal scales. Moreover, recognising that the quality of urban spaces influences and is influenced by the use of technology (Carmona, 2018), urban design and planning needs to consider sensing and perception of a place, in addition to improvements in the quantitative understanding of urban systems provided by Big Data.
Although there is interest in "data-driven urban morphology" (Erin et al., 2017 and Moudon, 1997 in Boeing, 2019, p. 2), Big Data remains an underexploited topic in urban morphology (Crooks et al., 2016). Most developments have been made in the interdisciplinary field of urban data science, aiming to translate big data into information that is relevant and understandable within and across disciplines. As a specific geographic application of Big Data, urban data science focuses on spatially explicit problems (Singleton & Arribas-Bel, 2019), such as the relationship between land rent and urban form (Wu, Wang, Zhang, Zhang, & Xia, 2019) and makes extensive use of spatial (or geographic) Big Data.
Spatial Big Data has emerged mostly through user-contributed open geographic data such as OpenStreetMap (OSM), for which gradually temporal information becomes available as well (Boeing, 2019). Combined with tools to query, analyse and interpret, spatial-temporal Big Data provides a promising base for applications in the field of urban morphology. It enables applications that were not possible with traditional geographic data, such as cross-city and cross-time comparisons for a better understanding of similarities and variations between urban environments on a global scale. Fuelled by spatial-temporal Big Data, the tools for urban analytics that can provide a networked and multi-scalar understanding of urban systems behaviour are quickly evolving to support decisions for short-term crises (Batty, 2013; Boeing, 2019). Examples include applications of the Space Syntax theory (Aschwanden, 2016) and tools such as OSMnx (Boeing, 2017), meant to help designers and planners to understand complex urban spatial configurations by means of street network analyses and data visualisation.
Data visualisation is an important component of data science, as it is essential for making sense of Big Data. As such, data science is well-aligned with the visual culture of urban design and planning (Boeing, 2019) tasked with the need to represent complex urban dynamics in a visually understandable way. Augmenting reality with multi-source, multi-scale, multi-time data combined in visual dashboards enabled by Big Data, facilitates better urban design and planning (Tunçer, 2020). Such visualizations can aid the integration of qualitative (interpretive or narrative) and quantitative (data-driven) perspectives in urban morphology (Boeing, 2019) necessary for urban design and planning decisions.

Big Data and urban resilience

Cities as complex socio-technical-environmental systems are subject to high levels of uncertainty and, in the face of climate change and accelerated worldwide urbanisation, they are exposed to disturbances of increasing frequency, magnitude, and variation. Hence, urban sustainability depends on urban resilience, that is, on the ability of cities to sustain, improve and innovate their key functions – through absorbing, reacting to, recovering from, adapting to or reorganizing – in response to chronic stresses, abrupt shocks, and disruptions (4TU Centre for Resilience Engineering, 2020). In this context, urban design and planning needs to consider resilience-building in shaping future urban environments. Yet systemic resilience targets are difficult to be translated into spatial transformations. Advanced urban analytics powered by urban Big Data can play an important role in attaining that translation by revealing patterns in urban systems dynamics that are difficult to understand otherwise (e.g. the level and location of extra spatial capacity needed in a network of public spaces to respond to certain disruptions, such as floods, pandemics or failures in transport services). It can reveal patterns in large volumes of data, and it can increase prediction capacity, thus providing better evidence on both acute shocks and chronic stresses occurring in urban environments.

With its 'three Vs'[1] – volume, variety and velocity – (Batty, 2016), Big Data can have a significant contribution to building urban resilience. First, with a better understanding of complex urban dynamics and evidence brought by Big Data, urban design and planning can be more integrated and strategic. By integrating a large volume of data, planners can gain a holistic understanding of the urban environment in question and can identify key spaces of vulnerability that require strategic intervention. For instance, a good understanding of emergent, self-organizing processes in relation to resilience in cities can inform decisions on where, whether, how much and what kind of actions should be taken, including inaction, subtraction or actions that have a desirable effect if a small amount of stress is induced (see the spatial-hormetic approach by Forgaci & van Timmeren, 2014a). Second, besides its quantitative advantages, Big Data can provide designers and planners with a wide range of qualitative insights about citizen behavior, needs and desires regarding urban space. Citizen science applications can improve response to disasters and increase the level of citizen participation in the resilience-building process. And third, if resilience-building has occurred on a long term in a trial-and-error fashion, advances in Big Data and urban analytics provide insight on resilience in much shorter time spans (e.g. mobile applications in which citizens and authorities interact to mitigate the impact of a disturbance, shock or chronic stress). The quality of spatial models, simulations and scenarios can be improved through Big Data, increasing preparedness to acute shocks and improving capacity to detect chronic stresses.

Resilience needs to be understood in a both place-based and generic manner. On the one hand, resilience needs a good understanding of location-specific threats to resilience and of the vulnerabilities of the local population and system. On the other hand, certain properties can be considered generic and transferrable from one urban environment to another. Redundancy, diversity, modularity and self-organization are examples of properties which are widely considered general properties of resilient systems (Carpenter et al., 2012) and which can be translated into properties of general urban resilience relevant for urban design and planning (Forgaci & van Timmeren, 2014b). Operationalizing general urban resilience properties requires data and tools of sufficient quality and size. Cross-city comparisons of spatial configurations enabled by spatial-temporal big data, such as OSM, can reveal morpho-dynamic properties that influence general urban resilience and hence can inform transferrable urban design and planning principles targeting resilience.

All in all, there is a growing need for planning support systems for urban resilience that are fully sentient (Deal, Pan, Pallathucheril, & Fulton, 2017). Established principles of resilience can be aligned with the emerging opportunities of the Digital (or Smart) City, but this will result in a wide range of challenges (Colding, Colding, & Barthel, 2020).

Challenges

The perspective presented in this paper recognizes the pervasiveness and ubiquity of Big Data and focuses on identifying opportunities at points of convergence between the Smart City and Resilient City paradigms. Moreover, it acknowledges that Big Data does present a number of technical, methodological and ethical challenges that must be considered in urban design and planning.

Technical and methodological challenges mostly revolve around integrating Big Data from multiple sources, multiple sectors, and multiple spatial-temporal scales, as well as ensuring computational capacity to store, analyse and interpret Big Data. Integrating Big Data from different sources and with traditional data sets remains a challenge. If *access* to data was a major concern until recently, that concern is gradually diminishing with the increased availability of open data (e.g. OSM data) and the greater challenge of *combining* and *making sense of* ever-growing and diversifying Big Data. Integrating different data sets can be potentially useful for understanding interdependencies between different systems (Deal et al., 2017) and responding to short-term disruptions (Batty, 2013).

Notably, integrating qualitative aspects of the urban environment and quantitative information originating from Big Data is a great challenge that require broad, inter-disciplinary approaches and advanced technical knowledge. Streamlining such approaches and making them more accessible to practitioners of urban design and planning is yet to be achieved. The visual culture of urban design and planning augmented with advanced (big) data visualisation presents a potential way forward for integrating mixed quantitative and qualitative perspectives in urban design and planning. All in all, urban design and planning require a better integration with the field of data science, either through interdisciplinary approaches involving data scientists or through increased awareness a widened skill set in the practice of urban design and planning that includes analysis of spatial Big Data. How much knowledge, skill and awareness is and could be achieved in practice requires further insight.

From a resilience perspective, a number of concerns arise in regard to the use of Big Data. As resilience is a complex systems property, how to *build* resilience remains

an open question. Many practices and technologies of resilience-building are 'low-tech', as they have developed through a slow evolution, based on knowledge gained through long-term processes of trial and error and adaptation (see, for instance, a discussion of social resilience in the case of elevated walkways during 'acqua alta' of Venice in Forgaci & van Timmeren, 2014a). Although advances of the Smart City paradigm, fuelled by Big Data, promise to provide evidence for resilience-building in considerably shorter timespans, urban designers and planner need to integrate those different modes of knowledge. The emergence of computational toolkits opens up a new era for evidence-based urban design and planning; yet synergies, complementarities and the transition between traditional and Big-Data-driven – or low- and high-tech – practices remain an important part of resilience-building. As there are several ethical concerns about the privacy implications of the transition to Big-Data-driven practices, there are considerable efforts to shift to an approach in which citizens are sensors rather than sensed (Doctorow, 2020), are well-informed, and have the freedom to decide whether and what data they share.
Another resilience-related concern is dependency on Big Data (Walloth, 2016). Even though Big Data may inform spatial decisions on resilience-building better, over-dependency on Big Data in the future might lead to reduced resilience in case of disruptions in data streams. Therefore, urban design and planning should make use of Big Data not only for systems optimisation and efficiency but to increase resilience to data stream disruptions as well. Moreover, from a social resilience perspective, continuous engagement, participation and co-creation allow for sustained citizen awareness, knowledge and innovation that do not depend on the availability of data. Augmented with Big Data, the collective knowledge and bottom-up decision power of citizens can contribute to better-informed, place-specific and timely resilience-building processes.

Conclusions

This article explored the potentials of Big Data in the field of urban design and planning in general and in its application in the domain of urban resilience in particular. As urban resilience is a *systemic* property – that of a city as a complex social-technical-environmental system – , it needs to be translated into resilience-building *spatial* transformations by urban designers and planners. The paper posits that Big Data can aid such efforts of 'translation' as it provides the possibility of an unprecedented evidence base for complex urban transformations. Potential advances in the field of urban morphology augmented by spatial-temporal Big Data as well as data visualisation were presented as potential additions to the toolset of designers and planners. Resilience-building aided by Big Data would benefit from more spatial-temporal evidence, would allow for easier modelling, prediction and testing in a short time, and could be carried out in a more strategic manner at multiple scales (e.g. by underpinning decisions on the locations of strategic interventions in a city, metropolitan area or region). The last part of the paper highlighted a number of challenges arising from the use of Big Data. Most challenges are a matter of integration: between different data sources, sectors and spatial-temporal scales, between quantitative and qualitative perspectives. Amongst the next steps in addressing those challenges, the paper recommends a better integration between urban data science and the fields of urban design and planning, seeking for complementarities, synergies, as well as a better understood transition between traditional and Big-Data-driven practices.

Acknowledgements

This work is part of the DeSIRE-programme (Designing Systems for Informed Resilience Engineering) of the 4TU Centre for Resilience Engineering (4TU.RE). DeSIRE is funded by the 4TU-programme High Tech for a Sustainable Future (HTSF). 4TU is the federation of the four technical universities in The Netherlands (Delft University of Technology, DUT; Eindhoven University of Technology, TU/e; University of Twente, UT and Wageningen University and Research, WUR).

Note

[1] A fourth 'V' of Big Data, veracity, is not included in this description of Big Data, but it is nevertheless essential for establishing the quality and authenticity of data (Ospina, 2018).

Bibliografia

4TU Centre for Resilience Engineering 2020, *DeSIRE mission statement*, Retrieved July 21, 2020, from https://www.4tu.nl/resilience/research/desire-mission-statement/

Allam Z. & Dhunny Z. A. 2019, *On big data, artificial intelligence and smart cities*, «Cities», *89*, 80–91. https://doi.org/10.1016/j.cities.2019.01.032

Aschwanden G. 2016, *Big Data for Urban Design: The impact of centrality measures on business success*, in *Proceedings of the 34th Education and Research in Computer Aided Architectural Design in Europe Conference*, *2*, 457–462.

Batty M. 2013, *Big data, smart cities and city planning*, «Dialogues in Human Geography», *3*(3), 274–279. https://doi.org/10.1177/2043820613513390

Batty M. 2016, *Big data and the city*, «Built Environment», *42*(3), 321–337. https://doi.org/10.2148/benv.42.3.321

Boeing G. 2017, *OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks*, https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Boeing G. 2019, *Spatial information and the legibility of urban form: Big data in urban morphology*, «International Journal of Information Management», https://doi.org/10.1016/j.ijinfomgt.2019.09.009

Carmona M. (2018), *Place value: place quality and its impact on health, social, economic and environmental outcomes*, «Journal of Urban Design» https://doi.org/10.1080/13574809.2018.1472523

Carmona M., Heath T., Oc T., Steve T. 2010, *Public Places Urban Spaces* (2nd editio), Architectural Press, London

Carpenter S. R., Arrow K. J., Barrett S., Biggs R., Brock W. A., Crépin A. S., de Zeeuw A. 2012, *General resilience to cope with extreme events*, «Sustainability», *4*(12), 3248–3259. https://doi.org/10.3390/su4123248

Colding J., Colding M., Barthel S. 2020, *Applying seven resilience principles on the Vision of the Digital City*, «Cities», *103*, 102761, https://doi.org/10.1016/j.cities.2020.102761

Crooks A. T., Croitoru A., Jenkins A., Mahabir R., Agouris,P., Stefanidis A. 2016, *User-generated big data and urban morphology*, «Built Environment», *42*(3), 396–414. https://doi.org/10.2148/benv.42.3.396

Data-Pop Alliance 2015, *Big Data for Resilience: Realising the Benefits for Developing Countries*, Retrieved from http://www.datapopalliance.org/projects

Deal B., Pan H., Pallathucheril V., Fulton G. 2017, *Urban Resilience and Planning Support Systems: The Need for Sentience*, «Journal of Urban Technology», *24*(1), 29–45. https://doi.org/10.1080/10630732.2017.1285018

Doctorow C. 2020, *The case for ... cities that aren't dystopian surveillance states*, Retrieved July 21, 2020, from https://www.theguardian.com/cities/2020/jan/17/the-case-for-cities-where-youre-the-sensor-not-the-thing-being-sensed

Forgaci C., van Timmeren A. 2014a, *A Spatial-Hormetic Approach to Urban Resilience*, in *Urban Regions under Change: Towards Social-Ecological Resilience*, HafenCity University, Hamburg.

Forgaci C., van Timmeren A. 2014b, *Urban Form and Fitness: A Space-Morphological Approach to General Urban Resilience*, in *Proceedings of the 20th International Sustainable Development Research Conference* (pp. 213–223), Norwegian University of Science and Technology, Trondheim.

Miller H. J., Tolle K. 2016, *Big data for healthy cities: Using location-aware technologies, open data and 3D urban models to design healthier built environments*, «Built Environment», *42*(3), 441–456. https://doi.org/10.2148/benv.42.3.441

Ospina A. V. 2018, *Big Data for Resilience Storybook: Experiences Integrating Big Data into Resilience Programming*, Winnipeg: International Institute for Sustainable Development. Retrieved from www.iisd.org

Scheutz M., Mayer T. 2016, *Combining Agent-Based Modeling with Big Data Methods to Support Architectural and Urban Design*, https://doi.org/10.1007/978-3-319-30178-5_2

Sim J., Miller P. 2019, *Understanding an urban park through big data*, «International Journal of Environmental Research and Public Health», *16*(20), 3816. https://doi.org/10.3390/ijerph16203816

Singleton A., Arribas-Bel D. 2019, *Geographic Data Science*, «Geographical Analysis», 1–15. https://doi.org/10.1111/gean.12194

Tunçer B. 2020, *Augmenting Reality: (Big-)Data-informed Urban Design and Planning*, «Architectural Design», *90*(3), 52–59. https://doi.org/10.1002/ad.2568

Walloth C. 2016, *Big Data or No Data: Supporting Urban Decision-Making with a Nested System Model*, pp. 65–78), Springer, Cham. https://doi.org/10.1007/978-3-319-30178-5_5

Wu J., Wang, S. Zhang, Y. Zhang A., Xia C. 2019, *Urban landscape as a spatial representation of land rent: A quantitative analysis*, «Computers, Environment and Urban Systems», *74*, 62–73. https://doi.org/10.1016/j.compenvurbsys.2018.12.004

Xu L., Yan Y., Huang S. (2017), *Social Restoration and Digital Relocalization - A Big Data Strategy for Spatial Self-organization*, «Landscape Architecture Frontiers», *5*(3), 60. https://doi.org/10.15302/j-laf-20170307

Yin L., Wang Z. 2016, *Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery*, «*Applied Geography*», *76*, 147–153. https://doi.org/10.1016/j.apgeog.2016.09.024

ricerche
research

A refined waste flow mapping method

Addressing the material and spatial dimensions of waste flows in the urban territory through big data: the case of the Amsterdam Metropolitan Area

Cecilia Furlan
Alexander Wandl
Bob Geldermans
Rusne Sileryte
Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands
c.furlan@tudelft.nl
a.wandl@tudelft.nl
r.j.geldermans@tudelft.nl
r.sileryte@tudelft.nl

Received: May 2020
Accepted: September 2020

Firenze University Press.
DOI: 10.13128/contest-11909
www.fupress.net/index.php/contesti/

keywords
urban metabolism
mapping
activity-based spatial material flow analysis
waste flows
circular economy

Fundamental changes in the societal use of biophysical resources are required for a sustainable transformation. Current (urban) metabolism research traces flows of energy and materials and products to capture resource use along value chains from resource extraction to production and consumption and the discharge of wastes and emissions. However, spatial relation, local carrying capacity and qualitative characteristics of the urban landscape are only featured in very few studies, even if they are becoming crucial elements towards future sustainable development. Simultaneously, spatial studies tend to neglect the dimension of processes of flows and the generated stocks that influence the construction and performance of space. Big data and GIS technologies have the potential to leverage the integration between the two fields of knowledge. Therefore, the article explores the development of an innovative method - Activity-based Spatial Material Flow Analysis -that integrates qualitative and quantitative flow specifications in material content and geographical space, starting from the analysis of waste flows relative to the Amsterdam Metropolitan Area (NL). Lastly, the article reflects on the results of the application of the AS-MFA method, namely a series of flow maps. Each flow map is a significant data-based network representation of a part of the urban metabolism within the AMA in a specific period of time.

Introduction

Since the industrial revolution, processes of urbanisation have become ever more resource-intensive (Girardet, 2017). Urban territories consume almost 75% of the world's resources and generate 50%–80% of the world's greenhouse gas emissions and half of all global waste (Lucertini and Musco, 2020). This unsustainable process is one of the significant societal and urban planning challenges. It requires a different approach to understand and interpret urban territories and their relation to resource production and consumption. The challenge mentioned above can only be addressed by significantly expanding our knowledge base, by integrating data from different sources and by renewing the current analytical and planning frameworks (Shiller et al.,2017; EEA, 2016). In this context, the project REsource Management in Peri-urban Areas: Going Beyond Urban Metabolism (REPAiR) was initiated to investigate and develop new methods aiming at the understanding of waste flows and at the quantitative reduction of the latter in specific dispersed territories. REPAiR, rooted in the European Horizon 2020 framework, applies – for the first time – a geodesign approach to reveal the space-specific flows and challenges of waste and resource management using life cycle analyses and Urban Metabolism (UM) frameworks.

UM frameworks are based on a metaphor that conceptualises cities as living organisms (Lucertini & Musco, 2020). With the aim of understanding resource processes of a hypothetical town, Wolman (1965) pioneered the UM concept. Only recently, Kennedy et al. (2007: 44) aptly broadened UM definition to "the sum total of the technical and socio-economic processes that occur in cities, resulting in growth, production of energy, and elimination of waste". Consequently, UM describes the continuous flows of resources *in* (e.g. energy, materials, water), *out* (e.g. waste, pollutants, materials) of and *within* (stocks) a given system boundary (city, territory, metropolitan area). The assessment of flows and stocks of materials within a chosen geographical boundary and temporally defined system is conventionally defined as Material Flow Analysis (MFA) (Brunner and Rechberger, 2004; Broto et al., 2012). The fundamental principle of MFA is the conservation of matter (Allesch and Brunner, 2015). Since the system has defined boundaries, the principle of mass conservation aids in the accounting exercise that follows: inflows equals to the outflows plus changes in stocks and depletion (Allesch and Brunner, 2015). Every system, as each process within the system, has to be balanced according to the mass balance principle (Brunner and Rechberger, 2004). To the processes within the system, usually, a black-box approach is applied, studying and balancing out only the inputs and outputs of the main processes. Until today, MFA is one of the most common methods in UM framework. The advantages of the current UM framework can be summarised as follows (Pincetl and Bunje,2009; Kennedy et al., 2014):

- Identification of system boundaries,
- Identification and classification of flows,
- Calculation of system inputs and outputs,
- Study of specific urban sectors concerning sustainability goals,
- Identification of adaptive approaches to solutions and their consequences
- integration of social dynamics with biophysical sciences/technology.

However, the current UM method includes some limitations (Shahrokni, 2015):

- Extensive data collection and resource requirements
- Lack of data for specific territories
- Lack of spatialisation of data
- The necessity to process a considerable amount of data
- Difficulties in identifying cause-and-effect relationships of the metabolic flows

This article addresses the first three limitations primarily. If we consider a system equal to a territory, with features that are continually interacting emerging and adapting to political, economic and natural environments, we should avoid doing an extensive large data collection and accountability exercise (Kennedy, 2012). Instead, we should aim to provide links with specific activities, actors and locations. Furthermore, urban system components (such as society, economy, built environment, politics) can influence material flow dynamics. Lastly, cities are reliant on a more extensive hinterland to meet their resource use and waste disposal, that can easily get disregarded in case system's boundaries are drawn to include the dense urban areas only. The recognition of inconsistencies in the current UM approach underscores the necessity to develop an alternative, interdisciplinary method to understand resource flows in urban territories (Williams, 2019; Zhang, 2013). Therefore the central question behind this article is how to add a more precise spatial dimension to material flows.

In the following sections, the article explains a refined UM approach to map material flows in which we address more explicitly environmental and spatial impacts (Minx et al., 2011; Schremmer et al., 2011; Pincetl et al., 2012). Firstly, the article explains the role and the use of big data in the elaboration of a refined flow mapping. Secondly, it describes the innovative methodology called Activity-based Spatial Material Flow Analysis (AS-MFA) and the test of the latter on the Amsterdam Metropolitan Area (AMA). Lastly, the paper discusses the limitations and benefits of the AS-MFA, starting from the results of the case study.

Role of big data in UM

At the heart of UM and MFA studies lies data, in particular highly multi-layered sets of information (such as amounts, geographical coordinates, material composition of products), not easy to understand and generally defined: big data (Ward and Barker, 2013). According to Xu and colleagues (2015), big data needs to be 'big' in complexity, not necessarily in size. The degree of complexity is constituted by the different types, lack of structure and mixed semantics that a database contains. Data may come from multiple sources in several non-interoperable formats and need to be processed and analysed. Another type of significant data complexity comes from ill-defined societal behaviour that the data describes.

