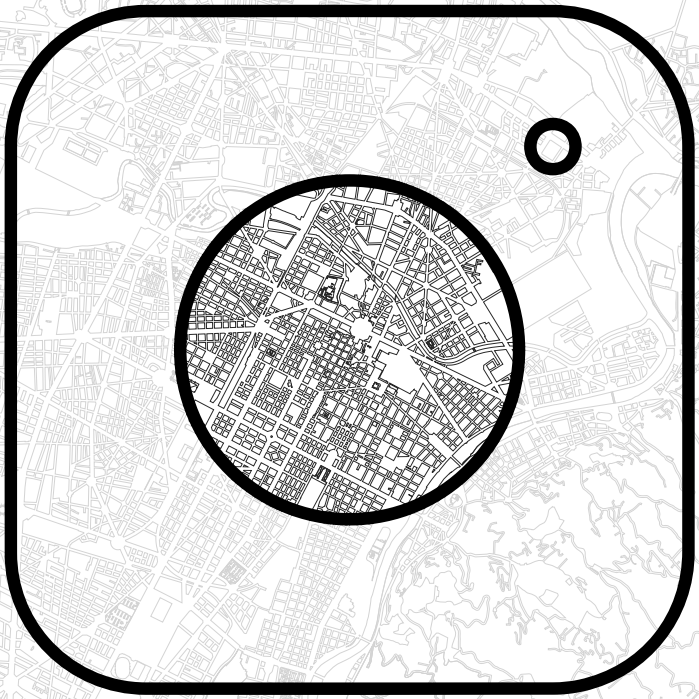
INSTA-TURIN



Revising the notion of space, community and proximity in the contemporary city through Instagram social urban data

Abstract

In the broad context of the social urban data mining applied in the study of the city, Instagram and its geotagged posts are used in this dissertation to understand urban dynamics, and to produce new visual representations and insights in relation to the city of Turin.
The idea of "community" and "city" as they adapted to the Information age are introduced in the first chapter as premises to validate the decision of using geotagged Instagram posts as a source of potential knowledge in the field of urban studies. The "eloquence" of geotagged Instagram posts is found in its intrinsic characteristics of social media where users filter through the selection of pictures the reality around them, making them a powerful proxy to understand the contemporary city in its complexity.
The user of the platform is a flaneur, experiencing and recording the everyday life of the city. His/her record is not a simple representation, but it is embedded of the meanings and the projections the user place in a venue. The geotagging is also a social action, shaping the user's identity. Geotagged Instagram posts are therefore an act through which an individual construct his/her social and individual identity through the reference to a physical space s/he immortalizes and sublimes.
A dataset of over 350 thousand geotagged posts has been collected with reference to 1071 venues in the city of Turin, insisting on a period of time of a year. Quantitative evaluations allowed to develop insights, confirming the validity of the data as indicators of the behavior of citizens in the city. A clustering of the venues according to the social affinity calculated in relation to the recurrence of same users in different locations permitted to rethink the notion of proximity and centrality beyond the border of the physical space.
The outcomes of this work of research consist in the depiction of the city with the paradigm of the network, defining element of contemporary society and tool to understand, visualize and plan the city.

Adriano Aimar

Master thesis

Politecnico di Torino
“Architecture Construction City”

relator
Matteo Robiglio - (Polito)

correlator
Marco Brambilla - (Polimi)
Elena Guidetti - (Polito)

INDEX

#1 MODERN TIMES

INTRODUCTION 13

1.1 THE SOFTARE - COMMUNITY/IES 21
- 1.1.1 Death to the community 22
- 1.1.2 From face-to-face to person-to-person 24
- 1.1.3 The subject as actor 25
- 1.1.4 Long live the communities! 27
- Conclusions 29

1.2 THE HARDWARE - CITIES 31
- 1.2.1 The Misunderstood City 33
- 1.2.2 The dynamic City 37
- 1.2.3 The Data City 41
  - 1.2.3.1 Social urban data – previous works 43
  - 1.2.3.2 Biases 47

#2 DE INSTAGRAM ELOQUENTIA

INTRODUCTION 57

2.1 THE USER 65
2.2 THE IMAGE 69
2.3 THE GEOTAG 77

CONCLUSIONS 81

#3 METHODOLOGY

INSPIRATION 87
3.1 MAIN IDEA 91
3.2 SOCIAL RELEVANCE AND BIASES 95

3.2 DATA COLLECTION 101
3.2.1 Venues listing 103
3.2.2 Venues location ID correspondence 104
3.2.3 Geotagged posts scraping 105
3.3 DATASET 107
3.3.1 Recategorization 108

#4 QUANTITATIVE EVALUATION

INTRODUCTION 115
4.1 Turin by Instagram: NUMBERS
4.1.1 Venues distribution 117
4.1.2 Population distribution 141
4.2 Turin by Instagram: TIME
4.2.1 Day distribution 145
4.2.2 Year distribution 149
4.3 Turin by Instagram: SPACE 153

#5 CLUSTERING

INTRODUCTION 167
5.1 GRAPH CONSTRUCTION 169
5.2 COMMUNITY DETECTION 177
5.2.1 Landmarks 181
5.2.2 Social life 201
5.2.3 Isolated networks 233

#6 CONCLUSIONS

INTRODUCTION 263
6.1 THE SHAPE: CENTRALITY AND PERIPHERY 267
5.2 THE DYNAMISM 271
5.2 FURTHER DEVELOPMENTS 276

ACKNOWLEDGEMENTS 280

#1

MODERN TIMES

Lots of people try to change human nature but it's a real waste of time. You can't change human nature, but you can change tools, you can change techniques. And that way you can change civilization.

Stewart Brand [1]

1. as cited in Cadwallar, 2013

Reflecting on the etymology of words can reveal itself as a fascinating activity. The historical evolution of the language can hide metaphorical meanings, which shaped through the centuries the way we think the word and, most importantly, the world.

The union of the Latin preposition *pro* ("forth") and the verb *iacere* ("to throw") – and the resulting idea of "throwing forth" – produced a series of words in European languages carrying a certain idea of future, but with two different nuances. On one side, there is the idea of wanting to do, of achieving, of inventing, of building something. On the other, the meaning is related with the speculation, the forecasting, the abstraction. In the English language is evident: let us consider the difference between the noun *project* and the verb *to project*, explicitly accented on different vowels. The Italian language generated even two different families of words, one embedded with the first idea of *progettare* ("to design") – *progetto, progettazione, progettista*… ("plan", "design", "designer") – and the other related with the second idea of *proiettare* ("to project) – *proiezione, proiettivo, proiettato*… ("projection", "projective", "projective") – despite the common Latin root.

Figure 1_1
Detail of the Lascaux prehistorical wall paintings with the contemporary tools highlighted

These two perspectives complement each other and form the attitude of the mankind towards the future: the action and the dream, the practicality to better reality and the abstraction to imagine a different one.

It recurs all over history, since the dawn of civilization till modern times. It will take more than a thesis to go through all the way the mankind has imagined and built its destiny and future, but let us take two examples: Lascaux prehistorical wall paintings and the futurist Antonio Sant'Elia *La città nuova*. They do not seem to share anything at all, but, instead, if we go deep into the motive behind the drawings, they are both representing a condition not contemporary to the time of the production of the image – history of art literature agrees that the hunting scenes were drawn *ex ante* (Argan, 2008). They do not represent a random event, but a forecast event, somehow predicted, surely desired – the dream – for which the authors were willingly to actively contribute to the success of it – the action. As far as regard the actual outcomes of these proposals, things change. While we can assume that our ancestors of the Stone Age quite likely succeeded in hunting one of those animals, we know by now that Sant'Elia's predictions did not see the light of the day.

Why? Although the spirit behind those projections was very similar, what differed is the distance from the targets they were aiming at. The men of the caves were looking at a very near future, while Sant'Elia was looking way further. Is it just that which made him to be "wrong" in the end?
Yes and no. Our ancestors and Sant'Elia share something else: both their projections of the future were based on the technologies of their time and their most likely evolution. And this is what failed Sant'Elia, and probably what would have failed our ancestors if the pressing necessity of providing food on the short period had let them imagine the way we would have hunt for food today. The mankind can imagine its future, but it cannot predict the next technological revolution.

Technological revolution does not simply mean a "new" technology, but a "a set of interrelated radical breakthroughs, forming a major constellation of interdependent

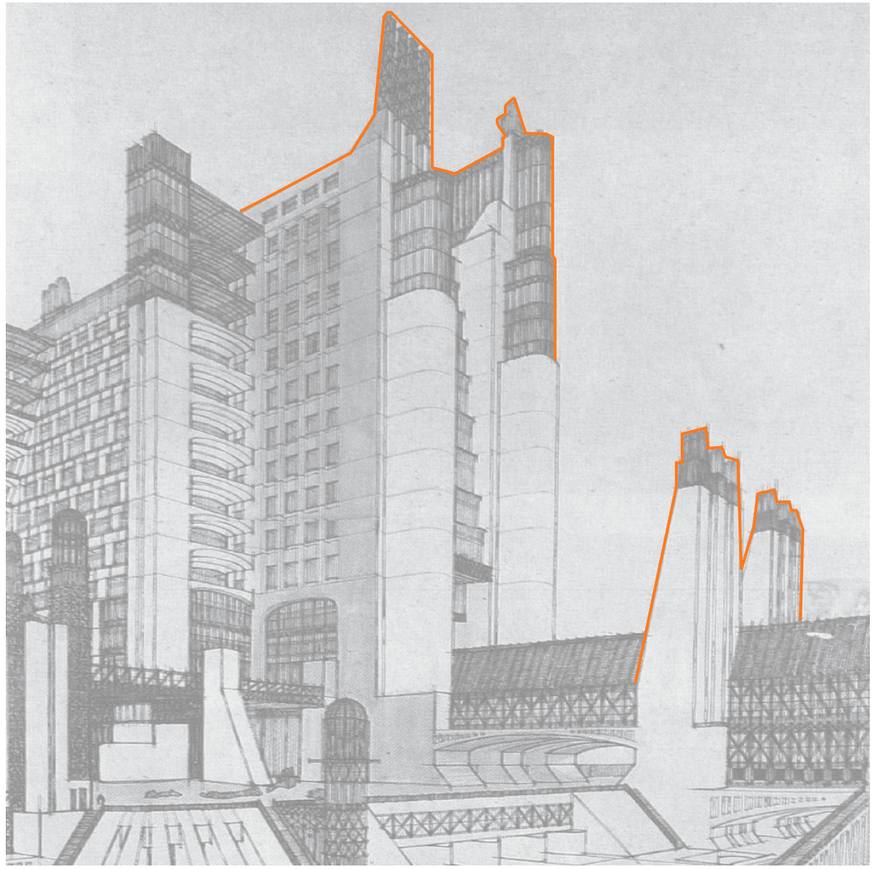


Figure 1_2
Detail of Antonio Sant'Elia "La città nuova" with the contemporary tools highlighted

technologies; a cluster of clusters or a system of systems", with the capacity to deeply impact the economy and, eventually, society *(Perez, 2009, p.8)*.
In the last analysis, while our ancestors were quite correct in imagining that in the next following years they would have still hunted in group, maybe in a better organized fashion, with slightly more refined tools, but still with the same organizational concept behind, Antonio Sant'Elia was trying to foresee the development of the car, the industry and all the wonders of his time, incapable of imagining that something else was going to develop and revolutionize the world: the computer and, most of all, the Internet.

The Digital Revolution is indeed a fact. As much as the Agricultural, the Industrial and all the other technological revolutions which occurred over the history of mankind, shaped it and changed the course of its destiny forever.

«Internet is not simply a technology. It is the founding mean of the information society, which make possible the illimited expansion of interactive networks in every sectors of our existence. It is not a fragment of the technological system: it is the heart forging and modelling the new social structure of everything» [2] (Castells, 2004, p. 20).

I advanced the hypothesis that man cannot predict the next technological revolution, but we can look at the context where they flourished, the deep instances that led the process.
The technological world is controlled today by an American WASP oligarchy made of Zuckerbergs, Bezoses, Gateses, Musks and many others, an elite able to influence the life of millions of people – if not even billions. However, those who, along the 1980s and 1990s, developed the technologies which led to the Digital Revolution, the actual "revolutionaries" had an idea of world they wanted, or at least a disruptive attitude towards the establishment. That establishment characterized by its fixity, its immobility, its adamantine ideologies brought to paroxysm of fanatism had led the world into the largest war of history. The tech community of the Silicon Valley was raised against that system, its philosophy had at its core the boycotting of borders, the creation of a unique free space where everything could flow, a philosophy against the immobility of

2. *"Internet non è semplicemente una tecnologia. È il mezzo tecnologico fondante della società dell'informazione, che rende possibile l'illimitata espansione di reti interattive in ogni settore della nostra esistenza. Non si tratta di un frammento del sistema tecnologico: è il cuore del sistema che forgia e modella la nuova struttura sociale di ogni cosa"*

the establishment and in favor of the movement (Baricco, 2018). The premises of the Digital Revolution were indeed disruptive, and they succeeded in changing the way people live by providing them with new tools.

Recent works have correlated the technical with the social and the natural to produce a more comprehensive and structured understanding of reality, as in, for example, the actor-network theory (Latour, 1999). In brief, the use of tools defines humans and this interrelation is fundamental to construct a theory of society and to read the transformation of it (Amin & Thrift, 2002). The comprehensive social transformation affects the space and, ultimately, the city, since «the space is not the reflection of society, but an expression of it, fundamental dimension inseparable from the general organizational and developmental process of society» [3] (Castells, 2004, p. 50).

3. *"Lo spazio non è un riflesso della società, ma un'espressione di essa, dimensione fondamentale inseparabile dal generale processo di organizzazione e sviluppo sociale"*

To understand the city of the Information Age we have to understand how the society changed and it is changing, and, consequently, how these phenomena affected and are affecting the urban space (Castells, 2004).

The chapter is divided in two main parts, following this consequentiality and withdrawing a metaphor from the digital world of computers. I will start with the software, the operating component of the system, society in its multiple aspects, to continue with the hardware, the more visible part, the solid one, the city and how to rethink the study and understanding of it.

Though we are still the heirs of *gemeinschaft* and *gesellschaft*, we are moving towards a different, more restless and more dispersed, vocabulary through a constant struggle over the three Rs of urban life: new social relationships, new means of representations and new means of resistance.

Amin & Thrift [4]

#1.1

THE SOFTWARE: COMMUNITY/IES

4. Amin & Thrift, 2002, p.48

The digital revolution, for its intrinsic nature of technological revolution, had a profound effect on society, mining the foundation of a consolidated system. In particular, it caused the redefinition of both the individual and the collective identity (Castells, 2004), attacking a fundamental dogma in social science, in general, and in the urban theory, in particular: the idea of community.

#1.1.1

Death to the community!

The twentieth century was founded on the idea of society as a series of coherent communities, concept introduced by the German social theorist Ferdinand Tönnies in 1887. More specifically, the realm of indirect interactions, formal values of the *Gesellschaft* – translated with "society" – incorporating the various communitarian spheres of the immediate and personal interactions of *Gemeinschaft* – translated with "community" (Forty, 2000).

The fascination for this idea of the community derive from the poetic view of a collectivity moving as one, inspired by a common set of values, basing relations and interactions on a face-to-face culture. In addition, this conception of community projects it in the past of a tradition and locates it in a place, forming a comforting idea, simple to understand and to communicate (Amin & Thrift, 2002).

With the diffusion of technologies enabling people to overcome the physical constraints for interactions, critics arose accusing ICTs of causing social woes. Those critics were based on the assumption that the online lives were virtual – and therefore fake – and on the cherishing of the face-to-face culture as the only meaningful type of encounter (Rainie & Welllman, 2012).

A sort of skepticism is even in the work of author like William J. Mitchell, whose book "E-topia" (see Mitchell, 1999) created a popular neologism. The apparent enthusiasm for the technological future is more a propositive attitude towards a dangerous situation which may be turned into a favorable one. The belief in the face-to-face culture and the deriving idea of community is strongly present in the mind of the author when he claims that the «so-called 'virtual communities' work best when they are allied with the possibility of occasional face-to-face encounters, and that online interaction actually stimulates demand for more familiar sorts of meetings and meeting places» (Mitchell, 1999, p.90). The old *Gemeinschaft* and *Gesellschaft* can resurface intact after the storm brought by the Digital Revolution and even purified by their previous flaws: «[…] you can live in a small community while maintaining effective connections to a far wider and more diverse world—virtual *Gesellschaft*, as we might term it, without tongue too far in cheek. Conversely, you can emigrate to a far city, or be continually on the road, yet maintain close contact with your hometown and your family—electronically sustained *Gemeinschaft*» (ib., p. 22).

Even the technology-enthusiast Carlo Ratti seems to be seduced by the comforting idea of the community seen as a mass of people moving as one. The last chapter of his "La città di domani" (see Ratti, 2017) is programmatically titled "Hacking the city" ("Hackerare la città", in Italian), showcasing a series of examples where the internet allowed the rise of bottom-up movements, reclaiming the city inspired by democratic sentiments. The resulting image is a modern version of Pellizza da Volpedo's "The Fourth Estate", a coherent community based on shared values.

To understand the real nature of society in the Information Age, it is necessary to overcome a few concepts that seems to be adamantine in the way we have read and interpreted society so far. First, it is important to eradicate the relevance given to the face-to-face encounters and resolve the inexistent dichotomy between the virtual and real with regards to human interactions. Second, the role of the individual needs to be considered in its unicity, not defined by a single ideal – and rhetoric – community, but defining a series of communities. Last, the society as the whole formed by the aggregation of the different typologies of human interactions can be understood in its complexity.

#1.1.2

From face-to-face to person-to-person

The electronic communications have been described as new forms of sociality (Castells, 2004) and although somebody saw the alienation of the individual in the interactions constructed with the mediation of the technological means, researches proved that internet activities were not affecting the local activity but instead they could even have a positive influence on it (Rainie & Wellman, 2010).

To understand the evolution of sociality and to overcome the supposed relevance of the face-to-face culture «it helps to think about communities as fluid personal networks, rather than as static neighborhood or family groups» (Raine & Wellman, 2010, p. 122). With this idea in mind, the door-to-door network characterizing the traditional society of the West had known a first transformation after the WWII and the spread of mass society, when individuals starting to build their sociality with place-to-place networks, transcending the boundaries of neighborhoods and, consequently, the so-called communities. With the spreading of technologies, the physical place to host interactions became unnecessary, further simplifying the place-to-place networks to person-to-person ones (Rainie & Wellman, 2010).

Under this perspective, the electronic communication no longer appears as a regression or a losing of meaning in human interaction, but as the final step of the historical evolution of it, in a process of liberation from constraints imposed by the physicality of space and the "peer pressure" of the group.
To go further on this line, this "postsocial forms" of sociality have been seen as a return to an authenticity and their diffusion explained because they «step into the place of social relations where these empty out, where they lose some of the thickness and meaningfulness they have had in earlier period» (Knorr Cetina, 2001, p.17).

This call for a redefinition of the role played by the subject, no longer influenced by external agencies like local or group boundaries: a new form of "networked individualism" rises, downloading «the responsibility – and the burden – of maintaining personal networks on the individual» (Rainie & Wellman, 2010, p.125).

«The conclusion» to be drawn «from the contemporary re-imagining of the individual contradicts postmodern social theories which tend to postulate the eclipse and death of the subject. As a first approximation, we can associate a postsocial environment with an expanding sphere of the subject, where 'subject' stands not only for mental or existential conceptions of individuals but for an open-ended series of individual-centered significations and processes. The remarkable rise of subjectivity thinking and the concomitant emptying out of a social imagination and of social principles and structures act in concert, so to speak, to create and unfold the space for this expansion» (Knorr Cetina, 2001, p.5).

#1.1.3

The subject as actor

The centrality of the individual might prefigure a sort of mass individualism, which to a certain extent is the exact opposite of society. Instead it is important to understand the role of the subject in the definition of itself and how this reflects, consequently, on society.

It is clear by now that the Information society calls for a redefinition of the identity of the individual out of the traditional schemes. This necessity can generate two different processes, varying according to the typology of individual involved. «In general, those

who have a high level of education and a wide possibility of access to cultural and personal resources, tend to build themselves as an individual, counting on themselves and their networks of friends. On the contrary, those who have less opportunity of survival, or of defense of their interests against the fluxes of the global capital, tend to seek refuge in collective identities. In both cases, however, the identity is redefined in a defensive sense, leading to a severe crisis of the political institutions and the social organizations»[5] (Castells, 2004, p.21).

The final part of the citation tragically depicts what it has been already described as the natural result of the liberation in social relations from external constraints and the inevitable end of traditional social structures. The rest of it, however, seems to cast a shadow over the construction of identity: on one side, the strong, succeeding in the fight for survival, on the other, the weak, succumbing to collective identities for desperation.

To rethink this perspective, Bauman's theorization of the liquid modernity may come in help. In his antiutilitarian philosophy the subject develops an understanding of him/herself, escaping the alienation of a world where the power, in its multiple manifestations, offers a unique dimension of the real. The subject becomes the actor of his/her own destiny. Thus, the crisis of the political institution and of social organizations need to be read in the context of the shifting from the modern ethic – where the law is given with a deriving sense of trust and security – to the postmodern moral. The latter configures itself freed from aprioristic ethical codes and based on the principle of ethical responsibility (Bauman, 2014). In other words, the Information age offers the individual with the possibility of choosing for him/herself, the privilege and the burden of the selectiveness.

To go back to Castells, it is true that the individuals risk succumbing in front of this freedom for their incapability to face it. However, the refuge into the abovementioned collective identities, remains a choice among what the author himself defines as an «infinite galaxy of cultural subsets» (Castells, 2014, p.60). The crisis of the modern

5. "In generale, chi ha un alto livello di educazione e una vasta possibilità di accesso a risorse culturali e personali, tende a costrure se stesso come singolo, contando su se stesso e sulle proprie reti di amici. Al contrario, chi ha poche oportunità di sopravvivenza, o di difesa dei propri interessi contro i flussi del capitale globale, tende a cercare rifiugio in identità collettive. In entrambi i casi, tuttavia, l'identità viene ridefinita in senso difensivo, il che porta ad una grave crisi delle istituzioni politiche e delle organizzazioni sociali."

ethic and the rise of Information age wiped away a societal model based on one dominant culture, on a unique community, which the individual was called either to accept or to reject. Instead, the contemporaneity is composed by a variety of other forms of social aggregations not reduceable to the traditional idea of community, because even when they present themselves as collective identities the individuality of the subject is preserved.

#1.1.4

Long live the communities!

Reducing the current state of society to a confuse composition of an indefinite number of parallel communities – intended in the traditional sense – would be misleading. The metaphors of the swarm compared with the herd will help to better understand the functioning of the systems around which modern society organized itself. While the latter is structured in its form, hierarchy and role of the individual within it – and can be ultimately assimilated to the function and functioning of the community in the traditional sense – the first is not. The swarm has an evolving shape, where the role of the unit in the whole is defined only by the temporary condition of the swarm itself. This explains the idea of the modern forms of communities – or better of "communitying" – which when they take shape in the urban space turn it into a space of «flow and mixture, promiscuous 'meshworks' and hierarchies of different relations, rather than patchworks of different communities, hybrids involving almost continuous improvisation in which the in-between' of interaction is crucial» (Amin & Thrift, 2002, p.81).

Those different relation were categorized by Amin and Thrift tried who identified five ty-

pologies of community peculiar of contemporary society and of contemporary cities. The first is the "planned community", constructed from the political/bureaucratic partitioning of the urban space in postcodes, census and poll area etc., which produce "spatial categorization" defining the individual from where s/he lives. The second Is the "post-social and post-human community", a typology arising from the modification of human interactions in the context of modern technologies (see chapter 1.1.2). The third is formed by "the new forms of human sociality": the "new means of 'light' sociality" of groups coming "coming together briefly for a purpose and then disperse" – like the people "involved in consumption at malls" or in the participation to festivals - , or the people joined because of common sentiments, political views or other kind of opinions, or friends, "the motivating force in cities […] the family we chose". The fourth typology of contemporary community is the "diasporic community, where the belonging and identification is anything but local", originated by the growth of global migration in some cities and allowing individuals to interact and to base elements of their social and cultural lives over long distances. The fifth – and we will see later, the most relevant for contemporary urban spaces – is the "everyday life itself". The explanation of what it actually consists in, which forms could it take, would be both simple and complex. Simple because the everyday life is nothing more than the ordinary life of the individual outside of all the "superior activities", but at the same time complex, because these leftovers of life are almost unattainable. The everyday life is «the community of the banal and the mundane, but also the community of improvisation intuition, play. The community of taking place, not place. The community that cannot be classified. The community without an identity in which humans co-belong without any representable condition of belonging» (Amin & Thrift, 2002, p. 43-48).

This categorization of contemporary human communities reinforces and proves the idea of the swarm and the privilege of selectiveness of the individual in the Information age. Apart from the "planned community", in all the other typologies the belonging of the individual is based on his/her preferences, aspirations, desire of identification, making him/her an active subject in the formation of social aggregations.

Conclusions

In this chapter we have gone through the fundamental elements characterizing contemporary societies. The redefinition of collective identities derives from the redefinition of the individual ones. Those, thanks to the spread of the internet, electronic communication and technologies, are no longer defined by place or group, and interpersonal networks are now based on the individual. The subject, thus, gains the role of actor in the definition of his/her identity outside the traditional institutions and one unique dominant culture, and s/he becomes the definer of the contemporary forms of social aggregations. The new communities are less structured than the traditional ones and intersect and overlap each others in the urban space. Among those categories of social aggregations, the everyday life plays a large role, involving all the individual of society indistinctively, and differently, for the most of their life both in spatial and temporal terms.

Rather than remaining within the field of a discourse that upholds its privilege by inverting its content (speaking of catastrophe and no longer of progress), one can try another path: one can analyze the microbe-like, singular and plural practice which an urbanistic system was supposed to administer or suppress, but which have outlived its decay.

De Certau [6]

6. De Certeau, 1980/1984, p.96

#1.2

THE HARDWARE: CITIES

In the 1990s the technology enthusiasts imagined a world where distance would no longer exist. As consequence, the city was deemed to disappear as an attraction element, since Internet would have freed the individual from the dependency from location. Although today the premises revealed themselves to be correct, the expected result was not: city are growing at different rate everywhere in the world. (Ratti, 2017).

Thus, the city of the Information Age still plays a crucial role, but its function and meaning deeply changed and they are still widely misunderstood. A correct reading of the contemporary city needs to include it in the context of the network society. This will shape a new concept whose final goal is to master the complexity of the city and develop correct interpretations for the study and the planning of it.

#1.2.1

The Misunderstood City

The main misconception about the city in the Information age can be grouped in two main trends: the failure to understand the revolutionary nature of the digital revolution – with the deep impact on society we already discussed –, and the failure to see the relation between the "virtual" and the "real", assumed to be either independent from each other, or the latter negatively affected by the first.

Starting with the first trend, many scholars as seen the technological revolution as merely a rising of new technologies, allowing humans to live more comfortably in an environment governed by the same dynamics of the one they used to inhabit. The logic behind this misconception, it is the same described in the introduction of the chapter about Sant'Elia's *La città nuova*: the imagination of the future based on the contemporary technological system.

Mitchell, for example, looks at the «cities of the twenty-first century as systems of interlinked, interacting, silicon- and software-saturated smart, attentive, and responsive places. We will encounter them at the scales of clothing, rooms, buildings, campuses and neighborhoods, metropolitan regions, and global infrastructures» (Mitchell, 1999, p.68). The misunderstanding of the new logics governing the physical space because of the digital one is clear when he attempts to make Pindaric flight in comparing the traditional, real, comforting architecture to the new technology: «In an online bookstore, for example, the home page is the equivalent of the facade, and you find it by means of a search engine or by following pointers from other pages rather than come upon it by strolling along a street. The online catalog corresponds to the shelves of actual books, search engines and software agents help you to browse, and the online order form executes the functions of the checkout counter and cash register.

Somewhere in the background, of course, there is a mega-warehouse, or a system of smaller warehouses distributed through the service area, where the books are physically stored, retrieved, packaged, and dispatched in the most traditional of fashions. And somewhere as well—maybe in a very different location, as determined by labor markets and telecommunications infrastructure— there are servers, call centers, and administrative offices» (ibidem, p. 126).

The attempt may be read in the context of the evident concerns and worries generated by the new world that was forming in front of scholars and researchers, and the difficulty of reading and face it. Mitchell reacts like a man the middle of a forest, lost and alone, facing a tiger and trying to comfort himself seeking the similarities between the wild animal and a kitten.

The other misconception is the thinking of the digital sphere as separate dimension from the real. Saffo, for example, explained that «Two parallel universes currently exist—an everyday analog universe that we inhabit, and a newer digital universe created by humans, but inhabited by digital machines. We visit this digital world by peering through the portholes of our computer screens, and we manipulate it with keyboard and mouse much as a nuclear technician works with radioactive materials via glovebox and manipulator arms. Our machines manipulate the digital world directly, but are rarely aware of the analog world that surrounds their cyberspace» (Saffo, 1997, p. 95). The digital space was not seen as neutral by some other scholars. Virilio, in the book programmatically titled "The Lost Dimension" (see Virilio, 1991), described the collapse of the three-dimensional public space into the two-dimensional screens of computer, the informational networks are seen as a sort of prison of illusion for the humanity.

The misunderstanding of the actual influence on society of the Digital Revolution lead to the consequent failure to understand how that society actually live the space of the contemporary city, rising problematic interpretation of it. While today nobody could possibly deny the intertwining networks at global level and the complex dynamics governing the world beyond the traditional scales of influence, it is still alive the belief of the city as an self-ruling, independent space, only partially influenced and touched by globalization.

In this perspective, the city is seen as the place of fixity, in contraposition to the world as the space of mobility. This misleading theory is based on conflicting dichotomies opposing the local to the global, the near to the far (Amin, 2001). The "new localism" produced by this concept rests on assumption based on ideas like proximity, on the culture of the face-to-face interaction etc (Amin & Thrift, 2002), all the wrong interpretations of contemporary society I pointed out in the previous chapter (see 1.2).