Within the environmental and geographical domains, the most direct application of big data is to improve the development of more realistic complex systems models to capture better essential features of environmental, geographical and behavioural dynamics (Axtell et al., 2001). The use of big data in UM studies is not a new practice. UM studies have long relied on the sectoral, industrial or public aggregated data to measure resource and energy supply and demand in urban areas. With big data, UM studies can amongst others develop more accurate models towards effective policy interventions; achieve a more precise definition of spatial-temporal resolutions in transportation-related areas through GPS travel trajectory and geotagged social media data; calculate the environmental footprint of human activities.

Opening up the black box: AS-MFA Methodology

A more accurate understanding of the urban system under scrutiny is crucial to pinpoint interlinkages between drivers, pressures and impacts, as well as response measures (Minx et al., 2011). The 'black box' approach of conventional MFA has clear limitations in that respect, as mentioned in the introduction. For that reason, a new approach has been initiated, aimed at linked actors and activities within - and extending beyond - that black box. Through this method, coined the Activity-based Spatial Material Flow Analysis (AS-MFA), specific activities relating to material flows and stocks in specific areas, the involved actors and their interrelations can be identified (Geldermans et al., 2017).

By means of the AS-MFA method, the region under scrutiny is investigated through a combination of grey-box and network approaches. See Zhang (2013) for a detailed difference between black box, grey-box and network process in urban metabolism. In a grey-box approach, the components of the system and their material inputs and outputs are accounted for. A network system identifies the links between the components and thereby highlights key players, processes, and locations.

Scope and approach of the AS-MFA

Within the framework of the *Resource management in Peri-urban Areas* project (REPAiR), where the method originates, the AS-MFA focused primarily on the output side, specifically regarding waste flows, while anticipating the future integration of inputs and throughputs into the model as well. Contrary to common MFA studies, mass balancing is not a priority within the AS-MFA approach. Instead, it builds on available data-points that are later aggregated in a multitude of different ways to facilitate data exploration and understanding of its complexity. The AS-MFA relies on the processing of big *bottom-up* data instead of disaggregating already aggregated *top-down* data sources. Another difference between the AS-MFA and a traditional MFA is that the latter requires strict delineation of system boundaries and choice of nodes, also in terms of geography. This way, everything that happens outside the chosen system boundary is simply disregarded, and what happens within the chosen node, gets aggregated. In the case of AS-MFA, the chosen geographical boundary only serves as a filter for the data points that do not fall within it. As long as at least one side of the flow falls within the boundary, the flow is represented on a map with the same level of detail both outside as inside the boundary.

The AS-MFA can be applied on any dataset that has the following critical components:

- a list of actors that have material output or material input;
- links between those actors, meaning that it is known whose output becomes which actor's input;
- content and quantity of the material input/output.

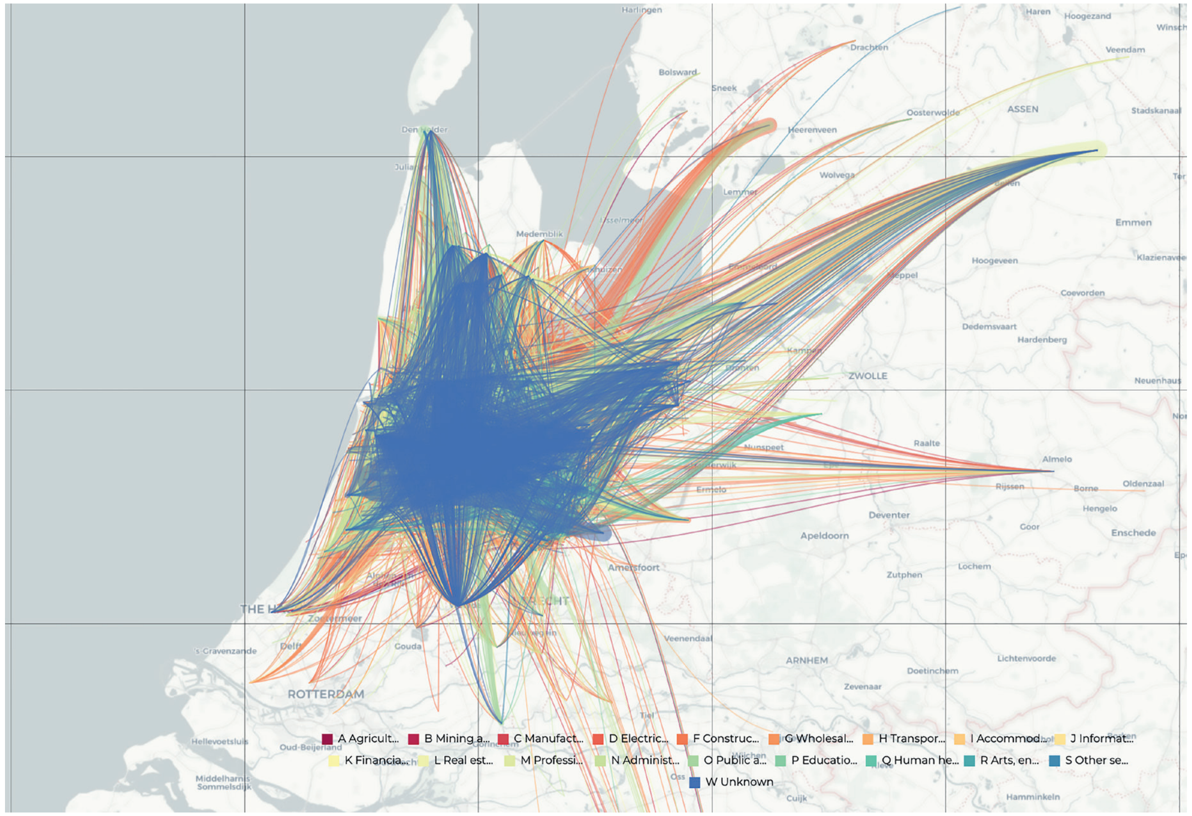
If the three components above are known, data can be enriched using the AS-MFA approach to create detailed maps and Sankey diagrams for any chosen geographical region, material or economic activity scope.

The AS-MFA consists of the following steps:

1. Data cleaning;
2. Geolocation of data points;
3. Matching actors with their economic activities using geospatial similarity check;
4. Semantic classification of material content;
5. Interactive visualisation.

NACE: section level	Description	CDW (tonnes)	Share of total (%)
A	Agriculture, forestry and fishing	23,728	0.06
B	Mining and quarrying	5,109	0.01
C	Manufacturing	2,039,225	5.31
D	Electricity, gas, steam and air conditioning supply	53,902	0.14
E	Water supply; sewerage; waste management and remediation activities	3,687,576	9.61
F	Construction	3,803,254	9.91
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	353,221	0.92
H	Transporting and storage	332,479	0.87
I	Accommodation and food service activities	5,957	0.02
J	Information and communication	10,083	0.03
K	Financial and insurance activities	214,126	0.56
L	Real estate activities	1,327,325	3.46
M	Professional, scientific and technical activities	1,402,586	3.65
N	Administrative and support service activities	63,565	0.17
O	Public administration and defence; compulsory social security	14,078,384	36.68
P	Education	35,407	0.09
Q	Human health and social work activities	124,217	0.32
R	Arts, entertainment and recreation	13,164	0.03
S	Other services activities	9,668	0.03
W	Unmatched	10,801,257	28.14
Grand Total	-	**38,384,233**	**100.00**

Amount of construction & demolition waste produced in AMA 2013-2018, aggregated per economic sector
Table 1



Economic activities of outwards flows from the AMA, relative to the year 2016

Fig. 1
Data and graphic Sileryte et al., 2020; (image out of scale)

Once all these steps are completed, the result of the AS-MFA is a digital model that can be kept in any relational database that supports geographic objects. The model is then used to generate a series of different visualisations to extract specific information:

- Reveal dominant material categories and flows in a defined territory, region or country;
- Quantify and compare material flows;
- Reveal economic activities and actors that are involved in the material flow system;
- Analyse magnitudes of contributions to the negative and positive impacts of the different activities and/or actors;
- Connect data of different material flows within a specific territory;
- Lay grounds for the assessment of environmental, social, spatial, and economic impacts caused by transportation and treatment of the (waste) materials;
- Reveal how territories relate with each other in terms of (waste) material exchange;
- Enable dynamic analysis that allows aggregating and disaggregating the data in multiple ways instead of generating a single analysis result.

In conclusion, the AS-MFA method enables the identification of key activities and actors, which reveals where responsibilities lie and therefore lays concrete points for policy or business (case) interventions. Mapping of the actors discloses their spatial location, thereby providing a spatial understanding of the regional actor-network, its spatial extent, clusters of specific activities and geographical scale of flows related to the different materials. Analysing these links and patterns allows designers, policymakers, investors and urban planners to seek optimum solutions (Moffatt and Kohler, 2008) and this way literally and figuratively provide more space for eco-innovation and circularity.

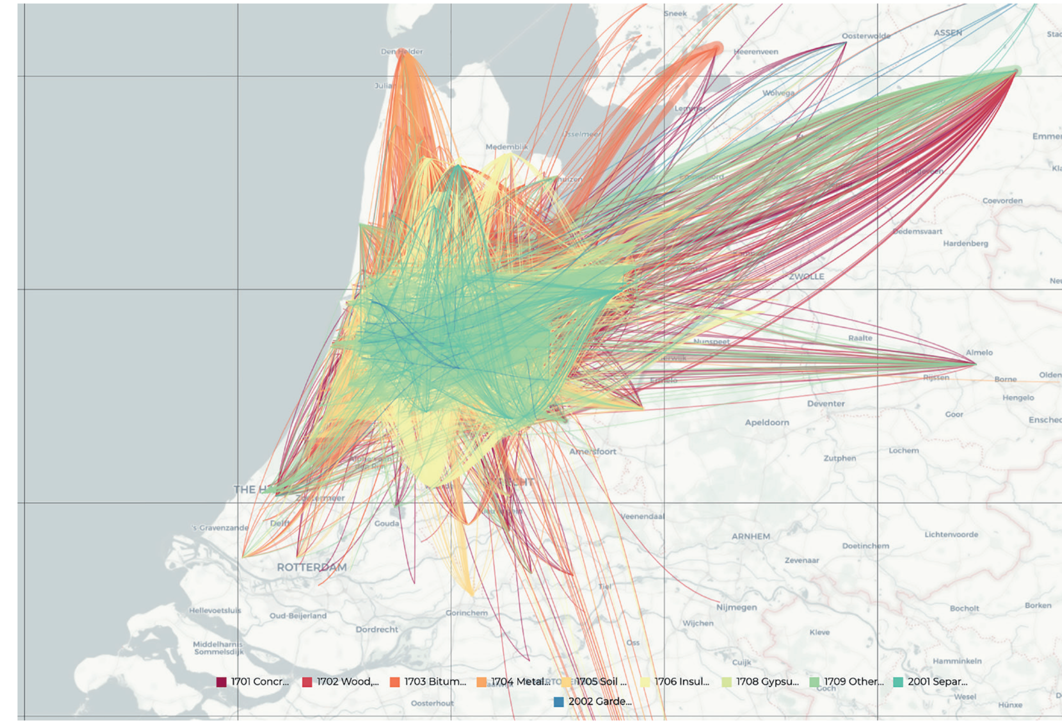
Data set and requirement

The AS-MFA provides a systematic way of analysing material flows within regions using the components: activities, materials, and actors and their interrelations. Concerning Data and databases, AS-MFA as much as possible adheres to EU-wide classification systems, ensuring the interoperability and possibility to compare different countries and regions (at least within EU). The Nomenclature des Activités économiques dans la Communauté Européenne (NACE) is used to classify the economic activities, the European Waste Catalogue (EWC) is used for the classification of waste and ORBIS[1] database due to its global coverage is used to related company names and addresses with their specific economic activities.

In order to enrich the AS-MFA model with additional information, a range of information-sources needs to be tapped concerning data on, for example, mass, energy, emissions, land use, monetary values, or social standards. The data collection specifications are then defined concerning the selected scope and key flows of materials. In general, the following data-sourcing steps should be followed, always starting from the highest level of data quality:

1. Obtain primary/company or actor specific (bottom-up) data;
2. Work with proxy data by e.g. disaggregating secondary data;
3. Use assumptions based on experts' interviews.

UM, studies on material flows typically use both top-down and bottom-up data (dis)aggregation. The level of data aggregation can significantly affect the representativeness of the mass flow simulation, and therefore, different aggregation options should be considered and, subsequently, compared (Beloin-Saint Pierre, 2017). While a model driven by data on a higher (e.g. national) aggregation level can be reliable, it loses reliability when scaled down. Contrary, a model based on bottom-up data is more reliable on a local level but will



Outwards flows from the AMA coloured according to the EWC codes, relative to the year 2016
Fig. 2
Data and graphic Sileryte et al., 2020
(image out of scale)

lose representativeness and reliability when scaled up (Roy et al., 2015). This observation suggests that if the primary goal of the study is to understand (spatial) patterns and flow specifics rather than construct the big picture, a bottom-up highly spatially disaggregated approach is more suitable than downscaling top-down/national data. The former method, however, is usually very resource-intensive as it requires extensive data collection efforts. Applying a hybrid system, using both top-down and bottom-up data, could mitigate some of the problems with the incompleteness or inaccessibility of detailed data.
The proposed AS-MFA method has been tested on the construction and demolition waste flow dataset for the Amsterdam Metropolitan Area (AMA).

Data and the Case-Study: Amsterdam Metropolitan Area
The Amsterdam Metropolitan Area (AMA) is a collaboration of 32 municipalities, two provinces (North Holland and Flevoland) and the Transport Authority Amsterdam. Around 2.4 million people live within the AMA. Moreover, Schiphol International Airport, the Port of Amsterdam, the Royal FloraHolland flower fair and the Tata Steelworks in Ijmuiden make the AMA the economic powerhouse of the Netherlands.

The waste dataset that has been used for the case study of AMA has been provided by the Division of Waste Management within the Dutch Ministry of Infrastructure and Public Works (Rijkswaterstaat). The dataset consists of reports submitted to the national online waste registration system (https://amice.lma.nl/Amice.WebAppHome/). The reports contain information about the waste pickup location, waste producer, collector, broker, processor, including their addresses, timestamp, waste quantity, collection method, EWC code and description of waste contents. These characteristics make the dataset ideal for the AS-MFA method.
For this paper, it has been filtered with the following criteria:

- The waste is either produced or treated in the AMA;
- The waste is classified according to EWC 17 as construction waste;
- The waste was reported within one calendar year.

Results
This section presents the results of the AS-MFA applied to the AMA.
The above-described selection of the national waste data set resulted in 64,905 individual data entries accounting in total for 42,108,577 metric ton of waste. See Table 1 for details about which economic sector contributes to which part of this amount.
The application of the AS-MFA resulted in flow maps, which are presented in Fig. 1 and Fig. 2. The flow map is a significant data-based network representation of a part of the urban metabolism of the region. Therefore, the map allows adding a spatial dimension to the discussion of the results. Some of these spatial questions are discussed as exemplary in the following subsection.

Discussion
This section discusses exemplary spatial questions that are relevant for policy making, working specifically in the field of Circular Economy (CE), aiming to demonstrate the added value of the AS-MFA. Furthermore, the advantage of a method that is both territorially extended but also precise in location is demonstrated.
In the Netherlands, the main actors of CE policies are municipalities and metropolitan regions administrations. Therefore, the question is, are the territories, which are the spaces of policymaking, the same or at least similar to the extent of the flow network, which are the spaces of urban metabolism and related economic activities? The flow map presented above gives a clear answer to the question. The network space is much larger than the territory of the AMA. For instance,

the waste flow network of the construction sector alone exceeds not only the AMA but also the Netherlands, expanding across the country into several European countries and even beyond. For policymaking, this can be interpreted two-fold, either policy on CE have to be integrated at territorial governance level in which waste is travelling now, or policies and spatial plans need to facilitate the definition of optimal scales and amount of land necessary to develop CE strategies. The latter leads to the spatial questions, which role in the network of urban metabolism do the economic activities in one specific sub-area play, be it neighbourhood, district or municipality? The question of localisation of the economy is of particular interest considering the redevelopment of an area, which often coincides with a significant programmatic change, in the context of the AMA mostly a shift towards residential and service-economy related land-uses (Gemeente Amsterdam, 2017). Alternatively, the (re)development of an area may take into consideration the use of nearby resources to reduce transportation costs and associated environmental impacts. In this case, two types of flows are considered:
1)one-off occurrences that convert an end-of-pipe flow into a stock, e.g. by reusing materials for construction;
2)constant occurrences that turn a constant end-of-pipe flow into a circular flow, e.g. by introducing a function which uses secondary materials for its economic activities.
We have a closer look at the development of

Comparison of waste flow movement within a portion of AMA between 2013 (top) and 2018 (bottom)
Fig. 3
Data and graphic Sileryte et al., 2020
(image out of scale)

the northern IJ-riverfront in Amsterdam to demonstrate the advantage of the AS-FMA. Circular area development is one of the fundamental principles here. For example, the neighbourhood of Buiksloterham, and most notably the 'De Ceuvel' location, has become a best practice of circular area development, widely known in the Netherlands and beyond. What if we look at this area from a network perspective?
Fig. 3 compares the situation of 2013 (top) with 2018 (bottom). Maps on the left show the relation between production on the treatment of CDW on a regional scale. The maps on the right provide insight into the production of CDW on the local scale.
Figure 3 and the data behind it demonstrate two aspects: (i) the amount of generated CDW nearly doubled, and (ii) there has been a shift in the locations of treatment. The first is an expected result in an area where in the same period, the number of dwellings doubled, while the number of companies shrank by around 25%(van Bakel, 2020). The second shows that the network space of CDW treatment became more local and followed in this sense the regional CE policy. Whether this shift leads to a higher rate of circularity, is, of course, dependent on changes in the treatment and recycling of the CDW. This question, as well as questions on how far policy changes had a direct influence on the spatial transformationsof the CDW network, goes beyond the scope of this paper.

Conclusion

Detailed data on material flows concerning urban territories is, by definition, big data - not only because of the extensive amounts of flows but also because of the complexity of the flow networks. This paper has explored Constructions & Demolition material flows, after becoming waste, by focusing on the AMA across five years. The proposed AS-MFA method allows constructing a digital model of flows on given space and juxtaposing with spatial components, such as infrastructural networks and landscape systems. Moreover, this method helps to reduce the complexity of interpreting big data, by selecting: territorial portions, temporal extent, economic activities and materials of interest. Lastly, the AS-MFA permits to explore each type of data separately from the system, and it works well with a single homogenous dataset. However, nowadays, it only includes waste materials. Although the AS-MFA method refined the current UM approach by adding a flow mapping through extended data analysis, it opens new future challenges and questions. The first one refers to the integration of a mass balance calculation in the current AS-MFA process and consequently, the necessary data required. However, the inclusion of material input and throughput pose a related challenge of data integration. For instance, if the datasets do not have the same granularity, taxonomy and semantics, then connecting input, throughput and output network nodes would require iterative mass balancing and semantic matching of material contents. This would also imply extensive and intensive manual efforts.
The second challenge relates to connecting the AS-MFA outcome with more common environmental, socio-economic indicators and assessment methods. Lastly, reflections on the AS-MFA through the lens of big data pose the question on how territories could better benefit from big data analysis and in particular, how stakeholders could integrate flow mapping in urban planning policies.

Acknowledgement

The authors would like to thank the whole REPAiR and CINDERELA Team, primarily the Research Team at Delft University of Technology, in the Netherlands, the Nulmeeting team and especially Vasileios Bouzas, Arnout Sabbe and Kozmo R. Meister for the technical development and LMA for providing the data.

The entire article, including the reviewed versions of it, has been written and approved by all of the authors. Specifically, C.F. wrote Section 1 and Section 2; B.G. and R.S. wrote Section 3.1 and Section 3.2; A.W. wrote Section 3.3; A.W and R.S. wrote section 4 and section 5; R.S commented on all the sections, all the authors contributed to Section 6.This research has been carried out within the framework of the European Horizon 2020 funded research REPAiR: REsource Management in Peri-urban AReas: Going Beyond Urban Metabolism. This project has received funding from the Europe-an Union's Horizon 2020 research and innovation programme under grant agreement no. 688920. This article reflects only the authors' view. The Commission is not responsible for any use that may be made of the information it contains.

Note

[1] https://www.bvdinfo.com/en-gb/our-products/data/international/orbis

Bibliografia

Allesch A., Brunner P. H. 2015, *Material flow analysis as a decision support tool for waste management: A literature review*, «Journal of Industrial Ecology», *19*(5), 753-764.

Axtell R. L., Andrews C. J., Small M. J. 2001, *Agent-based modeling and industrial ecology*, «Journal of Industrial Ecology», *5*(4), 10-14.

Backel van S. 2020, *Circular Area Development: Recommendations for creating a circularly functioning area*, Graduation thesis, Faculty of Architecture and the Built Environment, University of Technology, Delft

Beloin-Saint-Pierre D., Rugani B., Lasvaux S., Mailhac A., Popovici E., Sibiude G., Schiopu, N. 2017, *A review of urban metabolism studies to identify key methodological choices for future harmonization and implementation*, «Journal of Cleaner Production», *163*, S223-S240.

Broto V. C., Allen A., *Rapoport E. 2012, Interdisciplinary perspectives on urban metabolism*, «Journal of Industrial Ecology», 16(6), 851-861.

Brunner P. H., Rechberger H. 2004, *Methodology of MFA. Practical Handbook of Material Flow Analysis*, Lewis Publishers©, Boca Raton London New York Washington, DC, 34-166.

Geldermans B., Bellstedt C., Formato E., Varju V., Grunhut Z., Cerreta M., Amenta L., Inglese P., Leer J. van der, Wandl A. 2017, *Introduction to methodology for integrated spatial, material flow and social analyses, Resource Management in Peri-urban Areas (REPAiR)*, Delft University of Technology: The Netherlands, Delft.

Geldermans B., Wandl A., Steenmeijer M., Furlan C., Streefland T., Formato E., Cerreta M., Amenta L., Varju V., Inglese P., et al. 2019, *REPAiR: REsource Management in Peri-urban AReas: Going Beyond Urban Metabolism: D3.3 Process Model for the Two Pilot Cases: Amsterdam, The Netherlands & Naples, Italy*; Revised Version 30 April 2019; University of Technology, Delft

Gemeente Amsterdam,2017, *Haven-Stad: Transformatie van 12 deelgebieden,* Retrieved from https://assets.amsterdam.nl/publish/pages/841936/ontwikkelstrategie_haven-stad_vastgestelde_versie_web.pdf

Girardet H. 2017, Regenerative Cities, in S. Shmelev (Ed.), *Green Economy Reader: Lectures in Ecological Economics and Sustainability*(pp. 183–204), Springer International Publishing, Cham.

Kennedy C., Cuddihy J., Engel-Yan J. 2007, *The changing metabolism of cities*, «Journal of Industrial Ecology», 11(2), pp. 43-59.

Kennedy C., Pincetl S., Bunje P. 2011,*The Study of Urban Metabolism and Its Applications to Urban Planning and Design*, «Environmental Pollution», 159/8, 1965-1973: 1967.

Kennedy, C. (2012). Comment on article "Is there a metabolism of an urban ecosystem?" by Golubiewski. Ambio, 41(7), 765-766.

Kennedy C., Stewart I. D., Ibrahim N., Facchini A., Mele R. 2014, *Developing a multi-layered indicator set for urban metabolism studies in megacities*, «Ecological Indicators», *47*, 7-15.

Lucertini G., Musco F. 2020, *Circular Urban Metabolism Framework*, «One Earth», *2*(2), 138-142.

Minx J. C. and Baiocchi G. 2009, Time use and sustainability. *Handbook for Input–Output Economics in Industrial Ecology* ed S Springer, Suh Dordrecht.

Minx J., Creutzig F., Medinger V., Ziegler T., Owen A. and Baiocchi G. 2011, *Developing a pragmatic approach to assessing urban metabolism in Europe: A Report to the Environment Agency*, Technische Universität Berlin and Stockholm Environment Institute, Berlin.

Pincetl S., Bunje P. M. 2009, *Potential targets and benefits for sustainable communities research, development, and demonstration funded by the pier program*, http://sustainablecommunities.environment.ucla.edu/wpcontent/uploads/2012/10/BOA-99-207-P_FINAL.pdf. [Accessed 30 May 2020].

Pincetl S., Bunje P., Holmes T. 2012, *An expanded urban metabolism method: Toward a systems approach for assessing urban energy processes and causes*, «Landscape and urban planning», *107*(3), 193-202.

Remøy H., Furlan C., Wandl A., Dabrowski M., Amenta L., Arciniegas G., Muñoz Unceta P., Streefland T., Geldermans B., Heurkens E., Meister K., Craigen A., Šileryté R. 2018, *REPAiR Catalogue of solutions and strategies for Amsterdam, 2018*, Deliverable 5.2. REPAiR, Delft University of Technology.

Roy M., Curry R., Ellis G. 2015, *Spatial allocation of material flow analysis in residential developments: a case study of Kildare County, Ireland*, «Journal of Environmental Planning and Management», 58(10), 1749-1769.

Schiller G., Müller F., Ortlepp R. 2017, *Mapping the anthropogenic stock in Germany: Metabolic evidence for a circular economy*, «Resources, Conservation and Recycling», 123, 93–107.

Schneider S.H., Root T.L. 2004, *Climate change and ecology synergism of Encyclopedia of Biodiversity* 1, 709–725.

Shahrokni H., Årman L., Lazarevic D., Nilsson A., Brandt N. 2015, *Implementing smart urban metabolism in the Stockholm Royal Seaport: Smart city SRS*, «Journal of Industrial Ecology», 19(5), 917-929.

Sileryte R., Sabbe A., Meister K., Bouzas V., Wandl A. v., Timmeren A. 2020, *Nulmeting Amsterdam*, City of Amsterdam (unpublished)

Williams J. 2019, *Circular cities*, «Urban Studies», 56(13), 2746-2762.

Wolman A. 1965, *The metabolism of cities*, «Scientific American», 213, 179-190.

Xu M., Cai H., Liang S. 2015, *Big data and industrial ecology*, «Journal of Industrial Ecology», 19(2), 205-210.

Zhang Y. (2013), *Urban metabolism: A review of research methodologies*, «Environmental Pollution», 178, 463-473.

Snap4City
A big data platform for smart cities

Paolo Nesi
Michela Paolucci
DINFO, Università degli Studi di Firenze
Paola Zamperlin
Università di Pisa
paolo.nesi@unifi.it
michela.paolucci@unifi.it
paola.zamperlin@unipi.it

Received: May 2020
Accepted: September 2020

DOI: 10.13128/contest-11722
www.fupress.net/index.php/contesti/

keywords
big data
smart city
urban strategies
control room
citizen participation

Since the idea of smart city has become a practice to be implemented, urban planners and policy makers have needed tools that would allow them to exploit the information potential of the amount of data that the city itself produced in order to understand and manage its growing complexity. This article, after examining the close relationship between good governance based on a data-driven approach and the economic and social development of a city, describes the Snap4City Open Source solution, a Big Data Smart City Platform, adopted in European cities and region (Antwerp, Florence, Pisa, Santiago, France, Belgium, Finland, Tuscany, Sardinia, Croatia, Greece, etc.). Snap4City is GDPR compliant and designed to support city planners, combining strong tools for data integration, analytics, forecasting and visualization with the possibility of set Living Labs to enhance the collaboration among operators. Four scenarios have been selected and described (a. Florence Smart City Control Room; b. Traffic Flow Reconstruction and Air Quality predictions in Firenze, Pisa, Livorno; c. Air Quality in Helsinki for the citizens; d. People flows in Antwerp). The main KPIs have been briefly exposed, in order to evaluate the Snap4city impact in Smart City planning.

Introduction

The decision-making processes and operational tools which have been involved in the management of cities in recent decades have undergone a necessary redefinition which has gradually been adapted to the growing expansion of urban areas, the increase in the urban population and available technologies and data. The urban phenomenon for its quantitative dimensions, therefore, but also for the different qualitative meanings that it is assuming, obliges to push towards the adoption of innovative strategies capable of facing the expressed and potential needs of all the communities operating on a specific context, at different scales, ranging from hyperlocal, metropolitan and regional to global ones (Sassen, 1991).

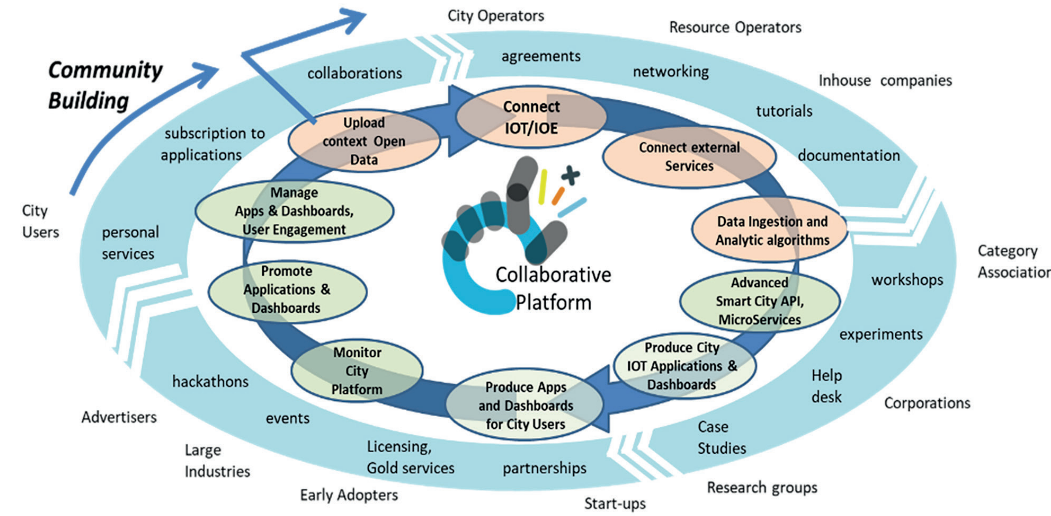
The recent and well-known estimates of the United Nations (UN, 2018) describe an urban growth that will involve 75% of the world's population by 2050.

If, as it seems, the forecast trend is verified by the facts, we must expect at least two macro-areas of related issues, which will see cities engaged in maintaining internal cohesion, efficiency in the provision of services and in the improvement of the quality of life of city users; and on the other in the need to maintain high international competitiveness, minimizing the factors of economic and social marginalization and maximizing those of attractiveness in global contexts. The state of solidity and well-being of a nation is increasingly measured through the well-being of its cities.

In Italy, as in many European countries, urban centers are delimited by a historical administrative perimeter that no longer corresponds to the area of the actual vital and productive settlement, but that conditions the adoption of adequate planning strategies. Without neglecting factors such as historical characteristics, morphogenesis and socio-economic structure, with the growth of its dimensions, an urban ecosystem also increases the degree of internal complexity and therefore requires the adoption of an integrated tool capable of understanding and managing its growing complexity (Mitolo et al., 2020).

The governance of such an urban fabric must take into account a fair distribution of land use (residential, recreational, service, productive, agricultural, etc.) and the network of connections between the different parts, so its primary objective is to make the mobility a well-organized and efficient system. Studies carried out in this regard have shown that an integrated transport system, flanked by agile and accessible pricing, for example, is a priority for territorial governance and a precondition for growth and development (OECD, 2015a).

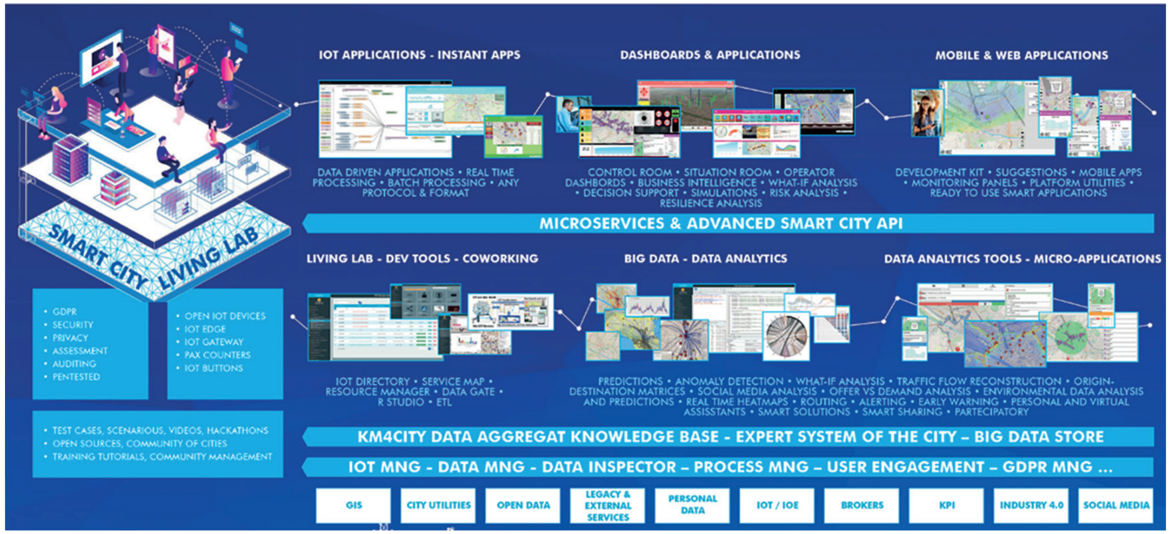
The successful characteristics of a city can be analyzed in three dimensions: population growth, economic performance and functional organization and considering the importance of good governance arrangements. "The fragmentation of a city's administration and the quality of its governance structure is directly reflected in its economic strength" (OECD, 2015a, p. 37 ff.), while, on the contrary, where the conditions of cohesion and integration of government structures are lacking, the city responds with an increase in inefficiency and consequently a decrease in wealth.
If it is true that the productivity of a city depends on a multiplicity of factors related to national economic policies, many others are due to local characteristics such as the quality of the infrastructures and the solidity of the market and local institutions. Good governance of urban agglomerations is essential for their functioning (Ahrendi et al., 2015) and it is becoming more automatic and based on the predictive power and anticipatory capacities of digital technologies and computational intelligence (Dodge, 2010). It is also appropriate to consider that good governance structures are not alone a guarantee of good policies, but it is very difficult to design and implement good policies without them. Some studies have highlighted how the presence of governance bodies can be an instrument of co-ordination within metropolitan areas, which are reflected, for example, in a better organization of land use and efficiently integrated transport system, in general a more forward-looking spatial planning (OECD, 2015b). In contexts where these assumptions are missing, on average, cities have higher levels of air pollution (i.e. as measured by the amount of particulate matters in the air), while where coordinated decision-making bodies are present the provision of the services across a metropolitan area could also improve their quality or reduce their costs, with positive spillover effects.
In *Triumph of the City*, Edward Glaeser (2011) states that "Not every city will succeed, because not every city has been adept at adapting to the age of information, in which ideas are the ultimate creator of wealth". This adaptability is based on two decisive conditions, namely the availability and access to relevant data and the time that separates the moment in which the data is produced from that one in which it's possible to analyze them and produce useful information for decision-making process.
"Innovation and technological progress are generally believed to be the main driver of long-term economic growth" (Aghion, Howitt, 2005), but technological innovation must be supported both on the strategic vision and on the possibility of triggering rapid and highly adaptable decision-making processes to sudden and unforeseen changes of state. For these reasons, human capital is crucial in fostering the capacity of a region in creating innovation and adapting that to new scenarios (Rodriguez-Pose, Crescenzi, 2008). In the planning and management processes of a complex ecosystem such as cities, where many distinct aspects coexist, and many kinds of actors interact – such as public administration, citizens, enterprises, stakeholders, research organizations, universities – therefore, the conditions to enhance innovation and technological development must be put in place as a precondition for improving the quality of life of all end-users.
In this continuous process of adaptation to information, applications that are based on multiple paradigms such as data-based processing, streams and batches, at the same time activate new work paradigms aimed at improving collaboration between all the heterogeneous actors involved (Mitolo et al., 2020). It follows from this that two are the key axes on which it is necessary to focus on in the design of a system that intends to manage urban complexity: data and people. An efficient architecture must be able to collect and analyze data for the areas identified as strategic and allow all stakeholders to be able to participate in the processes of selection, collection, integration, analysis, visualization of data.
The main areas on which the city planners are usually interested in are: Mobility, Environment, Energy, Weather, People (social and flows), Governance and Communications, Economy, Culture and Tourism, and Security. For each of these areas, data must be available that provide information on the status and established performance indicators that allow to evaluate the effectiveness of the solutions adopted or identify problems and alert thresholds. Obviously, these areas are non-independent of each other. Urban sprawl often accompanies increased costs in terms of traffic congestion and lack of adequate infrastructure that affect both the quality and quantity of urban traffic and the consequent spillover effects on pollution, quality of life and variations of the real estate market (OECD, 2015b, 29-54). Sensors for monitoring traffic and public lighting, for example, or variable direction signs used to steer traffic flows or to show parking spaces, together with an integrated public transport improve mobility within metropolitan area by lowering travel time and improving accessibility and quality of life.
Snap4City, the big data platform that we are going to describe in the next paragraphs, is currently at the last stage of development of an ecosystem made of integrated technologies that over the years has been modeling itself on the specific needs that smart cities were expressing. This architecture



Snap4City Collaborative Workflow Process

Fig.1:
Mitolo et al., 2020

brings together the results of important national and international projects, such as SELECT4Cities Pre-Commercial Procurement EU Project (https://www.select4cities.eu), REPLICATE EU project (https://replicate-project.eu), TRAFAIR CEF project (http://trafair.eu), Sii-Mobility National Smart City project, Italy, (http://www.sii-mobility.org), just to mention the main ones. It is therefore able to collect, integrate and manage data, analytics, predictions and applications related to the public transport system (operators schedule and paths, traffic flow sensors, parking status and positions, cycling paths, road graph, accidents and traffic events, etc., cfr. Bellini et al., 2018b); waste, pollution, pollination and air quality (PM2.5, PM10, CO, Benzene, NO, NO2, etc.); energy (recharging stations, consumption meters, smart light); weather (temperature, wind, humidity, rainfall, forecast and actual); social and cultural events (POI, entertainment, Twitter monitoring, sentiment analysis); triage status of major hospitals; people flows (from the Wi-Fi access points; Origin-Destination matrices). All this happens in a robust way, having Snap4City passed two vulnerability tests, efficient, capable of ingesting data from all types of different sources, interoperable with all IoT standards and respectful of data protection regulations, being fully compliant to the GDPR (CDPR)[1]. Planning tools need performance indicators that summarize the complexity of reality in constant updating with the changes that occur, creating a cognitive framework on which the decision-making process is based. Quantity, quality, availability and updating of data, especially for historic urban areas, are essential conditions for the development of effective policies (Cundari, 2015). The heterogeneity of data is parallel to the variety of actors (decision- makers, planners, stakeholders, end users) who, starting from these, intervene directly or indirectly on urban policies.
Therefore, the activation of collaboration environments is necessary, which are not mere virtual rooms, but real Living Labs where to define strategies and implement specific services in response to specific needs (Cosgrave et al., 2013).
The Snap4City architecture provides customizable tools and solutions to create a large range of smart city applications, created thanks to the collaboration of different kind of organizations (Universities, Public Administrations, Enterprises, Start-ups) and city users (operators, inhouse companies, tech providers, corporations, research groups, citizens, commuters, etc.) as part of a large and heterogenous community (see Figure 1).
The question of how technology and innovation influence social and economic development has always been debated among economists (Fagerberg et al., 2010). The ability to innovate and assimilate innovation are considered by many authors as key factors in the growth of an area and "it has now become widely accepted that innovation is a territorially-embedded process and cannot be fully understood independently of the social and institutional conditions of every space" (Rodrigues-Pose, Crescenzi, 2008). In light of this, Snap4City can be considered a space where geographically relevant information, innovative technologies and users with different profiles are made to coexist and operate in order to allow social innovation processes.
Big Data, due to their growing importance in terms of volume, velocity and variety, have determined a gradual shift towards a data-driven socio-economic model, where "data are a core asset for creating significant competitive advantages and for driving innovation, sustainable growth and development" (Ubaldi, 2013). It is fair to say that data and data analytics play a decisive role for the formation of knowledge-based capital, precondition for economic growth and sustainable development of society (OECD, 2013).
One of the expected impacts in the use of big data in urban contexts lies in the increase in resilience capabilities, given that the degree of effective response to both daily and catastrophic scenarios can measure whether a city is well-functioning. Identification of possible risks and adoption of prompt measures implies information sharing between different levels of government, public administration and final users of both the public and private actors.
There are also not secondary aspects, but to which we will dedicate only a brief mention, which concern the creation of a framework of awareness and trust – "subtle, yet important issue for the functioning of cities" (OECD, 2015a) – that generally triggers virtuous mechanisms of governance and public participation.



Snap4City Architecture: IoT management, data ingestion and processing, data analytics, Mobile and Wep Apps,
Fig. 2
Mitolo et al., 2020

A model like the one proposed here, which provides for data generation, data collection, aggregation and processing, data analytics and visualisation, and which is the expression of planned governance, has as its first effect that of improving government accountability, transparency, responsiveness and then of promoting citizens self-empowerment, social participation and engagement. Promoting innovation is therefore expected to increase efficiency and effectiveness in public services and generate value for the economy in general (Ubaldi, 2013).

Snap4City Solution

Snap4City is a Big Data Platform applied in the context of many European Smart Cities, providing innovative approaches flanked by consolidated technical solutions. Snap4City supports decision makers in their urban organization processes to involve citizens themselves, triggering a participatory process of evolution and innovation of their city. In this paradigm it is easy to recognize the combination of two complementary concepts of translation (how cities are translated into code), and transduction (how code reshapes city life) suggested by Robert Kitchin (2011). The solution proposed has been improved thanks to the experience had in the different geo-political contexts in which is applied: Italian (Firenze, Pisa, Livorno, Prato, Cagliari, Modena, Lonato del Garda, etc.) and European Cities (Antwerp, Helsinki, Santiago De Compostela, Valencia, Dubrovnik, etc.) and their surrounding geographical area (such as in Italy the region of Tuscany, Sardinia and Lombardia but also Belgium and Finland), (Azzari et al., 2018). The strength of Snap4City lies precisely in supporting the city's decision-makers in all the phases that involve the city's evolution processes, proposing alternative modes of communication to the classic ones: it promotes coworking activities and provides Living Lab services, which allow an immediate communication between the groups of people involved and an exchange of ideas supported by analysis and advanced visualization of concrete data in real time. Snap4City, as visible in Fig.2 in which its architecture is represented, manages a noteworthy amount of heterogeneous data in terms of protocols and formats, open and private data, static/periodical and real-time data, GIS and Maps, data stream, data driven, Industry 4.0., data coming from IoT Devices and networks, IoT Edges, gateways, IoT Brokers or via different protocols and formats, that are uniformly managed on IoT Directory via Knowledge Base for discovery. Another modality of data ingestion is the web scraping process that may be used for grabbing data from public web pages according to the exposed terms of use. Moreover, additional information comes directly from the final users (e.g., citizens, tourists, shopkeepers, students, etc.) via Mobile Apps and Social Media (Bellini et al. 2018b; Badii et al., 2017). For this reason Living Lab activities on the platform are fundamental and at the basis of a work process promoted and actuated that: i) starts from the identification of the final objectives that the smart city decisors want to achieve; ii) identifies the types of significant data and relevant stakeholders; iii) involves the different types of people providing effective technological tools for everyone; iv) realizing the final ad-hoc solution based on the needs expressed by the decision makers of the city and which can be a combination of various tools for visualization and interaction. For example, through the realization of dashboards or Mobile/Web Apps able to highlight the results obtained from the data analysis carried out, which translates in terms of visualization and interaction with: historical data, real time and future forecasts, connection with groups of decision makers, etc. (Nesi et al., 2018). Once the data have been identified, they are uploaded into the platform with the more effective ingestion tool among those available in the platform (e.g., IoT App via Node-RED for sensors and for real time data, ETL (Extract Transform Load) processes for Static or periodical Data, web scraping tools for web pages, scripts for complex data, etc.) and stored basing on their frequency of update. Moreover the data are semantically processed so that the information is enriched and connected each other, this aggregation process is managed in automatic or semi-automatic way by the platform, depending on the data type and on the methodology adopted to ingest the data and in compliance with the Km4City Smart City multi-ontology (https://www.km4city.org). All the data are stored in the Snap4City Knowledge Base on graphDB and in a Big Data store for data shadow based on Elastic Search and/or HBase noSQL database.