I would like to advance the theory that the fortune of the "new localism" derives by a unique nostalgic approach to the city. This takes shape, clearly, in the already mentioned misunderstanding of social changes, but it derives also from an older disease of urban theory: the never overcome death of the utopias of Modernism.

The modernists wanted to believe in a humanity to be rescued, they imagined the individual as a component of a wider system and, therefore, they designed buildings and cities as machines able to host these components. As it happens to big dreamers, they were disappointed. They actually miserably failed and that it is what some critics have been practically shouting since the 1970s.

Rowe wrote about the "Death of Utopia" in his famous "Collage city", advocating new cities conceived on pluralism, the bricolage etc. (see Rowe, 1978). De Certeau described the decaying of the "Concept city" with the powerful image of a man standing on the World Trade Center and looking at New York from that distant perspective, unable to grasp the real life of the man walking on the street at the bottom of the skyscraper (See De Certeau, 1980/1984). And I could go on with endless example of other scholars, urbanists and thinkers who said one simple thing: the Modernism was wrong, because a perfectly designed and functioning city – an utopian city – requires

a perfectly functioning society – an utopian society – and this kind of society does not exist. It never existed, probably. For sure, it did not exist at the time of Modernism – although the dynamics could have still been interpreted with of the categories of the local and the face-to-face culture of interaction – and it does not exist, more than ever, today.

Why is it still a popular conception, then, a perfect city for a perfect local community? Because it is a very simple construction: easy to understand, easy to master, easy to apply, easy to communicate. And, most of all, it is indeed beautiful and most certainly romantic. The contemporary city, now more than ever, needs to abandon those approaches and the surviving dream of the city as fortress of safety against the global menaces.

The idea of a "new localism" is tempting and seducing even for Manuel Castells, the theorist of the network society, described as structured «as the opposition between the global and the local»[7] where «cities, as systems of communication, should function as a link between the global and the local»[8] (Castells, 2004, p.56). This is seen by Castells as problematic as «the logics of the globality and of the localism are in conflict between each other and it is very difficult to manage to satisfy them both and at the same time»[9] (ibidem).

Instead, it is more reasonable to construct a theory about the contemporary city, where it is considered not in contrast with the world and its governing dynamics, but as an element involved and crossed by those dynamics, where «places are more than what they contain, and what happens in them is more than the sum of localized practices and powers, and actions at other 'spatial scales'» (Amin, 2001, p.395). A system which keeps a specificity not based on the notion of local, «on the powers of proximity/particularity in a world of displaced and multiscalar happenings and power geometries» (ibidem, p. 397), but, instead, claiming «a role for specific institutionalized practices within cities» (Amin & Thrift, 2002, p. 72).

7. *"[Dal punto di vista funzionale la network society si struttura] come opposizione tra globalità e localismo."*

8. *"Le città, in quanto sistemi di comunicazione, dovrebbero fungere da collegamento tra globale e locale"*

9. *"Le logiche della globalità e del localismo sono in conflitto tra loro, ed è molto difficile riuscire a soddisfarle entrambe e contemporaneamente"*

#1.2.2

The Dynamic City

Even though we might criticize Castell for ending up advocating the idea of the "new localism" – to stick to this definition –, nothing can be said to his premises, properly laid down and constructed. He called for «a new theory of shapes and spatial mechanisms, proving itself compliant with the new technological, social and spatial context we live in»[10] (Castells, 2004, p.49). However, the notion of the "space of fluxes" embedded in the physical but ruled by other and independent logics needs to be overcome.

The space of fluxes is a space where the physical and the virtual are neither absolute nor neutral: the virtual networks and the material spaces collide and melt, they influence each other. Individuals are surrounded by the space of fluxes not as passive subject, but as active agent in the construction and evolution of it (Ratti, 2017). In other words, the virtual cannot be understood without the physical and neither the opposite.

The "network" needs to be understood as the defining element of the contemporary city in all of its meanings. First, "network" as virtual networks, the Digital revolution, the influence of those technologies on all the aspects of the life and ultimately of the city. Second, "network" as the key aspect at the base of the social system of interactions and aggregations, the individual as member of multilayered, intertwining, different typologies of communities that "take place" in the spaces of the city. Last, "network" as organizing principle, last resort to reduce the complexity of the contemporary city to a scheme enabling to master that very complexity, an interpretational tool, which, however, still keeps the notion of flexibility proper of a net.

The "misunderstood" city becomes then suddenly understandable when paired with the notion of dynamism. The "dynamic" city moves while keeping still, it gets inputs, which it transforms into outputs by quickly putting into motion some elements of the

10. *"[Si ha bisogno di una] nuova una nuova teoria delle forme e dei meccanismi spaziali, che si dimostri conforme al nuovo contesto tecnologico, sociale, e spaziale in cui viviamo"*

system, and then all over again, all the time in a different way, with all the elements taking new roles, meanings, functions. A city not deductively defining its elements, but a city inductively created by them.

As pointed out by Amin and Thrift, the city needs to be seen more as a site, hosting some peculiar phenomena that do not take place elsewhere. First, the city is a site in "near-far networks". That means that the city – let us think of fields the research or the economic – is not independent in itself in the production of knowledge/revenue, but neither it is dispersed in the global flows: the two systems are integrated, and those fluxes take shape and strength in the city. This applies to all sort of interactions – technical, economical but also social – empowered by the possibility of distant connections, but still strengthened by local interdependencies. The second aspect is the city characterized by being the site of circular consumption. Although the urban space is nothing but economically autarchic and, actually, more and more intertwined with stretched networks providing goods and services, there are some fields – in particular the third sector, voluntary and community organizations – where the city consumes what it produces, and this aspect is not to be overlooked. Last, but not least indeed, is the presence in the city of "light institutions", varying from organized activities to smaller elements like, for example, the networks of personal contacts. These institutions are, first, the "urban services" of all kinds: transport, finance, third sector in general. Then, there are "meeting places", abundant in the city in a variety of forms. The city is also a site of "varied compositional knowledge", with the presence of cultural institutions, schools, universities, and of "unstructured activities", the system of all the advantages deriving from personal connections and local knowledge (Amin & Thrift, 2002).

The dynamic city is never a single city, but it is a city that changes, happens in a different way every day. Every day all the elements and dynamics described above combine and overlap, there is not an *a priori* order, but contingency is the rule. The everydayness is of vital importance for the life of the contemporary city and a fundamental element for a study and a reflection on the urban today.

Again, Amin and Thrift stressed this plural nature of the life in the city and the significance gained by the "banality of the everyday life". They propose three metaphors to capture what they call "recurring practices". «The first is transitivity, which marks the spatial and temporal openness of the city. The second captures the city as a place of manifold rhythms, forged through daily encounters and multiple experiences of time and space. The third notes the city as footprints: imprints from the past, the daily tracks of movement across, and links beyond the city» (Amin & Thrift, 2002, p. 9).

I said that the network is a defining element of contemporary city and most of all an interpreting tool for its complexity. The three metaphors can help in the understanding the intrinsic nature and characteristics of a network and, thus, facilitating the reading of the dynamics described by it. The network is in fact deeply transitive, it is not permanent, it arises from the sudden "lighting" of connections between points. It is not totally casual, it has an internal rhythm, a logic, a scheme to which it can be confronted and compared. Last, although it is not a permanent and lasting element, it does have effects, consequences, influences, impact, a footprint it leaves.

These three metaphors are suitable in describing both the network as a concept and the everyday as a system of practices. This similarity does not simply happen to be, but it accounts for the deep affinity of the two, despite their ontological evident difference. The everyday can be interpreted as a network, a system of individual, incoherent practices but still connected and linked to each other. The network, for its part, cannot be imagined as a fixed structure – as much as we cannot describe the shape of a net out of the instant when we see it – and the short time is the only temporal span in which a network can be understood in its ephemerality.

The everyday – in its totality and comprehensive meaning – is, therefore, not one of the elements to read, analyze and understand the city of the network society, but is the fundamental key of interpretation of it. Although it is not the first time the urban literature is fascinated by this concept, the everyday is returning with a renovated significance

and pregnancy, along with new tools to master it.

The everyday as the "space" where the actual life occurs is a theorization that dates back to Lefebvre, who defined it as the intersection of «illusion and truth, power and helplessness; the intersection of the sector man controls and the sector he does not control» (Lefebvre, 1947/1991, p.21). The everyday for him is not only composed by the reality and practicality of routine, but also by the imagination, the dream, the potential in all its forms. I will not go further through Lefebvre's theorization, because it has political and sociological implications not relevant to this dissertation. I will recall, instead, the already cited De Certau. He saw in the everyday life the reappropriation of individuals of their role of "artists"- used in the meaning of "producer" –, reinterpreting and transforming the models imposed by institutions of all sorts. He complains that science does not study this process, the actual life, but stay far away from it, as in the image of the man at the 110th of the World Trade Center. Instead, «the ordinary practitioners of the city live 'down below', below thresholds at which visibility begins. They walk – an elementary form of this experience of the city; they are walkers, Wandersmaenner, whose bodies follow the thicks and thins of an urban 'text' they write without being able to read it. These practitioners make use of spaces that cannot be seen; their knowledge of the is as blind a that of lovers in each other's arms. The paths that correspond in this intertwining, unrecognized pomes in which each body is an element signed by many others, elude legibility. It is as though the practices organizing a bustling city were characterized by their blindness. The networks of these moving, intersecting writings compose a manifold story that has neither author nor spectator, shaped out of fragments of trajectories and alterations of space: in relation to representations, it remains daily and indefinitely other» (De Certeau, 1980/1984, p. 93).

De Certeau seems both disappointed and yet fascinated by the intrinsic illegibility of the everyday life. He could not imagine how technologies would have developed into the recent "Big Data Deluge"[11], allowing us to have an insight of this everyday life in the city. An aspect of it, as I said, fundamental for its understanding.

11. *Simon, 2013*

#1.2.3

The Data City

The contemporary city defined by networks and by the everyday is for sure more complex to understand, but there are also many more means to attempt a reading of it.

In fact, another product of today's economic and informatic conjuncture is the so-called "Big Data Deluge", a legitimate "grand-son" of the Digital revolution, generated by a series of factors and other phenomena linked with the development of technology in today's society. These could be identified in the "always-on consumer" marketing approach, the plummeting of technology costs, the rise of data science (like Google and Infonomics), the development of the platform economy and the spread of social media. At present time, organizations, companies and other actors have the potential or actual availability of big, complex, generally constantly updated datasets on many aspects concerning logistics, production, distribution etc. Those datasets are composed by different typologies of data (unstructured data, semi structured data, metadata) which added up to the traditional structured data (Simon, 2013).

The most striking feature of Big Data is not their dimension. In fact, it is important to remember what has been theorized in 2005 in the so-called Kryder's Law: disk drive density, also known as areal density is destined to double every thirteen months, implicating the increase of the affordability of data storage systems. This means that what we define today as "big" with regards to the dimensions of the dataset will not be necessarily considered in the same way in the future relatively to upcoming bigger datasets. Instead, the intrinsic different quality of Big Data datasets is that they are characterized by a bigger volume compared to traditional ones both in length ("more rows") and, more importantly, in width ("more columns"), creating big and complex system of information (Simon 2013).

Let us leave on a side the potential – and current, actually – deployment of data made by companies, privates and other fields of research and let us focus on the city and urban theory, as if they were detached from the other sectors. Big data gives analysts both the opportunity to access larger statically sample and a finer grain of information, but, especially, they provide them with a greater variety of information which can be combined for a more comprehensive reading and understanding.

Definitions like "mediascape" (Appadurai, 1990) and "informational landscape" (Graham, 1998; Zook, 2000) explains the fusion between the physical and the digital space in the contemporary city, I already discussed in the previous chapters (Ratti, 2017; Castells, 2004). The ability to collect, store, read and interpret this information made by the city – in its comprehensive dimension of the technological and social aspect (see the actor-network theory – Latour, 1999) assigns a primary role to data analysis and data mining in the field of urban studies. This process has been defined as "urban computing" by Erica Paulos in 2004 (Eric Paulos et al., 2004).

Urban computing is the «process of acquiring, integrating and analyzing a large volume of heterogeneous data produced by various sources in urban areas, for instance, vehicles, sensors, and human beings, to help solve various problems that cities face such as traffic congestion and air pollution» (Silva et al., 2018, p. 3). A process, then, but also «a computer-mediated mean to understand the aspects of the urban phenomena and also provide estimates about the future of cities» (ibidem). Urban computing expands the field of research of urban studies as it involves «an interdisciplinary area resulting from the fusion of computer science with traditional areas such as economics, geography, transportation, and sociology in the context of urban spaces» (ibidem).

Urban computing deploys a variety of data coming from different sources, but whose actual and original producer is the user him/herself. Again, everything falls back to the idea of the individual as an actor, an active agent in the contemporary society and city. Although, then, all the data are, in the last analysis, user-generated and a further classification would seem pointless, I would like to make a distinction to give an overview of the current state of the art of the practice of urban computing in the literature.

The distinction is built on the different typologies of sources of urban computing. On one side, there are data coming from "indirect" sources in relation to the individual – sensors, infrastructural systems etc. – and which provide insights on the technical/managerial aspects of the everyday life – like commuting, energy consumption etc. On the other, there some user-generated contents accounting for the social aspects of the everyday life, coming from the interactions with other users on social media platforms. Although the analysis of urban data is a relatively recent practice, the concerned literature is quite vast, and a proper discussion would move the focus away from the scope of this research. In the following chapter I will discuss only the works related with the use of Location-Based social networks, deliberately overlooking the insights coming from other researches based on sensors and other similar indicators and sources of data.

#1.2.3.1

Social urbat daya: previous works

A definition of what social urban data exactly are may derive from the two adjectives defining this type of data. They are "social" because they are generated by people directly – like on social media – or indirectly through their actions – like mobile phone activities. They are "urban" because they are of some pertinence with regards to the city in the dimension of time and space. To sum up, social urban data are «generated either directly or indirectly from people and their actions; derive from emerging sources such as sensors, mobile phones, geo-enabled social media, and LBSNs; are multidimensional in nature, meaning that they are spatially and temporally referenced; can be used to infer spatial, temporal, and social aspects of human movement, activity, and social connectivity; but are less structured and more semantically ambiguous than traditional urban data» (Psyllidis, 2016, p.61).

Social urban data are increasingly popular because they can integrate and sometimes even substitute the traditional methods deployed to provide insights on society. Methods like surveys, questionnaires, interviews are characterized by intrinsic flaws concerning the dimension of the sample, the economic and temporal investment (Ciucarelli, 2014; Psyllidis, 2016) and their "refresh rate" does not keep up with the one of social media data (Psyllidis & Bozzon, 2015). In addition to that, social urban data are able to describe urban dynamics whose short temporality – like festival, cultural events, protests – does not allow traditional methodology to measure the influence on the city during the manifestation of these dynamics (Psyllidis & Bozzon, 2015).

In the contemporary social and technological conjuncture, social urban data are also relevant because they are based on the deployment of the action of the human being and his/her actions as proxies to understand urban dynamics (Psyllidis, 2016). This conception stresses the central role of the individual as a defining element of the system, whose human nature is combined with the technological aspects of his/her life. People are used and seen as "sensors" (Ciucarelli, 2014), "cyborgs" (Gandy, 2015), or "smart mobs" (Ratti, 2017), but no matter which metaphor is more appropriate: they all account for an implemented humanity. In this view the human is not divided from the digital more than the real city is not separated from the virtual one.

Understood the relevance of social urban data, an overview of the current state of art of the research in the field needs a further preliminary distinction concerning the sources, often overlooked by authors. Social urban data and related research are based on those social platforms allowing user to add some spatial information to the content produced. Among those platforms there are two typologies which need to be clearly separated, because of the different meaning of the spatial information provided by the user. The first typology is the one which has known a wider popularity in the literature: Location-Based Social Networks. LBSN are, evidently, exclusively based on location and the sharing of that spatial information – above all Foursquare. Users "check-in" in venues leaving tracks of their passage and movement in the city. The second typology refers to those non-location-based social media services – Twitter, Facebook, Instagram etc. – where geotagging and other localization services are just one of the features and sometimes not the most relevant. (Tasse et al., 2017) The following overview on previous works will be organized by social media platform to better understand the nuances of meaning in the act of geotagging in LBSN and in non-LBSN.

The most popular social platforms among LBSNs is Foursquare, an app where «users can 'check in' to locations to let their friends know where they are at that moment. When reporting their location, Foursquare users are shown a list of nearby places. They can also register new check-in locations for subsequent Foursquare users» (Quercia at al., 2014).

Many scholars saw the potential of using check-in to develop alternative methods for clustering venues in the city. Cranshaw et al. in the Livehood project proposed a new mapping of neighborhoods in some America cities, basing the process on social affinity (Cranshaw et al., 2012). Noulas et al. identified user communities in two metropolitan cities grouping them by the similarity of the locations visited and proposing a recommender system (Noulas et al., 2011). Users' preferences were the theme developed recently by Xie et al. who used check-ins to study the difference among travelers' preferences in different cities (Xie et al. 2018).

Other researches used check-ins as an indicator for the intensity or density of some human activities. Laman et al. «employed check-in data aggregated at the census tract level to develop a quantitative model for activity intensity as a function of land-use and built environments attributes for the New York City region» (Laman et al., 2019, p.737).

Others saw the potential even for social analysis, like Quercia et al., who deployed check-ins as a source of information to study urban deprivation in London (Quercia at al., 2014).

To move to non-LBSNs, the popular app Twitter has known a certain fortune, since users are able to geotag their tweets. Although the app is based more on the sharing of opinions, Cheng et al. used a large dataset of tweets to study urban mobility as well as some indicators more related to users' opinions like sentiment analysis (Cheng et al., 2011). Mohammady and Culotta paired Twitter information with county demographics, validating the census demographic data comparing them with the results of a linguistic analysis on the content of tweets (Mohammady & Culotta, 2014). Shelton and Porthuis explored the evolution of the neighboorhoods in Atalanta till their final stage, studied with the use of inhabitants' tweets (Shelton & Poorthuis, 2019). "Urban Stories", part of the project "Telltale" of the Polytechnic of Milan, produced maps to validate the use of Twitter for urban analysis in the city of Milan, while "Maps of Babel: The City of Languages" analyzed the linguistic scene of the city extrapolating demographics from tweets (Ciucarelli, 2014).

There were some attempts to combine the data of multiple social platforms for a variety and completeness of information. Del Bimbo et al. integrated Cranshaw (see Cranshaw et al., 2012) with geotagged data from Facebook, performing classification of venues in the city of Florence (Del Bimbo et al., 2014). Chen et al. combined two datasets of geotagged post from Twitter and Instagram for the study of human mobility prediction (Chen et al., 2019). Two datasets of Instagram and Twitter data were used too by Brambilla et al. during Milan Fashion Week to map and analyze the response of users to fashions shows and brands on social media (Brambilla et al., 2017). Similar datasets were used to study user behavior during a major cultural event on the Lake Iseo in Italy, extrapolating demographic and emotional insights from the content generated by users (Rahdari et al., 2017).

The work developed by Psyllidis et al. deserves a particular mention. They worked on SocialGlass, an integrated system combining public traditional records with different social media platforms (Twitter, Instagram, Foursquare, Flickr, Sina Weibo) and performed various analyses of urban dynamics in European cities (Psyllidis, 2016; Psyllidis & Bozzon, 2015).

The non-LBSN Instagram, the platform for the sharing of picture among the community of users, is the last I want to introduce because it is the scope of this research and it has been overlooked by the literature so far. Boy and Uitermark studied Instagram in multiple occasions to explore the potential of its social urban data, but they discarded the relevance of the results as they stressed the fact that users on the platform provide "selected" information users, therefore, not meaningful or representative of real life and dynamics (Boy & Uitermark, 2016; Boy & Uitermark, 2017).

I will argue in the following chapters the potential of these characteristics of Instagram of in the study of the city.

#1.2.3.2 Biases

The availability of a large amount of user-generated contents does not come without implications and difficulties. There are, of course, some general issues about privacy and ethic recurring in the literature (Ciucarelli, 2014; Mohammady & Culotta, 2014; Psyllidis, 2016; Olteanu et al., 2019 – just to name a few), but from a merely scientific perspective, the major concern interests the quality of the methodology and the presence of biases in the collection and analysis of datasets made of social urban data. However, it is important to remember the major issues not to disprove the approach in general but to be able to consider all the aspects that could negatively affect a research based of this typology of data while approaching it.

Almost all the publications concerned with social urban data mention the problem of biases (Cranshaw, 2012; Quercia, 2014; Mohammady & Culotta, 2014; Psyllidis, 2016; Shelton & Poorthuis, 2019). Despite this abundancy of sources, I will introduce the categorization proposed by Olteanu et al. in 2019 in a paper specifically dedicated to the issue (Olteanu et al., 2019).

First, there is the problem of the human representativeness of the sample in analysis, or in other word the "population biases". Social media are used generally from specific population groups and when the use is more socially distributed the mechanisms of use vary between younger and older users (Olteanu et al., 2019). Although this aspect is important to be considered when deriving conclusion from a dataset of social urban data, this should discourage the research as the use of social media is increasing and the sample in analysis is day by day more representative of the population.

There are also "behavioral biases" when the way users interact is not representative of their interests, aspects of their life or their actual behavior outside of the platform and differs from social media to social media, or from user to user (Olteanu et al., 2019). The presence of this kind of biases should be considered at the beginning of a research, wondering whether the final outcome of the project is in line with the behavioral aspects on the social platform of the deployed dataset/s. For example, we saw already – and we will go deeper in the matter in the following chapters – how the meaning of geotagging is different in LBSN and n-LBSN (see Tasse et al., 2014).

While the behavioral biases are distributed systematically among users of the sample on a certain social platform, some other kind of biases are more specific. The "content production biases", defined as «behavioral biases that are expressed as lexical, syntactic, semantic, and structural differences in the content generated by users» (Olteanu et al., 2019, p. 8) involve certain categories: users producing content in a different language, or users more likely to discuss certain topics because of cultural, religious, personal predisposition than other etc. Another typology of behavioral biases are the "linking biases", «expressed as differences in the attributes of networks obtained from user connections, interactions or activity» (Olteanu et al., 2019, p. 8), or, in other words, when the online interactions do not reflect the offline ones. The difference with behavioral biases in general is that, while the linking biases refer to the interaction of users, the first involves the behavior of users as individuals.

The "temporal biases" refers to many aspects of time in relation to social urban data. The changing of behavior, populations and system over time, for example, needs to be seen both on the long period – e.g. certain practices registered three years ago might have changed today – and on the seasonal level – e.g. how users behave and use the space in summer cannot be compared to the winter. In other words, the time granularity can be too defined to describe more complex dynamics, while it could also too coarse to interpret phenomena concentrated in a short span of time (Olteanu et al., 2019).

Last, the "redundancy" of some information can constitute another typology of bias in the realm of social urban data, due to the presence of fake accounts, bots, spam, repetition in the content posted etc. (Olteanu et al., 2019).

Apart from the typologies of potential biases not specific of social urban data but of data collection in general – like the acquisition and the processing of information – Olteanu et al. stresses some issues concerning sources at the origin, soome platform-specific biases: "functional biases", "normative biases", "external sources of bias", "non-individual accounts" (Olteanu et al., 2019). Although not to be overlooked, these biases are embedded in the platforms in a more neutral way. They become biases when the specificity of use of a social platform is misunderstood and the data used to infer conclusions based on wrong assumption: they are not biases in themselves, and they could actually constitute opportunities.

Social urban data need to be considered and weighted against all these different typologies of biases, as well as they need to be evaluated on the eight dimensions pointed out by Psyllidis of diversity, scale, timelines, structure, spatiotemporal resolution, semantic expressiveness, representativeness and veracity (Psyllidis, 2016, pp. 63-73) , in order to make the best employmet and correct interpretation of them.

BIBLIOGRAPHY

Amin, A. (2001). Spatialities of globalisation. *Environment and Planning, 34,* 385-399.

Amin, A., & Thrift, N. (2002). *Cities: Reimaging the urban*. Cambridge, England: Polity Press.

Argan, G. C. (2008). *Storia dell'arte italiana. Il Medioevo. Per le Scuole superiori*. Florence, Italy: Sansoni.

Baricco, A. (2018). *The Game*. Turin, Italy: Giulio Einaudi editore.

Bauman, Z. (2014). *Futuro liquido: Società, uomo, politica e filosofia* [Liquid future: Society, man, politics and philosophy]. Milan, Italy: Edizioni AlboVersorio.

Boy, J. D., & Uitermark, J. (2016). How to Study the City on Instagram. *PLoS ONE 11(6): e0158161. doi:10.1371/journal.pone.0158161*

Boy, J. D., & Uitermark, J. (2017). Reassembling the City Through Instagram. *Transactions of the Institute of British Geographers, 42*, 612-624.

Brambilla, M., Ceri, S., Daniel, F., & Donetti, G. (2017). Spatial Analysis of Social Media Response to Live Events: The Case of the Milano Fashion Week. I*nternational World Wide Web Conference Committee (IW3C2),* 1457-1462.

Cadwallar, C. (2013, May 5). Stewart Brand's Whole Earth Catalog, the book that changed the world. *The Guardian*, retrieved from https://www.theguardian.com/books/2013/may/05/stewart-brand-whole-earth-catalog#maincontent.

Castells, M. (2004). *La città delle reti* [The city of networks] (C. Rizzo, trans.). Venice, Italy: Marsilio Editori.

Chen, P., Shi, W., Zhou, X., Liu, Z. & Fu, X. (2019). STLP-GSM: a method to predict future locations of individuals based on geotagged social media data. International *Journal of Geographical Information Science, 33(12),* 2337-2362.

Cheng, Z., Caverlee, J., Lee, K., & Sui, D. Z. (2011). Exploring Millions of Footprints in Location Sharing Services. P*roceedings of the Fifth International AAAI Conference on Weblogs and Social Media,* 81-88.

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. P*roceedings of the Sixth International AAAI Conference on Weblogs and Social Media,* 58–65.

Ciucarelli, P., & Lupi, G., & Simeone, L. (2014). *Visualizing the Data City: Social Media as a Source of Knowledge for Urban Planning and Management.* Cham, Switzerland: Springer.

De Certau, M. (1984). *The practice of the everyday life* (S. Rendall, trans.). Berkeley, California: University of California Press. (Original work published as *L'invention du quotidien : Vol.1, Arts de faire* in 1980)

Del Bimbo, A., Ferracani, A., Pezzatini, D., D'Amato, F., & Sereni, M. (2014). LiveCities: Revealing the Pulse of Cities by Location-Based Social Networks Venues and Users Analysis. *Proceedings of the companion publication of the 23rd international conference on World wide web companion*

Forty, A. (2000). *Words and Buildings: A Vocabulary of Modern Architecture*. London, England: Thamse&Hudson.

Gandy, M. (2005, March). Cyborg Urbanization: Complexity and Monstrosity in the Contemporary City. I*nternational Journal of Urban and Regional Research, 29(1),* 26-49.

Graham, S. (1998). The end of geography of place? Conceptualizing space, place and information technology. *Prog Hum Geogr, 22,* 165-185.

Knorr Cetina, K. (2001). Postsocial Relations: Theorizing Sociality in a Postsocial Environment. H*andbook of Social Theory, 34,* 520-537.

Laman, H., Eluru, N., Yasmin, S. (2019). Using location-based social network data for activity intensity analysis: A case study of New York City. *The Journal of Transport and Land Use, 12(1),* 723-740.

Latour, B. (1999, May 1). On recalling ANT. T*he Editorial Board of The Sociological Review, 47(1),* 15-25.

Lefebvre, H. (1991) C*ritique of Everyday Life. Volume I: Introduction* (J. Moore, trans.). London, England: Verso. (Original work published as C*ritique de la vie quotidienne I : Introduction* in 1947)

Mitchell, W. J. (1999). *E-topia: "Urban Life, Jim But Not as We Know It"*. Cambridge,

Massachusetts: Mit Press.

Mohammady, E., & Culotta, A. (2014). Using county demographics to infer attributes of Twitter users. *Proceedings of the Joint Workshop on Social Dynamics and Personal Attributes in Social Media, June,* 7-16.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks. *International AAAI Conference on Web and Social Media, North America, July.*

Olteanu, A., Castillo, C., Diaz, F., Kiciman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. F*rontiers in Big Data, July 2019, 2,* art. 13.

Özkul, D., & Humphreys, L. (2015). Record and remember: Memory and meaningmaking practices through mobile media. *Mobile Media & Communication, 3(3),* 351-365.

Paulos, E., Anderson, K., & Townsend, A. (2004). UbiComp in the Urban Frontier. *Proceedings of Ubicomp '04 Workshops.*

Perez, C. (2009). Technological revolutions and techno-economic paradigms. *Working Papers in Technology Governance and Economic Dynamics, Working Paper No. 20*

Psyllidis, A., & Bozzon, A. (2015). Harnessing Heterogeneous Social Data to Explore, Monitor, and Visualize Urban Dynamics. 1*4th International Conference on Urban Planning and Urban Management (CUPUM 2015).*

Psyllidis, A. (2016). Revisiting Urban Dynamics through Social Urban Data: Method and tools for data integration, visualization, and exploratory analysis to understand the spatiotemporal dynamics of human activity in cities. *A+BE: Architecture and the Built Environment, 18.*

Quercia, D., Saez, D. (2014, April-June). Mining Urban Deprivation from Foursquare: Implicit Crowdsourcing of City Land Use. *PERVASIVE computing, April-June 2014,* 30-36.

Rainie, L., & Wellman, B. (2012). *Networked. The New Social Operating System*. Cambridge, Massachusetts: Mit Press.

Ratti, C. (2017). *La città di domani: Come le reti stanno cambiando il futuro urbano* [The city of tomorrow: How networks are changing the urban future]. Turin, Italy: Giulio Einaudi editore.

Rowe, C., & Koetter, F. (1978). *Collage City.* Cambridge, Massachusetts: The MIT Press.

Saffo, P. (1997, February). Sensors: The Next Wave of Innovation. *Communications of the ACM 40, 2,* 93-97.

Shelton, T., & Poorthuis, A. (2019). The Nature of Neighborhoods: Using Big Data to Rethink the Geographies of Atlanta's Neighborhood Planning Unit System. A*nnals of the American Association of Geographers, 109(5),* 1341-1361.