Smart City Control Room in the City of Florence: connections and What-IF analysis
Fig. 3

The ingestion and storage phase are at the basis of data Analytics processes and Data Transformation algorithms capable to analyze all the data present in the Snap4City KB, including personal data produced by users and respecting their privacy.
The platform is GDPR compliant and perfectly capable to manage the different level of privacy connected to each dataset, user, user role, group of users (called organizations). This aspect is fundamental for decision-makers who have to understand what kind of services provide to their citizens, what kind data can be viewed and by which category of people, what license has to be connected to each dataset.
Data Analytics and data transformations processes and tools can be created and shared in a reliable and scalable cloud infrastructure to provide advanced smart services, making prediction, signalling early warning, detecting anomalies, creating analysis, heatmaps, supporting decision makers with insights and reports, etc. Analytics and data processing exploit the Snap4City Smart City API, can be developed in ETL, R, Python, Java, IoT App Node-RED, etc., and can be executed on demand/on events, periodically and in real-time. Resulting data are also saved and indexed into the platform. Dashboards for city decision makers and mobile Apps for citizens, are the most suitable visualization tools both as decision support systems with high levels of interaction and to have a real time view of what is happening in the city.
The next section describes fours scenarios dealing with different issues and adopted in different European contexts. These scenarios highlight how the Snap4City platform has contributed in a relevant and innovative way in speeding up the urban development processes in different contexts.

Scenarios

The Big Data Snap4City Solution manages, elaborates data and provides services in different cities and geographical areas and it is applied in the contexts of Mobility, Environment, Energy, Weather, People (social and flows), Governance and Communications, Economy, Culture and Tourism. In this section, four scenarios are described, with the purpose to demonstrate the potentialities of the Solution and the advantages obtained. They are different in terms of City's needs and geographical areas, type of data managed, final users.

Scenario1: Florence Smart City Control Room

In the city of Florence, Snap4City is at the basis of the Smart City Control Room (SCCR), in the contest of the Replicate EU project (https://replicate-project.eu). The city's decision-makers and in particular the mayor and his staff, need to monitor the general state of the city in order to verify the usual daily course of the city's activities, to be able to intervene quickly in the event of catastrophic events and to make concrete short- and long-term forecasts. The Florence SCCR shows in Real Time data, statistics and graphics, provides services for drill down on data, shows real time and provisional maps, allow interactions, etc. The main thematic area covered are: Mobility, Environment, Energy, Weather, Social Media, People Flows, Governmental and Communications, Tourism and Culture. The Control room is organized with a starting view to which all the other dashboards are connected, as can be viewed in Fig. 3. In this way, the decision makers can navigate among the different dashboards basing on the contingent needs, can explore data and receive predictions or suggestions in real time. They can be put in contact with a restricted group of responsibles to analyze data, thus triggering shared decision-making processes. The Smart City Control Room in Florence is furthermore the result of the Snap4City's Living Lab support: some of the Dashboards and the available KPI, have been directly realized/estimated by the decision makers of the City of Florence. Users who do not have a strictly technical profile, have the possibility to work in full autonomy or in semi-guided mode. Having at disposal virtual spaces on which develop processes to ingest and manage data (IoT Applications) or services to share/download/view tutorials and simplified routines/IoT Applications/programs/etc., in addition to the possibility to make comparison with the work done in other Smart Cities, in which Snap4City is applied.
Some interactive tools, such as the What-If analysis, visible in Fig. 3 (on the right), play a fundamental role to answer to questions such as: what happens if I close a road or change it in terms of traffic and reorganize public transport? Which parking spaces are expected to be occupied in the next few days, also according to the cultural events that will take place? Looking at the SCCR, the city decisors can analyzes the situation in Florence knowing and visualizing: structure of the city, position of the services, paths of the public transportation, real time value of traffic (Bellini et al., 2018a), environmental variables, air quality (PM10, PM2.5, CO, Benzene, NO, NO2, O3, etc.), people flows, weather conditions and forecasts (temperature, humidity, wind, etc.), cycling paths, hospitals first aid, etc. And thus, defining restrictions to be applied, areas to be closed, the new paths for public and private vehicles and pedestrians, etc.

Traffic flow reconstruction, Air Quality predictions and historical data on NOx in the City of Pisa
Fig. 4

Scenario2: Traffic Flow Reconstruction and Air Quality predictions in Firenze, Pisa, Livorno

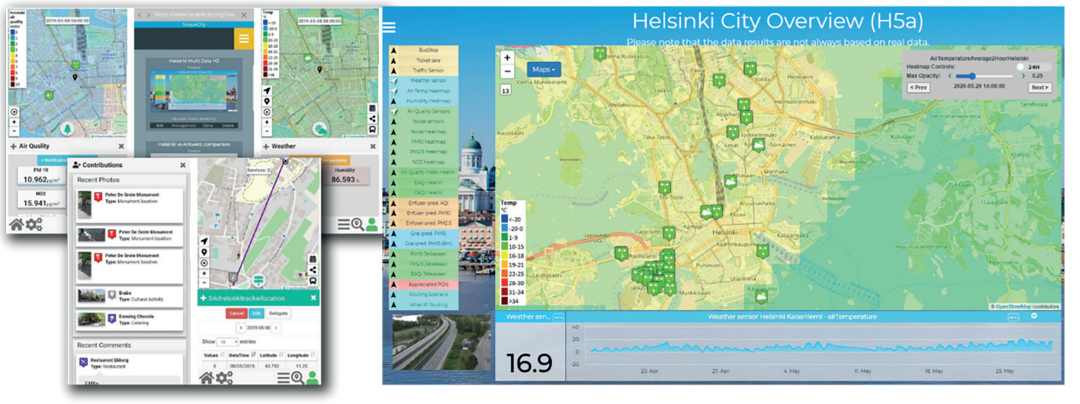
One of the needs that is increasingly growing for cities is to have a forecast of air quality in individual neighborhoods and areas of the city, to understand the reasons for any exceedances of permitted limits in order to amend city regulations to improve air quality. The source of pollution that has the greatest impact on air quality in cities is related to traffic. It is fundamental for the city operators, to monitor and have previsions on the air quality in each street of the city so that they may regulate urban mobility and providing the evidence at citizens that they are living in a city sensitive to the quality of life, predicting the overruns of law limits.
To this end, specific sensors and solutions become fundamental, such as: traffic flow sensors for monitoring urban mobility, traffic flow reconstruction, sensors for assessing air quality parameters, parking, traffic and RTZ sensors, weather forecasts, etc., together with the knowledge of the city structure, prediction model for environmental variables. Snap4city, thanks to the work done in the TRAFAIR CEF EU project (http://trafair.eu), is applied in three Tuscany cities, Pisa, Firenze and Livorno, in realizing air quality predictions, perfectly fitting the above described needs. In order to assess the air quality in each street of the city, the level of pollution aspects have to be measured, for example those regarding: SO2, NO, NO2, O3, CO, CO2, PM10, PM2.5, etc. and derived Air Quality Indexes. Most of the environmental pollutants are influenced by relevant traffic flows in different manners, while others are influenced by house heating, industries, waste, etc. Specific measures may depend on the sensor position and location context, on calibration, on the time of the measure, season, etc. A measure performed along a primary street in terms of traffic may strongly differ with respect to the actual values just in the garden of the house behind the primary street. To this end, mathematical methods have been set up to perform predictions of pollution diffusion and deductions.
In Fig. 4 the Pisa Dashboards with traffic in real time and air quality predictions are shown. The predictions are performed on NOX (depending on NO and NO2), for each hour in the next 48 hours, with a spatial resolution up to 4x4m, based on the GRAL dispersion Model (http://lampz.tugraz.at), both at 3 and 6 meters. The same Dashboards are available also for the cities of Firenze and Livorno and are all accessible from the main Tuscany Air Quality Dashboard (https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjY4MQ==). The air quality regarding NOX is primarily related to the production of pollution from the vehicles running in the city. The computation of the traffic flow in all primary and secondary streets of the city, realized with the DISIT Traffic Reconstruction tool (one of the results of the Sii-Mobility project, http://www.sii-mobility.org) is used to estimate the number of vehicles passing and producing pollutant in each road segment of the urban graph, by exploiting a limited number of measures of traffic by the city sensors. The production of traffic pollution is an input data for the predictive GRAL model to determine the air quality state in a city. Air quality predictions are visible in the dashboards that show the air quality in the individual streets and around the buildings, allowing not only the decision-makers but also the individual citizens to assess the air quality in their neighborhood and in their homes thanks to the color scale of values present to correctly interpret the colors on the map.

Scenario 3: Air Quality in Helsinki for the citizens

An environmental monitoring use case has been performed also in the of City of Helsinki, in a new smart district of Jätkäsaari: a small connected island to the South of the city. Jätkäsaari, is an expanding urban area, with more than 20.000 inhabitants, many are the works in progress for the construction of new buildings and services, including various hotels and office facilities. The city of Helsinki plans that at least 6.000 people who will live on the peninsula as a place to live or work. Jätkäsaari also encompasses the main part of Helsinki's passenger harbour. The large construction sites, the intensive and obstructed traffic, and the growing population create environmental challenges in Jätkäsaari. In this scenario, there was a strong involvement of citizens who collaborated with the public administration in the monitoring of pollutants. A set of low-cost sensors has been provided to citizens by the local administrations, in collaboration with city

innovation companies. Citizens placed them in their terraces, in ad-hoc selected areas: near to renovation works in progress, close to green areas or to the port. This process has allowed to receive targeted and localized data in areas relevant to the analysis of the specific problem, in addition to the official ones from the Finnish Meteorological Institute, but it also established greater communication among public authorities and citizens who feel active part of the city.
From a technical point of view, there was the need for Snap4City of integrating data coming from different sources and services, provide tools for data analytics about the current state of the air quality in different parts of Jätkäsaari both to the city decisors and to the citizens. Therefore, a number of Dashboards have been developed for the City Operators and ICT Officials (https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTQwNg==), while a Mobile App able to monitor in real time the different pollutants and fine particles produced, to have available basic information on commercial activities, events and public transport, etc. has been developed and published both on the Google Play Store and the Apple Store for the Citizens and Tourists, Fig. 5, (Badii et al., 2020). Moreover, the mobile App allows the city users to access the heatmaps, and to subscribe to notifications that are activated on their POI when the selected pollutant is above the critical values.

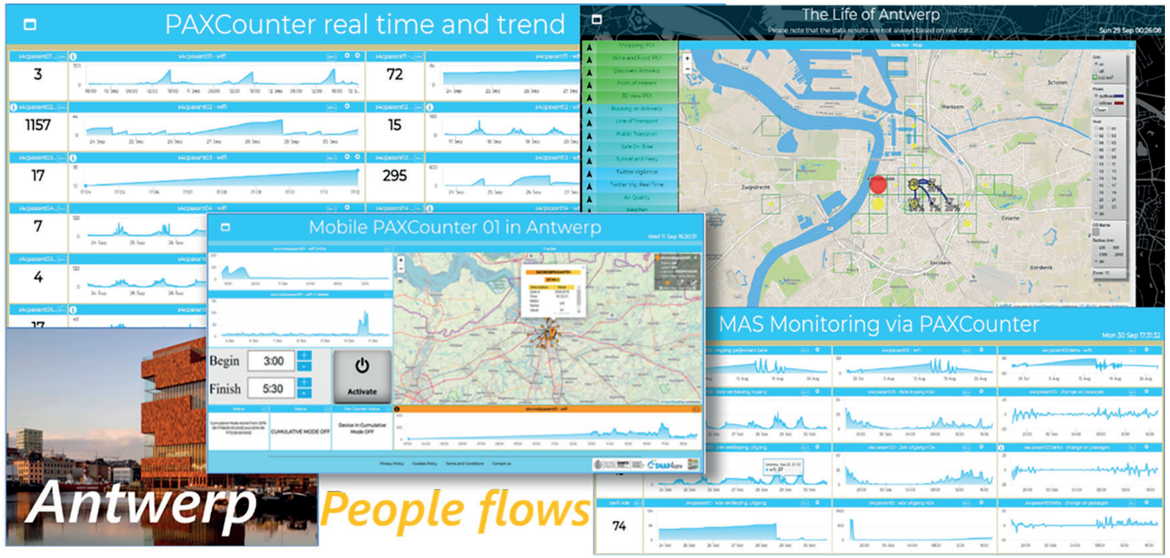
Scenario 4: People flow in Antwerp.
A further aspect at the basis of the organisation of a Smart City is to have tools to monitor people flows. The city of Antwerp is working on the construction of a smart zone in one of the city's districts. With this in mind, and in the context of the Select4City project (https://www.select4cities.eu), Antwerp city decisors expressed the need to monitor the influx of citizens and tourists into the smart zone to support local traders (museums, hotels, restaurants, etc.), to monitor events and to provide police support in road traffic management and to improve and make public transport services more efficient. The Snap4City platform worked with Antwerp for the positioning of pax-counters sensors, in order to monitor people and analyse the city flow (Badii et al, 2019). Obviously in this scenario, privacy aspects are fundamental, so the fact that the Snap4City platform is GDPR compliant was an essential starting point, without which the experimentation in the smart zone could not have been realised. The pax-counters can be located on fixed places, using WiFi to send the obtained measures including additional sensors or can also count the number of people on the move.
2 Mobile pax-counters are entrusted to local cops who take the sensors with them on daily patrol, are used to monitor the citizens movements on the streets in relation with sports or entertainment events. Each newly



Air Quality predictions: Mobile App and IoT Sensors
Fig. 5

measured count is associated with GPS coordinates and sent to the platform, thus obtaining trajectories with the associated values, as shown in Figure 6 (central dashboard), using the Snap4City Widget tracker. In 'The Life of Antwerp' Dashboards the flows of people entering and leaving the smart zone are highlighted, also indicating where people come from and where they go. A number of devices have been installed in Museums and other public facilities in Antwerp. In this case also a series of dashboards have been created, showing in real time the number of people who are in the facilities, producing statistics and data trends, etc. In Figure 6, the Dashboard presenting the time trends of the entrances and exits of the museum and derived data is shown (Mas Monitoring PaxCounter, below right). All the derived values of the initial real time data can be computed in real time, saved on personal storage device and shown on dashboard via an IoT Application exploiting MicroServices of Snap4City. At the same time as calculating the sums, the difference between the people entering the museum and those leaving to obtain the number of people inside the museum in another entity is also made.

Key Indicators to evaluate the Snap4city impact in Smart Cities
As seen above, through the dashboards within Smart City control rooms, the data collected and aggregated are summarized and made accessible for the decision makers and shared to the city operators. For each of the scenarios described, dashboards and/or Mobile Apps have been created, which can be public and consulted by all citizens or have restricted access to employees of municipalities. The data displayed do not only have a descriptive value of the state of the phenomenon observed, but in some cases measure the degree of efficiency of the solutions adopted, when they express through KPI (Key Performance Indicator) the present condition with respect to threshold values deemed sensitive by experts in the field of analysis.
Much of the KPI is representative of the resources deployed in the city and is not necessarily geo-localized, although the



People flows in Antwerp
Fig. 6

geographical component is increasingly important, considering that urban planning must combine a broad strategic vision with the specific needs of local contexts. Transportation and mobility, energy, weather, health, water, waste, security, civil protection, ICT are the main fields the city planners can monitor combining KPIs. In Snap4City, KPIs are saved in the Knowledge Base, specifically in the Dashboard Engine database. They can be values with which position and time or entity attributes are associated, such as shape (region), paths (lines, trajectories, etc.), administrative boundaries, etc. The principle on which they are based is to distribute data values over time and space, in order to follow the trend of different phenomena, correlating them to the geographical dimension in which they occur, a crucial objective for urban planners or decision-makers in general. Following this principle, cities must establish what are their strategic priorities and consequently define data to be acquired and related performance indicators. In this light, the analysis of dashboards of different cities shows in their own articulation what are the results of the local planning processes, of their integration with ICT infrastructures and of their ability to interpret citizens' requests or sentiment.
The possible examples are manifold, an examination and a punctual comparison would exceed the space of this article, so for the consultation of some of the dashboards currently created and publicly accessible, see the specific sections on Snap4city.org. To give an idea of the analytical potential of the platform, consider that for Florence and Tuscany Region described in scenario 1), that alone it is able to 'digest' over 1.5M complex events per day (static data, dynamic/real-time data, mobile app), which after being analyzed can be returned through tables, gauges or graphs of various types (Kiviat, pies, curved lines, barlines, etc.).
Being able to count on a synoptic framework that can show together, for example, structure of the city, position of the services, paths of the public transportation, real time value of traffic, environmental variables, people flows, weather conditions and forecasts, cycling paths, hospitals first aid, sets the best conditions to establish restrictions to be applied, areas to be closed, new paths for public and private vehicles or for pedestrians, etc.
Recent analysis on the behaviors of different users show that mobility and environment are the areas where attention is most concentrated (more than 6.000 accesses and 340.000 minutes, which is an average of 57 minute per access) (Mitolo et al. 2020). According to what is suggested by the European Resilience Management Guide[2], issued following the findings of the pilots of the RESOLUTE project (http://www.resolute-eu.org), the definition of performance indicators is an essential step if you intend to implement a life cycle for a resilient system, i.e. anticipate, monitor, respond, learn. "Monitoring resources generates information on resource allocation and the understanding of their flows, which represents one of the fundamental tools for planning activities, both as a primary input and as indicators for the potential need of planning revision or reassessment" (Gaitanidou et al., 2018 p. 20 and passim). When a KPI shows values lower than the threshold a priori defined, the response actions are prepared. This is only one, probably the most evident, of the effects produced on a city intended as a complex ecosystem. Some secondary effects are given by the fact that the aggregation of KPIs acts as a coordination mechanism between different parts of the system.
Furthermore, Snap4City allows anyone to manage, for example, their IoT devices by establishing personal KPIs (MyKPI, data, POI), which each user can decide whether to make it public or with a restricted use and manage their ownership. Whatever the choice, such a powerful architecture available to all constitutes a further step in the direction of public engagement and in the growth of civic awareness, that is the primary condition of a smart city.

Conclusions and future work

This article examines the influence had by the Snap4City Solution in sustaining and providing support to a good governance in a Smart City. This environment gives a clear insight on how data and digital technologies are changing the way people "understand and plan cities", "manage urban services and utilities", and "live urban lives" (Kitchin, 2011). In particular four different scenarios have been taken into account that have highlighted how the platform is on the one hand versatile because it is able to adapt to different political and geographical contexts. On the other hand, the solutions described the different use cases underline that Snap4City is not only an evolved ICT infrastructure, is 100% Open Source, secure and GDPR compliant solution designed to support city planners, combining strong tools for data integration, analytics, forecasting and visualization but it is also based on Living Labs strategies to support the development and actuation of urban strategies.

Acknowledgement

The authors would like to thank the European Union's Horizon 2020 research and innovation program for funding the 'Select4Cities' PCP project (within which the Snap4City framework has been supported) under grant agreement No 688196, and also all the companies and partners involved. Snap4City and Km4City are open technologies and research of DISIT Lab https://www.snap4city.org. The authors would like also to thank the TRAFAIR CEF project of the EC with grant AGREEMENT No INEA/CEF/ICT/A2017/1566782 also all the companies and partners involved. Snap4City and Km4City are open technologies and research of DISIT Lab.

Note

[1] This compliance is not only imposed by European regulations, but responds to a need to code ethically, focusing primarily on who has access to data, what should be collected, how should it be used, etc. (Kitchin, Dodge, 2011)

[2] https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5be08fd14&appId=PPGMS

Bibliografia

Aghion P., Howitt P. 2005, *Growth with quality-improving innovations: An integrated framework*, in Aghion, P., Durlauf (eds.) *Handbook of Economic Growth*, Vol 1A, Ch.2, pp. 67-111. DOI: 10.1016/S1574-0684(05)01002-6

Ahrendi, R., Farchyi, E., Kaplanisi, I., Lembckei, A.C. 2014, *What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries*, OECD Regional Development Working Papers, 2014/05, *https://doi.org/10.1787/5jz432cf2d8p-en*

Azzari M., Garau C., Nesi P., Paolucci M., Zamperlin P. 2018, *Smart City Governance Strategies to better move towards a Smart Urbanism*, in *The 18th International Conference on Computational Science and Its Applications (ICCSA)*, July 2-5, 2018 in Melbourne, Australia in collaboration with the Monash University, Australia.

Badii C., Bellini P., Nesi P., Paolucci M. 2017, *A smart city development kit for designing Web and mobile Apps*, IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 10.1109/UIC-ATC.2017.8397569.

Badii, C., Bilotta, S., Cenni, D., Difino, A., Nesi, P., Paoli, I., Pantaleo, G., Paolucci M. *Real-Time Automatic Air Pollution*. Accepted paper in 10th IEEE ISCC MoCS Workshop 2020. https://mocs20.gforge.uni.lu/

Badii C., Bellini P., Cenni D., Difino A., Nesi P., Paolucci M. 2017, *User Engagement Engine for Smart City Strategies*, in *IEEE International Conference on Smart Computing*, IEEE SMARTCOMP, Hong Kong.

Bellini P., Bilotta S., Nesi P., Paolucci M., Soderi M. 2018a, *WiP: Traffic flow reconstruction from scattered data*, *IEEE International Conference on Smart Computing*, SMARTCOMP, art. no. 8421365, pp. 264-266. DOI: 10.1109/SMARTCOMP.2018.00052

Bellini P., Nesi P., Paolucci M., Zaza I. 2018b, *Smart city architecture for data ingestion and analytics: Processes and solutions*, in *Proceedings IEEE 4th International Conference on Big Data Computing Service and Applications, BigDataService*, pp. 137-144, March 26 - 29, 2018. DOI: 10.1109/BigDataService.2018.00028

Cosgrave E., Arbuthnot K., Tryfonas T. 2013, *Living Labs, Innovation Districts and Information Marketplaces: A Systems Approach for Smart Cities*, in *Procedia Computer Science*, Volume 16, 668-677.

Cundari C. 2015, *Il rilievo urbano per sistemi complessi. Un nuovo protocollo per un sistema informativo di documentazione e gestione della città*, Edizioni Kappa, Roma. ISBN:978-88-7890-668-6

Dodge M. 2010, *Code/space*, in «Urbis Research Forum Review», Vol. 1, Issue 2, pages 15-25.

Duraton G., Puga D. 2014, *The growth of cities*, in Aghion, P., Durlauf (eds.), *Handbook of Economic Growth*, vol. 2, ch. 5, pp. 781-853, Elsevier. DOI: 10.1016/B978-0-444-53540-5.00005-7

Fagerberg J., Srholec M., Verspagen B. 2010, *Innovation and Economic Development*, in *Handbook of the Economics of Innovation*, vol. 2, ch. 20, pp. 833-872. DOI: 10.1016/S0169-7218(10)02004-6

Glaeser E. 2011, *Triumph of the City. How our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and*

Happier. New York, Penguin.

GDPR: General Data protection Regulation, *https://eur-lex.europa.eu/eli/reg/2016/679/oj*

Kitchin R., Dodge M., 2011, *Code/Space: Software and Everyday Life,* MIT Press, Cambridge MA.

Kitchin R., 2011, *The programmable city*, in «Environment and Planning B: Planning and Design», Vol. 38, No. 6, pp. 945-51.

Mitolo N., Nesi P., Pantaleo G., Paolucci M. 2020, *Snap4City Platform to Speed Up Policies*. Accepted paper in SSPCR conference, *Smart and Sustainable Planning for Cities and Regions*, 2020. *https://www.sspcr.eurac.edu/*

Nesi P., Paolucci M. 2018, *Supporting living lab with life cycle and tools for smart city environments. In DMSVIVA: 24th International DMS Conference on Visualization and Visual Languages*, 10.18293/DMSVIVA2018-019

OECD 2013, *Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data"*. OECD Digital Economy Papers, No. 222, OECD Publishing, Paris. *https://doi.org/10.1787/20716826*

OECD 2015a, *The Metropolitan Century. Understanding Urbanisation and its Consequences*. OECD Publishing, *https://doi.org/10.1787/9789264228733-en*

OECD 2015b, *Governing the City*. OECD Publishing, *https://doi.org/10.1787/9789264226500-en*

Rodríguez-Pose A., Crescenzi R. 2008, *Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe*, in «Regional studies», 42 (1). pp. 51-67. DOI: 10.1080/00343400701654186

Sassen S. 1991, *The Global City*. New York, London, Princeton University Press, Tokyo.

Ubaldi B. 2013, *Open Government Data. Towards Empirical Analysis of Open Government Data Initiatives*, OECD Working Papers on Public Governance No. 22. *https://doi.org/10.1787/19934351*

United Nations 2018, *World Urbanization Prospects 2018. Highlights*. Department of Economic and Social Affairs, *https://population.un.org/wup/Publications/Files/WUP2018-Highlights.pdf*

Gaitanidou E., Bellini E., Ferreira P. 2018, *European Resilience Management Guide*, RESOLUTE D3.6.

Digital Territories and energy transition
the limits to growth

Cécile Diguet
Urban Planner and Design Director at l'Institut Paris Région
Fanny Lopez
University Gustave Eiffel, School of Architecture, Paris
cecile.diguet@institutparisregion.fr
fanny.lopez@paris-est.archi.fr

DOI: 10.13128/contest-11725
www.fupress.net/index.php/contesti/

keywords
data centers
digital infrastructure and urbanism
digital territories and energy transition
energy systems
data centers and public authorities.

Internet is not magical territory in which flows circulate without any physicality, nothing consumes more space, energy and resources than digital technologies. The cornerstone of the digital technical system, data centers are present everywhere and their construction is accelerating. They are found in city centers, in suburbs, in rural territories, as well as isolated and deserted regions. Whether they are connected to the existing electricity networks, autonomous or incorporated into energy exchange circles with variable perimeters (block, neighborhood, city, territory), they redefine, each time, the energy project of the territories in which they are sited because their electricity consumption is very high. By the yardstick of a pioneering and flourishing American digital industry, but with a weakened energy system, field studies on the West Coast (Oregon, Silicon Valley) and the East Coast (New York) have enabled us to shed light on the challenge of integrating data centers into territorial energy systems in Europe. In this article we propose bringing the energy and spatial excess of the digital world back to the center of the debates on the subject. The article is divided into 3 parts which describe The data center siting strategies and its various location ; 2) How local public authorities grappling with a complex and opaque sector ; 3) The excessive energy consumption of data centers.

Introduction

During the lockdown period faced with Covid-19, working at home and streaming have put digital infrastructures to the test. Google, Amazon and Netflix have opted for a lower resolution to reduce throughput. Despite these efforts, the projections on the evolutions and the growth of online videos, but also all the other digital activities will further increase demand and energy consumption, where the 5G infrastructure is being rolled out. The battle of numbers is constantly trying to lessen or complicate the energy impacts of the digital,[1] even while no serious study has made it possible to demonstrate that the digitizing of the world will permit us to meet the energy transition and environmental goals necessary for our species' survival. As a certain Cédric Price said: "The answer is technical. But what was the question?"

What are the paradigms that founded this technological choice? Paradoxically to the imaginary dimension of dematerialization anchored in the tradition of the cybernetic utopia that made the Internet a sort of magical territory in which flows circulate without any physicality, nothing consumes more space, energy and resources than digital technologies. In 2015, one of the most recognized and cited studies evaluated the consumption of the IT sector (equipment, networks, data centers) at 7% of global electricity. The projections reach a maximum of 13% of global electricity used by the data centers in 2030, and between 20 and 51% for the IT sector[2] in its totality (Andrae and Edler, 2015).

If these energy consumption figures are complex to measure, the spatial impact is more visible. The cornerstone of the digital technical system, data centers[3] are present everywhere and their construction is accelerating. They are found in city centers, in suburbs, in rural territories, as well as isolated and deserted regions (Diguet and Lopez, 2019). Whether they are connected to the existing electricity networks, autonomous or incorporated into energy exchange circles with variable perimeters (block, neighborhood, city, territory), they redefine, each time, the energy project of the territories in which they are sited because their electricity consumption is very high. The corollary of the exponential growth principle of data since the explosion of the digital sector is an increase in space and energy needs. As a new stage of the network urbanism (Dupuy,1991), the digital city has often been analyzed in terms of uses and practices, services and events, the cultural meanings, social effects and environmental impacts (Townsend, 2013) (Picon, 2015) neglecting the materiality and energy consequence of its physical infrastructures on territories.

The data center program raises strategic spatial and energy questions, even though the environmental footprint of the smart city and our digital lives has not been adequately discussed (Tinetti et al. 2016). The data center industry seems to have knowingly maintained an effacement that limits or pushes back questions on the viability of our practices and their energy mirror. We can ask if the unmasking of these storage infrastructures and their energy functioning participates in revealing the untenable character of the digital technical system and its globalized economy, which functions on invisibility and distance. By the yardstick of a pioneering and flourishing American[4] digital industry, but with a weakened energy system, field studies on the West Coast (Oregon, Silicon Valley) and the East Coast (New York) have enabled us to shed light on the challenge of integrating data centers into territorial energy systems in Europe. In this article we propose bringing the energy and spatial excess of the digital world back to the center of the debates on the subject[5].

The article is divided into 3 parts which describe

1. The data center siting strategies and its various location (urban, peri-urban and rural)
2. How local public authorities grappling with a complex and opaque sector
3. The excessive energy consumption of data centers

The data center siting strategies

If they are present worldwide, the siting of data centers is concentrated in the United States, Asia and Europe. Amsterdam, Dublin, Frankfurt, London and Paris have a very large number of them. Two location criteria are genuinely structuring for data centers. First, being located near the optical fiber Internet network, notably the principal hubs – the Internet backbones – to connect to Internet Exchange Points, platforms on which all the networks that transmit information throughout the world are connected. Next, sufficient electrical power for its installation and possible development, in a stable and reliable manner, and preferably inexpensively, must be made available. This can be teamed with a nearby substation. An unweighted criterion until now, this could change in the future.[6] We should also mention a certain number of complementary location criteria: security, absence of nearby inhabited zones, reactivity of the local hosting administration in administrative actions, availability of renewable energies and specific rates for data centers, various tax incentives, abundant and affordable land, with the fewest restrictions and easements possible. These siting criteria make it possible to better understand how new digital territories are consolidated. We can distinguish three of them: the rural world, the metropolitan outskirts and city centers

Big Tech out to conquer the rural world

The rural world and the peri-urban territories are of great interest to large data center operator because of their isolated nature, their available land, but also for the tax advantages that the local authorities, in search of a new economic impetus, offer them. In terms of development, the result is often that of territorial sprawl, in a sometimes unbalanced relationship between the operator and the local administration. On the west American coast, Oregon is a prime location for data centers, thanks to its climate, land availability and abundant and inexpensive energy thanks to the hydroelectric dams on the ColumbiaRiver. In Prineville, in a rural and deserted territory in Oregon (US), Facebook built its first data center in 2009 and Apple closely followed in 2011. The two digital giants have continued to develop their installations: a third Apple storage space is being completed and Facebook will soon have a total area of 200,000 m^2. The data centers arrived there concealed. The code name of Facebook was Vatas. The local administration only knew that the project needed 80 to 120 hectares, a great deal of water and a large amount of electricity. It had however no idea of this operator's activity at startup and until rather late in the project's progress. The secrecy culture of GAFAM can also work against the territory's need for anticipation and urban and economic planning. Betty Rope, the mayor of this town of 10,000 inhabitants, estimates however that Facebook "is a good, very good neighbor"[7] that, with Apple, redynamized an economy destroyed by the 2008 crisis and the decline of the timber industry (21% unemployment), first with construction sites that have been ongoing, next with the resulting direct and indirect jobs whose number rose, according to the mayor, to 500 at the end of 2017. Facebook provided a great many services to its employees. Moreover, no one could have imagined a decade earlier that the electrical power needed in the town would rise from 10 to 500 MGW. Consequently, Facebook and Apple co-invested in five solar farms of 15 MGW each, not far from their sites. And each pays annual 'project expenses' to install, then develop, and helps invests in local infrastructures (substations, water production plants, sewerage systems). Here, the digital giants are considered above all a godsend, a vector of jobs and economic development.

If this digital rural model does not yet concern all the European countries, Swedish, Irish and Finnish rural territories have dealt with these situations because they propose inexpensive land and energy as well as various tax incentives. GAFAM questions here the European digital strategy in this area, since French professionals have put forward tax dumping techniques in particular.
In France, territories like Saclay, south of Paris, could be concerned by dynamics such as these due to their connectivity (backbone along the A10 highway), the presence of research centers, universities and internationally known higher institutions of learning, a good electricity supply (St. Juste substation, another being built on the Plateau de Saclay), but also real estate opportunities with generous footprints. These territories will be able to draw up, with the energy operators, strategies to make the most of these installations in terms of energy (pooling) and space (preferential occupancies of abandoned spaces, for example).

The outskirts and the digital: planning to be built

Metropolitan outskirts, often industrial in the past, comprise ideal sites for colocation and Cloud data centers, and for local companies that only need low latency. There is a common development model: a formerly industrial territory, often serving the metropolis, offering large areas of land and abundant electrical power, but also good connectivity – data centers then replace factories. Each territory uses their arrival differently. Certain cities have taken hold of the development of these buildings by adopting proactive positionings to attract them and obtain direct benefits from them. In Silicon Valley, the city of Santa Clara concentrates the most data centers on its territory. Fifty or so of them consume 70% of the energy supplied by the municipal public company, Silicon Valley Power, created in 1896, offering the least expensive electricity in all of California. This advantage, undeniable for the data center industry, pushed them to concentrate there starting in the 1990s, transforming former electronics factories then constructing their own buildings. Santa Clara also provides a dark fiber[8] network for companies, completing a perfect offering for the digital world. If it perpetuates in a certain way its role of a territory serving the metropolitan heart and decision-making centers, Santa Clara (just like San José) now hosts many activities: the service sector, company headquarters, universities, offices, large-scale sports and cultural facilities. Today, land is starting to become scarce in Santa Clara and the space dedicated to data centers is competing with housing needs. Whereas they were often in industrial zones, they are now found in office zones, which illustrates their mutation in terms of urban integration (as in Hillsboro in Oregon). The infrastructures are more compact and architectural, like the Vantage campus, built higher on a densified plot on Walsh Boulevard, which now has six data centers.

In France, Saint-Denis and Aubervilliers, just north of Paris, seem to have had this development foisted on them. Gathered in the intermunicipal structure Plaine Commune, the presence of data centers dates back to the late 1990s. The presence of enormous plots of land available at affordable prices in the immediate vicinity of Paris, good electricity availability, the quality of connectivity with the presence of one of the main lines of the Internet network along the A1 highway (which permits the link between the different European data centers) has favored the development of these infrastructures. The data centers present on Plaine Commune are mostly large structures with an average area of 1,000 to 5,000 m². If a slowdown in the number of data center constructions has been observed since 2010,[9] large-scale construction sites are underway: the Interxion data center on the former Eurocopter site[10] in La Courneuve, as well as the Interxion data center extension on the rue Rateau, which should reach 9,000 m² of IT rooms and the Equinix data center in Saint-Denis with 13,000 m² of IT rooms on a plot of over 6 hectares. This therefore amounts to over 88,000 m² that will be dedicated to digital storage with over 360 MGW available. Nonetheless, the Plaine Commune metropolitan area recalls that data centers strongly impact the territory and raise concerns about the durability of the local electricity system; about possible problems raised by citizen groups (noise, danger linked to heating oil storage, electromagnetic waves); and, in a context of competition between uses in the center of the metropolis, and the very large impenetrable zones – the data centers – correspond to a very small number of jobs. Emphasizing the mono-functional character of the territories, similar in terms of landscapes of logistics or commercial zones, these digital activity zones are moving toward urban fragmentation with digital enclaves, often protected by defensive fences without much urbanity.

The hyper-urban Gateways, strategic hubs of the networks

Metropolitan centers, in the United States and Europe, and international cities such as London and New York even more so, are strategic hubs where Internet cables are found and connected; global and centralized decision-making centers where consumed, distributed products of multiple informational contents (cultural, financial, communicational, commercial, etc.) are also found. The former telephone exchanges have often become

Internet Exchange Points, combined with greater or lesser areas dedicated to data storage.
In New York, several buildings in Lower Manhattan are continuing, with the Internet, their communicational destinies and were able to adapt the historic real-estate heritage, formerly the head offices of the major telegraph and telephone companies – 32 Avenue of Americas (headquarters of AT&T) and 60 Hudson Street, (headquarters of Western Union), respectively built in 1914 and 1928. In the late 1990s, the two buildings shifted to the digital and carried out an impressive molting, getting rid of telephone installations to install Internet Exchange Points, data storage spaces and cutting-edge telecommunication infrastructures in them. Starting in 2000 and after the 9/11 terrorist attacks in 2001, data centers began to relocate in nearby New Jersey, more advantageous than New York in terms of the price of energy as well as that of real estate, but also less climate vulnerability. In 2012, Hurricane Sandy shut down several data centers in Manhattan. Many data centers however remain in New York City, like the Sabey tower built by the New York Telephone Company in 1975 at the foot of the Brooklyn Bridge, and subsequently a Verizon telephone exchange. This 32-story building offers 102,000 m^2 of available floor space and has continued to develop.

Three subjects are particularly interesting here: the recycling of existing major service sector building in the heart of the metropolis; the infrastructural compactness of these machine-buildings, containing all the infrastructures to handle their autonomy (heating oil, water, cogeneration, heat storage); the mixed uses in the buildings.
For example, at 32 Avenue of the Americas, offices, radio studios, data centers and 70 telecom operators (30% of the floor space) and the head office of the Tribeca Film Festival coinhabit in 186,000 m².
The public authorities are sometimes helpless or lost faced with the arrival of these new space and energy-guzzling infrastructures. Positionings and strategies differ from one country to another.
In France, in Paris, it is above all former industrial buildings with a metal structure built for the textile industry, the press or department stores that have been mobilized for data center use. It is also a part of the telecom real-estate heritage that now hosts data. The Sentier ['garment district'], also once called Silicon Sentier, is the headquarters of several large data centers that were able to be installed in former industrial buildings (Zayo, Telehouse). The Telehouse operator also was installed in the 1990s in a former department store, on the boulevard Voltaire, near the town hall of the 11th arrondissement, with a total floor space of 7,000 m^2 and available power of 5 MGW. The Parisian industrial heritage therefore made it possible to handle large-scale connectivity and a digital hosting capacity in the heart of Paris and in the immediate vicinity of its business districts (central business district and La Défense) in particular. The buildings specifically erected to house data centers in very dense urban fabrics are much rarer, giving the complexity of building in them, and the priority often given to housing and other uses, but especially the ease, in comparison, of developing in better adapted territories, with fewer urban and environmental constraints and those with neighboring homes and businesses.

Local public authorities grappling with a complex and opaque sector

Faced with the creation of these digital territories, several positionings have been observed. The great autonomy of American cities has permitted them to regulate and locally profit from the data center industry, which is the case for Santa Clara, whereas in the rural zone, the town of Prineville in Oregon illustrates the imbalance between the municipality and Big Tech.
In France, north of Paris, development took place without any urban planning and without technicians or local elected officials being able to anticipate what came next, or not having the determination or desire to better integrate data centers. In Europe, the case of the Netherlands is unique where the cities of Amsterdam and Haarlemmermeer took the lead with a moratorium whose aim was to stop their construction for a while in order to better think about their development.

Putting up with but promoting: Prineville (Oregon), a small rural town faced with Big Tech

If the GAFAM data centers in particular often replace aluminum foundries and sawmills, which also needed a great deal of electrical power, the power consumed by data centers is much greater and as are the potential imbalances on the territories. Furthermore, the secrecy surrounded the siting of data centers and which does not differ much from what usually concerns military installations or national security, makes it difficult if not impossible to control the development and planning of a territory.
This imbalance in information does not permit local administrations to sufficiently anticipate a positioning on the data center question, to consolidate expertise on the subject, on the alternatives in terms of urban and architectural siting and energy integration, for example. This is all the truer in rural territories like Prineville or nearby Umatilla, where some of the technical departments are not necessary equipped to deal with these subjects, where the political or demographic weight of the territories is not of the sort to

offset the power of GAFAM in particular and, lastly, where the often very difficult economic situation in terms of employment pushes the territory to favor job creation, even at a minimal level. Likewise, these departments are not prepared to deal with urban, energy, environmental or economic questions in the long term.

Welcoming and regulating: the 'one-stop shop' of Santa Clara in California

In the early 1990s, data centers, first small ones, in existing buildings notably those of the telecoms, were installed in Santa Clara. Then, and notably with the development of environmental labels and energy performance requirements, new larger and larger buildings were erected. Whereas the city initially approved these projects without too many technical studies, especially focusing on water consumption for cooling (and the sizing of the sewerage networks), it quickly decided to organize to develop a one-stop shop for data centers and consequently to anticipate and administer the development of these digital infrastructures proactively[11] by setting up a one-stop shop.

This one-stop shop brings together:

- Silicon Valley Power (SVP), the municipal electricity company created in 1896. One person is specifically in charge of the data center clients (which represent 80% of the company's revenues). The company also developed and manages an optical fiber network (and free Wi-Fi in the city) and is responsible for managing the water and sewerage networks.
- The firefighters: Santa Clara is a rare exception in Silicon Valley – the city's firefighters are not dependent on the county, but are a municipal corps. This makes it possible to more easily find consensual solutions with the data centers, more in terms of an obligation of results rather than that of means.
- The urban planning and environmental departments.

Technical reviews of the projects are therefore done in three to four weeks, after which the permits are granted or rejected. All the partners, in particular Silicon Valley Power, are involved as much as possible upstream in order to rapidly incorporate all the technical constraints.

Moreover, the benefits Silicon Valley Power accrues through the large electricity consumptions of the data centers are reinvested by the Santa Clara municipality in the city's public facilities. One example is the imposing football stadium built in the northern part of the city, but more largely the quality of the public spaces, parks and local facilities. We can note here the interest in pooling certain departments for an integrated and connected approach to siting complex facilities – the data centers – but also to make use locally of the benefits derived from them.

Putting up with: the example of Plaine Commune north of Paris, a rapid and uncoordinated development

On the Plaine Commune territory, north of Paris, the first data center arrived in 1999 (Interxion), then, in about 10 years, the territory became a major digital hub for data centers. This new corporate real estate, still little known, consequently developed on the Saint-Denis plain, and notably in La Courneuve. The subject did not seem strategic and was only questioned when the ALEC (local energy agency) published a study in 2013, whereas numerous installations had already been built. However, in the absence of strong political support on the subject, an alliance between the communes concerned and the metropolitan area and the creation of expertise on data centers, no authority or strategy on the subject existed. The conclusions of the ALEC study therefore were not followed by actions, notably on heat recovery.

Most of the sites have of course received an ICPE[12] authorization delivered by the state, but this does not constitute a lever for genuine negotiations. Moreover, only Enedis (but not the communes or the metropolitan area), which manages the electricity network in France, has a real vision of the electricity consumption of data centers and reserve capacity or queue management. This lack of knowledge is a major problem for treating projects and urbanism authorizations with a global vision while the energy consumption of these newcomers sometimes disturbs the territories.

Controlling and planning: Haarlemmermeer and Amsterdam's moratorium

Data centers in the European Big 4 – London, Amsterdam, Frankfurt and Paris – continue to develop, but the Dutch capital is the first to have stopped this growth in 2019 with a moratorium prohibiting the siting of data centers for one year. At the moment when the growth of data centers reached a peak in electrical power, the cities of Amsterdam and Haarlemmermeer stopped the development of this industry. It was a one-year shutdown to rethink the siting policy for the next three decades. The startup will be done in the framework of a reworked planning with all the actors concerned.

Today, 75% of all the country's data centers are in the region around Amsterdam.[13] Nearly 2 million square meters of floor space for data centers are likely to be dedicated to digital storage in the years to come, which would almost double the current capacity according the Dutch association dedicated to this industry. In total, the metropolitan area houses 47 colocation data centers and 53 installations with a sole user. They are located in Amsterdam, Almere, Aalsmeer, Haarlem, Hoofddorp/Schiphol (Haarlemmermeer municipality) and Purmerend. In June-July 2019, the cities of Haarlemmermeer and Amsterdam decided to temporarily prohibit, for 12 months, the construction of new data centers on their territory. Aiming at the long term, a new data center policy is being drawn up for Amsterdam and for Haarlemmermeer.[14] The Data Center Spatial Strategy, Roadmap 2030 (REOS) presented a new national spatialization of data centers in three illustrated scenarios: metropolitan concentration, regional clusters and decentralization throughout the territory (Ruimtelijke, 2019). The goal of the moratorium was not to prohibit data centers, but to obtain more rational regulation through an urban siting policy better in order to facilitate their growth in the long term. Before modifying the zoning plan or totally prohibiting construction through this tool, the objective is to stabilize a policy capable of issuing mandatory conditions when the building permit is filed. The flexibility of the conditions applicable to certain zones in the municipalities is under discussion.[15] In Haarlemmermeer, it seems that the problem of energy availability prevails over the other aspects. Nonetheless, the desire to better control data centers in terms of durability and to further restrict their placement is strong.[16] What is happening in the Netherlands illustrates the spatial and energy tension caused by the digital industry on the territory. It could prefigure a European trend in the densest sectors.