Silva, T., Carneiro Viana, A., Benevenuto, F., Villas, L., Salles, J., Loureiro, A., & Quercia D. (2018). Urban Computing Leveraging Location-Based Social Network Data: a Survey. A*CM Computing Surveys, 0(0).*

Simon, P. (2013). *Too Big to Ignore: the Business Case for Big Data*. Hoboken, New Jersey: John Wiley & Sons.

Tasse, D., Liu, Z., Sciuto, A., & Hong, J. (2017). State of the Geotags: Motivations and Recent Changes. I*nternational Conference on Web and Social Media (ICWSM),* 250-259

Virilio, P. (1991). *The Lost Dimension* (D. Moshenberg, trans.). New York City, New York: Semiotext(e). (Original work published as *L'espace critique* in 1984)

Xie, R., Chen, Y., Xie, Q., Xiao, Y., Wang, X. (2018). We Know Your Preferences in New Cities: Mining and Modeling the Behavior of Travelers. *IEEE Communications Magazine, November 2018,* 2-9.

Zook, M. (2000). The web of production: the economic geography of commercial internt content production in the United States. *Environ Plan A, 32,* 411-426.

#2

DE INSTAGRAM ELOQUENTIA

« The world we live in today is characterized by a new role for the imagination in social life. To grasp this new role, we need to bring together: the old idea of images [...]; the idea of the imagined community [...]; and the French idea of the imaginary (imaginaire), as constructed landscape of collective representation[...], now mediated through the complex prism of modern media.»

Arjun Appadurai [1]

1. as cited in Amn & Thrift, pp. 115-116

This chapter will explain why Instagram should be considered relevant for urban analyses, because of its intrinsic characteristics that well align with the configuration of society in the Information age and the necessity to shift the focus on the everyday in urban theory. The theories advanced in this chapter will serve as premise to the following parts of the dissertation where geotagged posts in the city of Turin are deployed to understand some dynamics of the city and propose a new mapping of it.

Instagram is a social media launched in October 2010, allowing users to share photo and video. Posts (the image or video) can contain tag of other users or of locations, and the caption under the post can contain "hashtag", words preceded by the symbol #, enabling users to search posts by hashtag used as keywords. The relationships between users could be unilateral, based on the logic of "following" – a user can follow another and have his/her more recent posts displayed in his/her feed, but the other user is not obliged to follow back (Hu et al., 2014). The diffusion of Instagram among population is rising and in Italy has reached in 2019 19 million of users, equal to a third of the population– I will go deeper into the demographics in the chapter of analysis of the dataset. (We Are Social, 2019).

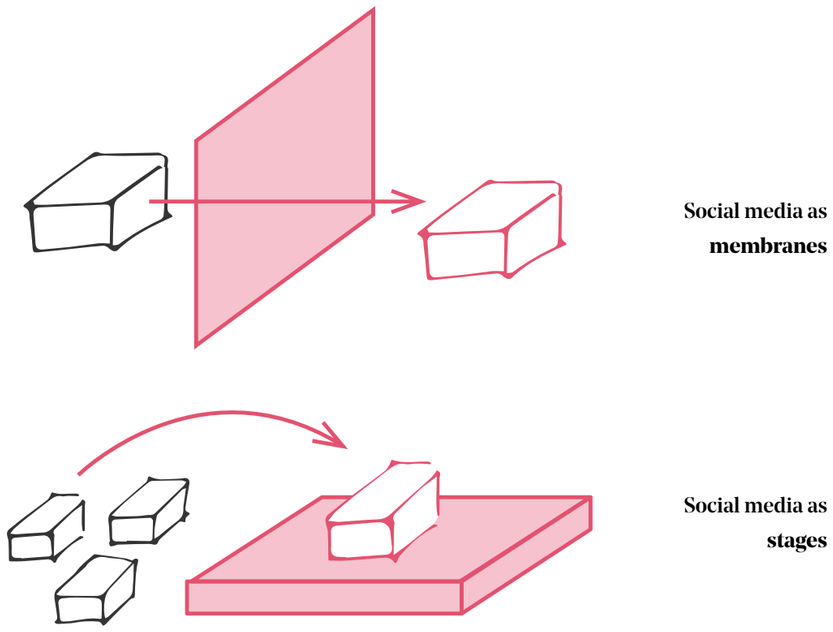


Figure 2_1
Conceptual representation of social media

Despite its growing popularity, Hu et al. in 2014 registered that little researches had been conducted on Instagram till that time (Hu et al. 2014). This is surprising when realizing that most of the literature has been focusing – and still is – on Twitter (see chapter 1.2.3.1 Social urban data – previous works), whose diffusion has been inferior to Instagram's at least for the last three years (We Are Social, 2017).
The conclusions of the already mentioned researches carried out by Boy and Uitermark (Boy & Uitermark, 2016; Boy & Uitermark, 2017) reveal the reason behind the minor interest of the literature towards Instagram: the information depticted on the platform do not reflect the reality. Users tend to select, filter, edulcorate the reality of their lives before posting content on Instagram and therefore the information potentially obtained are biased and can not be used as proxies of people's behavior in the city.

This perspective underrating the value of the information available on Instagram is founded on the idea that social media are nothing but a network of sensors – the humans behind the screen. Let us imagine a sensor in a city registering the air quality: it if the sensor would only keep record of the air quality at certain time, over or below a certain threshold, nobody would consider reliable the data provided. In the same fashion, if social media are conceived as a record of human activities, it is not scientifically relevant a social platform where only some of these activities are recorder and stored.

In the following chapters I propose, instead, a different reading and a change of perspective for why Instagram has indeed the potential for a certain *eloquentia*[2]. Some authors have already described social media as something else rather than a network of human sensors: some had seen them as "membranes" (De Waal, 2014) mediating and filtering the city reality and perception of users, others imagined them as "stages" (Rettberg, 2014), enabling users in a creative and productive process.

The scope of this research is to see Instagram beyond the dimension of a registering tool, but as an interpretive tool of ongoing dynamics and trends in society. These phenomena may be partially caused by the tool itself – people share their filtered reality because they have a tool to do so – but the diffusion of the tool prove that it responds

2. The title of the chapter plays with Dante Alighieri's "De vulgari eloquentia" [On the eloquence of the vernacular], an essay advocating the relevance for literature of the proto-vernacular Italian in literature. The eloquence is the quality of delivering a message and meaning and by this pun I would like to stress the potential of the information available on Instagram

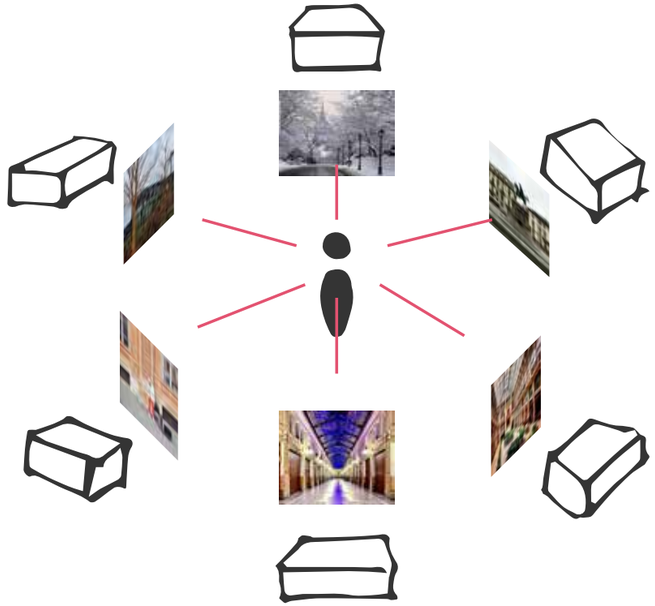


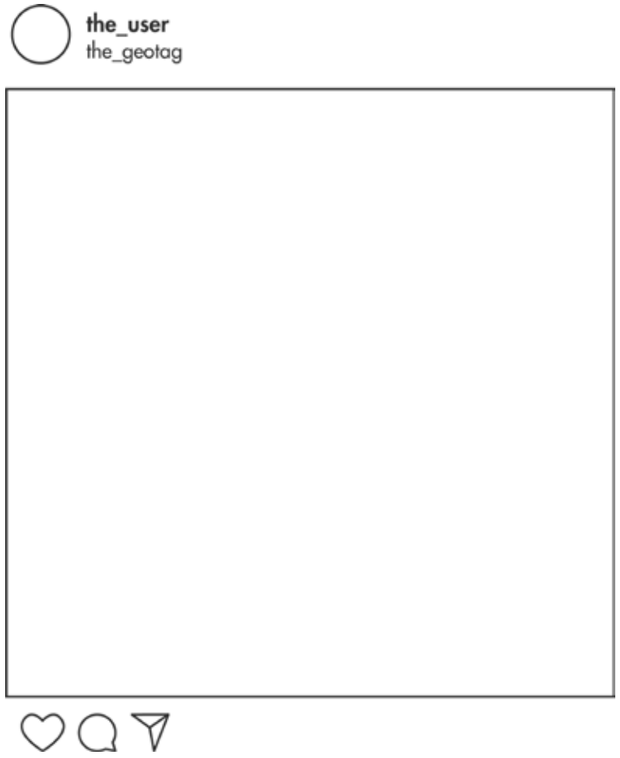
Figure 2_2
Construction of one's personal view

to a human predisposition of the spirit that probably always existed, but it was not easy to record before.

Alessandro Baricco described the experiences humans have with technology as "post-experiences" ("postesperienza" in Italian), meaning that all the apparent new experiences technologies have enabled humans to have are nothing else that something that used to be done in the past in a different way (Baricco, 2018). To limit the discourse to social media, the self-representation of individuals taking place on them has a written, a visual and a quantitative dimension and all of them have a precedent in the pre-digital history (Rettberg, 2014).

I will give a personal example to prove the same point. I have once dated a guy who, according to some common friends, would look at your number of followers on Instagram to decide whether to hang up with or not. As shallow as it may look like – and this guy, to be frank, was not indeed the brightest mind I ever met – it tells us something not about the world we live in, but how it adapted and took other forms. I cannot think of a century where the fascination of charism, money and popularity did not give extra points and a considerable advantage to any human being in the chasing of a partner whatsoever. But this is nature, female deer would reproduce with the strongest male deer, male peacocks show their tail to female, etc and all the other animals are born to do something similar. We already discussed how the alleged un-meaningfulness of digital interactions is not true, how they are actually real interactions between humans and how technology critics always saw a barbarization in the upcoming generation (Rainie & Wellman, 2010), but this is also a trend that always existed. Let us look back at the times f the Republican and the Imperial Rome, where commentators were referring the golden period of the *mos maiorum* ("way of the ancestors") to criticize the relaxation of morals of their contemporaries, and we will see the same dramatic complaining attitude crossing centuries till present times.

In this perspective, the filtering and the selection of some aspects of life on Instagram is a part of a process proper of the human being and the phenomenon is today ac-



cessible with datasets accounting for it. The kind of study I propose still remains in the field of urban computing and particular the one of "opportunistic sensing". This last concept has been defined as the use of data gathered for a specific purpose to analyze them in another context to come to new conclusions (Ratti, 2017): geotagged Instagram posts could not be seen as data, but in a broader sense data is what we look at to derive information and build knowledge.

The intrinsic characteristics making Instagram relevant for urban studies are the image-oriented approach of the platform and the publicness of it.

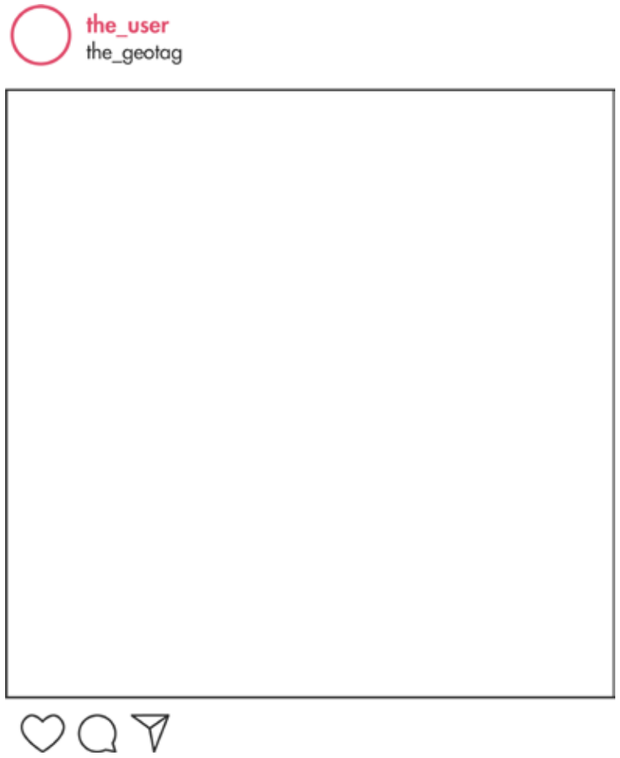
The image carries multiple meanings on the platform that I will explore int the following chapters, which partially accountable to the new role of the image in the Information age in general. In fact, it has an interactive and programmable dimension on screens, so it «no longer functions as a (political and iconic) representation, but plays a vital role in synchronic data-to-data relationships» (Hoelzl & Marie, 2015, p. 3-4). We are surrounded by images, images created for all kind of purposes – advertising, art, representation, memory – and displayed on a huge variety of "hard" screens – like posters, books, magazines, ban – and "soft" screens – television, telephones, computers, tablets. The image is not like an opinion like on Twitter, a text that was elaborated, needs more context and to interpreted, the image on Instagram is immediate.

The aspect of publicness is embedded in the platform. As I said, by default profiles are public and images are shared to and are accessible by the all world. It is different from the logic of Facebook where the information have different degrees of accessibility according to the relation between users. Clearly, it is impossible that some users have chosen the option to make their account private, but I will discuss this aspect in the chapter dedicated to the user.

The following chapter go deeper into the relevance of geottaged Instagram posts for urban studies by breaking down and analyzing the three main components of a post: the user, the image, and the geotag.

#2.1

THE USER



I stressed in the chapter on cities, the relevance of the everyday in the contemporary urban framework, but also the difficulty to capture this everyday (Amin & Thrift, 2002). In fact, it happens “on the street level” and the fascination of the mystery shrouding the walking man is a topos in the literature (see for example De Certeau, 1980/1984).

The recording of the aspects characterizing the everyday life of cities on the street level have a tradition in the practice of what is known as “flânerie”, the idle walk through the streets and the recording of the phenomena happening around the observer. Walter Benjamin referred to flânerie, stressing the receptiveness of this kind of user of the city and the overcoming of the normal “distract use of architecture (Amin & Thrift, 2002), but there many accounts outside of the nonfiction literature. There are Joyce’s “Ulysses” for Dublin, Salinger’s “The Catcher in the Rye” for New York, the very idea of flânerie is Baudelaire’s and he was referring it to Paris.

The flânerie in itself is practice that could technically well describe the dimension of the everyday in the city but, evidently, these series of impressions are always a person’s account, biased by his own view – yes, “his” because these accounts were produced mainly by men – and the subjective perspective does not authenticate his findings (Amin & Thrift, 2002). Therefore, as much as poetic and yet scientifically relevant the art of flânerie could have seem, it was always overlooked for its obvious intrinsic bias.

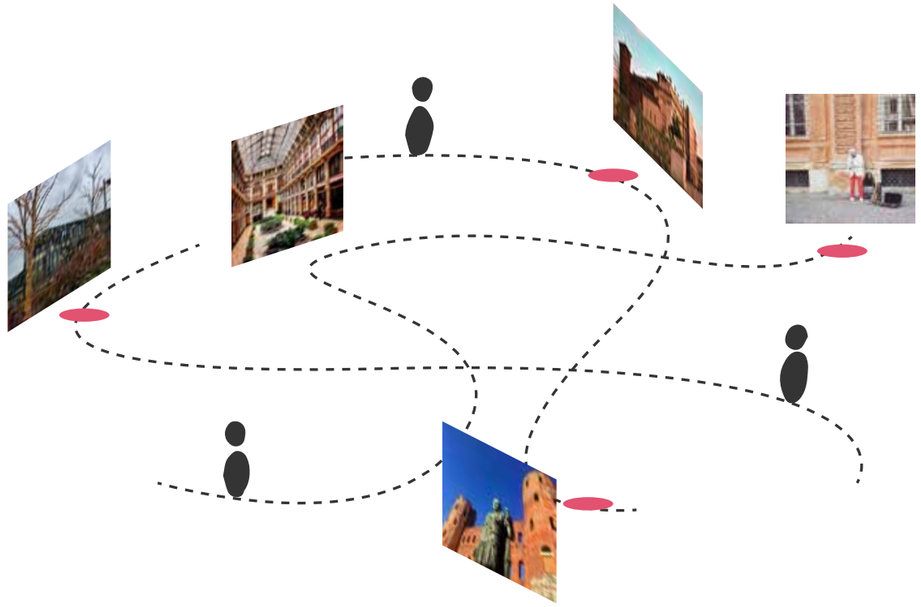
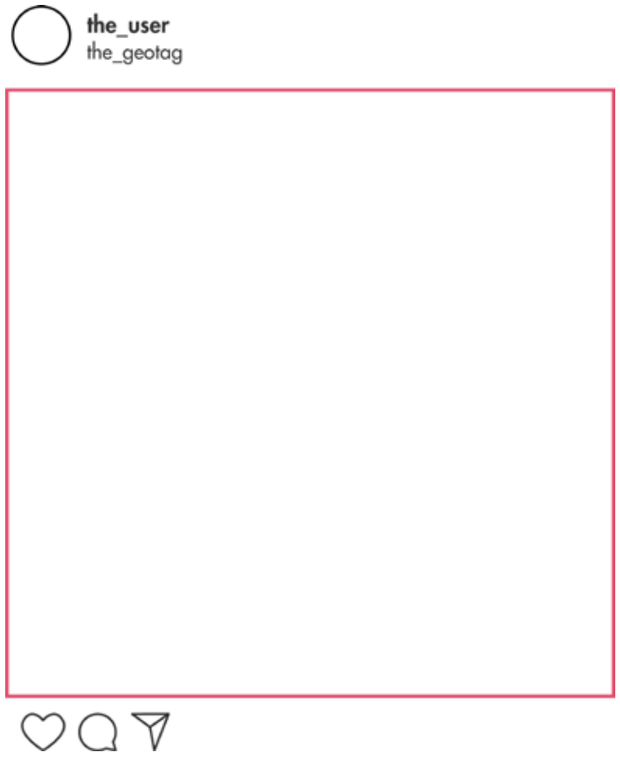


Figure 2.1_1
Objectivity deriving from multiple accounts

The practice of flânerie, however, has a corresponding, an equivalent "postexperience" – to use Baricco's definition – in contemporary society: the use of Instagram. If we think about it, the users of Instagram are flaneur: they experience the city during their everyday life till their attention is attracted by a specific aspect of the urban life, they record it and share it with the community of users. The biases deriving from the one-sided perspective of traditional flaneurs' accounts are now absorbed by the pervasiveness of the practice among the whole audience of the social platform. The intrinsic subjectivity of flânerie becomes objectivity because this «can be pursued only by multiplying the point of observations. The more numerous and partial are the perspectives from which a phenomenon is considered, the more objective and impartial will be its observation» (Venturini, 2010, as cited in Ciucarelli, 2014).

For Venturini objectivity is, then, directly proportional to the number of points of observation – and Instagram users are indeed a lot – and to the partiality of these perspectives. I will discuss the relation between Instagram images and imagination in the following chapter, but I would like to anticipate that this process of pervasive flânerie described above is in line with the concept of the individual as an actor and a producer in the contemporary society. In the postsocial environment the imagination is no longer an imposed category, but it has been replaced by an imagination centered on individual, self constructing significations and processes (Knorr Cetina, 2001). The sum of the contents produced by subjective self-centered users is therefore relevant and can be taken as an objective account.

With a political and historical metaphor, we could see this process as the "Storming of the Bastille" of urban representation and story-telling. The common people, after centuries of imaginaries imposed by experts, artists, the ruling class in general take over and reappropriate themselves of the right to tell their own story. Again, as I said for the Castell's general discourse about collective identities (see the chapter 1.2.3 The subject as actor), the user might not be completely independent in the development of his/her aesthetic preferences – even as far as regard urban flânerie and urban story-telling – however s/he has access to a wide variety of aesthetic systems among which to choose his/her own.



#2.2

THE IMAGE

From a more general perspective, the image has its the relevance in the network society particularly when it is produced and share on social platform. In fact, it «coincides with the screen, which is its current display platform but also its form», being at the same the «moment of network access» and the «actualization of the networked data» (Hoelzl & Marie, 2015, p.6, p. 3). This is line with the idea of social media as "membrane" between the virtual and the real (De Waal, 2014).

To return to the specificity of this chapter, the previous identified the user posting a geotagged image on Instagram as a flaneur, sharing his/her own perception of the everyday life of the city. This chapter will analyze the relevance of the product of this experience of the city: the image.

As recalled before, the major critic towards the use of Instagram data for the interpretation of the urban phenomena concerns the selection operated by the user in the choosing of the content to post, the consequent distorted representation of reality and the biases in the deployment of this type of data (Boy & Uitermark, 2016; Boy & Uitermark, 2017).

Does it make sense, however, to underrate the power of information on Instagram to a fake, distorted reproduction of reality? What is indeed reality? Is the world we actually live in, with its misery, sordidness or should we rather believe in a world we imagined

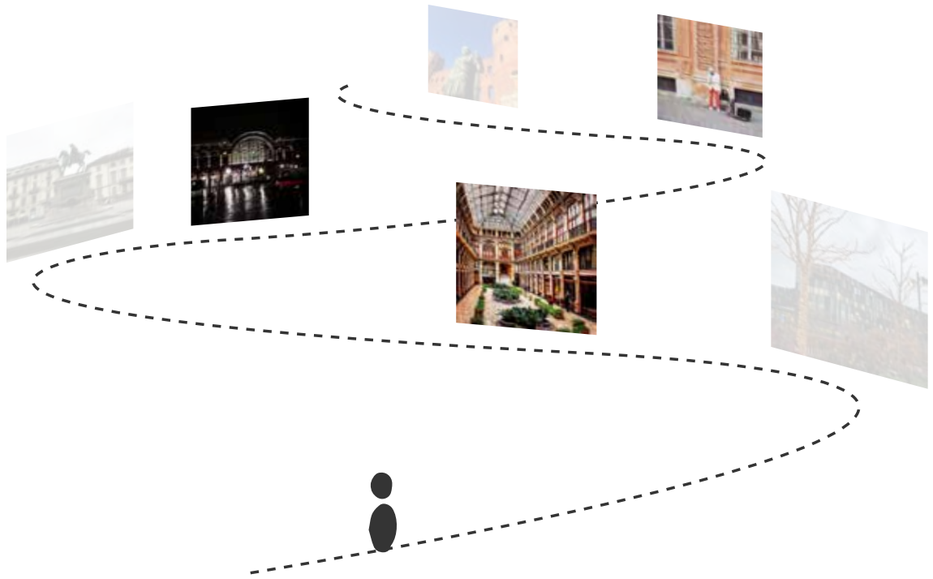


Figure 2.2_1
The importance of selection

starting from reality? Is not it the root of the urban project the power of imaging from reality something beyond reality, the project in its literally meaning of projection towards something else?

The theorical assumption of this dissertation is the relevance of geotagged posts for urban considerations and the disproval of the critics is founded on two main claims, validating the intrinsic meaning of georeferenced Instagram contents. The first regards the construction of the image of the city in itself and the second is overcoming of the alleged dichotomy virtual-real from a philosophical perspective.

To start with the construction of the image of the city, Lynch and his famous "The image of the city" are of the utmost relevance. He describes "environmental images" as «the result of a two-way process between the observer and his environment» (Lynch, 1960, p. 6). The selection of certain element is at the root of the forming of these images as they are made by three main components. The first is identity, stressing the individuality behind the construction of that image, the second is the structure, as images should put in relation according to some logic the producer and the object, and the third is meaning, practically or emotionally related to the observer (Lynch, 1960). It emerges clear that selection is the foundation of the construction of the image of a city for an observer. Therefore, the blamed selectiveness of Instagram contents is indeed what makes them extremely even more relevant for analysis, rather than if users shared every single space they encounter during their daily lives.

The selection in the representation of the city is linked also with the idea that the reality is transfigured in more pleasant fashion by users, deliberately excluding the unpleasant elements of experience. (Boy & Uitermark, 2017). This could as well be extended to all the typology – according to Hu et al. they are eight: friend, food, gadget, captioned photo, pet, activity, selfie and fashion (Hu et al., 2014) – of photos shared on Instagram. Under this perspective the content shared – the virtual – would be "fake", because it does not depict the nature of the object – the real.

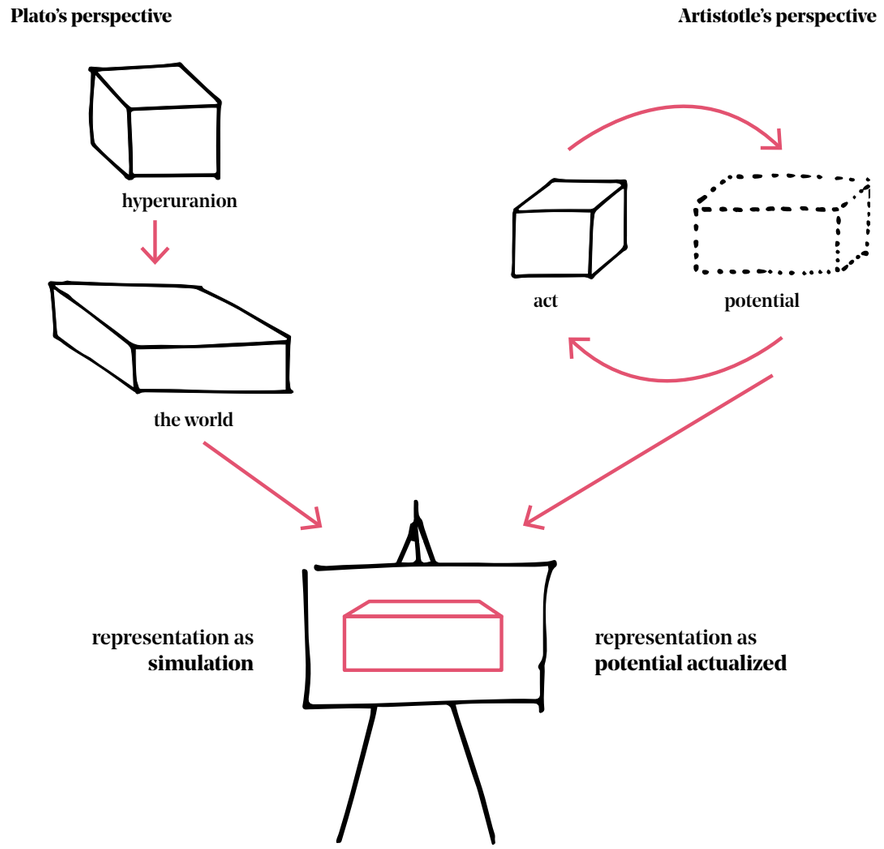


Figure 2.2_1
Virtual vs real in Plato's and Aristotle's conceptions

Focusing only on the representation of venues, de Souza e Silva and Sutko developed a new interesting perspective on the meaning of "virtual" for location-based technology. The philosophical conception that dichotomize the virtual – seen as a representation – from the real is based on a tradition that dates back to Plato and it was later developed by Baudrillard. For Plato the representation is a copy of the real and has less value than it – given the fact that in Plato's perspective the world is also a reproduction of the hyperuranion, the representation is a copy of second generation. However, the digital representation, for example in the case of augmented reality, adds information to the venue that are not present elsewhere, making the Platonic perspective of doubtful utility to describe the relation between virtual and real in the Information age. The Platonic concept was expanded by Baudrillard, who saw the virtual as a simulation, a wall concealing the reality. This could seem reasonable while thinking of GPS navigation system, making drivers follow a series of direction alienating themselves from the reality around them. However, the system works in connection to reality, adapting to changes of it along the route, a dynamicity and a bilateral dependency of virtual and real absent in the Baudrillardian perspective. The Platonic and Baudrillardian philosophies, reducing the virtual to either simulation or representation, evidently fail to describe the complex relationships occurring between the real and the virtual in locative media (De Souza e Silva & Sutko, 2001).

The position theorized by Aristotle and further developed by Deleuze, however, could better explain those dynamics. Aristotle's philosophy is built on the ideas of "act" – what an entity is – and "potential" – what an entity could be and the passing from the first status to the second happens through the process of "becoming": the entity remains the same and just changes its attributes. Under this perspective «while the virtual corresponds to the potential state, which is actualized through the act of becoming, the actual is realized potential» (De Souza e Silva & Sutko, 2001, p. 32). This was further developed by Deleuze who «considers the virtual as desire, as a potential for creation. […] there is more beyond what we see in the world (reality), […] the real encompasses

both our actual world and the potential to produce new realities» (ibidem, p. 33) The virtual, then, is considered real and there should not be any dichotomy between the two. «The real is pure immanence, containing virtual (potential) realities and actual realities. The potential to be actualized and differentiated into diverse realities is what makes the virtual part of the real rather than opposed to it» (ibidem).

What does this mean for geotagged Instagram posts? «Following a Deleuzean logic, every location-based application can be envisioned as a virtuality, which is differently actualized by each viewer. Therefore, they are per se fully potential entities that can only be completed through user interaction. Each user, in turn, actualizes the application differently, revealing some (but not all) aspects of its potentiality» (De Souza e Silva & Sutko, 2001, p. 33).