The excessive energy consumption of data centers

Data centers will be one of the largest consumers of electricity in the 21st century, buoyed by a multiplication of Internet traffic, the explosion of data exchanges, the growth of the Cloud and the projections of 100 billion connected objects by 2030 (Cisco, 2016). Their frenzied and irrational growth questions the programmed obsolescence of their IT systems as well as the squandering of energy linked to the globalized digital economy.
A large amount of electricity is necessary not only for their functioning but also to air-condition the machine rooms and racks. In terms of equipment, each data center – generally connected for security reasons to two electrical power supplies – is equipped with backup generators and battery rooms providing autonomy in the event of a cutoff on the network. As these are relatively rare, this infrastructure multiplication creates a redundancy whose usefulness raises questions today.
Whereas the search for energy restraint is put forward to reduce CO_2 emissions, the growing electricity consumption of data centers is excellent news for the energy actors. Data centers are perfect customers for the electricity companies: no peaks to manage, a large stable consumption day and night, all year long, and constantly increasing…

Energy and digital operators: dependency and disruptions

The American energy market and the history of its deregulation varies from one state to another and is very different from the European situation. However, it can be noted that on both sides of the Atlantic, common characteristics regarding the requirements of the energy transition and the resistance of the utilities that see their markets disrupted and their activities transformed in depth. In the United States, in 2014, data centers consumed about 70 TWh, which represented 1.8% of the country's total energy consumption. Their consumption had increased by 4% between 2010 and 2014 (Shehabi et al. 2016). In the United States, in a general manner, the condition of the electricity transmission networks was not as good as in Europe where the electricity transmission and distribution companies has heavily invest in infrastructure. In the three American areas studied (California, Oregon and New York), the energy territories are unquestionably specific but the production of renewable energies in large quantities is a constant.
In California, in Silicon Valley, where despite high electricity prices (except in Santa Clara), there are a great many data centers and they take part in the development of renewable energies (27% of the electricity production). In Oregon, a territory that serves California, which welcomed the first large Google data center in 2006, the climate is cooler and electricity less expensive, thanks to the dams on the Columbia River, and where there are strong land purchase possibilities (71% of Oregon's electricity production is generated by traditional hydroelectric power plants and other renewable energy resources).
The urban density of the city of New York and the gradual transformation of emblematic telecom buildings into data centers, in a post-Hurricane Sandy context, favored very strong energy resilience policies, notably on renewable energies and micro-networks (23% of New York State's electricity production comes from renewable sources).
For Gary Cook of Greenpeace, author of the Clicking Clean report,[17] if GAFAM makes efforts in using renewable energies, notably by promoting commitments to be supplied by 100% renewable energies, there is still a long row to hoe because this energy is not local, additional or really used by their data centers: it is partially bought from the electricity companies through certificates (RECS in the United States, GO in Europe). GAFAM is increasingly particular and demanding

and has gone as far as forcing the energy transition from certain old utilities with a very high-carbon energy mix. This is notably the case in certain American states where land is inexpensive and the energy mix high in fossil fuels (with coal-fired power plants whose industrial lobby has remained very strong). In Wyoming, Microsoft bought over 230 MW of electricity from a wind farm from a local utility. In Virginia, Microsoft negotiated an agreement with Dominion Virginia Power and the state government for the production of a 20 MW solar power plant. The company is doing the same in Europe where it signed a contract to reserve all the electricity production of a General Electric wind farm, which had just been started up in Kerry County in southwestern Ireland. This purchase aims at greening the energy mix of its large data center west of Dublin dedicated to Cloud services for all it European corporate clientele. On the other hand, the most recent Clicking Clean report (2019) very specifically targets the leader of the Cloud, Amazon Web Services, all of whose new data centers built in the Ashburn cluster, near Washington, DC, are supplied by coal. The author therefore stresses that the exceptional growth of data centers on this enormous site justifies, for Dominion, the local electricity company, supporting the development of the Atlantic Coast pipeline whose environmental effects are considered devastating on the Appalachian Mountains and regarding CO_2 emissions and pollution.

On the West Coast, GAFAM seems to exert strong pression on the energy system to move toward the use of renewable energies, and would like to end the monopolies of electricity companies that it considers an obstacle to reaching this goal because the utilities first want to make their current installations profitable before investing in new, cleaner ones. Apple consequently created its subsidiary Apple Energy, which enables it to sell the surplus energy produced by four of its American installations, including the Newark wind farm in California. These renewable energy productions (solar, hydraulic, wind and biogas fuel cells) enable the company to claim "100% renewable energy for 100% of these installations"[18], notably for its data centers. In Oregon, Apple bought the 45-Mile hydroelectric dam to supply its Prineville data center. Certain GAFAM companies are also increasingly envisaging developing their energy autonomy and their own infrastructures either onsite or nearby.

The autonomy aspirations of data center operators

In France, the weak presence of GAFAM limits a form of competition on the electricity market, moreover locked up by the Enedis monopoly for transmission. However, the capital resources of the major colocation data center operators permit them to invest in the construction of electricity infrastructures and dark fiber networks. They then replace, in a certain way, the historic electricity and telecom transmission operators.
The large data centers follow the geography of the transmission stations and keep a large quantity of electricity in reserve. They have three options for transmission. The first is to locate near the existing transmission stations. The second is to ask Enedis to build new ones. The third is to build them themselves. In Île-de-France, for Enedis, the data center projects that are known correspond to a doubling of the network capacity north of Paris over the next five years. For data center industrialists, there is a great deal at stake concerning transmission (that is, the moment, when the level of the voltage is lowered from 250,000 V à 20,000 V). Access to a substation is a major element of the business model of these industrialists. Historically, the construction of a substation by Enedis took five to seven years. The data center industry is trying to exert pressure to lower this period to three years and sometimes doesn't hesitate to circumvent Enedis to turn to the RTE (transmission system operator) directly, when the power of the buildings exceeds 50 MW (regulatory ceiling above which this is possible). They will therefore directly connect to a RTE high-voltage line. This is consequently a new siting criterion. This is notably the case for Interxion. Thus, Fabrice Coquio, president of the France group states: "We already began in Frankfurt where we had a 100 MW substation built and we are in the process of building one in Stockholm. I will soon file a permit in Marseille because I grabbed everything that was left: 90 MW, there isn't any more. […]. You have to be a real expert in energy management and infrastructure management to know how to work a substation and you have to have the means. For 73 MW in Frankfurt, it cost 25 million euros. That will immediately thin out the ranks of who can do what"[19]. Moreover, data centers must now pay 100% of the cost of the connection to the substation whereas it used to be Enedis who incurred this cost.

The situation is the same in the United States. The Sabey data center heavily invested in offsite infrastructure for its Intergate Manhattan site, with the construction of four electrical substations ($25 million), which ensures it a lower energy price over the long term (currently 14 cents per kW/h[20]). Sabey uses an ESCO[21] to buy energy – Constellation Energy – and only pays Con Ed for transmission. The maximum power available is 40 MW, but only 18 MW are actually consumed. The company has not positioned itself on renewable energies.
On the West Coast, the GAFAM companies also increasingly envisage developing their own energy autonomy and their own

production infrastructures or ones that are nearby via electricity micro-networks. This is the case for Microsoft, heavily committed in the development of autonomous micro-networks, and Apple, which created its subsidiary Apple Energy.
Actors like Interxion and Equinix area also creating their own telecom installation to replace operators that, according to them, no longer have the money to do so. "When you're a telecom operator, to make a hole in a sidewalk to connect to a data center, it costs €1,000 the meter. Then, to install a PoP, a point of presence, an active telecom element to deliver services, it's an investment of between €500,000 and €1,700,000. So with prices that are decreasing by 30% each year you ask yourself how many meters you have to do. If you have 30 km of networks to create, which however on the country-wide scale is nothing, you're not going to do it"[22]. Consequently, Interxion is building its considerable capacities that are sold to telecom operators – SFR and Orange, among others. There is a substitution game that is created between the historic operators and the data center operators that without having infrastructure licenses can install them since it is a matter of connecting buildings belonging to the same data center operator.

The organization of the territory and the disrupted network

Everywhere in Europe, urban planning departments and electrical transmission and transmission operators have observed that the demands of data centers can block the territories development. Electricity demand can be summed up as follows: electricity is allotted on a first come/first served basis, so queues are created on different sites. The reserve capacity (60 MW for 5 years, 10 years) blocks other customers even though this power corresponds to a maximal consumption hypothesis (when the data center is full, which sometimes takes several years).
In France, the ALEC study already raised questions on this subject in 2013. An example of this is Marseille, where the mayor Jean-Claude Gaudin had to negotiate with Interxion to recover 7 MW "because they had forgotten to reserve them for their electric buses."[23]
The colocation market disturbs readability and calculations. It must be recalled that many colocation data centers are charged on average at 30-40%, some at less than 20%, and colocation is 30% of the market. This is somewhat like a headlong rush, the operators having built data centers in anticipation of the clientele.

The territories are all the more disrupted as they now must participate in the investment in electricity infrastructures. The law on Solidarity and Urban Renewal modified investment conditions in these infrastructures, for all electricity transmission. The developer and the local authority must participate financially, whereas before, EDF strived to structure its network and reinforce it. How can data centers be pushed to play more collectively? Could the local authority, transmission system operator, even distribution system operator, in the name of the general interest, force them and impose installation sites, as is the case for the city of Stockholm? The Stockholm Data Parks Initiative requires data centers to set up on campuses where the heat produced must be reused. This program was launched in partnership with the urban cooling and heating operator Exergi, the electricity operator Ellevio and that of dark fiber Stokab. This experimentation and engagement of the city of Stockholm seems relatively unique. As David Rinard, sustainable development director of Equinix[24] recalls, whether it is in the United States or in France, any increase in energy efficiency and any new capacity in captured renewable energy makes it possible above all to consume more and at a more stable price. This is the rebound effect, otherwise called the Jevons paradox. As technological improvements increase the efficiency with which a resource is used, total consumption of this resource can increase instead of decreasing. In other words, the use of technologies that are more efficient in the use of energy and emit less CO_2 does not guarantee a drop in total energy consumption, on the contrary...

A part of the world of architecture and urban planning still tends to minimize, today, the energy and spatial impact of the digital world on cities, the territory and climate. The digital sector does not increase, it transforms. It is not an urban exoskeleton that can be worn and taken off as one wishes, but a pervasive system the is gradually modifying urban forms, using an infrastructure that increasingly mobilizes resources and space (energy production units, storage centers, underwater cables, land-based networks, but also electronic equipment production plants, digital waste disposal, etc.). If the materiality of the digital technical system and its environmental impact (on energy as well as on rare minerals) has begun to be mentioned more often, the frenzied race to achieve never-ending technological innovation continues to limit the urban digital imagination to a single discourse that makes the digital, progress, extremely sophisticated technology and 'green growth' inseparable, whereas many alternatives exist.

Support of the territories in accessing a free and open Internet is as indispensable as a broader reflection on the data center object and the associated digital system: to better measure the environmental impact of the technical choices regarding the expected social added value; and to move toward more reasoned and restrained digital practices that stress degrowth.

Bibliografia

Andrae Anders S. G., Edler T. 2015, *On Global Electricity Usage of Communication Technology: Trends to 2030*, «Challenges», *6*, pp. 117-157.

Cisco, 2016, *Network Traffic Forecast*, Report.

Diguet C., Lopez F. 2019, *From the Cloud to the Ground, L'enjeu éco-systémique des infrastructures numériques*, Report Ademe.

Dupuy G. 1991, *L'urbanisme des réseaux, théories et méthodes*, Armand Colin, Paris. [2008, *Urban Networks-Networking Urbanism*, Techne Press].

Picon A. 2013, *Smart cities: théorie et critique d'un idéal auto-réalisateur*, Paris : B2 ; 2015, *Smart Cities a Spatialised Intelligence*, Wiley, Chichester.

Ruimtelijk economische ontwikkel strategie. 2019, *Ruimtelijke Strategie Datacenters Routekaart 2030 voor de groei van datacenters in Nederland*, March 2019

Shehabi A. 2016, *United States Data Center Energy Usage Report*, Lawrence Berkeley National Laboratory.

Tinetti B. et al. 2016, *Potentiel de contribution du numérique à la réduction des impacts environnementaux: Etat des lieux et enjeux pour la prospective* – Final Report, ADEME.

Townsend A, M. 2013, *Smart Cities, Big Data, Civic Hackers, and the Quest for a New Utopia*, W. Norton & Co, New York.

Note

[1] In March 2020, Carbon Brief attempted to show that the carbon footprint of viewing videos via streaming on Netflix was as much as 57 times lower than the conclusions of the report written in 2019 by the celebrated French think tank The Shift Project.
[2] In this study, electricity consumption in 2013 was estimated at 21,000 TWh (7% represents 1,470 TWh) and the projections for 2030 reach 61,000 TWh. An annual production of 7 TWh for a nuclear reactor is the basis of these projections.
[3] A data center is a hosting building that holds a group of digital infrastructures (data computing, storage and transport equipment). It is equipped with cooling and heat recovery systems as well as backup equipment: batteries, undulators, generators.There are different types of data centers and uses can vary. There are two categories. First, that of company or administration data centers, which host and manage their own data servers in a building reserved for them. Second, that of colocation data centers, in which different uses are possible: hosting of client company digital equipment hosting (the operator provides the space and electricity); the availability of the host's IT servers and equipment for its clients (the clients can make temporary physical reservations for the server, the disk rack and network equipment in order to benefit from a guarantee and unshared use of the infrastructures); Cloud reservation (the clients can reserve virtual machines on data center servers).
[4] United States is the first country in terms of data center installations.
[5] Diguet C., Lopez F. 2019, From the Cloud to the Ground, L'enjeu éco-systémique des infrastructures numériques, Report Ademe.
[6] Microsoft's unexpected experiment in Clondalkin, near Dublin, shows that GAFAM can also take on their own energy production. Faced with the incapacity of the transmission network to incorporate the enormous rise in load of the hyperscale data centers, the company has been building, over the last three years, its own installation equipped with 16 gas-powered generators, for a total power of 18 MW.
[7] Interview with Betty Rope, mayor of Prineville, in Prineville, Oregon, October 27, 2017.
[8] Dark fiber designates an optical fiber infrastructure (cables and repeaters) installed but not yet used.
[9] Three principal reasons: land is no longer as available or inexpensive as before; the colocation sector has been consolidated (purchases of companies and existing buildings); the operators had anticipated the growth of the market by building enormous spaces, a part of them are still waiting to be filled by clients.
[10] Today, the metropolitan area would like to improve the insertion of the data center into the city (fences, entrances, planting, urban façade with windows), but the complicated and few exchanges with Interxion, which has already completed its architectural project according to its own interests, does not let any major change to the benefit of the community be envisaged.
[11] Interview with Yen Chen, urban project planner of the city Santa Clara, Santa Clara, November 2017.
[12] ICPE: Installation Classified for the Protection of the Environment
[13] Interview with the city of Amsterdam, urban planning department, Amsterdam, October 2020.
[14] A roadmap is also expected on the scale of the metropolitan region with specificities for each municipality.
[15] For the port or the industrial zones that are not in the immediate vicinity of urban transformation sectors.
[16] Until now, the city could better control the data centers located in the activity zones of the Schiphol Area Development Company, 25% of which is owned by the city. The data centers that are located in it are considered more satisfactory than those outside it. For the municipality, this explains the necessity of extending control.
[17] http://www.greenpeace.org/usa/global-warming/click-clean/
[18] https://www.apple.com/fr/environment/climate-change/
[19] Interview with Fabrice Coquio, president of the Interxion group France, Paris, March 2018.
[20] Compared, for example, with 22 cents per kW/h by 365 data centers, also located in Manhattan. Interview with Jim Grady, January 2018.
[21] ESCO: Energy service company
[22] Interview with Fabrice Coquio, op.cit.
[23] Ibidem.
[24] Interview with David Rinard, San José, California, October 2017.

Advance smart cities through Digital Twins

expanding the knowledge and management capacity of public buildings stock for energy efficiency rehabilitations

Antonella Trombadore
Gisella Calcagno
DIDA – Università degli Studi di Firenze
Giacomo Pierucci
DIEF – Università degli Studi di Firenze
antonella.trombadore@unifi.it
gisella.calcagno@unifi.it
giacomo.pierucci@unifi.it

Received: May 2020
Accepted: September 2020

DOI: 10.13128/contest-12015
www.fupress.net/index.php/contesti/

keywords
smart governance
energy efficiency
digital twin
public buildings retrofitting

In a new vision of smart cities, one of the ambitious challenges for Public Administrations is to expand the knowledge framework (smart data/smart governance) on public building stock supporting a new advanced management capacity for renovation and energy efficiency rehabilitation. The paper shares the experience of the international project Med-EcoSuRe (Mediterranean University as Catalyst for Eco-Sustainable Renovation), specially the methodology of Digital Twin experimented in the Pilot Project of Santa Verdiana, merging both the building performance with indoor environmental data and well-being characteristics. The vision of the project is very ambitious and aims to represent a comprehensive methodology and tool dedicated to PA for a sustainable and smart governance - programming, planning, managing - of energy in public buildings, supporting the evaluation of retrofitting scenarios, decision making on cost-effective and innovative energy rehabilitation actions in Med contexts.

The challenge: how to expand the knowledge and management capacity of public buildings stock

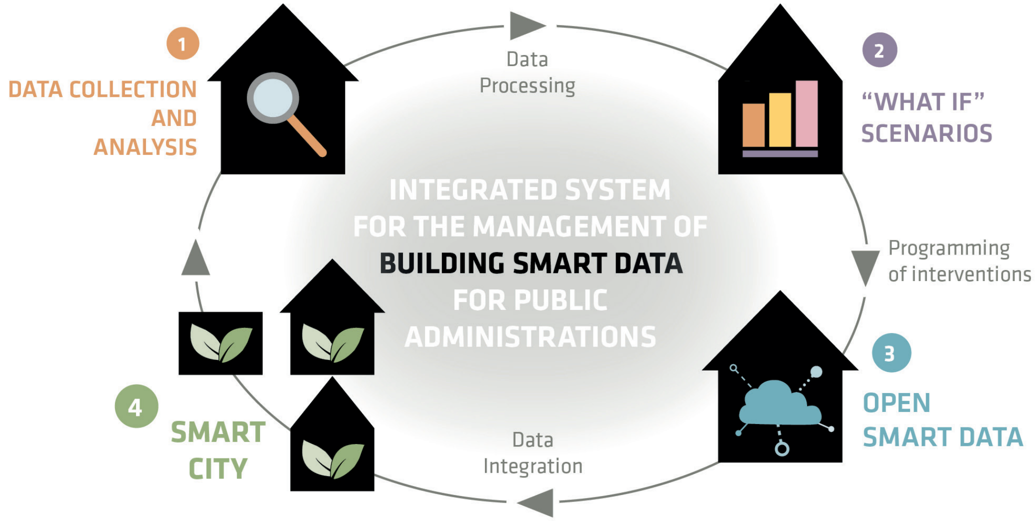
Reflecting the conditions of the surrounding built environment, public buildings in the Mediterranean area are often obsolete, also in terms of Energy Efficiency (EE): they are outdated, sometimes historical, in any case inadequate to the growing comfort demand of the occupancy, such as the cooling one, and to the sustainability targets. The result is uncomfortable public buildings using a large amount of fossil energy, with a high impact on the environment in terms of CO2 emissions. Moreover, the energy status of public buildings is often unknown, such as its potential of improvement.

The high fragmentation of public buildings data that influence EE (geometry, envelope, MEP systems, consumptions, end-users' feedbacks, etc.) is the main barrier to implement strategic EE plans and to intercept funds for large retrofitting and rehabilitations of public buildings stocks: knowledge and integrated management is indispensable for the strategic programming and planning of energy rehabilitations.

To address this challenge, many years ago an international and interdisciplinary research group, led by DIDA/ABITA researchers of the University of Florence, shared the vision that digital advancements in the building sector (such as BIM and Digital Twin) represent a unique occasion for Public Administrations to manage public buildings stocks, driving towards an overall innovative and open approach to EE, based on institutional capacity building, knowledge&skills and pro-active participation of all the stakeholders.

The first idea was to design a smart platform for building data collection and management [I_Building Platform], creating a 'building identity card' where all the information and data of existing public buildings (the project focused on schools) are collected according to the same criteria and validation process and managed in a more integrated, collaborative and transparent way, serving more efficient processes. The platform becomes a tool to manage the knowledge framework and could work as decision support system, very useful to guide the process of enhancing the existing building stock, aiming at a more efficient data management, with a logic of access and sharing of 'open data' on a user-friendly platform (Fig. 1).

The Med-EcoSuRe project (financed by ENI CBCMED Programme) offers the opportunity to implement and investigate this challenge in the framework of Mediterranean context, experimenting the methodologies of the Digital Twin in the local pilot. The project is setting collaborative learning schemes to foster scientific progress on innovative energy renovation solutions within the university's immediate neighbourhood, which is the university building. The main concept behind the project is that working



with stakeholders can produce more effective innovative solutions: the idea is to stimulate participatory processes, supporting university building manager with predictive tools and enhancing their capacity to plan and implement sustainable energy mix strategies and technologies for the Mediterranean climatic and social contexts. Based on the Living Lab approach, the purpose is to create a collaborative platform for research, development, and experimentation of product and service innovations in real-life contexts, based on specific methodologies and tools, and implemented through concrete innovation projects and community-building activities. Pilot buildings in different Mediterranean countries are becoming Local Living Labs, connected by the overall Cross-border Living Lab.
In charge to manage the Med-EcoSuRe Cross-border Living Lab, and within the Space BEXLab[1], the research proposes, through its outputs, significant and vibrant changes in the MED region:

1. optimize/innovate the knowledge framework and management of the public buildings stocks, as basis for cost-effective energy mix plans, with toolkit for BIM/Digital Twin to support the digital transformation of the building sector and Public Administrations;
2. reinforce the cooperation among stakeholders (universities/research centres, companies, public authorities, end-users) both at local and Cross border level, opening EE processes in Living Labs for more inclusive and participated processes;
3. promote awareness and a more pro-active social behaviour, thanks to real data interaction on EE in pilots/Living Labs and experimenting the best comfort experiences.