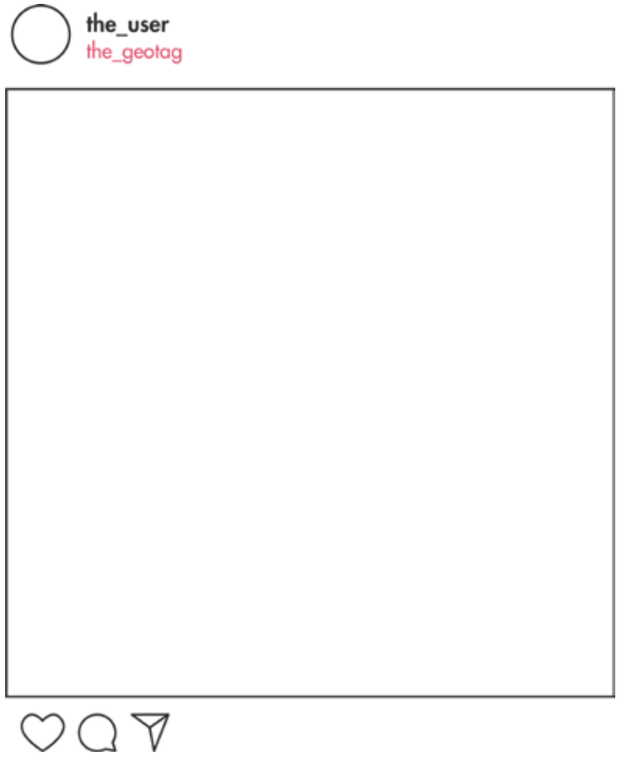
In conclusion, geotagged posts are images of locations with a clear representative purpose, but they are also project/projection of that location by the user and what has been described as a representational bias, it is actually the representation of a potentiality, unveiled by the user's imagination.

Many authors stressed the role of imagination in contemporary social life and Instagram geotagged posts could constitute «a constructed landscape of collective representation» (Appudarai, as cited in Amin & Thrift, 2002, p. 115). Again, the constitution of worlds beyond reality could constitute another of Baricco's "postexperiences". Foucault's eterotopias (Foucault, 2005/2006), for example, are the juxtapositions in a real place of places that would be separated or incompatible, "other places" that every society created in its own way to escape from reality. Foucault could not imagine something like Instagram where an eterotopia is created in parallel and from the existing reality. A magical realm of vacations, stylish ways of life, beauty, a wonderland for the human being, where everyone can create through the logic of following its own world of ambitions, envy and self-approval. A fantastic world that encounters the real one, the city, the spaces of it and that uses them, makes them part of it, unveiling the possibilities of potential they have.

This imagination is «no longer mere *fantasy* (opium for the masses whose real work is elsewhere), and no longer simple *escape* (from a world defined principally by more concrete purposes and structure) and no longer mere *contemplation* (irrelevant for the new forms of desire and subjectivity), the imagination has become an organized field of social practices, a form of work (both in the sense of labour and of culturally organized practice) and a form of negotiation between sites of agency ('individuals') and globally defined fields of possibility.» (Appudarai, as cited in Amin & Thrift, 2002, pp. 115-116). This new subject-centered imagination (Knorr Cetina, 2001), as far as concern the city, has its epiphany in the representation users give of it with Instagram posts.

#2.3

THE GEOTAG



In the previous chapters it was explained how the act of sharing a picture of a city venue on Instagram is an act of flânerie on the account of the user, producing his/her version of the intrinsic potential state of that specific venue. A question is unanswered, though: why do users decide to geotag their posts? Could they not simply post it without any geotagged reference, since the meaning of the act would be unaltered?

As we saw, some social platforms are Location-Based and therefore the sharing of the spatial information is embedded in the system: Foursquare, for example, could not work without its check-ins. As far as regard non-LBSNs like Instagram, the question is more complex.

A recent research conducted by Tasse et al. proved by surveying a sample of social media users that geotagging contents on non-LBSN is not casual, but is a conscious choice. The reasons behind the act are mainly two: first, showing off to people where the user has been, and second, to keep family and friends updated (Tasse et al., 2017). Considering the fact that family and friends are, presumably, people too, and the vaguely defined "people" to which users want to show off where they have been do get updated – while probably green with envy –, I will consider the outcome of this research as one: the act of geotagging is a way of distinguishing oneself from others.

The act, far from being of the utmost vanity and shallowness, has social and relational

implications and motivations. As pointed out by Bourdieu, distinguishing oneself from others is part of the attitude of humans. The aesthetic choices made by individuals have a classist dimension, as they are made in opposition with the choices made by individuals belonging to other social classes (Bourdieu, 1979/1996). The aesthetic choices in Bourdieu's perspective are influenced by society, and consequently the reason behind geotagging one's posts could be seen outside of the user's sphere of self-determination. However, as we saw, the imagination in Information are subject-centered and the identification in collective identities a conscious choice: distinguishing oneself with marking one's passage in a place should be still interpreted as a social act, but through which the individual testifies his/her belonging to a certain social milieu.

Beside the aspect of distinction, however, Özkul and Humphreys introduced another meaning to the act of geotagging one's content on social media: memory. The user shares locational information in order to keep track of his/her passage in certain venues, associating them with his/her life and deploy them in writing of an autobiography. A push, then, to preserve one's history, to store memories connected to places to recall them in the future (Özkul & Humphreys, 2015).
Even the aspect of memory, generally referred to a more intimate and personal sphere of the self, when takes place with the sharing of those information can be regarded as a social act. Memory, in fact, has been theorized as a collective construct and associated with the dimension of the group identity (Halwbacks, 1950).

In conclusion, the act of geotagging oneself in a venue with an Instagram post is, in the last analysis, a social act in the construction and definition of one's identity, through the category of distinction and memory. The claim does not seem to bold, since, as we saw in the previous chapters, digital interactions among individuals have been proved to be consistent and the dichotomy between virtual and real resolved.

To validate the claim outside of the field of self-representation on social media, the definition and construction of one's identity through space has been proven in the realm of psychology in general.
Place can be considered, indeed, a social category and it is subjected to the same rules of social identification. The four principles of identity in Breakwell's model – continuity, self-esteem, self-efficacy and distinctiveness – were examined in relation to environment and it was proved that place is an element of identification, and this latter always has location implications (Twigger-Ross & Uzzell, 1996). The experience of the environment has the double dimension of a place and of a social group of inhabitants, the attachment to a place and the identification in it varies among individuals according to factors like the educational level, the participation in local activities and other aspects related to pure sociality (Rollero & Piccoli, 2010).

The act of geotagging one's posts is, then, a conscious and deliberate choice, made in relation to other individuals and part of a wide and complex process of construction of one's identity through environment.

#2.3

CONCLUSIONS

In the previous chapters we saw how the geotagged Instagram posts can constitute a powerful source to understand the contemporary city in its complexity.

The user of the platform is a flaneur, experiencing and recording the everyday life of the city. His/her record is not a simple representation, but it is embedded of the meanings and the projections the user place in a venue/location. The geotagging is also a social action, shaping the user's identity.

Geotagged Instagram posts are therefore an act through which an individual construct his/her social and individual identity through the reference to a physical space s/he immortalizes and sublimes.

These considerations makes therefore unrelevant the fact the some users have private profiles: their sharing limited to their community of friends, even if relevant for them, does not have relevance for the scope of this research if is not shared with the rest of community, if it is not a public and social act.

BIBLIOGRAPHY

Amin, A., & Thrift, N. (2002). *Cities: Reimaging the urban*. Cambridge, England: Polity Press.

Bourdieu, P. (1996) *Distinction : A Social Critique of the Judgement of Taste* (R. Nice, trans.). Cambridge, Massachusetts: Harvard University Press. (Original work published as *La Distinction:Critique sociale du judgement* in 1979)

Boy, J. D., & Uitermark, J. (2016). How to Study the City on Instagram. *PLoS ONE 11(6): e0158161. doi:10.1371/journal.pone.0158161*

Boy, J. D., & Uitermark, J. (2017). Reassembling the City Through Instagram. *Transactions of the Institute of British Geographers, 42, 612-624*.

Castells, M. (2004). *La città delle reti* [The city of networks] (C. Rizzo, trans.). Venice, Italy: Marsilio Editori.

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media, 58–65.

Ciucarelli, P., & Lupi, G., & Simeone, L. (2014). *Visualizing the Data City: Social Media as a Source of Knowledge for Urban Planning and Management*. Cham, Switzerland: Springer.

De Certau, M. (1984). *The practice of the everyday life* (S. Rendall, trans.). Berkeley, California: University of California Press. (Original work published as *L'invention du quotidien : Vol.1, Arts de faire* in 1980)

De Souza e Silva, A., & Sutko, D.M. (2001). Theorizing Locative Technologies Through Philosophies of the Virtual. *Communication Theory, 21,* 23–42.

De Waal, M. (2014) *The City as Interface: How New Media Are Changing the City*. Rotterdam, Netherlands: NAi010 Publishers.

Foucault, M. (2006). *Utopie. Eterotopie* [Utopias. Eterotopias] (A. Moscati, edited by). Naples, Italy: Cronopio. (Original work published as Die Heterotopien/Der utopische Koerper in 2005)

Halbwachs, M. (1950). *La mémoire collective* [The collective memory]. Paris, France: Les Presses universitaires de France.

Hoelzl, I., & Marie, R. (2015) *Softimage : Toward a New Theory of the Digital Image*. Bristol, Englang : Intellect.

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What We Instagram: A First Analysis of Instagram Photo Content and User Types. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media,* 595-598.

Knorr Cetina, K. (2001). Postsocial Relations: Theorizing Sociality in a Postsocial Environment. *Handbook of Social Theory, 34,* 520-537.

Özkul, D., & Humphreys, L. (2015). Record and remember: Memory and meaningmaking practices through mobile media. *Mobile Media & Communication, 3(3)*, 351-365.

Ratti, C. (2017). *La città di domani: Come le reti stanno cambiando il futuro urbano* [The city of tomorrow: How networks are changing the urban future]. Turin, Italy: Giulio Einaudi editore.

Rettberg, J. W. (2004). *Seeing ourselves through technology: How We Use Selfies, Blogs and Wearable Devices to See and Shape Ourselves*. Basingstoke, England: Palgrave Macmillan

Rollero, C., De Piccoli, N. (2010, June). Place Attachment, Identification and Environment Perception: An Empirical Study. *Journal of Environmental Psychology, 30(2),* 198-205.

Tasse, D., Liu, Z., Sciuto, A., & Hong, J. (2017). State of the Geotags: Motivations and Recent Changes. *International Conference on Web and Social Media (ICWSM)*, 250-259

Twigger-Ross, C. L., & Uzzell, D. L. (1996). Place and Identity Processes. *Journal of Environmental Psychology, 16,* 205–220.

We Are Social. (2017). *Digital in 2017*. Retrieved from https://wearesocial.com/it/blog/2017/01/digital-in-2017-in-italia-e-nel-mondo

We Are Social. (2018). *Global digital in 2018*. Retrieved from https://wearesocial.com/it/blog/2018/01/global-digital-report-2018

We Are Social. (201). *Digital in 2019*. Retrieved from https://wearesocial.com/it/blog/2019/01/digital-in-2019

#3
METHODOLOGY

INSPIRATION

This dissertation is the final work of five years dedicated to the study of architecture and urbanism in the Polytechnic of Turin, city where I lived and studied. Design is a skill every student is supposed to develop and I had the chance to have a variegated learning experience: I participated as a student in design studios both in my university and abroad during two Erasmus programs, I took part to international workshops, I also had the chance to teach as an assistant in three design studios. In all these expe-riences there has always been a haunting theme that seemed impossible to escape: the context.

I will not go deep into the development of this concept, because it will take a long space and it is not really the point. The context in its multidimensional aspects is – evi-dently – a set of important elements to consider in the design of a building or anything in relation to the urban space. What I really could not understand was the veneration of the context, the conception of it as a gold mine of information with which support even the most absurd design hypotheses, in particular when referred to sociological aspects of a design,

My doubts reached their apical points when I had to battle a few months ago with the students I was tutoring for a design studio. They were supposed to develop an urban analysis of the context around the area of their case study to help them with the design (the design studio was in the San Salvario neighborhood in Turin). Apart from some questionable patriarchal conceptions of society ("We think a park is needed as there are some schools around: MOMS can wait there for their children to go out from school" – and this group was one of the most feminist), there were some conclusions extremely wrong to me.

"We saw there are some medical centers in the area… the users for the apartments we are designing will be then doctors". Apart from the fact that I still have to understand how a doctor's house is different from another professionals', the main problem behind it was the misconceptions I discussed in the previous chapters: the new localism, the face-to-face interaction culture, the idea of community linked to neighborhood etc. To the majority of students seemed natural to create a "village" of doctors, or students, or young professionals, or commuters – everyone had his/her own vision of the main type of human fauna populating San Salvario – because it seems right that an neighborhood is kept together by a cohesive and unison typology of community of users.

My questioning of the idea of the context defining communities was based also on some empirical and personal observations.

First of all, I visit places around the city of Turin, places that are sometimes really faraway from each other. Still, I manage to find there people I know for different reasons and there is no social event I attended where I did not have to shake a hand, give a kiss to someone or make dreadful small talk. Am I part of a community? It would seem so. And, although my supposed community is not located in a neighborhood of the city or a village, still there are some venues where the people I know go and some where they do not: I am part of a network of individuals eradicated from a single physical area but taking place into a series of physical spaces. Could these venues be part of a network too?

Second, I see juxtaposed venues all around me that do not fit with the idea of proximity I tried to uproot from my students. Let us take as an example Piazza della Repubblica in Turin. On one side, there is the historical market of Porta Palazzo, with its exotic and local products, its variegated community of customers made of immigrants, old local people, young pretentious hipsters etc. On the other side, there is the new Mercato Centrale, where fashionable high-class – and expensive – restaurants opened their trendy street-food – and cheaper – spin-off. If you go there you will be able to see the audience of a shopping mall: young families with noisy kids, people from other cities "who wanted to try something new that night", another kind of pretentious hipsters etc. Are the Porta Palazzo market and the Mercato Centrale close? Yes, they are next to each other. Do they share the same community of people? I do not think so.

My question was then: could there be a method, a way to go beyond the concept relating venues and communities on the idea of proximity and to map the actual community and neighborhood of the city?

“Where do people geotagged their posts?”,
“When?”, “How much?” etc.
In short where and when this flânerie act
through which individuals build their
identity takes place?

#3.1

MAIN IDEA

In the previous chapters I illustrated the relevance of social urban data in the study of the city and, in particular, the potential of using Instagram geotagged posts as a privileged point of observation of city dynamics.
The purpose of my research was the collection of data concerning the city of Turin and the visualization of what we could refer to as “the Instagram city”. In short, starting from the theoretical sociological assumptions of the previous chapter, the main idea was to reveal – or better “to unveil” – ongoing dynamics and structures organizing the life of individuals in relation to the physical space and venues of the city, with particular reference to Turin. The research goals were mainly divided into two main fields, as presented later in the dissertation.
On one side, there is the more quantitative dimensions of data visualization. The questions to be answered were something like “Where do people geotagged their posts?”, “When?”, “How much?” etc. In short where and when this flânerie act through which individuals build their identity takes place.

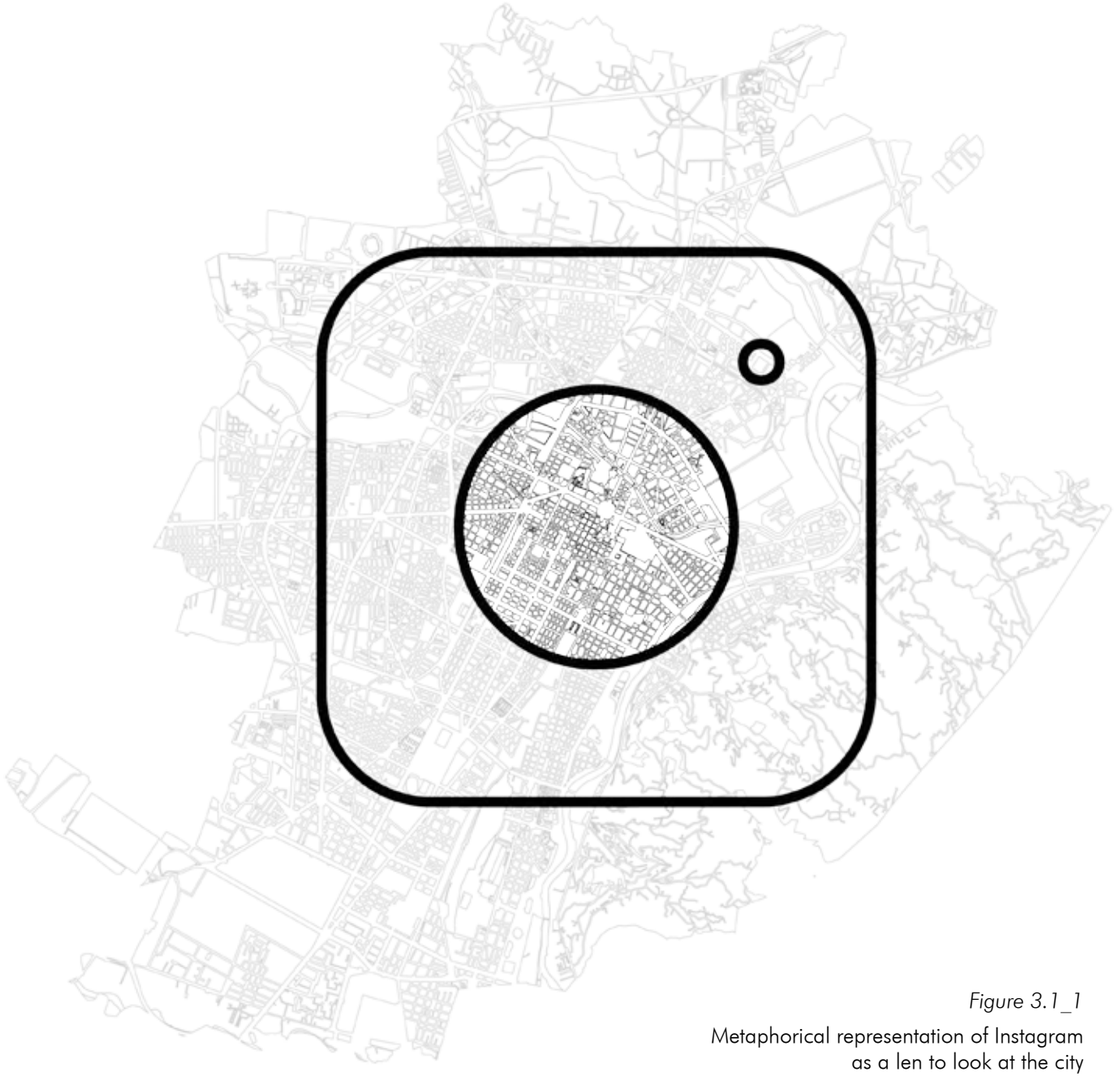


Figure 3.1_1
Metaphorical representation of Instagram
as a len to look at the city

On the other, there is the dimension of the networks linking different venues. We saw in the introductory chapters how society organizes itself on more fluid and dispersed communities and the contemporary city should be regarded as the site where this network society take place rather than the place of localism and face-to-face social interactions.

Previous works like Cranshaw's Livehoods project (Cranshaw, 2012) investigated the dimension of neighborhoods basing the belonging of their venues on social affinity, but still including the physical proximity in the parameters. My assumption was the possibility to map "diffused neighborhoods of venues" held together only by social affinity of the individuals identifying themselves into them and derive from this process relevant information for the understanding and the design of the city of Turin.

#3.2

SOCIAL RELEVANCE AND BIASES

Among the typologies of biases in the use of social data presented by Olteanu et al. (Olteanu et al., 2019) the only one potentially relevant to this research is the "population bias". All the other categories either do not apply or constitute the actual core of the theoretical premises: in the case of "behavioral biases", for example, the way users interact on Instagram through geotagged posts was described in the previous chapter as what makes the social platform relevant to urban studies.

Instead, the population biases could constitute a problem for the relevance of the information derived from the social platform. It is a conception universally accepted and shared among people that Instagram – and other social media like Facebook – are used mainly by young individuals: the knowledge would be, then, built on data produced only by a sample of population limited in size and not representative of the whole.

It is not possible to determine precise demographic statics on a specific set of Instagram users: the information is not available, and it can be only inferenced from the short bio provided by the users him/herself and/or deploying machine learning technics analyzing the profile picture. Despite this, however, it is possible to evaluate general demographic statics and to assume that a smaller sample of the whole population of a social platform would be a scaled representation of it.

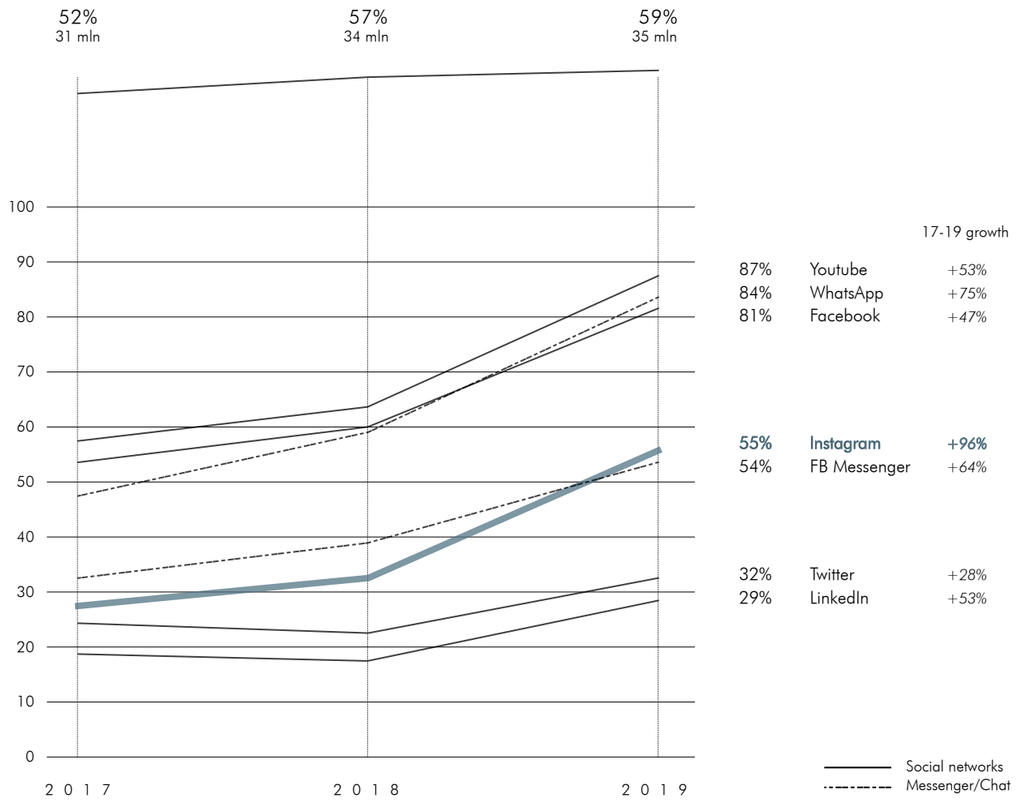


Figure 3.2_1
Growth of social media in Italy in the last two years

1. Data retrieved from the latest three reports of the agency We Are Social. (We Are Social, 2017; We Are Social, 2018; We Are Social, 2019)

As it can be seen in *[Figure 3.2_1]* [1], the spread of social platforms in Italy increased consistently in the last period, reaching almost the 60% of the population of country in 2019 – approximately 35 million people. Among the most popular social platforms, Instagram is the second most used social media after Facebook, but its spread knew the highest increase in the last two years, almost doubling the percentage of users declaring to deploy it on a regular basis. In 2019 the 55% of Italians using social platforms declared to be on Instagram: 19 million of total users accounting to a third of the Italian population.

The other common perception about the usage of social media like Instagram and related data, it is that it would be representative only of the younger age groups of the population. Although no data on the specific demographics of Instagram users are available, We Are Social in its latest report of 2019 provided the segmentation according to sex and age group for a combine audience of Facebook, Instagram and Facebook Messenger. (We Are Social, 2019)

The comparison between the percentage of the total Italian population and the audience of social media segmented in different age groups leads to interesting considerations. The 0-24 age group is apparently less represented on social media, but in this case the reason is that the children have not access to these platforms and if the segment of population above 14/15 years old was taken alone, it would give the same proportion of the following age groups. In fact, the age groups 25-34 and 35-44 tend to be more represented on social media, being respectively twice and one and half larger than the corresponding segment of population. It is more interesting to see what happens in the following two age groups. The segment of population aged 45-54 is in percentage more present on social media than in the actual society and people aged 55-64 are only slightly less numerous. Instead, as expected, the presence of users aged over 65 years old is not consistent in comparison to the actual number of people.

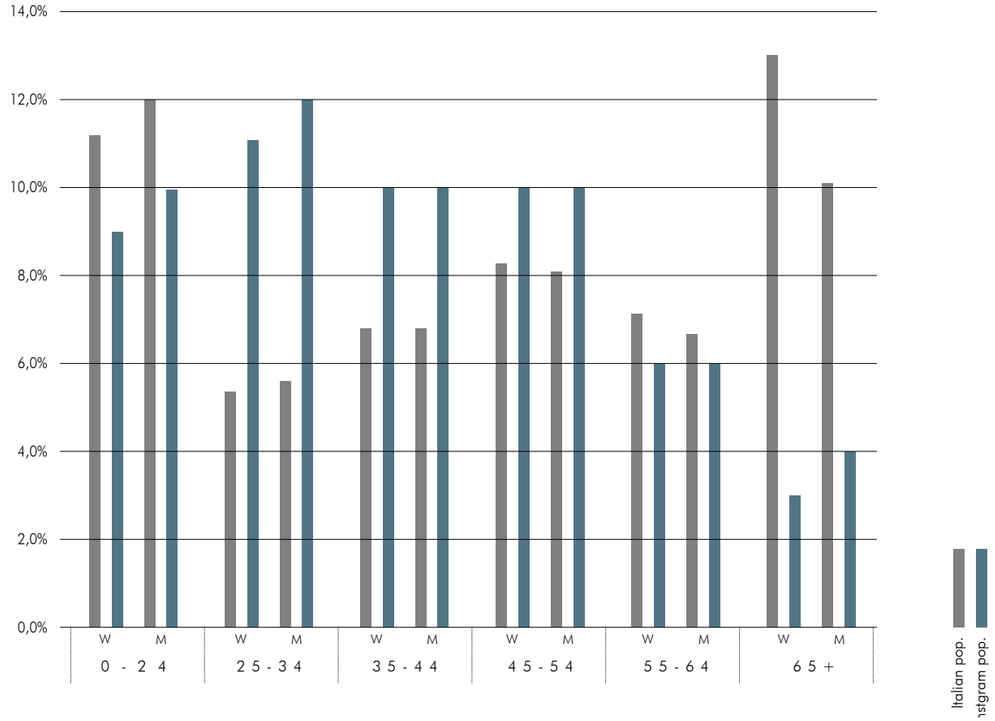


Figure 3.2_2
Italian population for age range compared to Instagram's

Although it is not possible to state that the population of Italy is equally present on social media, it is also a mistake to believe in the opposite. Younger age groups tend to be more represented but, surprisingly, the "over-representativeness" stretches up to not digital natives aged 54 years old, and the next age group is almost corresponding to the actual situation.

In conclusion, two main statements can be derived to validate this research work. First, the work does not aim to be comprehensive and absolute for the present times, but it wants to propose a methodology to be implemented in a – near? – future and stimulate the discussion on different ways to look at the city from new perspectives. Second, despite the population biases are still determinant in the evaluation of the information deriving from social media and Instagram, even today we can look at these data and see them as a representation of the active part of the population with a imbalance towards younger age groups. *[Figure 3.2_2]* [2]

2. Data retrieved from the latest report of the agency We Are Social and ISTAT demographics. (Tuttitalia, 2019; We Are Social, 2019)

#3.3

DATA COLLECTION

The aim of this research was to gather a dataset of Instagram posts geotagged in the city of Turin over a consistent period of time, in order to derive information and knowledge from these data.

Geotagged posts are accessible both from the mobile application and from the desktop version of the social media through the search option, allowing to look for "Accounts" – showing profiles according to the name users give to themselves –, and for posts, sorted by either "Tags" – the key words preceded by a # symbol to categorize the pictures in captions – or "Places" – the already mentioned geo-reference scope of this research.

Focusing on the latter, in the desktop version of Instagram – once accessed to private navigation with a profile's credentials – the displayed URL for a random location like the city of Turin has the following format:

https://www.instagram.com/explore/locations/215689286/turin-italy/

The URL is composed by:

- The website https://www.instagram.com/
- The search paramenters explore/locations/
- The location ID in a numeric value 215689286/
- The location name in a text format turin-italy

The search displays the nine “Most popular posts” at the top of the webpage, followed by a collection of the “More recent” ones, displayed in chronological order and the page can be refreshed indefinitely to go backward in time and to see older and older posts.
To perform a research from the URL it is not necessary the location name, but the location ID is sufficient, and the browser will automatically fill in the missing part.

The main steps in the collection of the dataset have been, then:

- Selecting a list of venues in the city of Turin for which collect geotagged posts
- Finding the corresponding location ID for each venue
- Scraping the posts geotagged in those venues and related information

#3.3.1

Venue listing

The selection of venues was performed deploying the database of the platform www.yelp.com. Yelp is a crowd-sourced website gathering and reviewing local businesses and locations in cities across the world. The platform allows users to perform a research in a city for a specific venues/business or a more generic one for typology. Venues are, in fact, classified into 22 broad categories[1], further subdivided in more specific typologies. (see *https://blog.yelp.com/2018/01/yelp_category_list*)

The typology of venues present on the platform is vast. It goes from generic public spaces – square, streets, monuments – to businesses with an actual position – restaurants, shops, bars etc. – to even typologies of businesses which cannot be referred to a specific location – plumbers, electricians etc.

The list of venues was scraped using Parsehub [2], focusing on those categories of venues that were, first, related to a physical place and, second, characterized by an urban and social relevance. [3] On Yelp multiple subcategories can be tagged to a venue – for example a restaurant could be tagged to “Italian cuisine”, but at the same to “Vegetarian”, “Regional cuisine” etc. – but only the first one was collected and associated to the venue in the scraping.

The outputs of the scraping process were lists of locations defined by their name, their address and the subcategory associated to it on Yelp. The final number of venues was approximately around 6300, although the exact number of discrete locations was not defined in this phase, since some of them were listed twice because they belonged to multiple categories – cocktail bars resulting both in “Nightlife” and “Restaurants”, take-away and sit-in restaurants resulting both in “Restaurants” and “Food” etc.

1. Active Life, Arts & Entertainment, Automotive, Beauty & Spas, Education, Event Planning & Services, Financial Services, Food, Health & Medical, Home Services, Hotels & Travel, Local Flavor, Local Services, Mass Media, Nightlife, Pets, Professional Services, Public Services & Government, Real estate, Religious Organizations, Restaurants, Shopping.

2. Parsehub is a web scraper collecting large amount of data from webpages.

3. The categories of Automotive, Event&Planning, Financial Services, Home Services, Pets, Professional Services and Real Estate were excluded from the research.

#3.3.2

Venue location ID correspondence

Once identified the list of venues to be potentially scraped on Instagram, the challenge was to associate to each of them the corresponding location id, in order to perform the correct "search for places" on the social media platform.

I deployed the technic described by Oh (see Oh, 2017 September 11; Oh, 2017, September 13) and the related Instagram scraper. [4] Modifications were performed in order to adapt it to the needs of the research and several empirical attempts were performed in order to obtain the correct location ID for each venue.

It was noticed, in fact, that while for the less popular venues the correspondence to a location ID on Instagram is direct – there is one unique location ID for each venue – the same cannot be said for the more geotagged ones. There are multiple location names corresponding to different location IDs for famous places like squares and museums: some use only the name – "Piazza Castello", for example – some add the name of the city in Italian – "Piazza Castello Torino" or "Piazza Castello-Torino" – some are partially or completely in English – "Piazza Castello – Turin" etc.

However, the correspondence with posts is not exclusive: the posts geotagged with the variation from the "official" location ID – "Piazza Castello Torino", for example – show in the search using the "official" one – "Piazza Castello" – but not the other way around. It was therefore important to associate to each venue the appropriate location ID in order to able to collect the majority of posts users had geotagged in that location.

The variations made to the scraper consisted in performing multiple researches readjusting the search parameters according to the results.
Initially, the venue was searched with its name as obtained from Yelp without adding any reference to the city of Turin. The first results were verified checking if that Instagram location was likely to be in Turin by verifying the associated coordinates – latitude either 44 or 45, longitude above 7. The very first result for "Teatro Regio", for example, is Parma Opera House, while Turin's shows after in the results. If the first search did not produce any relevant result, it was performed again adding first "Torino" and then "Turin" to the name of the venue.

The list of venues retrieved from Yelp was updated to a new list where the information about venues were obtained from Instagram – location ID, location name in full, address, geographic coordinates – and integrated with the Yelp subcategory.

Not all the venues scraped from Yelp found a correspondence on Instagram and their number went down to circa 2500, further reduced to 2178 when the data cleaning operation were performed and deleted the doubles.

4. https://github.com/rarcega/instagram-scraper

#3.3.3

Geotagged posts scraping

Using the location IDs associated to each venue the related geotagged post on Instagram were obtained using the scraper implemented by Mattia Gasparini.[5] The original scraper was created to search posts by user ID and hashtag and it was modified in order to perform the research for locations.

The original scraper was implemented to collect a specific number of posts per query and it was modified to this research introducing the time variable. As it was previously said, the posts are displayed in descending chronological order, therefore the scraper was set to continue gathering posts till a set date was reached.

5. https://github.com/gaspa93/instagram-scraper

The decision was to collect a year of posts for each location in order to perform analysis on a relevant period of time sensible to seasonal changes in the behavior of users.

The gathering process was consequently a time-consuming activity, which took place in three moments during December 2019. The first data collection took place between December 8th and 11th, when posts geotagged in the venues gathered from Yelp and converted into Instagram location IDs were collected. During the first preliminary phase of analyses, it emerged that some categories had been overlooked in the previous steps and this led to two integrations, the first between December 18th and 19th, and the second between December 26th and 27th. Every time the scraper was launched, the date up to which collect posts was consequently updated to an year before, therefore, although some of the data in the final dataset were collected with a temporal shift of two weeks, the information are consistent.

#3.4

DATASET

The final dataset consists in 354 503 posts geotagged in 1 071 different locations by 140 363 users.

The reduction in the number of scraped locations is justified by the fact that some venues have a correspondence on Instagram but there were either no post geotagged there or they were dated before the period in analysis and therefore they were skipped by the scraper.

Each entry of the dataset corresponds to a post, for which the following information were collected during the scraping process, while other were added from the information of the query.

Therefore, for each venue, each related post keeps information about:

- the location ID — e.g. “17206”
- the location name — e.g. “Piazza Castello”
- the latitude — e.g. “45,07090249”
- the longitude — e.g. “7,684981983”
- the category — e.g. “Piazze pubbliche[6]”

6. “Public squares” in Italian

Each post is then defined by:

- the ID of the post — e.g. “5df09a8cacda9648f0b8c9d6”
- the link of the post — e.g. “https://www.instagram.com/p/B57HMVBoxCw”
- the shortcode of the post — e.g. “B57HMVBoxCw”
- the date and time of the post — e.g. “11/12/2019 08:23:00”
- the ID of the user — e.g. “273059499”
- the number of comments, likes
- whether the post is a video or not
- the URL of the image — e.g.“https://scontent-mxp1-1.cdninstagram.com/v/t51.2885-15/e35/77245772_2485231591759653_806402917161532400 2_n.jpg?_nc_ht=scontent-mxp1-1.cdninstagram.com&_nc_cat=109&oh=7c99b81fa2954a6db5b1891324de05cf&oe=5E68CC8F”

For each post the caption was scraped as well, but as that information was not considered relevant for the goals of the following analyses and because some characters in the texts were complexifying the processing of data, it was decided not include it in the dataset.

#3.4.1
Recategorization

Each entry of the dataset was associated to a category derived from Yelp, which was, as stated before, the first of a series of multiple categories a venue can be associated to on the platform. The consequence was that the 1054 scraped locations were labeled with a multitude of attributes too broad for further analyses.

The categorization system was therefore manually adapted, organizing venues according to nine new categories, associated with the nature of the location:

- PUBLIC — public outdoors spaces likes streets, squares and monuments
- LEISURE — venues for the recreation from cultural activities (museum, theaters) to other typologies (stadiums, events venues)
- FOOD — restaurants in different typologies, but also street food and take away
- NIGHT — all the venues connected to recreational night activities
- SHOOPPING — shops, markets, outlets and other locations related to purchase of goods
- TRAVEL — hotel and other typologies of accommodation, but also venues related to commuting and travelling
- EDUCATION — institutions related to education at all levels and degrees
- WELLNESS — locations where to take are of one’s body, from hospital and medical centers to hairdressers and gyms
- RELIGION — religious organizations and related buildings

For each category, every venue was sorted more specifically according to its specific typology, grouping the multitude of Yelp’s labels to simplify the reading and understanding of the dataset.

BIBLIOGRAPHY

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 58–65.

Oh, B. (2017, September 11). Data Mining: Instagram Scraper (1). *Medium*, retrieved from https://medium.com/@bruceoh/data-mining-instagram-scraper-1-d557e7ae4264.

Oh, B. (2017, September 13). Data mining: Instagram Scraper-location tag w/ yelp_data_set (2). *Medium*, retrieved from https://medium.com/@bruceoh/data-mining-instagram-scraper-location-tag-w-yelp-data-set-2-94728064a4c0.

Olteanu, A., Castillo, C., Diaz, F., Kiciman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data, July 2019, 2*, art. 13.

Tuttitalia. (2019). *Popolazione per età, sesso e stato civile 2019*. Retrieved from https://www.tuttitalia.it/statistiche/popolazione-eta-sesso-stato-civile-2019/

We Are Social. (2017). *Digital in 2017*. Retrieved from https://wearesocial.com/it/blog/2017/01/digital-in-2017-in-italia-e-nel-mondo

We Are Social. (2018). *Global digital in 2018*. Retrieved from https://wearesocial.com/it/blog/2018/01/global-digital-report-2018

We Are Social. (2019). *Digital in 2019*. Retrieved from https://wearesocial.com/it/blog/2019/01/digital-in-2019

#4

QUANTITATIVE EVALUATION

The following chapter will perform some analysis on the dataset in order to get insight on the urban dynamics and the way the city is represented on Instagram. The chapter is divided in three main parts exploring different dimensions.

The first part takes into consideration the quantitative aspects of the dataset focusing on the number of posts in the different categories and the distribution of them among users and venues.

The second part focuses on the dimension of time both on a day and a year level.

The third explores the dimension of space, plotting the dataset on the city.

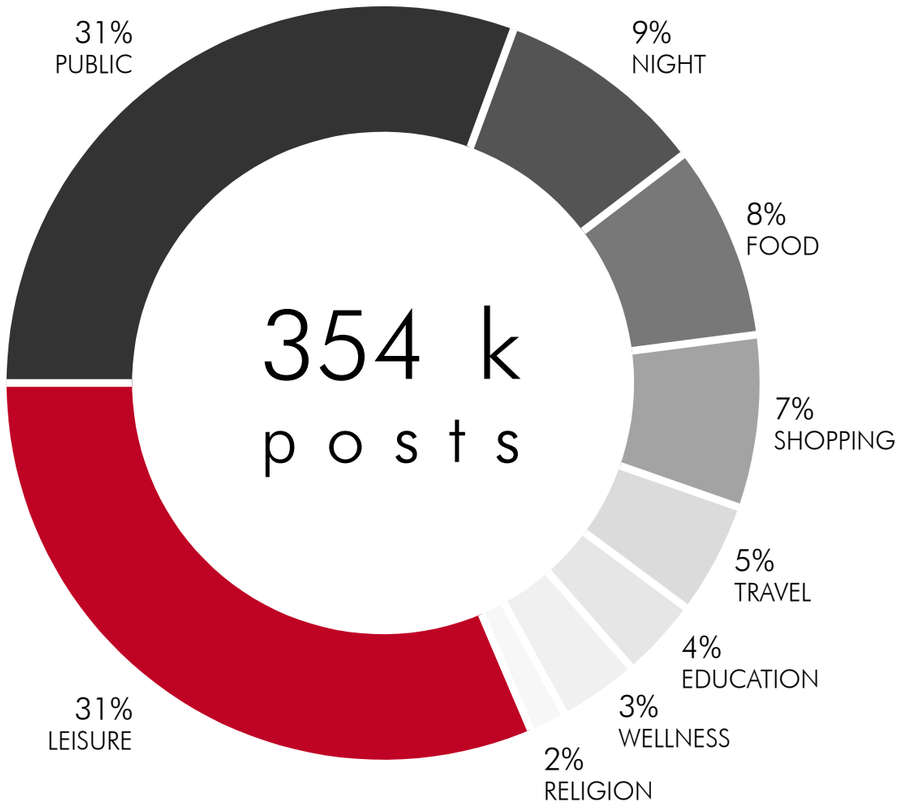


Figure 4.1.1_1
Posts percentage in each category

#4.1

Turin by Instagram NUMBERS

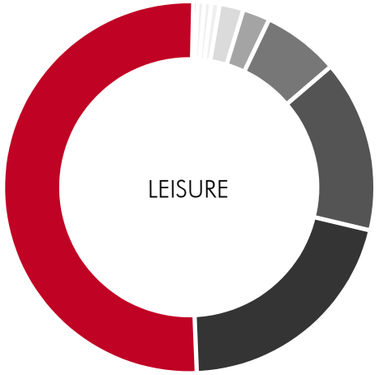
#4.1.1

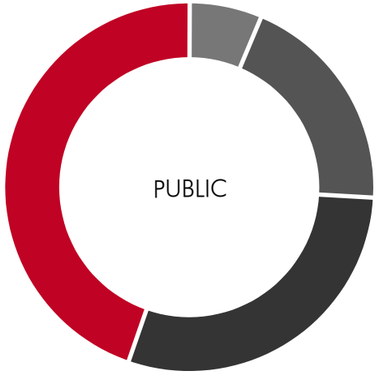
Venues distribution

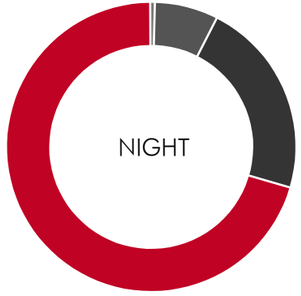
The dataset is composed by 354 503 geotagged posts in the city of Turin referred to one solar year.

What emerges from a first analyses on how the posts are distributed is the prevalence of two categories over the others, roughly dividing the dataset into three parts of similar dimensions. The majority of posts are geotagged in public spaces (31%) and venues connected to leisure activity (31%), while the rest (38%) is distributed among the other categories. [Figure 4.1.1_1]

When looking at subcategories, it is clear how some of them are strongly predominant in their percentage within the category they belong to and this is generally reflected in the most tagged venues in each group as well. [Figure 4.1.1_2] As far as regards he distribution of posts per venue, it does not seem equal even at visual evaluation, and this is confirmed by the calculation of the Gini coefficient[1] equals to 0.88, proving that some venues are widely popular, while others are not. [Figure 4.1.1_3]







LEISURE

- 51% Stadium/Concert halls
- 21% Event venues
- 15% Museums
- 7% Theaters
- 2% Art galleries
- 2% Cultural centers
- *2% Others*

FOOD

- 35% Italian cuisine
- 12% Pizzerias
- 10% Taverns/Bistrots
- 10% Wine & apetizers
- 9% Cafés
- 4% Japanese cuisine
- *20% Other cuisines*

EDUCATION

- 88% Universities
- 3% High/middle schools
- 3% Various education
- 2% Language schools
- 2% Boarding schools
- 1% Children schools
- 1% Driving schools

PUBLIC

- 43% Monuments/historical places
- 30% Public squares
- 20% Parks
- 7% Streets

SHOPPING

- 59% Shopping malls
- 21% Clothing
- 10% Markets
- 3% Coworking spaces
- 2% Various articles
- 2% Furniture
- *3% Others*

WELLNESS

- 26% Sports clubs
- 20% Hospitals
- 18% Fitness centers
- 12% Hairdressers
- 12% Tattoos
- 8% Bodycare
- *4% Others*

NIGHT

- 70% Nightclubs
- 22% Cocktail bars
- 7% Pubs/Bier bars
- 0,5% Wine bars

TRAVEL

- 50% Railway stations
- 27% Hotels & accommodations
- 21% Airport
- 1% Other transports
- 0,7% Subway stops
- 0,5% Travel agencies

RELIGION

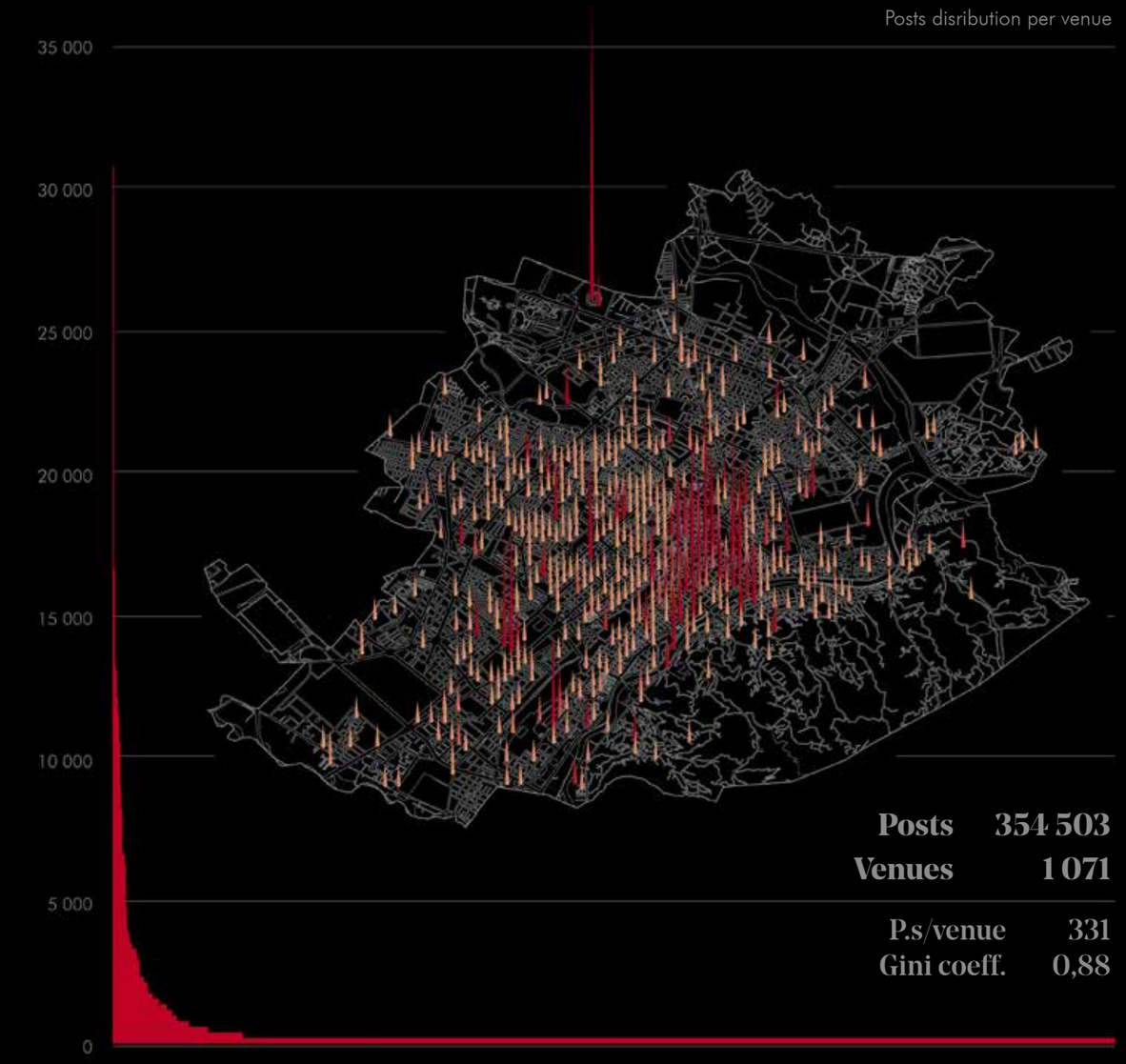
- 99% Churches
- *1% Others*

Figure 4.1.1_2

Subcategories percentage per category

Figure 4.1.1_3

Posts disribution per venue



1. The Gini coefficienti is « […] a statistical measure of distribution developed by the Italian statistician Corrado Gini in 1912. It is often used as a gauge of economic inequality, measuring income distribution or, less commonly, wealth distribution among a population. The coefficient ranges from 0 (or 0%) to 1 (or 100%), with 0 representing perfect equality and 1 representing perfect inequality. Values over 1 are theoretically possible due to negative income or wealth» (Chappelow, 2019). In this case, it evaluates the distribution of posts – seeing as the "income" – among venues or, later, users.

Starting to evaluate categories individually, half of the posts belonging to "Leisure" are geotagged in "Stadiums/Concert Halls" and the first three most popular venues belong to that subcategories as well. "Event venues" and "Museums" are both second and third in the overall percentage of posts and in the ranking of single venues. The distribution of posts per venue is similar to the general one, with a Gini coefficient of 0.85. [Figure 4.1.1_3.1]

To move the other larger group, "Public", there, apart from "Streets", posts are more evenly distributed among the remaining subcategories. Although the similar evenness seems to be seen in the ranking of the most tagged venues and in the curve of distribution as well, the Gini coefficient is equal to 0.84 and close to the overall score of the dataset. [Figure 4.1.1_3.2]

To generalize the discourse for the remaining smaller categories, on one side, some of them follow what we could call the "Leisure venues behavior": when a subcategory is predominant in percentage in the overall count (above 50%), the distribution of posts per venue is uneven too and the average is always a triple digit (see "Travel", "Education", "Religion", "Shopping" and related figures).

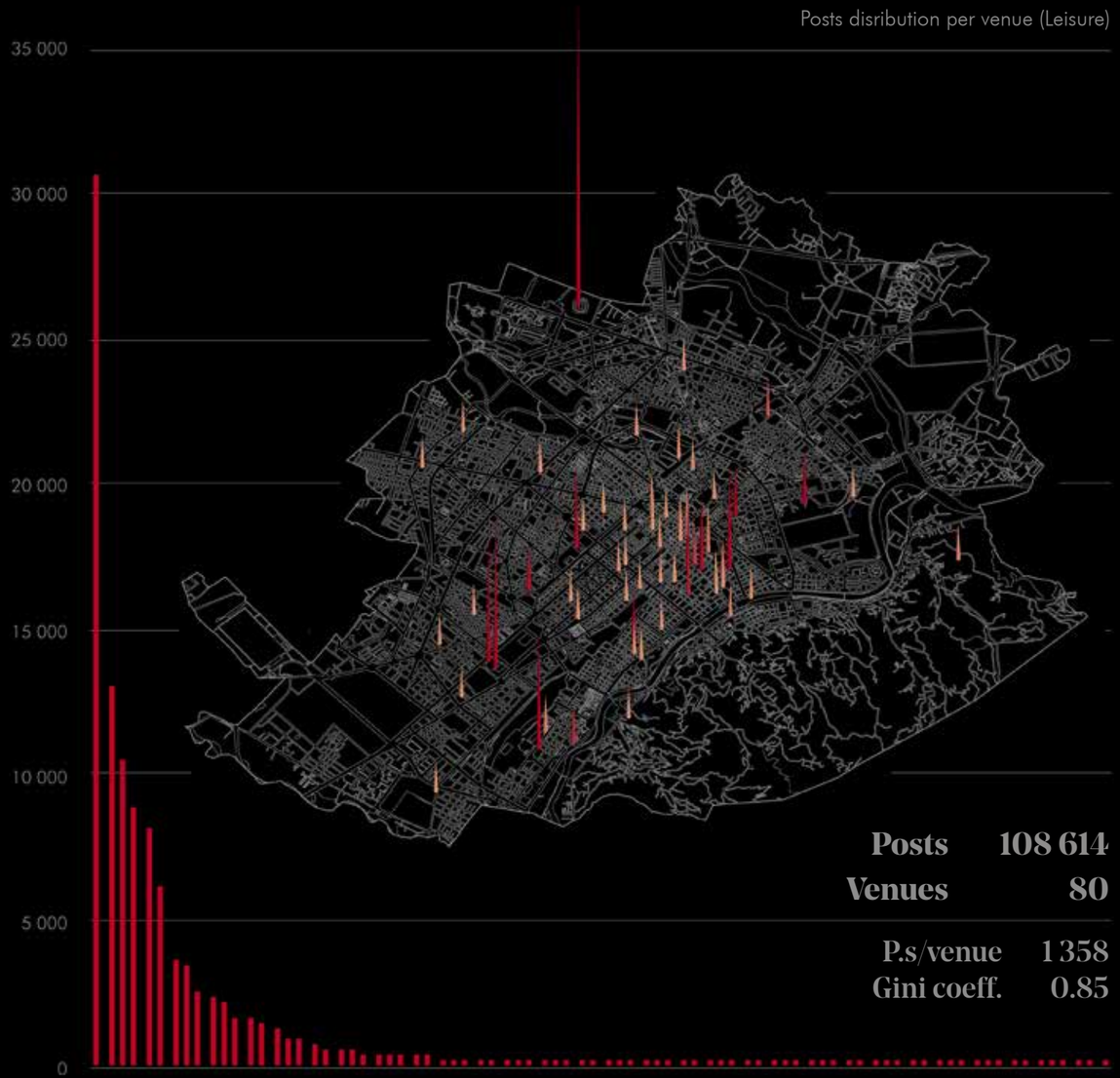
Instead when there is not a prevalence on one subcategory over the others, the distribution of posts per venue is more levelled and the average number of posts per venue as well. Both "Food" and "Wellness", in fact, have relatively low Gini coefficient compared to the average – respectively 0.70 and 0.74. [Figure 4.1.1_3.7] [Figure 4.1.1_3.8]

An exception is the category "Night", where although the majority of posts are tagged in "Clubs" and the percentage of post per venue is the fourth largest, posts are more evenly distributed among venues, with a Gini coefficient of 0.75. [Figure 4.1.1_3.5]

The conclusions that can be derived about the visibility of venues on Instagram are partially in line with what explained by Boy and Uitermark: Instagram marks spatial inequality allowing some venues to be more present on the scene while others less (Boy&Uitermark, 2017.) However, as stated by (Psyllidys, 2016) social data are a proxy of human behavior. That means that like in real life some locations get more visibility than others. What we can add to that is the fact that this cannot be generalized and, as we saw, some locations have a more even access to social visibility like restaurants and wellness.

Figure 4.1.1_3.1

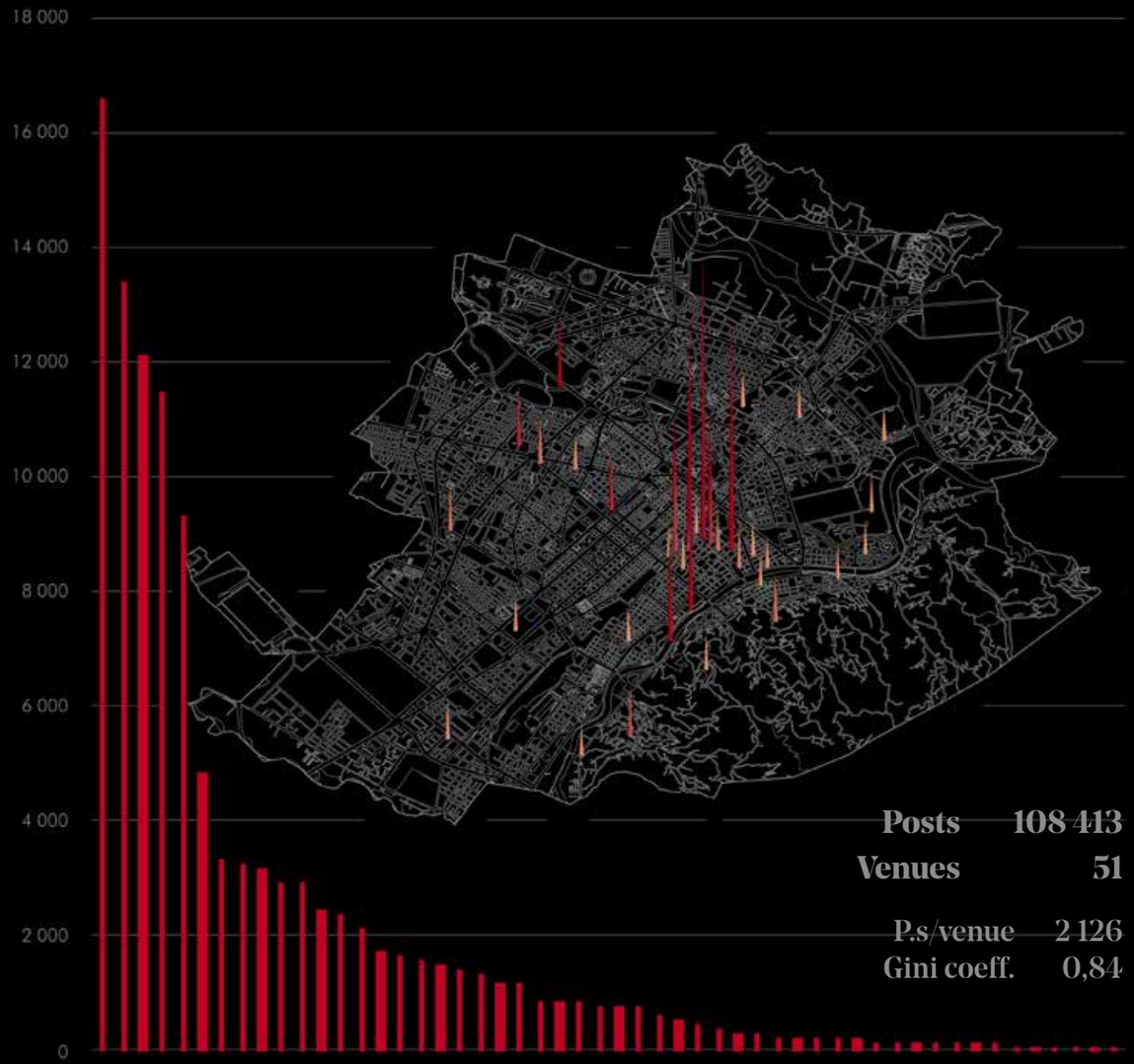
Posts disribution per venue (Leisure)



LEISURE

Figure 4.1.1_3.2

Posts disribution per venue (Public)