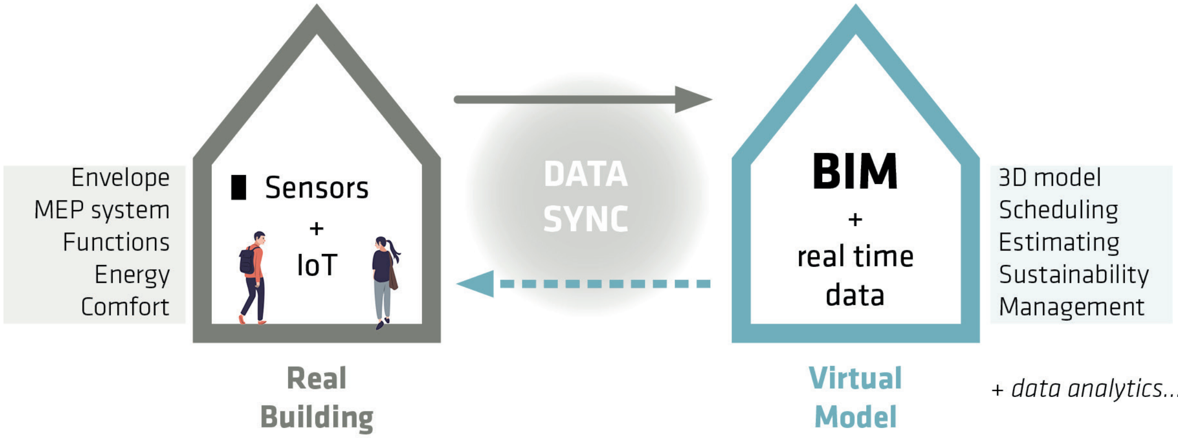
Vision for a smart management of public buildings stocks
Fig. 1

Energy efficiency through digitalisation: the renovation of public buildings
Within the global race to reduce the effects of climate change, the European Union is operating for a transition towards a clean energy society: for the next ten years, the ambitious target is to reach at least 32,5% of energy efficiency, 40% cuts in greenhouse gas emissions, and 32% renewables in energy consumption (European Commission, 2019a).
Improving the energy performance of the construction sector is crucial, since buildings are the larger energy consumers, accounting for approximately 40% of the energy consumption and 36% of carbon dioxide emission, performances that can be explained by the fact that three quarters of the EU's building stock were built when EE requirements were limited or non-existent and the rate of renovation is just of 0,4-1,2% per year (Agora Energiewende, 2019).
A relevant role in this challenging process is played by the public sector: even if publicly owned or occupied buildings occupy about the 10-12% by area of the EU building stock, they have a strategic role as best practice for the overall sector, where innovation can be collectively tested and proposed to the wider public. Public bodies should set the example and give impulse for the creation of the necessary know-how for the rest of the country, with the implementation of public nZEB (nearly Zero Energy Building) both in the case of new buildings and for the renovation of existing ones (EUROSAI, 2018).
Both the two main directives driving the building sector's EE transition, the Energy Efficiency Directive (2012/27/EU - EED) and the Energy Performance of Buildings Directive (2010/31/EU - EPBD), consider the leading role of public buildings. Article 5 of the EED sets binding renovation targets for public buildings[2] and stresses that governments shall undertake an exemplary role in the energy retrofit of their countries' building stock. In the same direction the EPBD, which fixes minimum energy performance requirements for all buildings, considers additional conditions for the public sector: mandatory energy performance certification and public display of certificates, as well as an earlier date at which all new buildings owned and occupied by public authorities should be nZEB.
To reach the ambitious clean energy objectives, the amended EPBD (2018/844/EU) clearly indicates to turn on the renovations of existing building, to be transformed in nZEB by 2050, and on digital technologies, promoting smart systems and digital solutions as a means to achieve energy savings in a cost-efficient manner[3].
Digitalization is one of the EU priorities[4], but even if important progress has been made in digitalising Europe, still a lot needs to be done to ensure the EU industry will fully seize the

digitalisation opportunities. This evolution is globally relevant for the construction sector, second less digitalized sector after agriculture, which is largely dominated by Small and Medium Enterprises (SMEs), characterised by decreasing innovation, low rates of technological adoption and diminishing efficiency (European Commission, 2019b).
From a policy perspective, the EU Directive on public procurement (2014/24/EU) is the only policy instrument that clearly refers to the use of digitalisation technologies in the construction sector, in particular to BIM (Building Information Modelling). The Article 22 of the Directive stipulates that "for public works and design contests, Member States may require the use of specific electronic tools, such as of building information electronic modelling tools or similar". The transposition of the EU Directive at national level varies greatly across Europe: eight Member States (Austria, Denmark, Finland, Germany, Italy, Luxembourg, Spain and the UK[5]) have made the use of BIM mandatory in the procurement of public works.
At international level, where the use of this digital technology is accelerating in the global construction market, and in particular in the North America, BIM is defined as "shared digital representation of physical and functional characteristics of any built object (including buildings, bridges, roads, etc.) which forms a reliable basis for decisions" (ISO 29481-1: 2010).
According to the European Commission (2019b), and beyond BIM, the digitalisation of the construction sector encompasses at least six digitalisation technologies, which should be integrated in BIM: additive manufacturing, robotization, drones, 3D scanning, sensors and IoT (Internet of Things). Yet, the diffusion of BIM is very thin because several gaps in the construction sector are still in place, such as the limited obligatory digitalisation targets at national level, the lack of investment for SMEs, of trained employees, of investment in research and development and, more importantly, a lack of a holistic vision for the research needs, not only related to BIM but to the full set of digital technologies and potentials (ibid.).
As for EE, the public sector plays a leading role also in the digitalisation of the construction/building sector, as sustained by the EUBIM Task Group. According to them, government policy and public procurement methods are powerful tools to support the positive change in the building sector. As non-competitive, transparent and non-discriminating, the public sector can encourage to fully grasp the digital opportunities, providing better public services, with an increased transparency of building performances, and in turn better value for public money and better quality of the built environment (EUBIM, 2017).

The Digital Twin

To face the ambitious clean energy objectives related to the built environment, it is even more recognised as strategic to take full advantage of the several technologies spreading in the digital domain, overcoming the single BIM digital technology. Nearby BIM in fact, digital technologies are advancing at an ever-increasing pace, exploiting the even more smaller, cheaper and integrated sensors, the Internet of Things (IoT) and more experimentally, Artificial Intelligence (AI) and its recent advancements (big data analysis, semantic web, machine learning, deep learning) (Boje et al., 2020). Sustained by adjacent digital technologies (i.e. sensors and IoT), a new concept of Industry 4.0. is circulating in the building sector, born in the field of product lifecycle management, and firstly expressed by Micheal Grieves in 2003: the Digital Twin (DT). Grieves (2014) describes the DT as consisting of "three main parts: a) physical products in Real Space, b) virtual products in Virtual Space, and c) the connections of data and information that ties the virtual and real products together" (p. 1). The idea is to associate a digital representation to a physical product and to create a bi-directional data connection between the physical and the virtual, in order to improve the physical twin.
Expanding the DT concept to the physical world, this means real entities, such as buildings, with data collected through sensors and, in the digital side, models and tools where data and information can be analysed and processed, to be used to improve the physical assets. It is easy to capture the potential of such twinning (which is the act of synchronising): exploiting digital technologies, DTs can optimise the understanding of buildings' functioning as basis for decision-making. The interaction of the built environment with semantic 3D models, enriched with real-time data fed by sensors, provides the opportunity for a real-time monitoring and data acquisition, processing and re-introduction of data in the buildings' lifecycle. According to Khajavi et al. (2019), the DT of a building can be used for predictive maintenance, resource efficiency improvement, enhancement of tenants' comfort, what-if analysis for optimization of the building design and enabling closed-loop design to transfer learnings from a building to the future ones.
Due to this functionality, DTs are strategic for the challenges related to the energy renovation of existing buildings, with the possibility to capture and elaborate real-time data on EE, also related to the human experience (through IoT), i.e. to predict and simulate the most cost-effective building interventions. In this field, BIM has already evolved in BEM (Building Energy Modelling) to simulate energy performance, evaluate



energy needs and optimize buildings' physical and technical assets, but principally referring to the design stage. The DT possibility to manage energy data in the operation stage and to share them with all relevant stakeholders (from energy agencies to citizens) represents an opportunity for a better EE management. The structured and functional amount of data on buildings can overcome the traditional approach to renovation design, opening towards an evidence-based and data-driven decision making across all the building's lifecycle (and in particular in the use phase), supporting more collaborative schemes.
Yet, if BIM is seen as the starting point for the DT, BIM development towards DT (by the addition of sensing capabilities, big data and the Internet of Things from site to building operation) is still very low, as noticed before, due to the delay in the BIM general adoption. According to the DT maturity spectrum developed by IET et al. (2019) in the evolution of BIM towards DT, going further the nD BIM models, it is time to handle real-time data from sensors and to integrate IoT, in order to improve the operational efficiency also in relation to human behaviours (Fig. 2).

Experimenting the Digital Twin in the Med-EcoSure project

As an experimentation within the Med-EcoSure project, the Italian team is developing a Digital Twin of the pilot university building under renovation, tackling the open digital challenges in the construction and building sectors.
The selected pilot building/Living Lab is located in the Santa Verdiana building complex, School of Architecture of the University of Florence, which sits in the UNESCO city-centre (*quartiere* of Santa Croce). Transformed into educational centre in 1986, Santa Verdiana was a former convent (1395) and female prison (1865), which required over the years a series of radical restoration and adaptation interventions to fit the new use. In particular, the pilot is a more recent building block added in the north side of the historical cloister, study room in

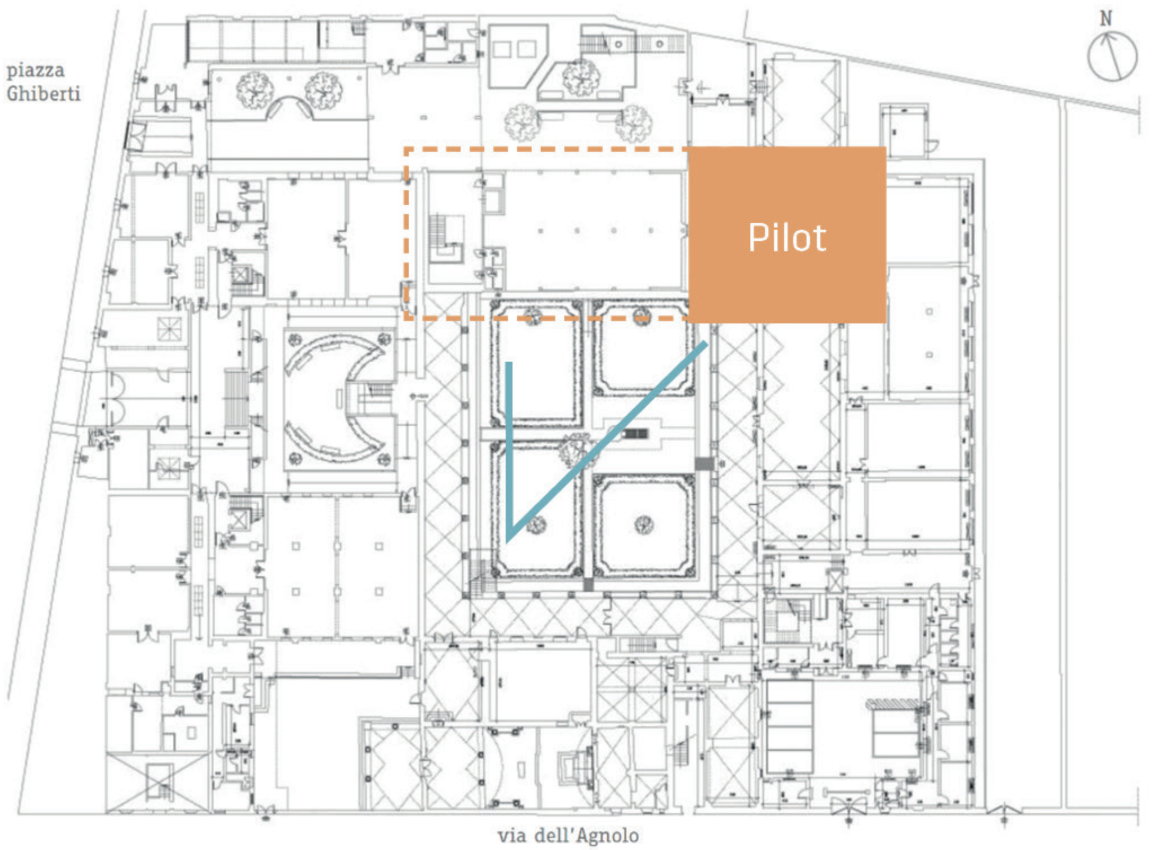
Functioning of the Digital Twin for buildings
Fig. 2

the ground floor and classroom/laboratory in the first (Fig. 3).
Starting from the survey data, collected by the Space BExLab, a 3D BIM model has been developed in order to fulfil the BIM uses for 4D (time), 5D (costs), 6D (facility management), 7D (sustainability) and for communication and interaction with the stakeholders (academics, Public administration entities and technicians, professionals, economic operators and end-users) along the development of innovative strategies for the building energy renovation. The survey information loaded into the model contains, apart from the geometrical and physical information(materials), the energy performance information collected on the energy audit. Both geometrical and information level of development/detail is LOD C (Level of Detail corresponding to 'definite object', according to the UNI 11337-4), level of development needed for the information transfer for the energy analysis and energy simulations as BEM (Building Energy Model). For this scope, and thanks to interoperability (gbxml format), simulations in dynamic regime will be carried out using dedicated software (e.g. TRNSYS, Design Builder and EnergyPlus).
Based on the Living Lab principles, the team is introducing a collaborative approach which exploits the potentialities of BIM/DT for the decision making along the design, construction and operation phases, since the pre-measurement of the intervention (e.g. data on envelope and technical system, monitoring, sensors, IoT, end-users feedbacks on behaviour and comfort, etc.). Supported by the DT, the objectives of the local pilot renovation are:

- to assess that the building operates according to quality standards on sustainability (e.g. energy performances) and comfort (e.g. indoor quality);
- To highlight the criticalities and gaps and planning to modify the operational systems (physical and technical assets, but also behaviours) in order to improve their performances;
- To create n. 'what if' scenarios and test strategies and materials to evaluate how the different economic, environmental and social conditions influence the energy performance of the public building, in order to reduce its negative impact.

Monitoring indoor environmental quality and well-being characteristics

At this stage of the project, the team is working on the monitoring set-up of the pilot building, process that serves both the pre and post renovation of the site for the refurbishment and use phase.
A complex measurement system will be developed to create an information map of 2 rooms used for different scopes at the



First floor plan of Santa Verdiana and the pilot building
Fig. 3

School of Architecture. The collected data will describe the thermo-physical, the air quality and other well-being characteristics that change as a function of time and boundary conditions.
Some sensors of the same type will be installed in different spots, covering a uniform grid with the objective to detect the spatial distribution of monitored parameters. The rooms were properly selected as a significant sample of 'common' academic living place relating to their orientation, urban context, building manufacture, usage. The first measurement layout is meant to be redundant, at least, because physical phenomena need to be verified from different points of view since no background data are available. Furthermore, the sensors' correct installation has to be confirmed with multiple crossed measures.
The first year of data monitoring is fundamental for understanding the site behaviour and the response of sensors' set-up. In this context the measurement layout could be also optimized suggesting the minimum number of devices which is useful to characterize the environments in an exhaustive way.
It is important to mention again that one of the aims of project is defining procedures and methods for the Digital Twin process (experimental part) which would be repeatable in a very large range of different possible applications (public and private building stock). The measurement set-up should be adaptive, modular and plug&play to facilitate the research without compromising accuracy. This led to consider the overall signals chain starting from the sensitive elements, that must be non-intrusive and easy to connect and supply, up to the data concentrators and transmitters. Even the communications protocols are chosen with the same principles taking into account the need of cables and the suitable wireless publishing tools. At the moment, commercial solutions are preferred for the ease and speed in installation, but in parallel, the overall control and data management logics, which are going to be implemented in detail, will be also approached towards open source programming languages. In the next future, another room at the same boundary conditions could be equipped with a new full customised measurement system aiming at a comparison with the commercial one in terms of capability, flexibility and costs.
As well known, the intended use and the level of occupancy play a key role in the energy performance and well-being of a building.

From one side, human behaviour modifies the operative conditions strongly (for instance the temperature through the heating plant); from the other side the occupants experience the continuous variation of the internal parameters in terms of comfort or discomfort (illuminance, air exchange, etc..).
It is important to underline that people, in general, are rarely informed about the connection of those concepts at an upper level beyond temporary feelings. No complete tools are available for the managers that handle buildings, with the lack in defining strategies of intervention wherever is needed and plans for long term administration and maintenance.
A revolution in terms of awareness is necessary with the development of decision methods, primarily derived by quantitative parameters extrapolated on field, and then used for validating predictive models. The living laboratory makes a complex measurements database available to the researchers, which represents the springboard

Implementation of the pilot Digital Twin
Fig. 4

for the design of post-processing procedures. They will let to convert the acquired values from sensors into useful information at different levels of aggregation and synthesis. On a microscopic point of view, the analysis of the signals helps in understanding the details of actions-reaction relationships between the building system and the external forcing (weather conditions as well as occupants). As a following step, an overall macroscopic view is needed to summarize results for a clear and effective dissemination up to the draft of guidelines for a smart management and usage.

Open challenges and future works: the strategic vision of augmented cities

The Pilot Project of Digital Twin/Living Lab will be helpful to demonstrate its scalability to the entire heritage building stock, providing Public Administrations with a powerful method for the energy rehabilitation of public buildings to be supported by Digital Twin, with monitoring data on indoor quality, real performance of building envelope as well as the contribution of end user behaviour.
This will also make it possible to outline a procedural model that can be extended on a large scale and that can be used throughout the territory (bottom up) and thus be able to design an intelligent urban network; it will be possible to create a large-scale global database/dashboard on which information regarding buildings (public/private), as well as the services connected to them, can be stored and analysed. Such possibility opens towards a qualitatively better and more efficient government of the territory, from the scale of the single building to the entire city. The development of buildings' DTs, and the flux of data they bring, can intercept, and be intercepted by, the smart city's objectives (e.g. data for smart grids) to address the SDGs on affordable, reliable, sustainable and modern energy for all and on make cities and human settlements inclusive, safe, resilient and sustainable.
The implementation of DTs in fact allows the integration of real-time data on buildings in even more hyper-connected digital systems, which can support a better management of the economic, environmental and social resources interacting in the built environment.
The possibility to manage even more high quality and integrated data represents an opportunity for public sector to be more socially responsible for the global climate change challenges: firstly, the availability of integrated energy data can to enhance the capacity of PA to plan more innovative and performative buildings, such as NZEB, leading to the best cost-effective renovations (e.g. scenario simulations); secondly, managing several DTs consent to prioritize EE interventions in public buildings stocks

(e.g. schools), neighbourhood or even at city scale; finally, the human interaction on data (i.e. living labs, IoT, apps or simple monitors) is strategical to educate citizens/end users to a more conscious behaviour in relation to energy savings (understanding needs and benefits) and environmental protection. Nevertheless, as mentioned above, DT is not a static model based on a single digital technology, but rather a responsive system made of a "constellation, or ecosystem of technologies that work and connect" (IET, ATKINS, 2019), which need to be systematised and transferred to the large use, starting from the public one: it is emerging the need of methodologies and user-friendly tools for the management of DTs.
Looking more forward, DTs have the potential to evolve into autonomous systems with less human intervention, through AI-enabled design and control. Through data and feedbacks, both simulated and real, a DT can develop capacities for autonomy and to learn from and reason about its environment: in this perspective, "a digital twin will not only lead to better decision making, but it will make better decisions" (ARUP, 2019, p. 16).

Acknowledgementct

Med Eco-SuRe Project Partner: Department of Architecture (DIDA), University of Florence. Team: Saverio Mecca, Antonella Trombadore, Carla Balocco, Maurizio De Lucia, Gisella Calcagno, Giacomo Pierucci, Juan Camilo Olano Salinas, Marco Paolini.

Space BExLab/DIDALABS, University of Florence: Saverio Mecca, Antonella Trombadore, Carla Balocco, Maurizio De Lucia, Roberto Bologna, Elisabetta Cianfanelli, Debora Giorgi, Gisella Calcagno, Giacomo Pierucci, Juan Camilo Olano Salinas, Marco Paolini.

Note

[1] The Space BExLab (Real and Virtual Space for Building Environmental Experience) has been established as a laboratory of the Department of Architecture (DIDALABS) of the University of Florence with the aim to experiment with augmented Digital Twins for interdisciplinary research and didactics on predictive planning and design.

[2] The article states that Member States shall ensure, since January 2014, that the 3% of the total floor area of heated and/or cooled buildings owned and occupied by its central government is renovated each year to meet at least the minimum energy performance requirement, or alternative measures (such as behavioural change of occupants of public buildings or installation of energy management systems).

[3] The amended Directive also refers to "smart readiness indicators" to measure the capacity of buildings to use information and communication technologies and electronic systems to adapt the needs of the occupants and the grid and to improve the energy efficiency and overall performance of buildings.

[4] The 2015 "Strategy for a Digital Single Market" outlined the path for the EU to build the right digital environment.

[5] Due to the recent Brexit, UK is not more an EU State Member, but it is still a reference for the BIM national adoption.

The paper is written jointly by Antonella Trombadore, Gisella Calcagno, Giacomo Pierucci. Referring to the individual chapters, they have been written as follows:
Trombadore A.: 1°, 4°
Calcagno G.: 2°, 3°
Pierucci G.: 5°
Trombadore A. & Calcagno G.: 6°

Bibliografia

Agora Energiewende 2019, *European Energy Transition 2030: The Big Picture. Ten Priorities for the next European Commission to meet the EU's 2030 targets and accelerate towards 2050*, < https://www.agora-energiewende.de/fileadmin2/Projekte/2019/EU_Big_Picture/153_EU-Big-Pic_WEB.pdf> (08/20).

ARUP 2019, *Digital Twin. Towards a meaningful framework*, <www.arup.com/digitaltwinreport> (08/20).

Boje C., Guerriero A., Kubicki S., Rezgui Y. 2020, *Towards a semantic Construction Digital Twin: Directions for future research*, «Automation in Construction», vol. 114, 103179, pp. 1-16.

EU BIM Task Group 2017, *Handbook for the introduction of Building Information Modelling by the European Public Sector. Strategic action for construction sector performance: driving value, innovation and growth,* < http://www.eubim.eu/downloads/EU_BIM_Task_Group_Handbook_FINAL.PDF > (08/20).