PUBLIC

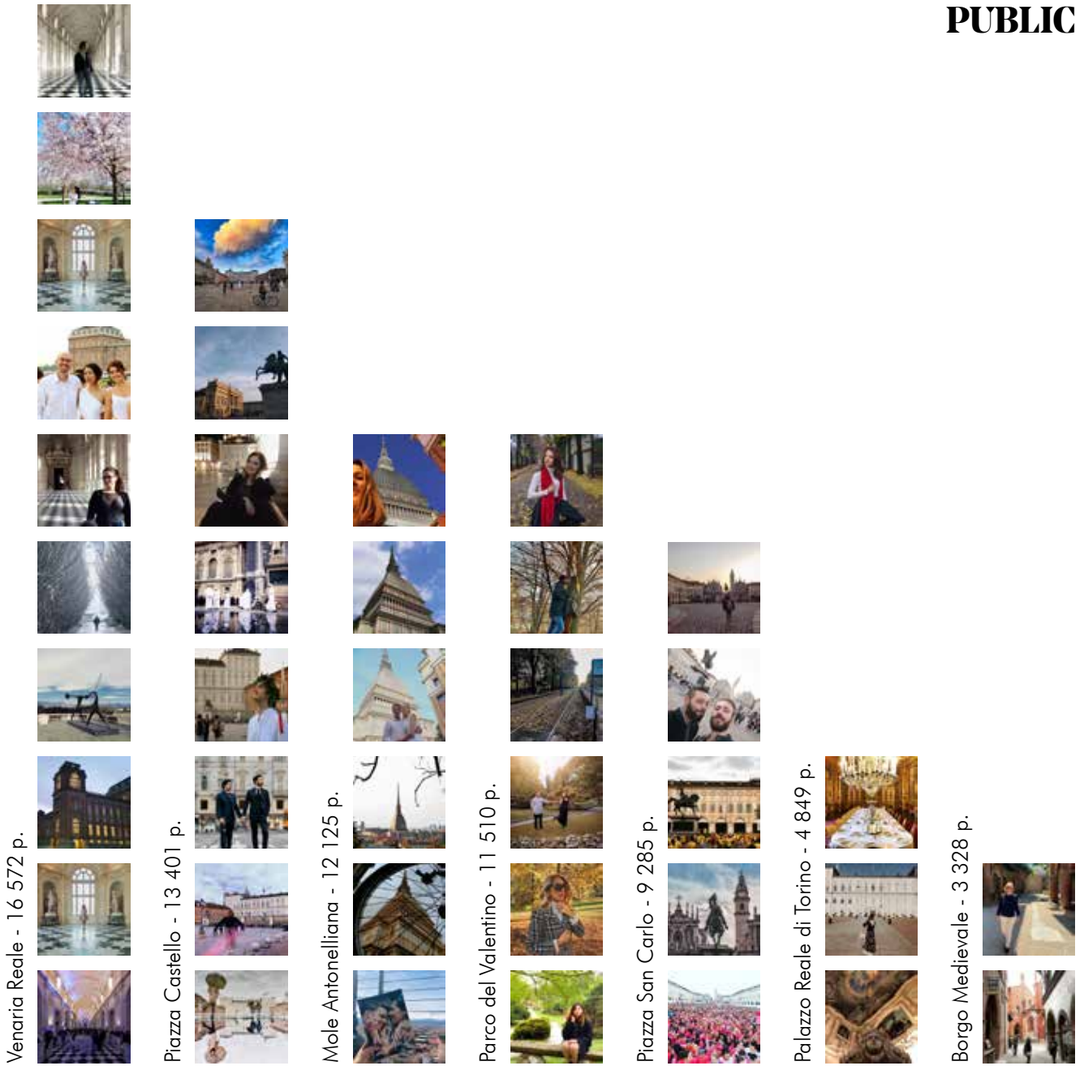
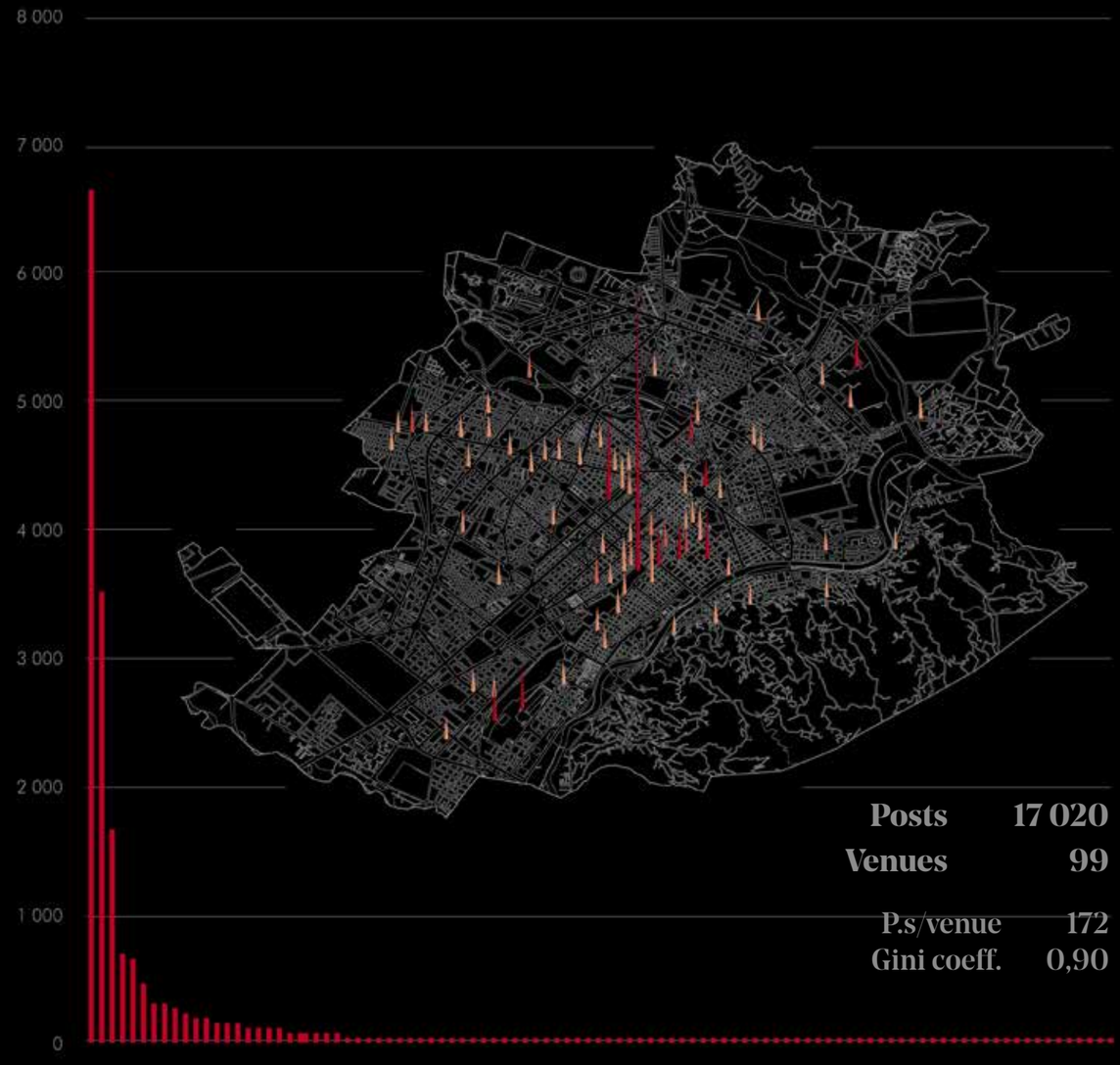
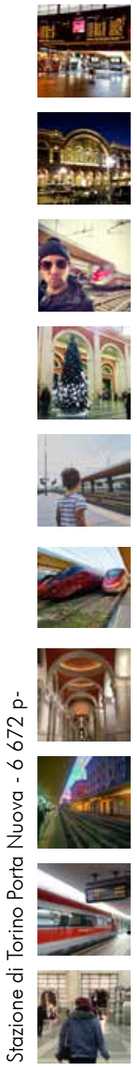


Figure 4.1.1_3.3

Posts disribution per venue (Travel)



TRAVEL



Stazione di Torino Porta Nuova - 6 672 p-

Aeroporto di Torino-Caselle - 3 523 p.

Stazione di Torino Porta Susa - 1 641 p-

Turin Palace Hotel & Les petites Madeleines - 1 641

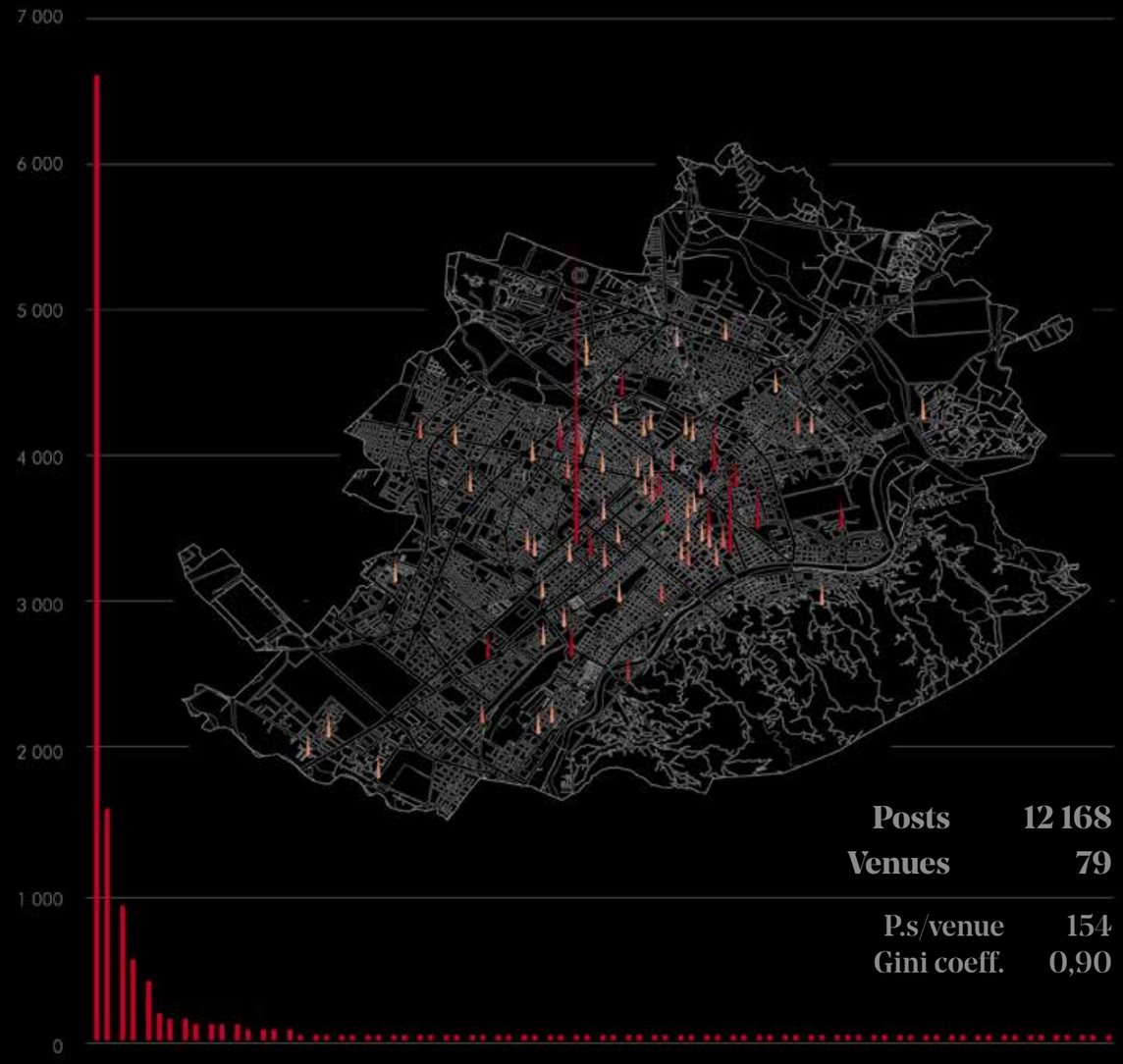
NH Collection Torino Piazza Carlina - 670 p.

DoubleTree by Hilton Turin Lingotto - 453 p.

Stazione di Torino Lingotto - 290 p.

Figure 4.1.1_3.4

Posts disribution per venue (Education)



EDUCATION










Politecnico di Torino - 6 595 p.




Università di Torino - Palazzo Nuovo - 1 587 p.



Scuola Holden - 941 p.


Università di Torino - Campus Luigi Einaudi - 571 p.

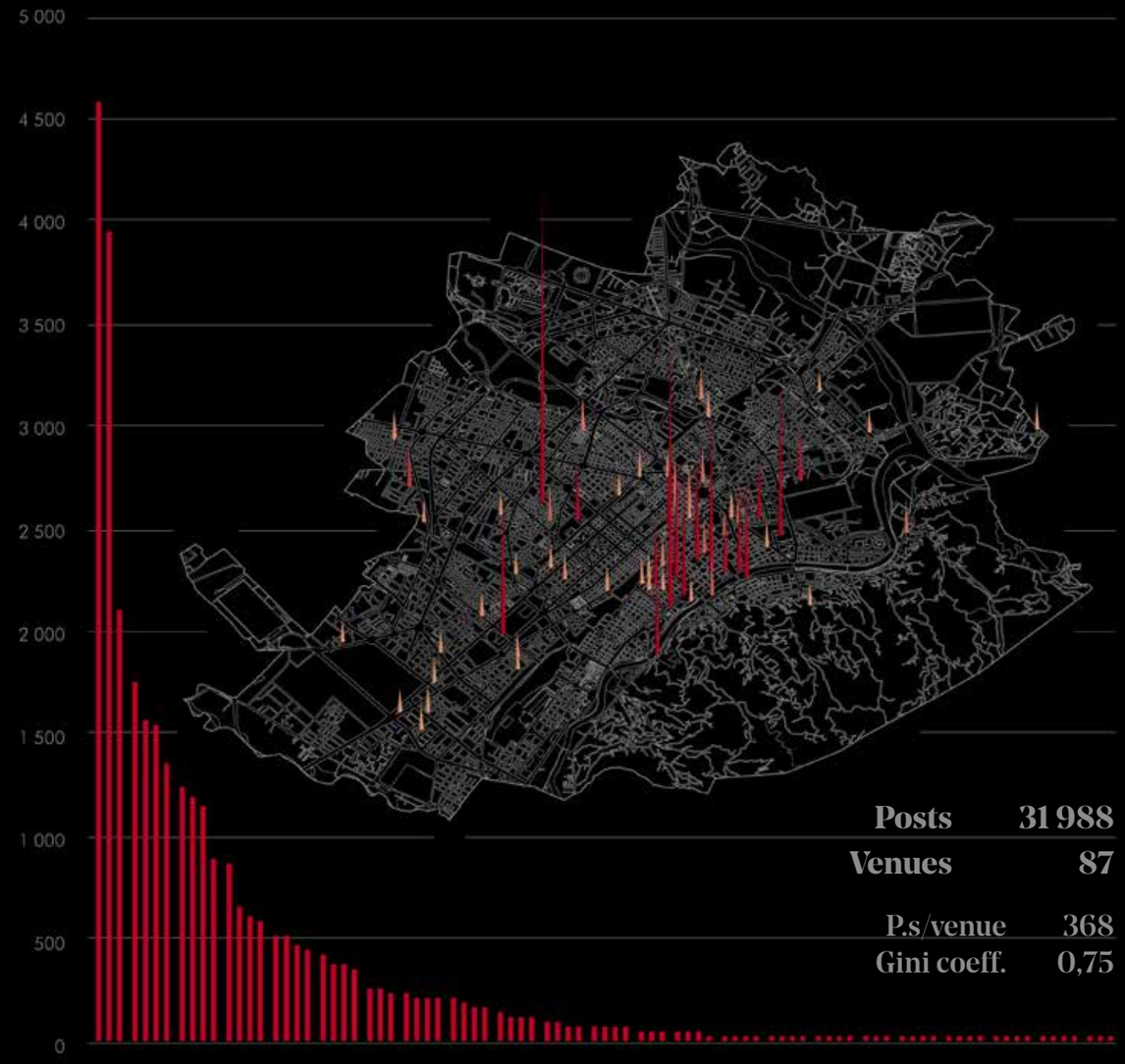

Università di Torino - Palazzo del Rettorato -408 p.

Gypsy Musical Academy - 190 p.

IAAD - Istituto d'Arte Applicata e Design - 148 p.

Figure 4.1.1_3.5

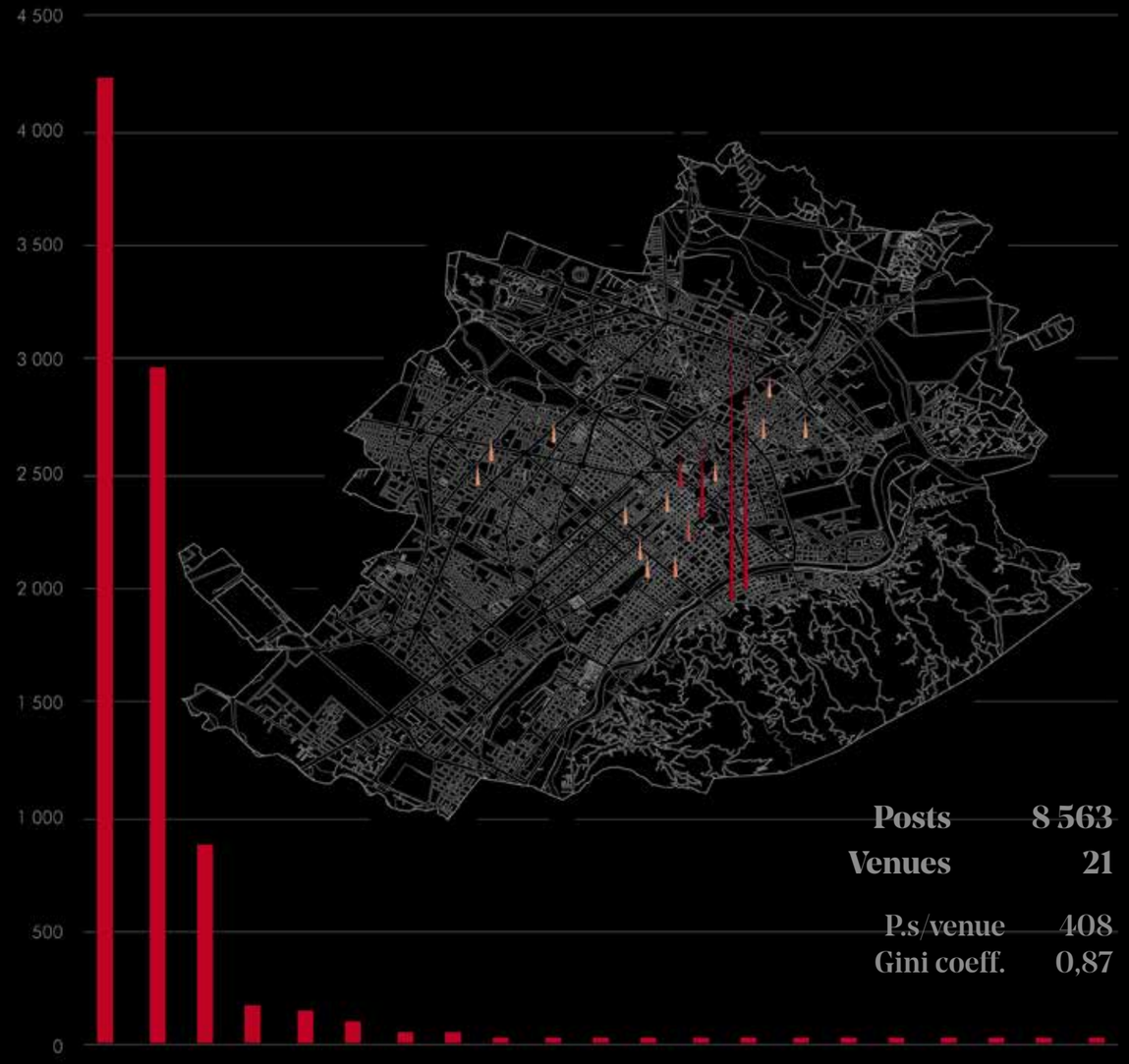
Posts disribution per venue (Night)



NIGHT

Figure 4.1.1_3.6

Posts disribution per venue (Religion)



RELIGION

Monte dei Cappuccini, Turin - 4215 p.

Chiesa della Gran Madre di Dio - 2950 p.

Duomo di Torino - 873 p.

Santuario della Consolata - 156 p.

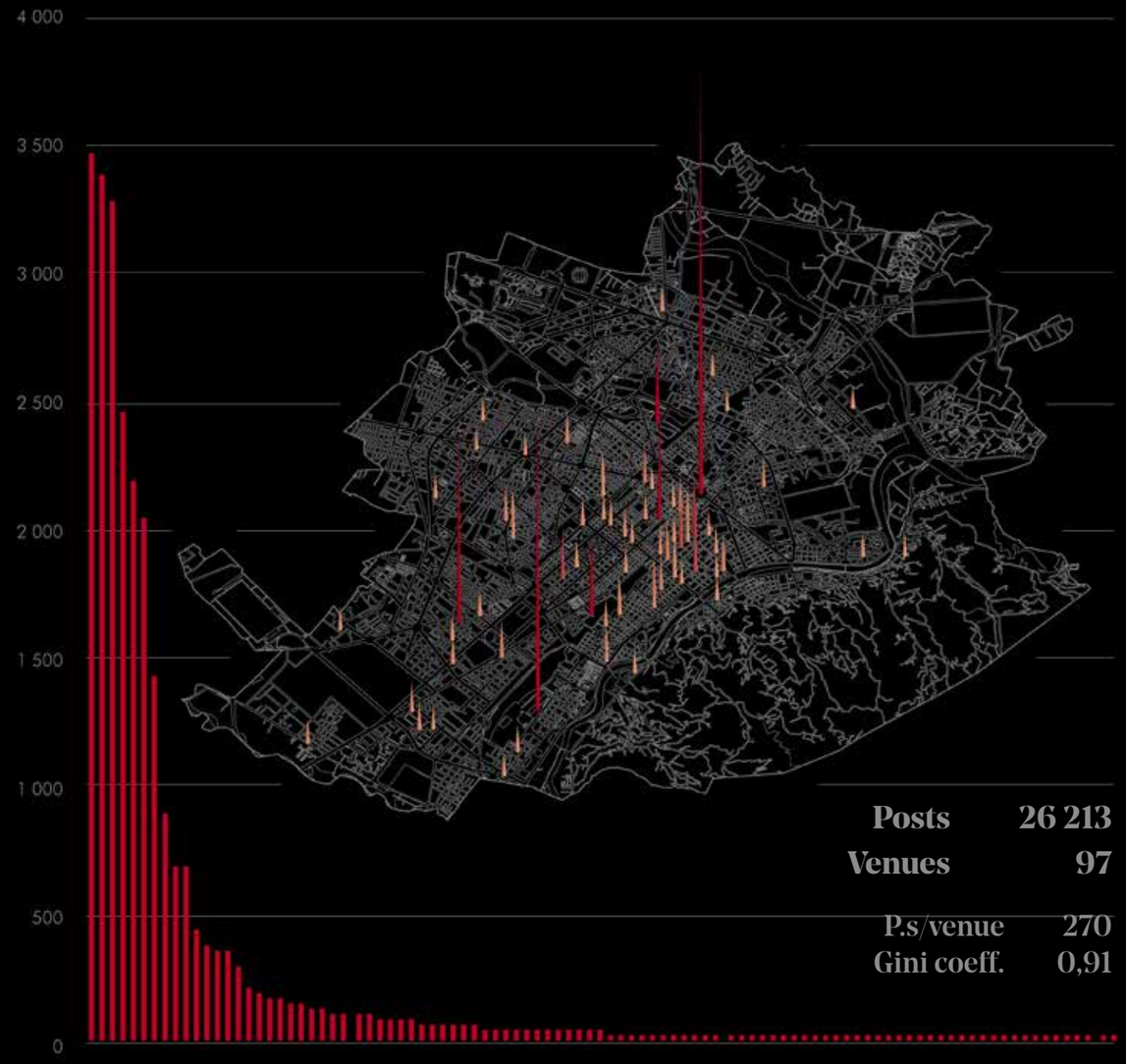
Santuario di Sant'Ignazio - 140 p.

Chiesa di San Filippo Neri - 95 p.

Tempio Valdese - 46 p.

Figure 4.1.1_3.7

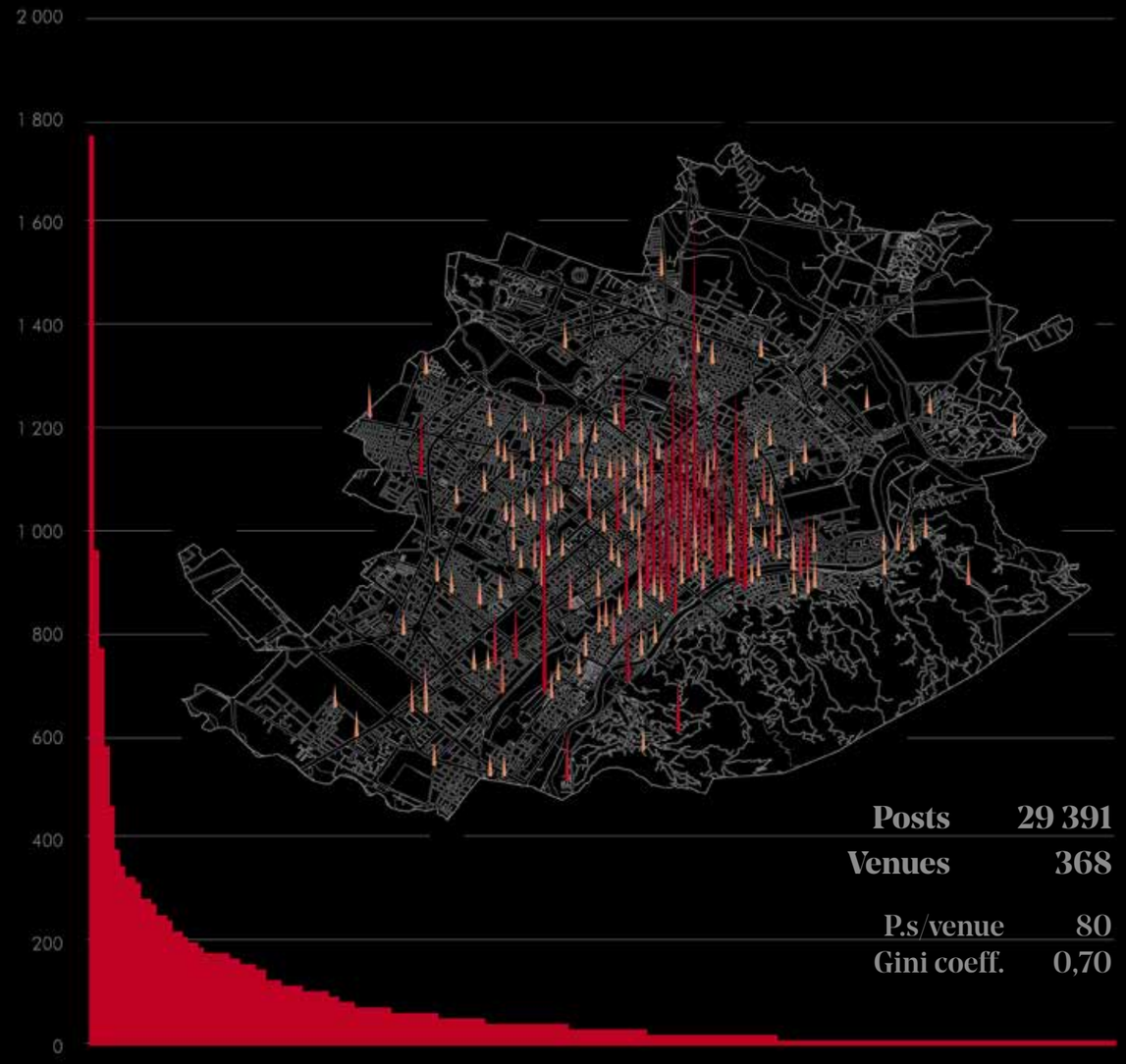
Posts disribution per venue (Shopping)



SHOPPING

Figure 4.1.1_3.8

Posts disribution per venue (Food)



FOOD

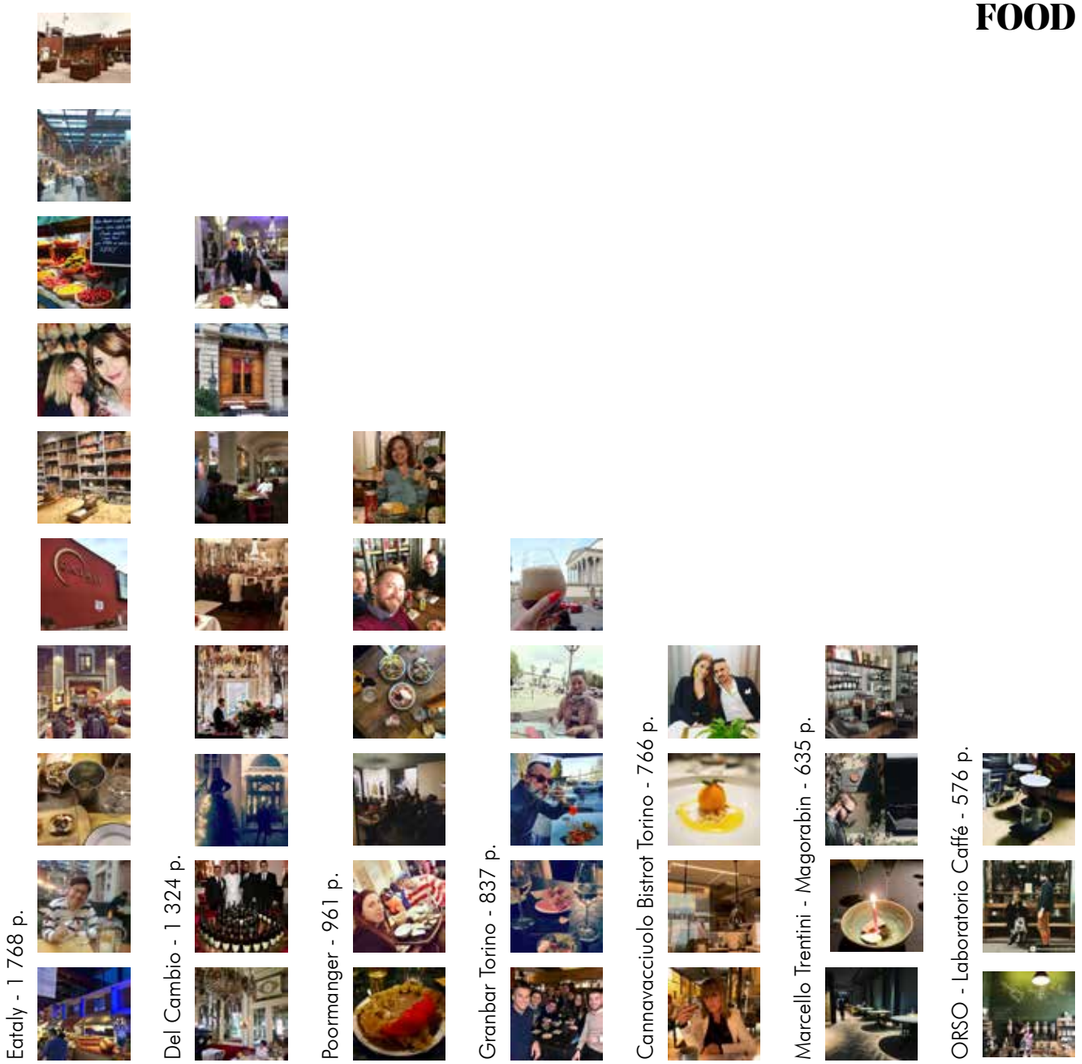
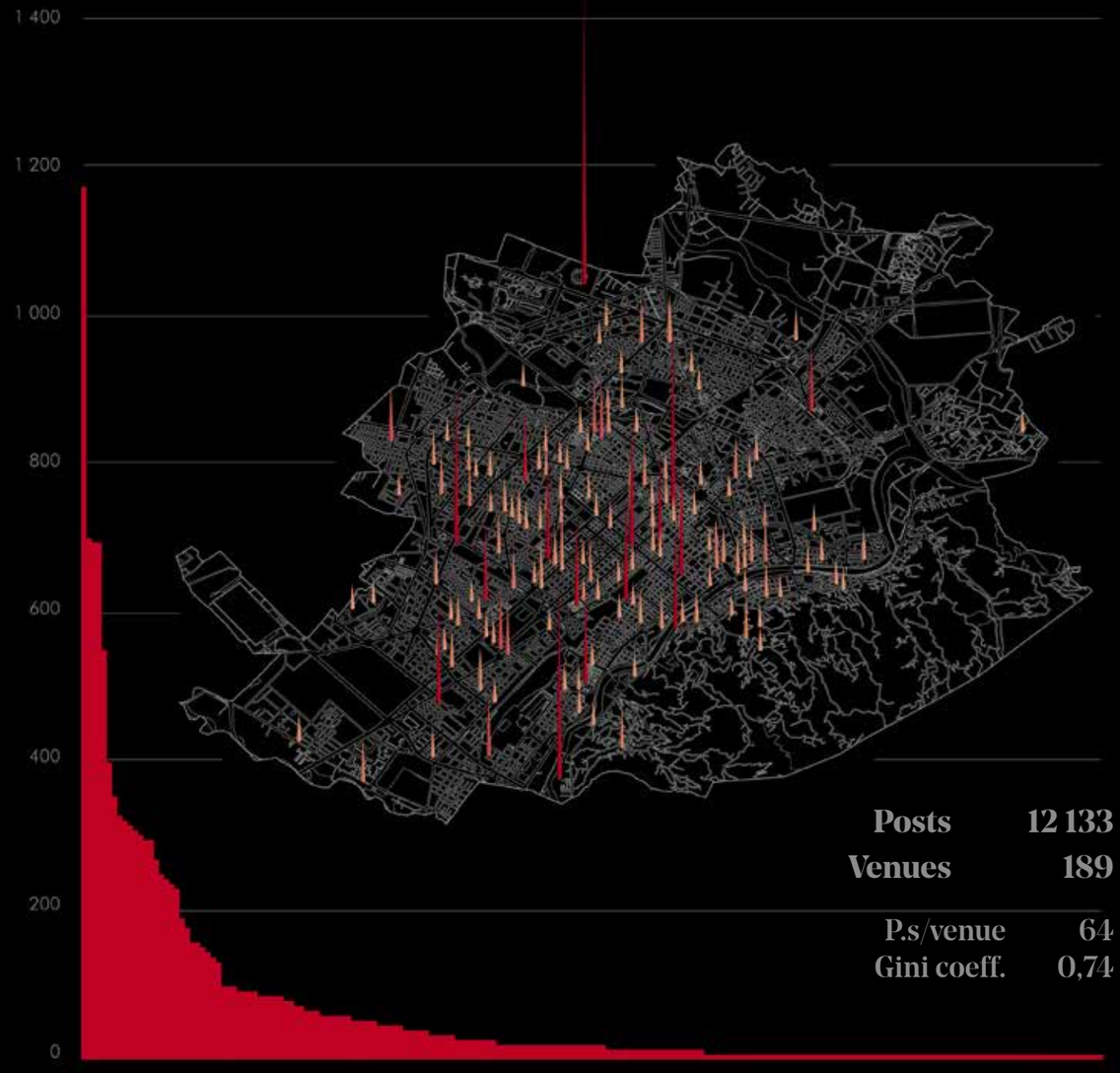
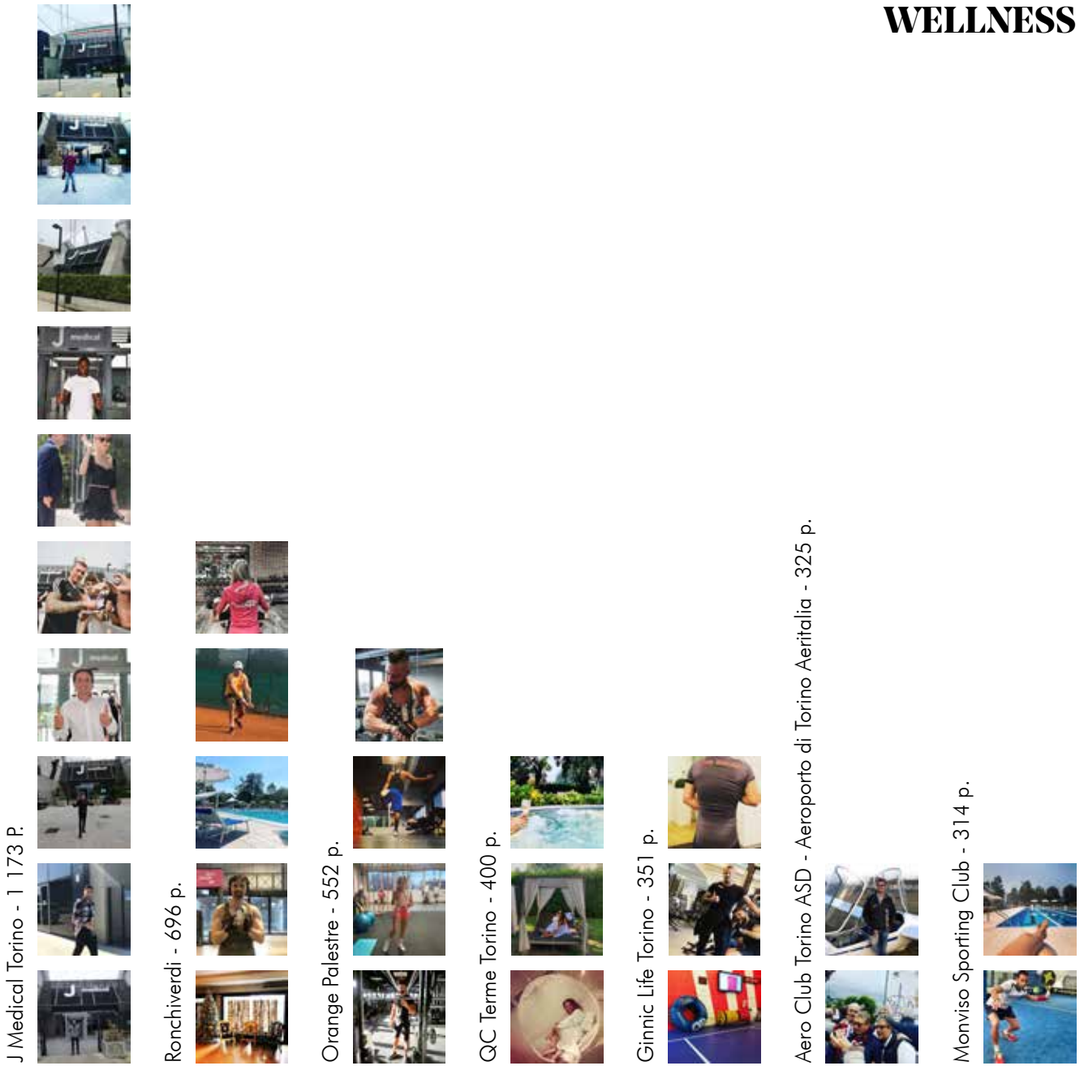


Figure 4.1.1_3.9

Posts disribution per venue (Wellness)



WELLNESS



#4.1.2
Population distribution

The distribution of posts per users as shown in [Figure 4.1.2_1] is referred to the whole datasets. On average every user has posted 2,5 picture and the distribution according to the Gini coefficient is overall quite even as this index is equal to 0,52.

It is interesting to notice how the Gini coefficient varies among categories in [Figure 4.1.2_2].

In fact, the distribution of posts per user is less equal in the categories of “Night”, “Shopping” and “Wellness”. The phenomenon could be explained in relation to two intrinsic characteristics of the venues listed in those categories. On one side, most of the locations are associated to businesses which tend to have a personal profile posting pictures geotagging the business itself. On the other side – with particular reference to “Night” and “Wellness” – locations like clubs, hairdressers, sports clubs and gyms are visited by people on a regular basis, explaining why some users are “richer in posts” than others according to the Gini coefficient.

As far as regards the other categories, it emerges clear how the distribution of posts per user is more even and the average of posts as well, explaining how those venues

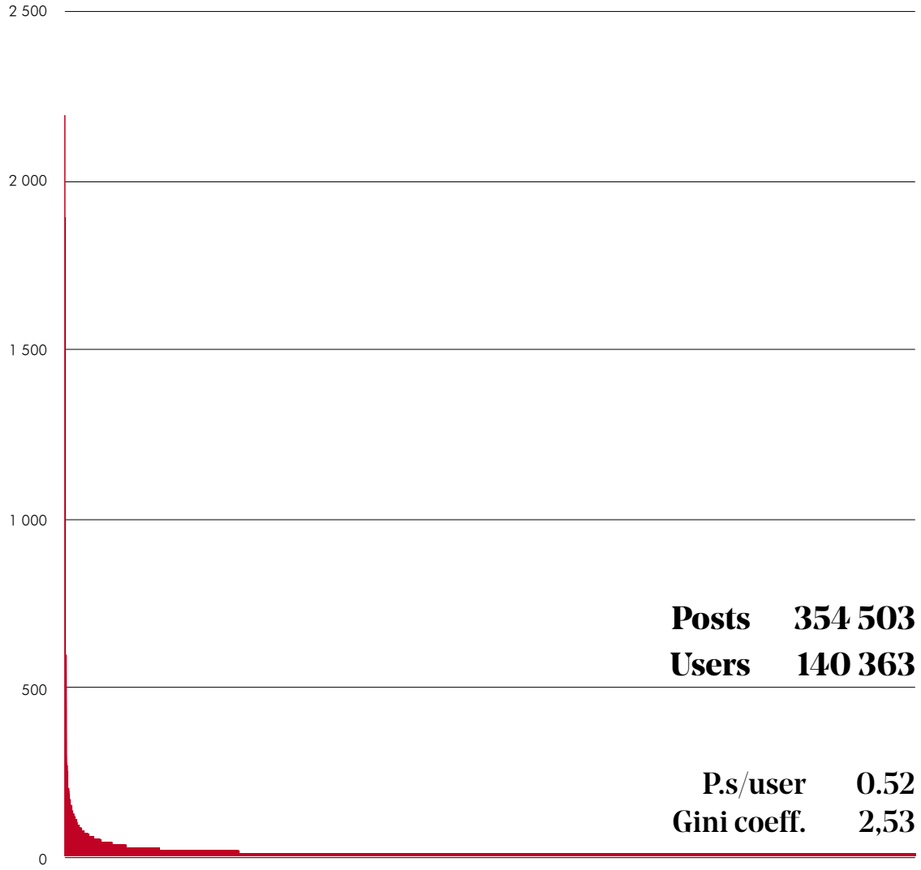


Figure 4.2.1_1
Post distribution per user

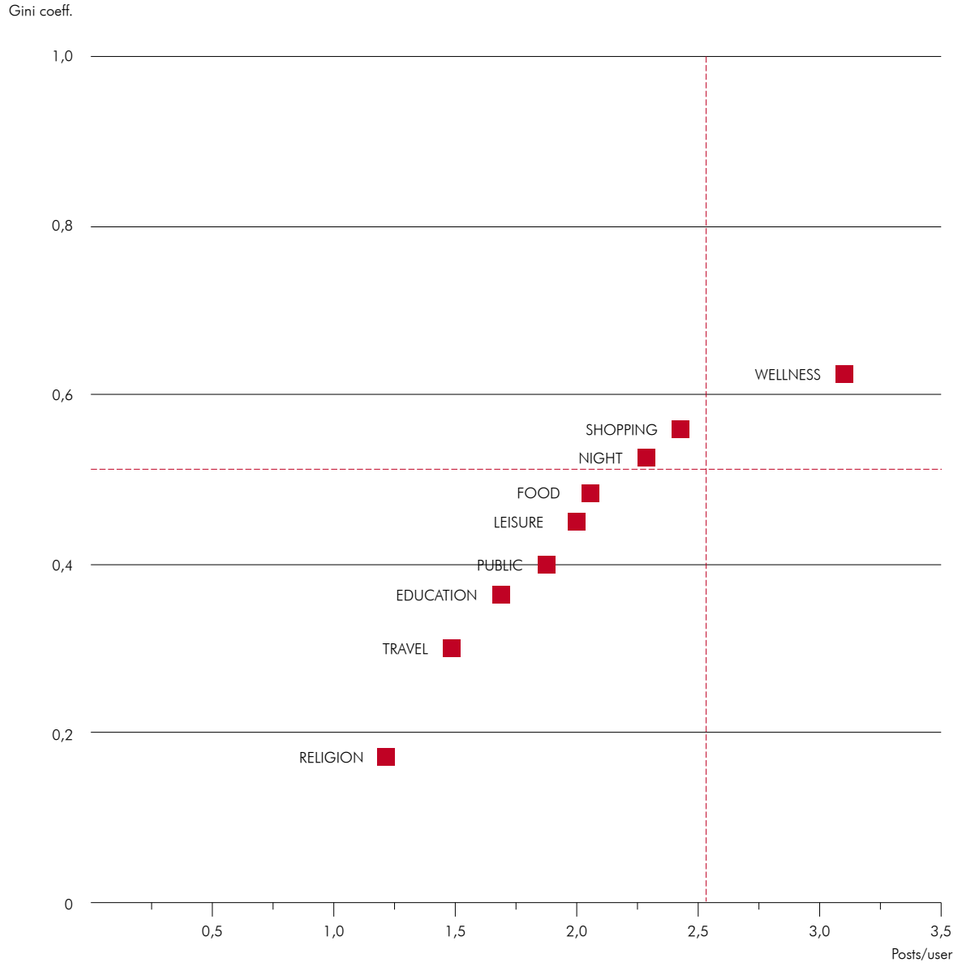


Figure 4.2.1_2
Average and Gini coefficient confronted with the values of the dataset

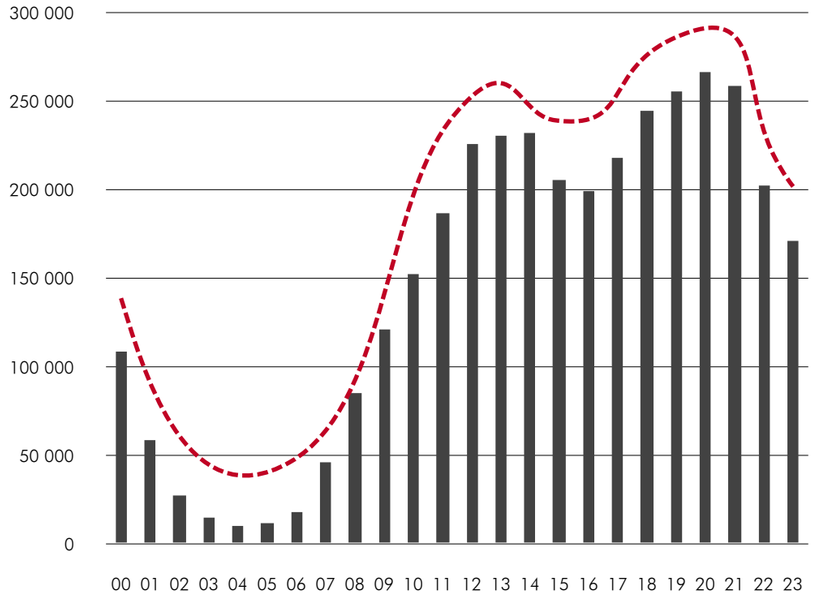


Figure 4.2.1_1
Distribution of post during the day

#4.2

Turin by Instagram TIME

#4.2.1

Day distribution

The analysis of the distribution of post during the day at different hours follow a curve, reaching its lowest point between 4 AM and 5 AM and its peaks during lunch and dinner hour. [Figure 4.2.1_1]

The pattern of distribution is similar in in general similar in all the categories in analyses, allowing to advance the theory that the act of posting a picture is not contemporary to the actual moment when the picture has been taken. [Figure 4.2.1_1.1-9]

The theory is further confirmed if comparing the curve obtained from the dataset with a realaboration of the same typology of data retrieved from the study performed by Yang et al. in some major European cities. (Yang et al., 2016). [Figure 4.2.1_2]

It can be derived that the act of posting happens when the users are not engaged in other activities and they can spend their free time interacting with their networks.

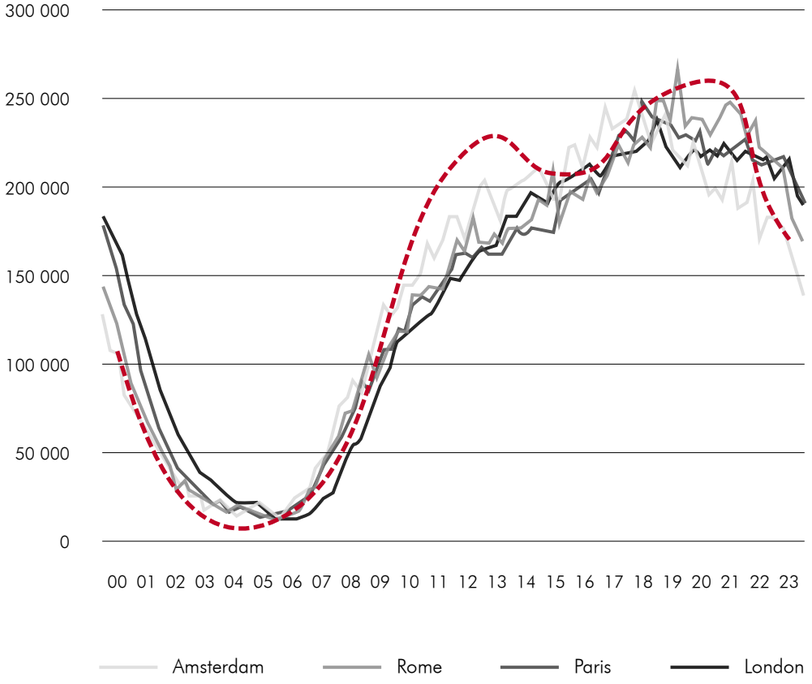


Figure 4.2.1_2

Results confronted with previous researches[v]

2. Data retrieved from Yang et al. (Yang et al., 2016).

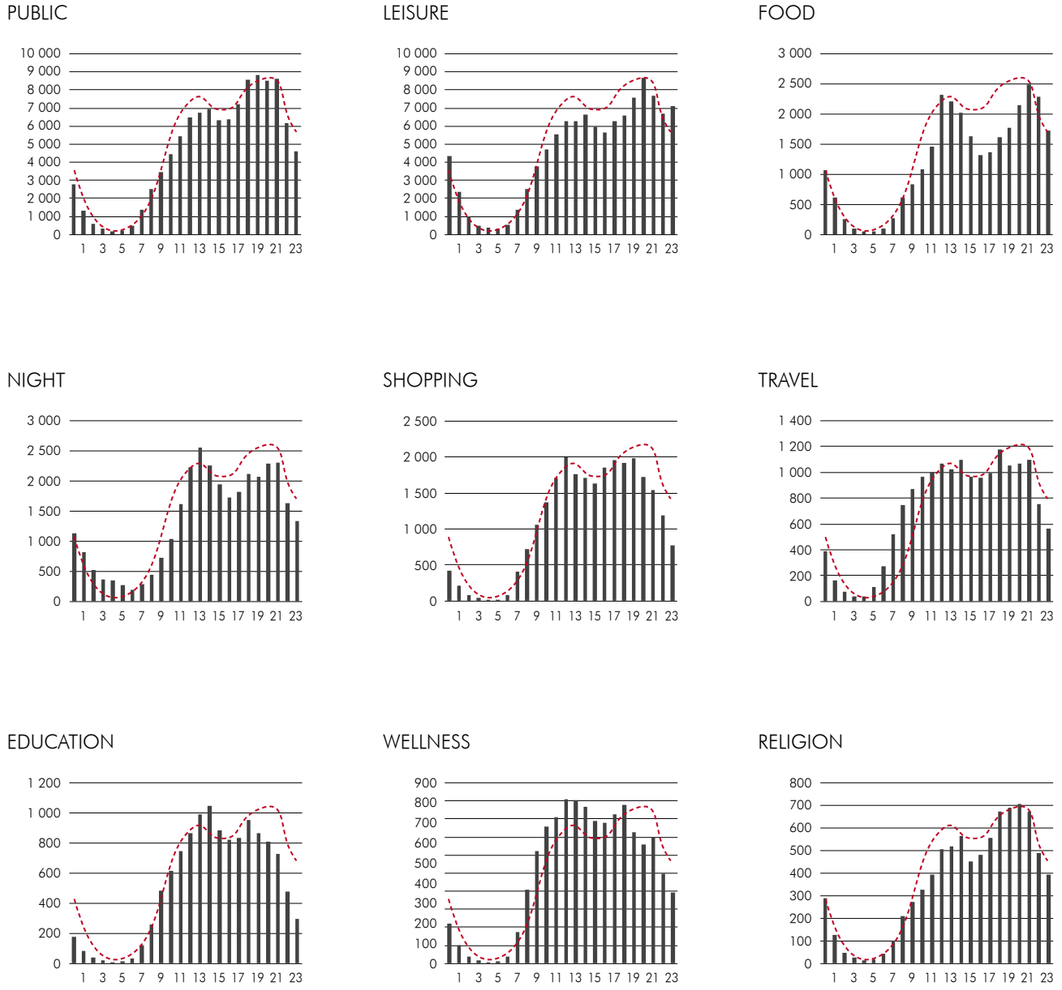


Figure 4.2.1_1.1-9

Day distribution curve confronted in each category

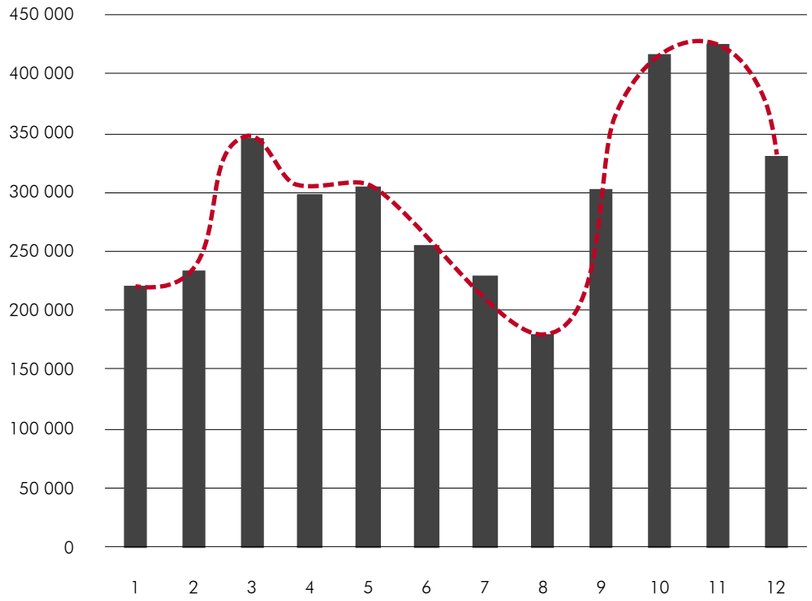


Figure 4.2.2_1
Distribution of post during the year

#4.2.2
Year distribution

The conclusions derived for the distribution of posts during the day cannot apply to the distribution of posts during the year, as there are not previous studies to compare this distribution to and categories behave differently from the overall dataset.

While seeing [Figure 4.2.2_1], it is clear how the peaks of posting activities happen during the spring and autumn month while there is a consistent decrease during summer. This phenomenon could account for the actual presence of resident (and even tourists) during the year in the city. The city of Turin has not a seasonal typology of tourism and although their presence in the city probably increases during festivities and holidays, it could be assumed as a stable presence during the year. As far as regard the residents, the winter and summer periods are spent usually outside the city and this could explain the rise and fall of geotagged posts during the year.

To go deeper into the categories, the distribution of posts varies, although the lowest number of posts is almost always during the month of August.

The distribution in the category of "Public" and "Night" follow approximately the behavior of the whole dataset. "Travel" has a similar distribution as well, although more levelled than the general dataset, and "Shopping" has the spring peak shifted of a couple of months.

The category "Leisure" follows the overall behavior, although the autumn peak is dramatically out of scale in comparison to the other months. This can be explained with the fact that the major cultural and social events of the city occurs during that period and the venues where they are located get more popular on the Instagram.

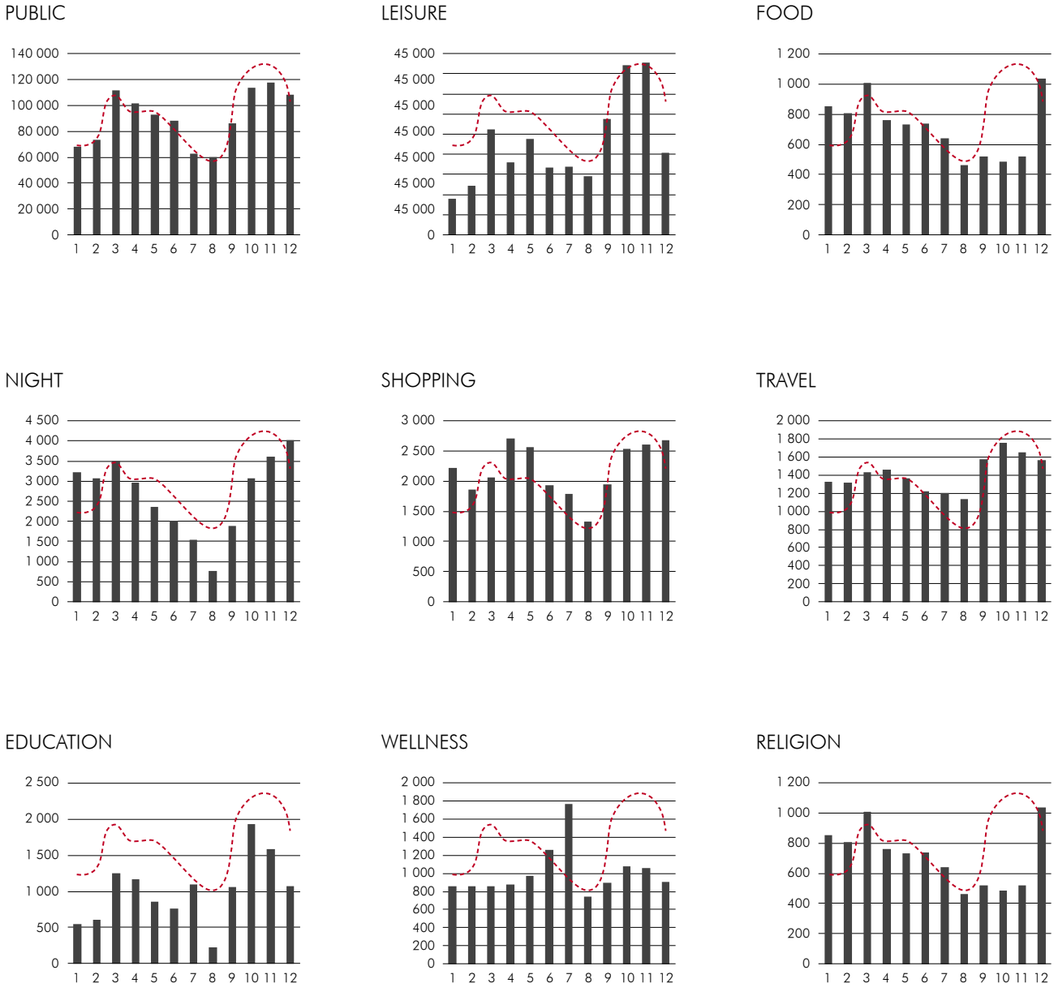


Figure 4.2.2_1.1-9

Year distribution curve confronted in each category

This could partially explain why the category "Food" "misses" the autumn peak: during that period of the year people are more likely to geotag themselves at event to "show off" their social life rather than at restaurants.

The category "Education" knows the more dramatic fall in the number of posts in the month of August and this could be explained with the closing of university premises, as the majority of posts are geotagged at universities. The peaks in October, March/April, and July are in parallel with the degree ceremonies in the various faculties.

As far as regards "Wellness", the diffusion of posts is overall stable but knows a sudden increase in proximity with summer. Remembering how sports clubs, hairdressers and gyms are listed in that dataset, the rise in the posting activity could be linked with the coming of summer. People are more engaged in sport activity and tend to get physically ready for summer and holidays.

The category "Religion" is mainly composed by churches and although the number of posts and venues is not wide enough to make generic assumption, it may be noticed that the peaks in the posting activity coincides with the major catholic festivity, being them in March – Easter – and December – Christmas.

Psyllidis and Bozzon proved using social urban data how in the city of Amsterdam tourists tended to concentrate in touristic area, while residents and commuters in other more linked to everyday life activity. They did not discover something that could not be inferenced from common knowledge and sense, but the discovery was the possibility of using social urban data as a proxy of human behavior. (Psyllidis & Bozzon, 2015).
In the same way, what the distribution during the year of this dissertation dataset entries proves is not something novel but reflects an actual behavior of citizens in the city.

#4.3

Turin by Instagram SPACE

The spatialization of the entries of the dataset on a physical map of Turin allows to see how the city is depicted through Instagram geotagged posts and how it changes according to different parameters.

Three typology of maps were performed, using Kernel Density estimation as provided in the QGIS software, setting three different parameters.

The first map is the “Density map”, measuring the density of venues according to their coordinates and showing the concentration of the venues present in the dataset. [Figure 4.3_1] shows how the majority of them are located in Turin city center and from there spreading mainly in the area south from the Dora river. In the northern areas of the city some spots light up, but their density is scarcer.
Despite the number of venues in each category does not allow to have a comparable reading, what is striking to notice in [Figure 4.3_1.1-9] is that the situation is comparable to the general one in each category.