European Commission 2019a, *Clean Energy for all Europeans,* < https://ec.europa.eu/energy/topics/energy-strategy/clean-energy-all-europeans_en> (08/20).

European Commission 2019b, *Supporting digitalisation of the construction sector and SMEs. Including Building Information Modelling,* < https://ec.europa.eu/growth/content/report-supporting-digitalisation-construction-sector-and-smes_en> (08/20).

EUROSAI 2018, *Energy Efficiency of Public Sector Buildings,* <https://www.eurosai.org/handle404?exporturi=/export/sites/eurosai/.content/documents/Energy-Efficiency-of-Public-Sector-Buildings.pdf> (08/20).

Grieves M. 2014, *Digital twin: manufacturing excellence through virtual factory replication*, <https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication> (08/20).

IET (The Institution of Engineering and Technology), ATKINS (Member of the SNC-Lavalin Group) 2019, *Digital twins for the built environment. An introduction to the opportunities, benefits, challenges and risks,* < https://www.theiet.org/media/4715/leaflet-digital-twins-for-the-built-environment-iet-atkins.pdf> (08/20).

Khajavi S., Motlagh N., Jaribion A., Werner L., Holmstrom J. 2019, *Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings*, «IEEE Access», vol. 7, pp. 147406-147419.

Trombadore A. 2015, *Mediterranean Smart Cities. Innovazione tecnologica ed ecoefficienza nella gestione dei processi di trasformazione urbana*, Altralinea Edizioni s.r.l., Firenze.

letture
readings

The need for rationality

Christopher Alexander

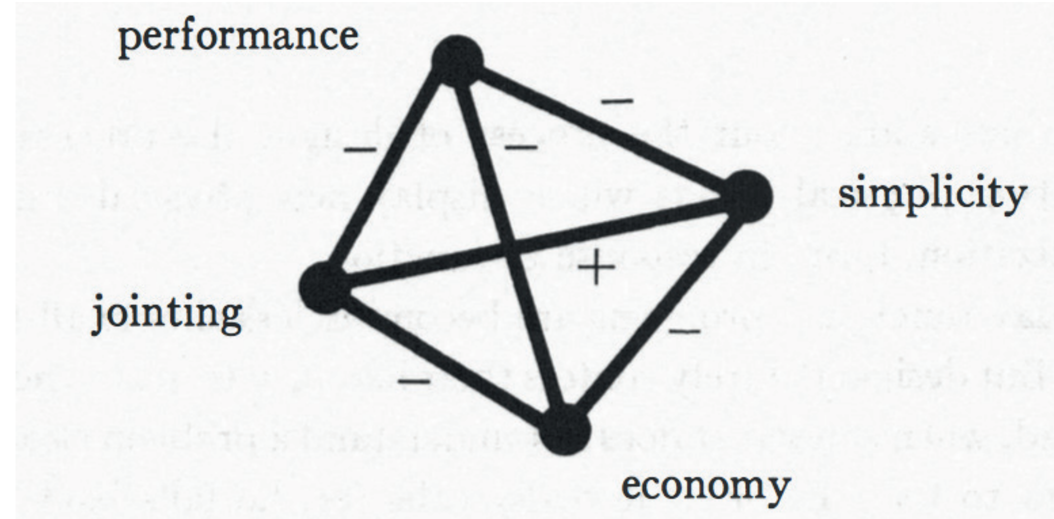
Introduzione al volume *Notes on the synthesis of form* (1964), pubblicato da Harvard University Press, Cambridge, Massachusetts, edizione 1973, pp. 1-11.

DOI: 10.13128/contest-12100
www.fupress.net/index.php/contesti/

These notes are about the process of design; the process of inventing physical things which display new physical order, organization, form, in response to function. Today functional problems are becoming less simple all the time. But designers rarely confess their inability to solve them. Instead, when a designer does not understand a problem clearly enough to find the order it really calls for, he falls back on some arbitrarily chosen formal order. The problem, because of its complexity, remains unsolved.

Contesti dedicates the Readings section of this issue to the introduction of Christopher Alexander's book Notes on the synthesis of form published in English by Harvard University Press in 1964 and translated into Italian in 1967 and included in the Saggiatore's series entitled Urban structure and form founded in that same year by Giancarlo De Carlo. In Notes on the synthesis of form Alexander pursues the ambitious task of proposing an alternative design method, not linked to individual intuition, a position heavily criticized in the introduction, but based on deductive reflection assisted by the use of the electronic calculator.

Consider a simple example of a design problem, the choice of the materials to be used in the mass production of any simple household object like a vacuum cleaner. Time and motion studies show that the fewer different kinds of materials there are, the more efficient factory assembly is - and therefore demand a certain simplicity in the variety of materials used. This need for simplicity conflicts with the fact that the form will function better if we choose the best material for each separate purpose separately. But then, on the other hand, functional diversity of materials makes for expensive and complicated joints between components, which is liable to make maintenance less easy. Further still, all three issues, simplicity, performance, and jointing, are at odds with our desire to minimize the cost of the materials. For if we choose the cheapest material for each separate task, we shall not necessarily have simplicity, nor optimum performance, nor materials which can be cleanly jointed. Writing a minus sign beside a line for conflict, and a plus beside a line for positive agreement, we see that even this simple problem has the five way conflict pictured below.



This is a typical design problem; it has requirements which have to be met; and there are interactions between the requirements, which makes the requirements hard to meet. This problem is simple to solve. It falls easily within the compass of a single man's intuition. But what about a more complicated problem?

Consider the task of designing a complete environment for a million people. The ecological balance of human and animal and plant life must be correctly adjusted both internally and to the given exterior physical conditions. People must be able to lead the individual lives they wish for. The social conditions induced must not lead to gross ill-health or to gross personal misery, and must not cause criminal delinquency. The cyclical intake of food and goods must not interfere with the regular movements of the inhabitants. The economic forces which develop must not lead to real-estate speculation which destroys the functional relation between residential areas and areas supporting heavy goods. The transportation system must not be organized so that it creates a demand that aggravates its own congestion. People must somehow be able to live in dose cooperation and yet pursue the most enormous variety of interests. The physical layout must be compatible with foreseeable future regional developments.

The conflict between population growth and diminishing water resources, energy resources, parklands, must somehow be taken care of. The environment must be organized so that its own regeneration and reconstruction does not constantly disrupt its performance.

As in the simpler example, each of these issues interacts with several of the others. But in this case each issue is itself a vast problem; and the pattern of interactions is vastly complicated. The difference between these two cases is really like the difference between the problem of adding two and two, and the problem of calculating the seventh root of a fifty digit number. In the first case we can quite easily do it in our heads. In the second case, the complexity of the problem will defeat us unless we find a simple way of writing it down, which lets us break it into smaller problems.

Today more and more design problems are reaching insoluble levels of complexity. This is true not only of moon bases, factories, and radio receivers, whose complexity is internal, but even of villages and teakettles. In spite of their superficial simplicity, even these problems have a background of needs and activities which is becoming too complex to grasp intuitively.

To match the growing complexity of problems, there is a growing body of information and specialist experience. This information is hard to handle; it is widespread, diffuse, unorganized[1]. Moreover, not only is the quantity of information itself by now beyond the reach of single designers, but the various specialists who retail it are narrow and unfamiliar with the form-makers' peculiar problems, so that it is never clear quite how the designer should best consult them[2]. As a result, although ideally a form should reflect all the known facts relevant to its design, in-fact the average designer scans whatever information he happens on, consults a consultant now and then when faced by extra-special difficulties, and introduces this randomly selected information into forms otherwise dreamt up in the artist's studio of his mind. The technical difficulties of grasping all the information needed for the construction of such a form are out of hand - and well beyond the fingers of a single individual[3].

At the same time that the problems increase in quantity, complexity, and difficulty, they also change faster than before. New materials are developed all the time, social patterns alter quickly, the culture itself is changing faster than it has ever changed before. In the past - even after the intellectual upheaval of the Renaissance - the individual designer would stand to *some* extent upon the shoulders of his predecessors. And although he was expected to make more and more of his own decisions as traditions gradually dissolved, there was always still some body of tradition which made his decisions easier. Now the last shreds of tradition are being torn from him. Since cultural pressures change so fast, any slow development of form becomes impossible. Bewildered, the formmaker stands alone. He has to make clearly conceived forms without the possibility of trial and error over time. He has to be encouraged now to think his task through from the beginning, and to 'create' the form he is concerned with, for what once took many generations of gradual development is now attempted by a single individual[4]. But the burden of a thousand years falls heavily on one man's shoulders, and this burden has not yet materially been lightened. The intuitive resolution of contemporary design problems simply lies be yond a single individual's integrative grasp.

Of course there are no definite limits to this grasp (especially in view of the rare cases where an exceptional talent breaks all bounds). But if we look at the lack of organization and lack of clarity of the forms around us, it is plain that their design has often taxed their designer's cognitive capacity well beyond the limit. The idea that the capacity of man's invention is limited is not so surprising, after all. In other areas it has been shown, and we admit readily enough, that there are bounds to man's cognitive and creative capacity. There are limits to the difficulty of a laboratory problem which he can solve[5]; to the number of issues he can consider simultaneously[6]; to the complexity of a decision he can handle wisely[7]. There are no absolute limits in any of these cases (or usually even any scale on which such limits could be specified); yet in practice it is clear that there are limits of some sort. Similarly, the very frequent failure of individual designers to produce well organized forms suggests strongly that there are limits to the individual designer's capacity.

We know that there are similar limits to an individual's capacity for mental arithmetic. To solve a sticky arithmetical problem, we need a way of setting out the problem which makes it perspicuous. Ordinary arithmetic convention gives us such a way. Two minutes with a pencil on the back of an envelope lets us solve problems which we could not do in our heads if we tried for a hundred years. But at present we have no corresponding way of simplifying design problems for ourselves. These notes describe a way of representing design problems which does make them easier to solve. It is a way of reducing the gap between the designer's small capacity and the great size of his task.

Part One contains a general account of the nature of design problems. It describes the way such problems have been solved in the past: first, in cultures where new problems are so rare that there are no actual designers; and then, by contrast, in cultures where new problems occur all the time, so that they have to be solved consciously by designers. From the contrast between the two, we shall learn how to represent a design problem so that it can be solved. Part Two describes the representation itself, and the kind of analysis the representation allows. Appendix 1 shows by example how the method works in practice.

The analysis of design problems is by no means obviously possible. There is a good deal of superstition among designers as to the deathly effect of analysis on their intuitions -with the unfortunate result that very few designers have tried to understand the process of design analytically. So that we get off to a fair start, let us try first to lay the ghosts which beset designers and make them believe that analysis is somehow at odds with the real problem of design.

It is not hard to see why the introduction of mathematics into design is likely to make designers nervous. Mathematics, in the popular view, deals with magnitude. Designers recognize, correctly, that calculations of magnitude only have strictly limited usefulness in the invention of form, and are therefore naturally rather sceptical about the possibility of basing design on mathematical methods[8]. What they do not realize, however, is that modem mathematics deals at least as much with questions of order and relation as with questions of magnitude. And though even this kind of mathematics may be a poor tool if used to prescribe the physical nature of forms, it can become a very powerful tool indeed if it is used to explore the conceptual order and pattern which a problem presents to its designer.

Logic, like mathematics, is regarded by many designers with suspicion. Much of it is based on various superstitions about the kind of force logic has in telling us what to do. First of all, the word 'logic' has some currency among designers as a reference to a particularly unpleasing and functionally unprofitable kind of formalism[9]. The so-called logic of Jacques François Blondel or Vignola, for instance, referred to rules according to which the elements of architectural style could be combined[10]. As rules they may be logical. But this gives them no special force unless there is also a legitimate relation between the system of logic and the needs and forces we accept in the real world. Again, the cold visual 'logic' of the steel-skeleton office building seems horribly constrained, and if we take it seriously as an intimation of what logic is likely to do, it is certain to frighten us away from analytical methods[11]. But no one shape can any more be a consequence of the use of logic than any other, and it is nonsense to blame rigid physical form on the rigidity of logic. It is not possible to set up premises, trace through a series of deductions, and arrive at a form which is logically determined by the premises, unless the premises already have the seeds of a particular plastic emphasis built into them. There is no legitimate sense in which deductive logic can prescribe physical form for us.

But, in speaking of logic, we do not need to be concerned with processes of inference at all. While it is true that a great deal of what is generally understood to be logic is concerned with deduction, logic, in the widest sense, refers to something far more general. It is concerned with the form of abstract structures, and is involved the moment we make pictures of reality and then seek to manipulate these pictures so that we may look further into the reality itself. It is the business of logic to invent purely artificial structures of elements and relations. Sometimes one of these structures is dose enough to a real situation to be allowed to represent it. And then, because the logic is so tightly drawn, we gain insight into the reality which was previously withheld from us[12].

The use of logical structures to represent design problems has an important consequence. It brings with it the loss of innocence. A logical picture is easier to criticize than a vague picture since the assumptions it is based on are brought out into the open. Its increased precision gives us the chance to sharpen our conception of what the design process involves. But once what we do intuitively can be described and compared with non-intuitive ways of doing the same things, we cannot go on accepting the intuitive method innocently. Whether we decide to stand for or against pure intuition as a method, we must do so for reasons which can be discussed. I wish to state my belief in this loss of innocence very clearly, because there are many designers who are apparently not willing to accept the loss. They insist that design must be a purely intuitive process: that it is hopeless to try and understand it sensibly because its problems are too deep.

There has already been one loss of innocence in the recent history of design; the discovery of machine tools to replace hand craftsmen. A century ago William Morris, the first man to see that the machines were being misused, also retreated from the loss of innocence. Instead of accepting the machine and trying to understand its implications for design, he went back to making exquisite handmade goods[13]. It was not until Gropius started his Bauhaus that designers carne to terms with the machine and the loss of innocence which it entailed[14]. Now we are at a second watershed. This time the loss of innocence is intellectual rather than mechanical. But again there are people who are trying to pretend that it has not taken place. Enormous resistance to the idea of systematic processes of design is coming from people who recognize correctly the importance of intuition, but then make a fetish of it which excludes the possibility of asking reasonable questions.

It is perhaps worth remembering that the loss of intellectual innocence was put off once before. In the eighteenth century already, certain men, Carlo Lodoli and Francesco Algarotti in Italy and the Abbé Laugier in France, no longer content to accept the formalism of the academies, began to have serious doubts about what they were doing, and raised questions of just the sort that have led, a hundred and fifty years later, to the modem revolutionary ideas on form[15]. Oddly enough, however, though these serious doubts were clearly expressed and widely read, architecture did not develop from them in the direction indicated. The doubts and questions were forgotten. Instead, in late eighteenth century Europe, we find evidence of quite another atmosphere developing, in which architects based their formal invention on the rules provided by a variety of manners and 'styles' like neo-Tudor, neoclassicism, chinoiserie, and neo-Gothic[16].

It is possible to see in this course of events a desperate attempt to ward off the insecurity of self-consciousness, and to maintain the security of innocence.

Lodoli and Laugier wanted to know what they were doing as makers of form. But the search for this knowledge only made the difficulty of their questions clear. Rather than face the responsibility of these difficult questions, designers turned instead to the authority of resurrected 'styles'. The architectural decisions made within a style are safe from the nagging difficulty of doubt, for the same reason that decisions are easier to make under tradition and taboo than on one's own responsibility. It is no coincidence, in my opinion, that while the Renaissance had allowed free recombinations of classical elements, the neoclassicism which replaced it stuck as closely as it could to the precise detail of Greece and Rome. By leaning on correctness, it was possible to alleviate the burden of decision. To make the secession from responsibility effective, the copy had to be exact[17].

Now it looks as though a second secession from responsibility is taking place. It is not possible today to escape the responsibility of considered action by working within academic styles. But the designer who is unequal to his task, and unwilling to face the difficulty, preserves his innocence in other ways. The modern designer relies more and more on his position as an 'artist', on catchwords, personal idiom, and intuition - for all these relieve him of some of the burden of decision, and make his cognitive problems manageable. Driven on his own resources, unable to cope with the complicated information he is supposed to organize, he hides his incompetence in a frenzy of artistic individuality. As his capacity to invent clearly conceived, well-fitting forms is exhausted further, the emphasis on intuition and individuality only grows wilder[18].

In this atmosphere the designer's greatest gift, his intuitive ability to organize physical form, is being reduced to nothing by the size of the tasks in front of him, and mocked by the efforts of the 'artists'. What is worse, in an era that badly needs designers with a synthetic grasp of the organization of the physical world, the real work has to be clone by less gif ted engineers, because the designers hide their gift in irresponsible pretension to genius.

We must face the fact that we are on the brink of times when man may be able to magnify his intellectual and inventive capability, just as in the nineteenth century he used machines to magnify his physical capacity[19]. Again, as then, our innocence is lost. And again, of course, the innocence, once lost, cannot be regained. The loss demands attention, not denial.

Note

1 Bullivant D.,"Information for the Architect" in Architect's Journal, 129 (April 1959), pp. 504-21; Serge Chermaye and Rene d'Harnancourt, "Design for Use" in Art in Progress (New York, 1944), pp. 190-201.

2 For some practical suggestions as to how this might be improved, see Christopher Alexander, "Information and an Organized Process of Design" in National Academy of Sciences, Proceedings of the Building Research Institute (Washington, D.C.), Spring 1961, pp. 115-24.

3 Cook T. W., "The Relation between Amount of Material and Diffculty of Problem-Solving" in Journal of Experimental Psychology, 20 (1937), pp. 178-83, 288-96; E. J. Archer, L. E. Bourne, Jr., and F. G. Brown, "Concept Identification as a Function of Irrelevant Information and Instructions" ibid., 49 (1955), pp. 153-64.

4 This feeling has been expressed in many quarters, ever since the beginning of the Modern Movement. See, for instance, L. Moholy-Nagy, The New Vision: From Material to Architecture, revised trans. by Daphne Ho man (New York, 1947), p. 54; Walter Gropius, The New Architecture and the Bauhaus, trans. P. Morton Shand (London, 1935), pp. 17-20.

5 Karl Duncker, "A Qualitative (Experimental and Theoretical) Study of Productive Thinking (Solving of Comprehensible Problems)" in Journal of Genetic Psychology, 33 (1926), pp. 642-708, and On Problem Solving, trans. Lynnes Lees, American Psychological Association, Psychological Monographs, No. 270 (Washington, D.C., 1945); Max Wertheimer, Productive Thinking (New York, 1945).

6 George A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information" in Psychological Review, 63 (1956), pp. 81-97; D.B. Yntema and G.E. Mueser, "Remembering the Present States of a Number of Variables" in Journal of Experimental Psychology, 60, pp. 18-22 (July 1960).

7 Alex Bavelas and Howard Perlmutter, classified work done at the Center for International Studies, M.I.T., quoted in "The Relation of Knowledge to Action" by Max Millikan, in The Human Meaning of the Social Sciences, ed. Daniel Lerner (New York, 1959), p. 164.

8 In fact there are cases where a form has been uniquely determined by its requirements, but such cases are very rare. One striking example is the crane hook. See L. Bruce Archer, Design, No. 90 (June 1956), pp.12-19, esp. p. 16; H. J. Gough, H. L. Cox, D. G. Sopwith, "The Design of Crane Hooks", Proceedings of the Institute of Mechanical Engineers (England), 1935; also Annual Report of the British Iron and Steel Research Association, 1954.

9 A typical collection of paintings based on such a kind of 'logical' formalism is to be found in Karl Gerstner, Kalte Kunst, published by Arthur Niggli (Teufen AR, Switzerland, 1957).

10 Jacomo Barozzio Vignola, Regola delli cinque ordini d'architettura (Rome, 1562; Jacques-François Blondel, Cours d'architecture (Paris, 1771), Book IV.

11 Another example of this 'logically' inspired formalism is to be found in Ludwig Hilbersheimer, The New City (Chicago, 1944), pp. 106-21.

12 Whether we like it or not, however rational we should like to be, there is a factor of judgment in the choice and use of a logical system which we cannot avoid. Logical pictures, like any others, are made by simplification and selection. It is up to us to see which simplifications we wish to make, which aspects to select as significant, which picture to adopt. And this decision is logically arbitrary. However reasonable and sound the picture is internally, the choice of a picture must be, in the end, irrational. For even if we can give reasons for choosing one logical scheme rather than another, these reasons only imply that there is another decision scheme behind the first (very likely not explicit). Perhaps there is still another behind this second one. But somewhere there are decisions made that are not rational in any sense, that are subject to nothing more than the personal bias of the decision maker. Logical methods, at best, rearrange the way in which personal bias is to be introduced into a problem. Of course, this 'at best' is rather important. Present intuitive methods unhappily introduce personal bias in such a way that it makes problems impossible to solve correctly. Our purpose must be to repattern the bias, so that it no longer interferes in this destructive way with the process of design, and no longer inhibits clarity of form.

13 The relevant part of William Morris' thinking is to be found in volumes 22 and 23 of the 1915 London edition of his complete works. See also Nikolaus Pevsner, Pioneers of Modern Design (New York, 1949), pp. 24-30.

14 Ibid., pp. 18-19.

15 Their work and ideas are fully discussed by Emil Kaufmann in Architecture and the Age of Reason (Cambridge, Mass., 1955) , pp. 95-99 and 134. No writings of Lodoli's remain, but see F. Algarotti, Saggio sopra l'architettura, in Opere, vol. II (Livorno, 1764); Marc-Antoine Laugier, Essai sur l'architecture, 2nd ed. (Paris, 1775), and Observations sur l'architecture (The Hague, 1765).

16 Nicolaus Pevsner, An Outline of European Architecture, Penguin Books (London, 1953), pp. 242-62.

17 In denying the possibility of understanding reasonably the processes of form production, the fetish of intuition is closely parallel to other famous attempts to shelter from the loss of innocence under the wings of magic and taboo; see, for comments, Sigmund Freud, Civilization and Its Discontents, trans. James Strachey (New York, 1962), or K. R. Popper in The Open Society and Its Enemies (Princeton, 1950).

18 For some recent protests against the willful nature of modern intuition in design, see Serge Chermayeff, "The Shape of Quality" in Architecture Plus (Division of Architecture, A. & M. College of Texas), 2 (1959-60), pp. 16-23.

19 The possibility of amplifying intelligence has already been hinted at in W. Ross Ashby, "Design for an Intelligence Amplifier," in Automata Studies, ed. C. E. Shannon and J. McCarthy (Princeton, 1956), pp. 215-34. See also M. Minsky, "Steps towarsArtificial Intelligence", Proceedings of the Institute of Radio Engineers, 49:8-30 (January 1961).

ISSN 2035-5300