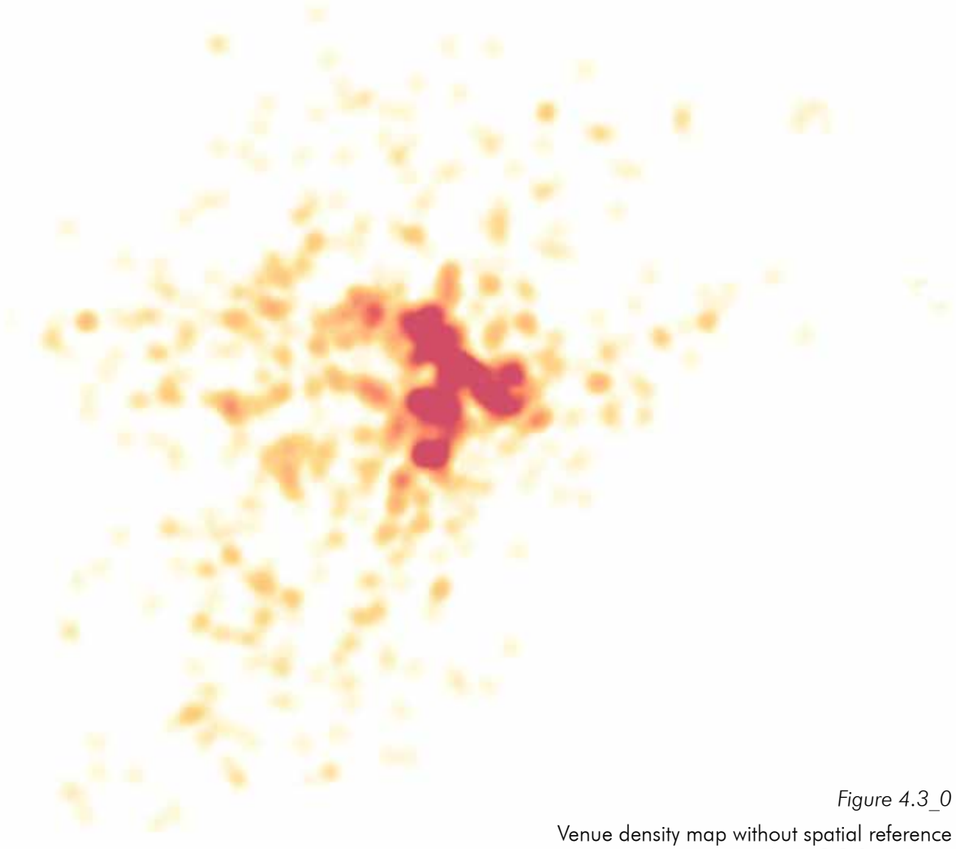


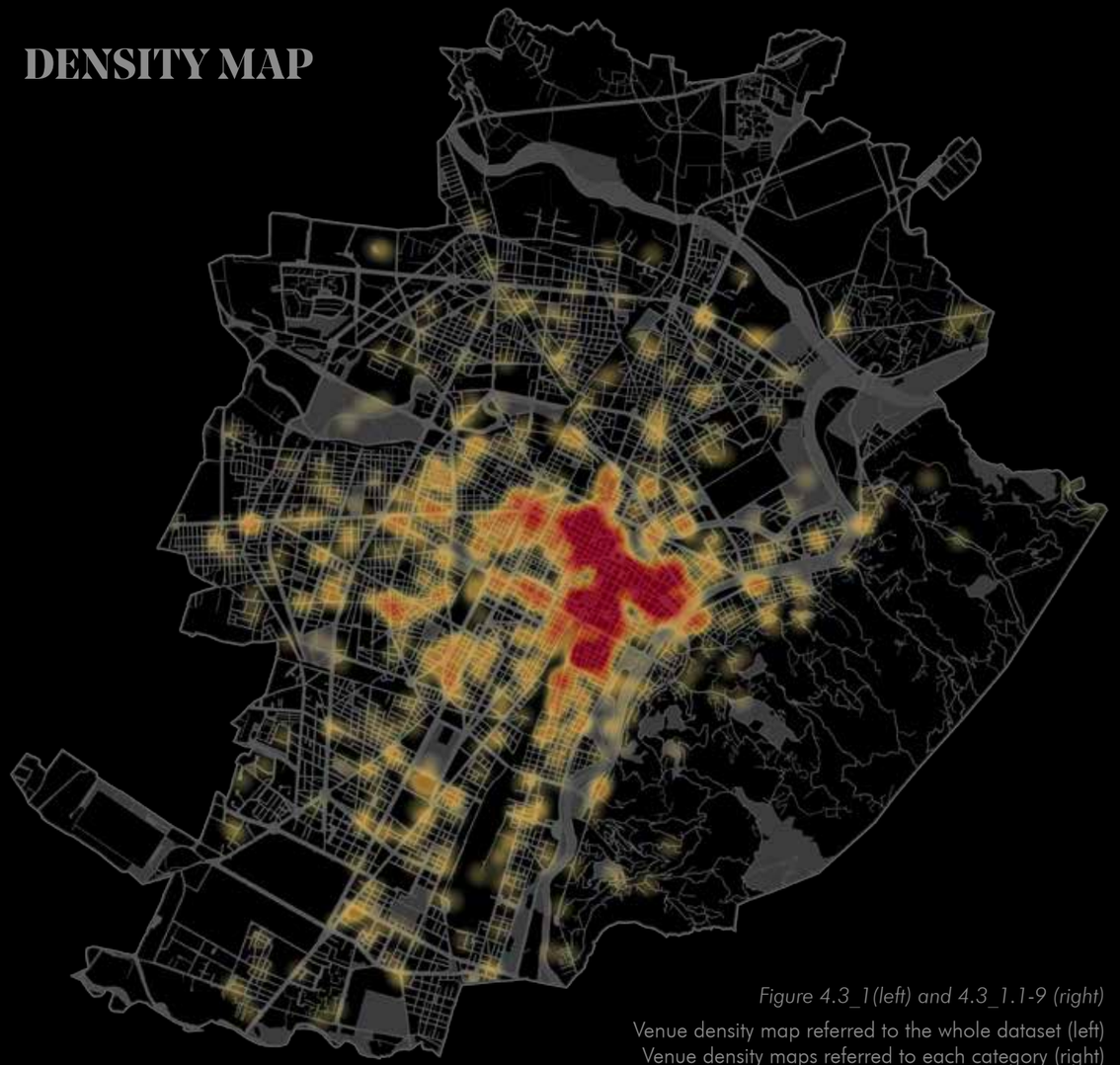
Figure 4.3_0

Venue density map without spatial reference
to show the raw output of the QGIS program

The second map is the “Popularity map” as the density of venues is not equal to the value users give to it and a proxy for this can be sought in the analysis of the number of posts per venue. The heatmap of this parameter in [Figure 4.3_2] shows how the hierarchy of visibility changes: while the city center is still very visible, some major attractions points are scattered through the city, particularly in correspondence of event venues.

The same reading can be performed in the maps referred to each category [Figure 4.3_2.1-9] where the major attractions points light up in red in different part of the city. The third map is the “Fidelity map”. A popular venue can be occasionally visited by a large number of users, while another visited more frequently by fewer individuals who geotagged themselves there on Instagram. If the post per venue could be accounts for the popularity of a venue, the number of posts per user in each location could is seen as a representation of the “fidelity” of people towards a place. The radius of each dot is weighted against the general number of posts and the resulting map [Figure 4.3_3] tells a different story from the other two. The popular spots of the Fidelity map vanish in yellow, the city center is still very vivid but peripherical lights up, particularly in the south-west part of the city.

DENSITY MAP



EDUCATION

FOOD

LEISURE

NIGHT

PUBLIC

RELIGION

SHOPPING

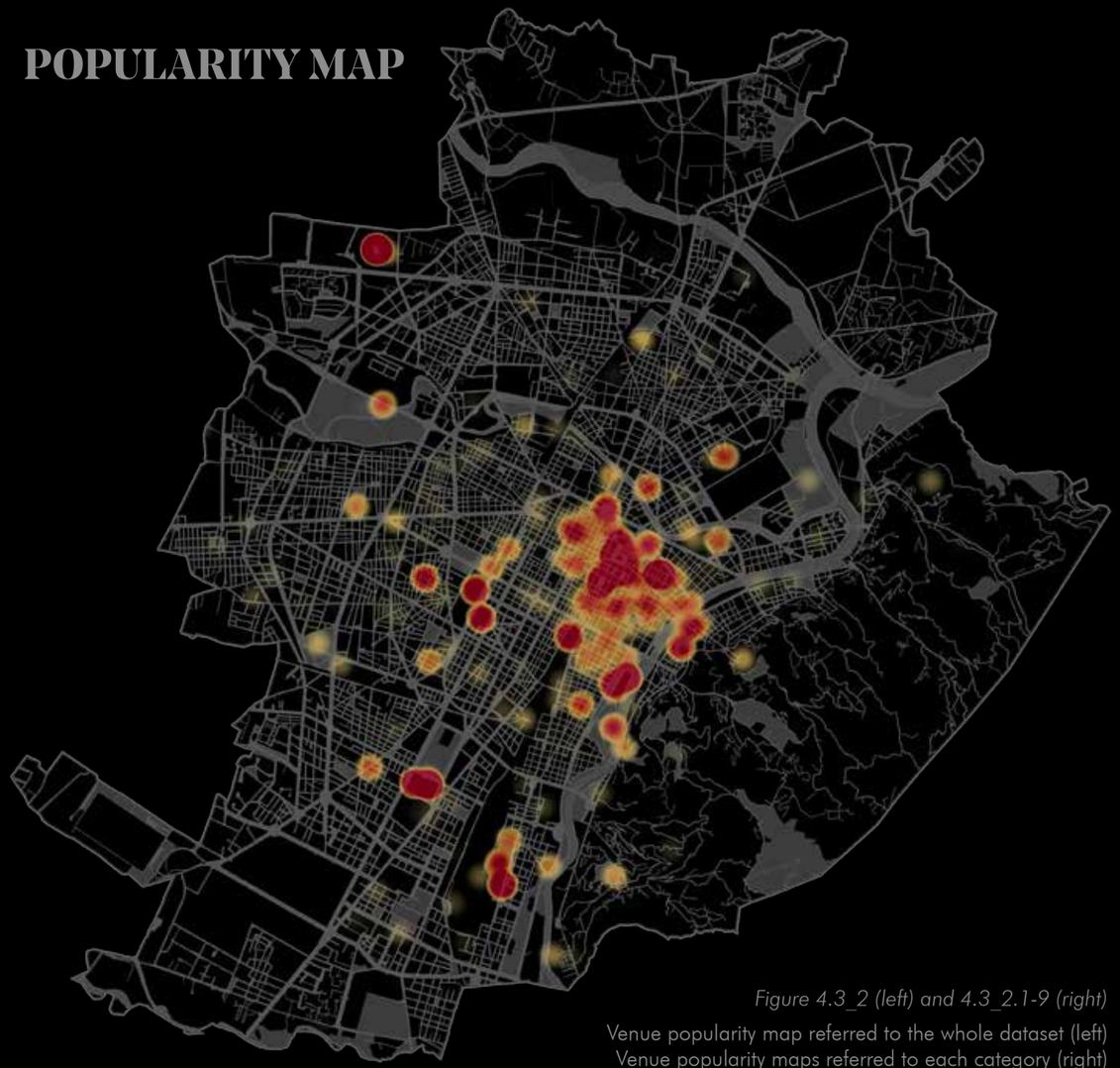
TRAVEL

WELLNESS

Figure 4.3_1(left) and 4.3_1.1-9 (right)

Venue density map referred to the whole dataset (left)
Venue density maps referred to each category (right)

POPULARITY MAP



EDUCATION

FOOD

LEISURE

NIGHT

PUBLIC

RELIGION

SHOPPING

TRAVEL

WELLNESS

Figure 4.3_2 (left) and 4.3_2.1-9 (right)

Venue popularity map referred to the whole dataset (left)
Venue popularity maps referred to each category (right)

FIDELITY MAP

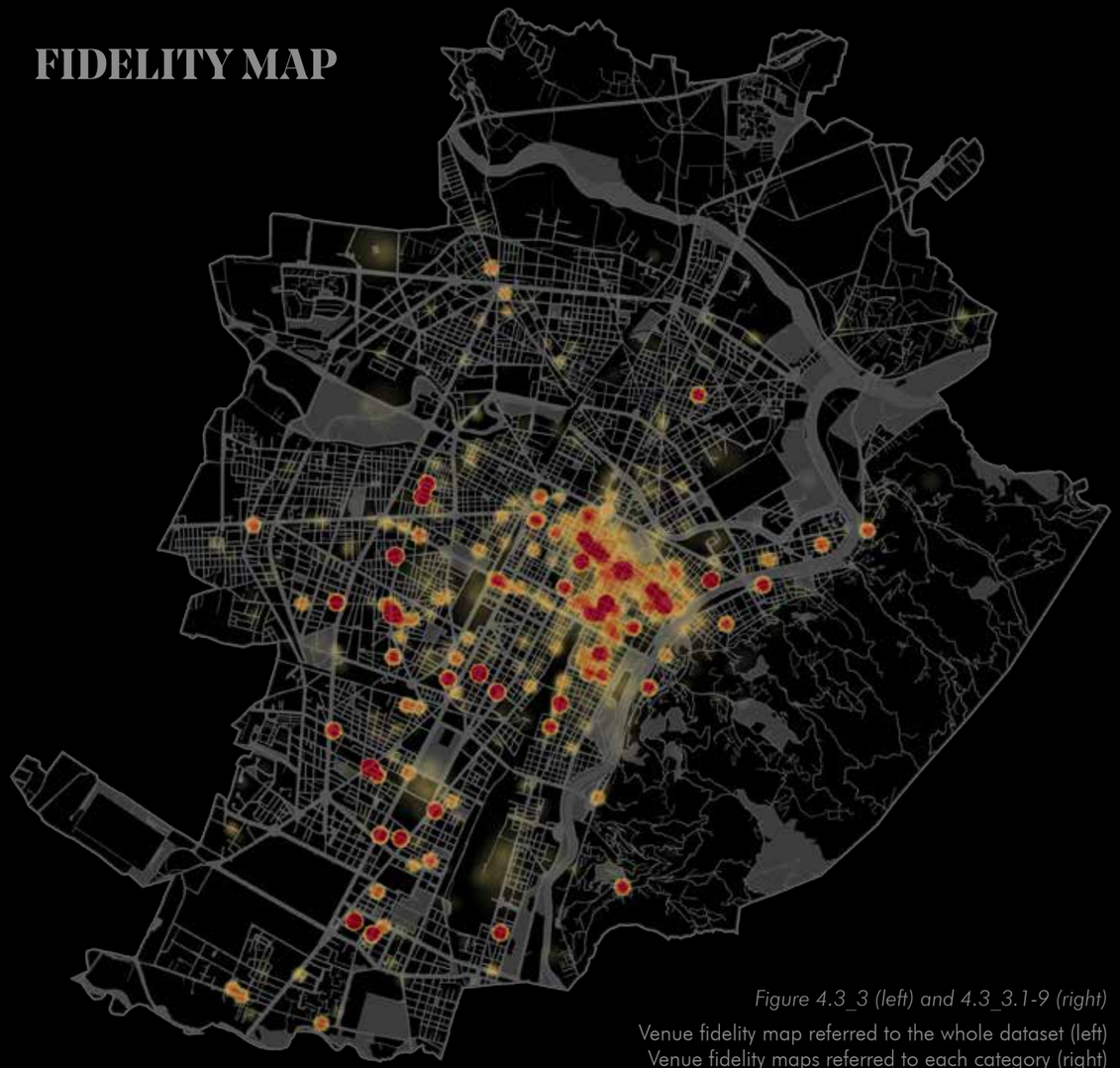
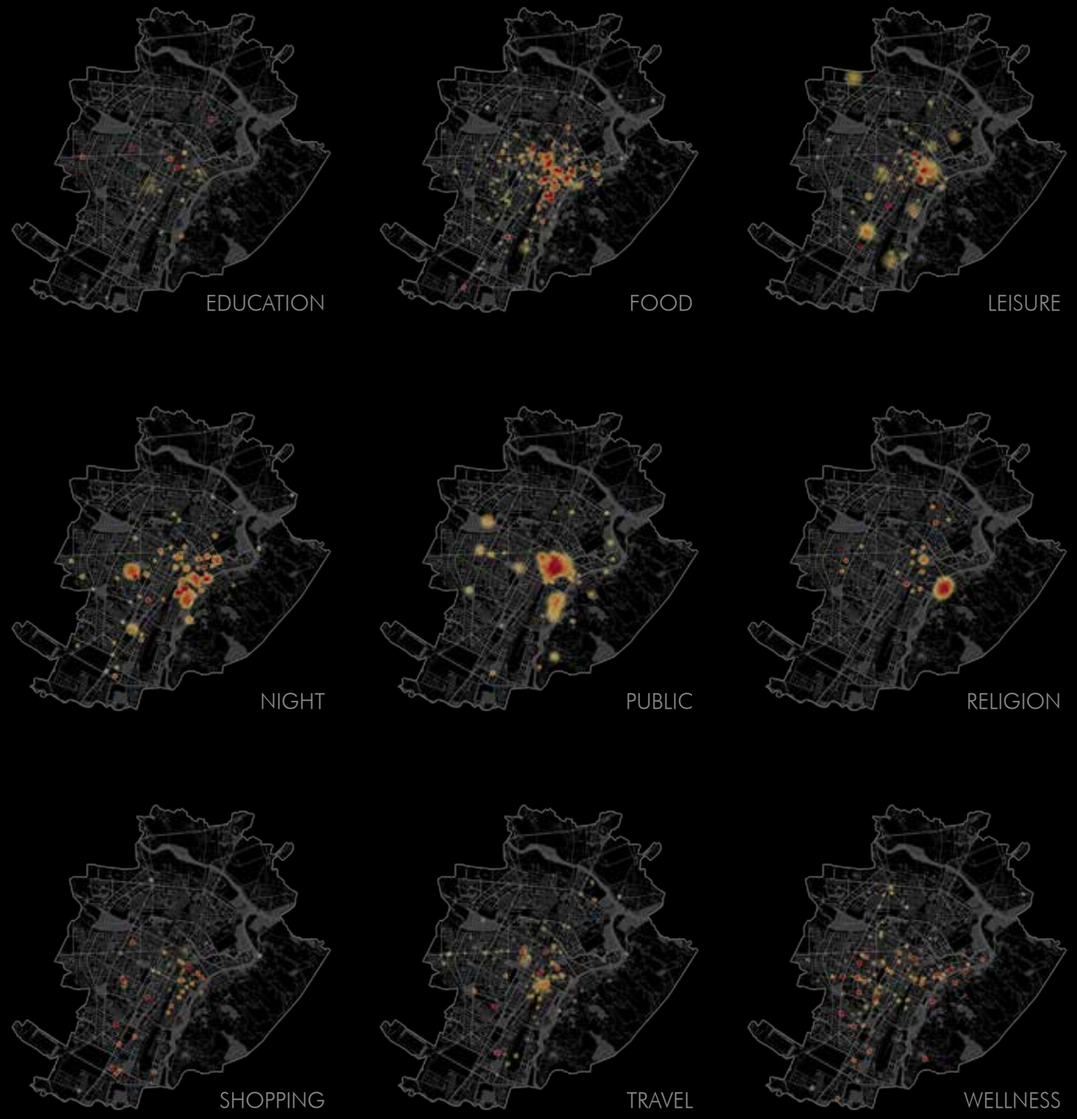


Figure 4.3_3 (left) and 4.3_3.1-9 (right)

Venue fidelity map referred to the whole dataset (left)
Venue fidelity maps referred to each category (right)



BIBLIOGRAPHY

Boy, J. D., & Uitermark, J. (2017). Reassembling the City Through Instagram. *Transactions of the Institute of British Geographers, 42,* 612-624.

Chappelow, J. (2019, Sep 18). Gini Index. *Investopedia,* retrieved from https://www.investopedia.com/terms/g/gini-index.asp

Psyllidis, A., & Bozzon, A. (2015). Harnessing Heterogeneous Social Data to Explore, Monitor, and Visualize Urban Dynamics. *14th International Conference on Urban Planning and Urban Management (CUPUM 2015).*

Psyllidis, A. (2016). Revisiting Urban Dynamics through Social Urban Data: Method and tools for data integration, visualization, and exploratory analysis to understand the spatiotemporal dynamics of human activity in cities. *A+BE: Architecture and the Built Environment, 18.*

Yang, J., Hauff, C., Houben, G.J., & Bolivar, C. T. (2016). Diversity in Urban Social Media Analytics. *Web Engineering: Proceeding of 16th International Conference (ICWE 2016),* 335-352

#5
CLUSTERING

The considerations derived from the quantitative evaluation of the dataset show in general how the dynamics represented on the Instagram corresponds to actual dynamics of the city, validating the relevance of the information deriving from this kind of dataset.

A new topography has been shown in the previous chapter allowing new considerations in the way the city can be pictured and understood. Would it be possible to use the entries of the dataset to perform further enquiries in the city? In a world where networks are ubiquitous and connect individuals – and places – beyond physical borders, is it possible to map these networks?

	VENUE 1	VENUE 2	VENUE 3	VENUE 4
USER 1	0	2	0	1
USER 2	1	0	1	0
USER 3	3	0	1	0
USER 4	0	1	0	0
USER 5	0	0	0	1
USER 6	0	1	1	0
USER 7	1	2	1	0

Figure 5.1_1

Construction of vectors recording for each venue the number of post from each user in the dataset

#5.1
Graph construction

The study performed by Cranshaw et al. started from a similar question: are the geographical/topographic neighborhoods the areas within which people actually live? They mapped the "livehoods" of same American cities based on Foursquare check-ins, clustering venues according to the "social affinity" of them. (Cranshaw et al., 2012)

In this research I started from the model proposed by Cranshaw to evaluate the social similarity between the venues in the dataset.

The approach consists in considering the set U of the n_U Instagram users, the set V of n_V venues and the set P of geotagged posts.

1. Cranshaw, 2012, p. 59

Each venue v is represented with the number of posts geotagged in v, computing «an n_U dimensional vector c_v, where the u^{th} component of c_v is the number of times user u» geotagged a post in « v. Under this representation, we can compute a social similarity $s(i; j)$ between each pair of venues $i; j \in V$ by comparing the vectors c_i and c_j».[1]

The social similarity $s(i; j)$ is calculated with cosine similarity as:

$$s(i; j) = \frac{c_i * c_j}{||c_i|| \; ||c_j||}$$

2. *ibidem*

Resulting in an n_V x n_V «affinity matrix»[2] $A = (a_{i;j})_{i;j=1;\ldots;nV}$, where:

$$a_{i;j} = s(i; j)$$

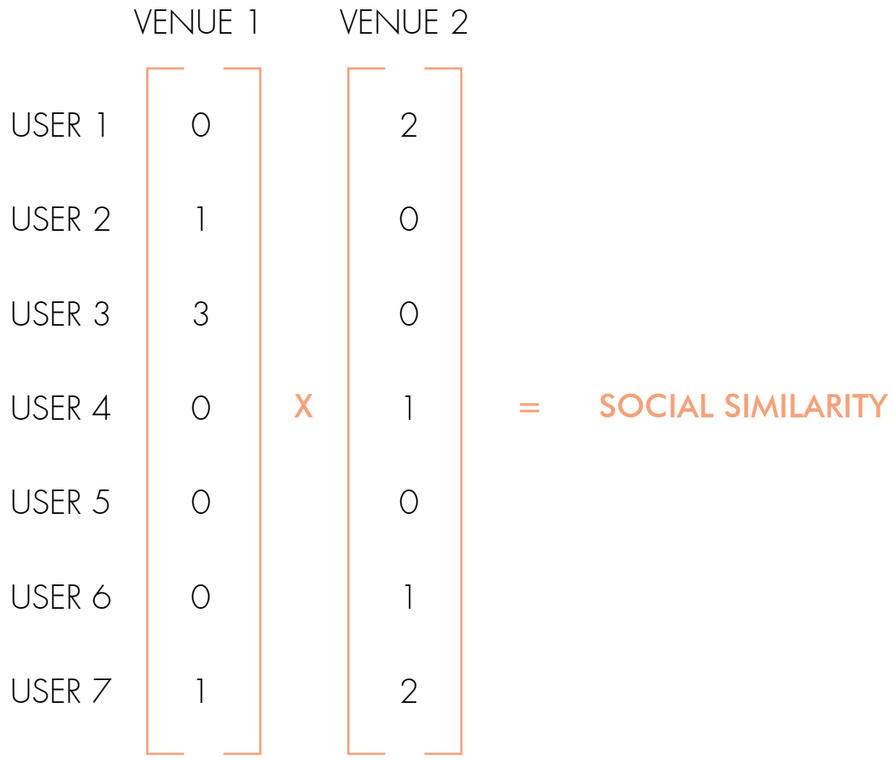


Figure 5.1_2
Calculation of the affinity between venues

To avoid redundancy and noise the dataset was filtered keeping only the venues that had more than 52 posts – approximately one for each week of the year – and the users who had geotagged at least 3 posts in three different locations. The resulting filtered dataset consisted in 117 042 posts geotagged in 228 venues.

The subtraction from the 228 x 228 affinity matrix calculated as explained above of the corresponding identity matrix, allowed the creation of a graph with 228 nodes and 15 495 edges weighted according to social affinity, with an average degree of 135.92°.

The graphical representation of this graph in [Figure 5.1_3] depicts the edges, filtered for visual clarity, with a thickness proportional to the weight of them, and the nodes with a radius proportional to the betweenness centrality[3] within the network, or, in other words how often a node is on the shortest path between other nodes.

As we can see in [Figure 5.1_4] where the dimension of nodes is proportional to the number of posts, "popularity" does not guarantee the centrality within the network of venues, connected by the sharing of same users.

These graphs allow to look at the hierarchy of places within the city in a new way, and advance some consideration on the dimension of centrality and periphery.

3. Betweness Centrality measures how often a node appears on shortest paths between nodes in the network

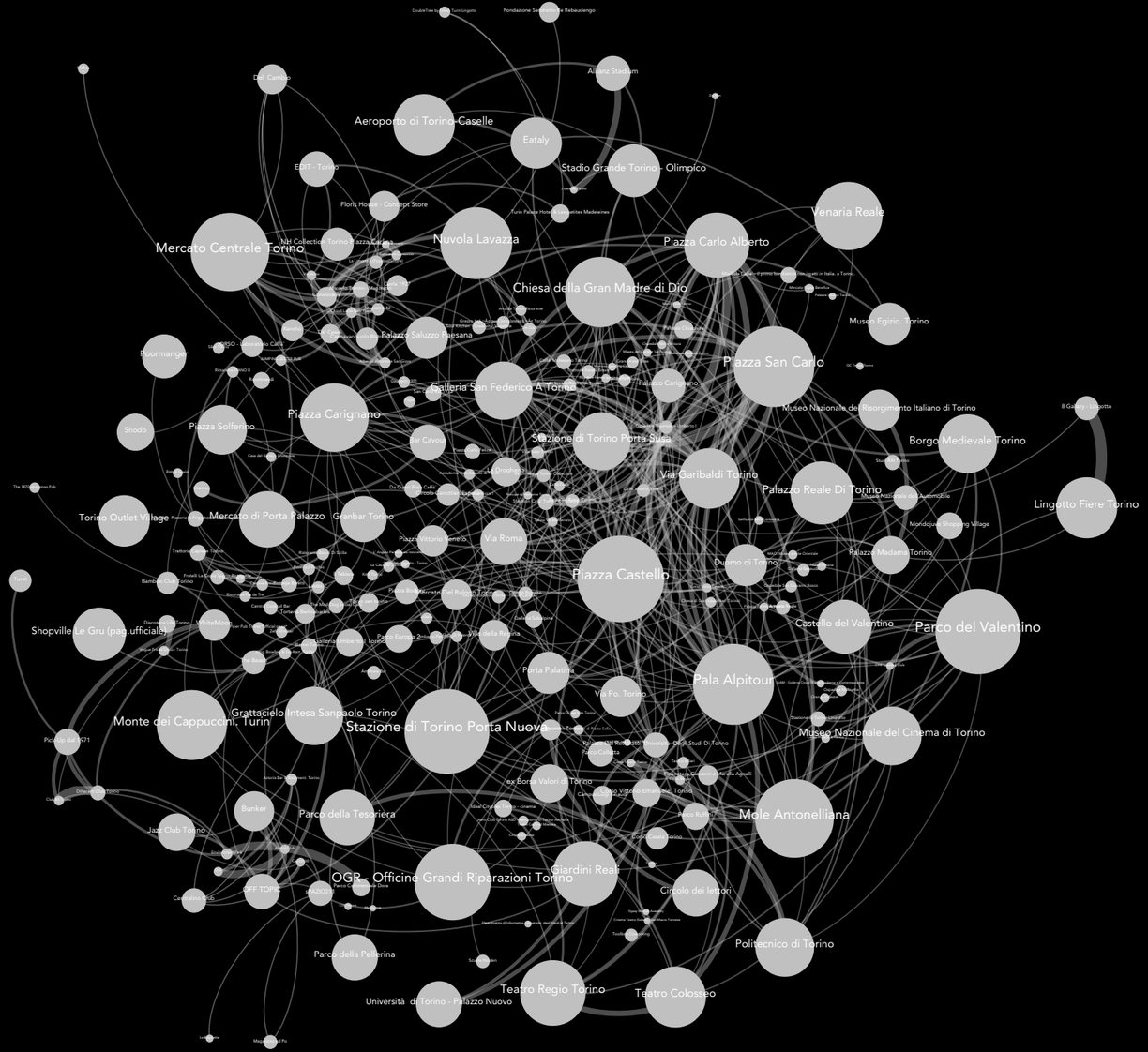


Figure 5.1_3

Network of places weighted with betweeness centrality
(graph constructed with Gephi and Force Atlas 2 as layout)

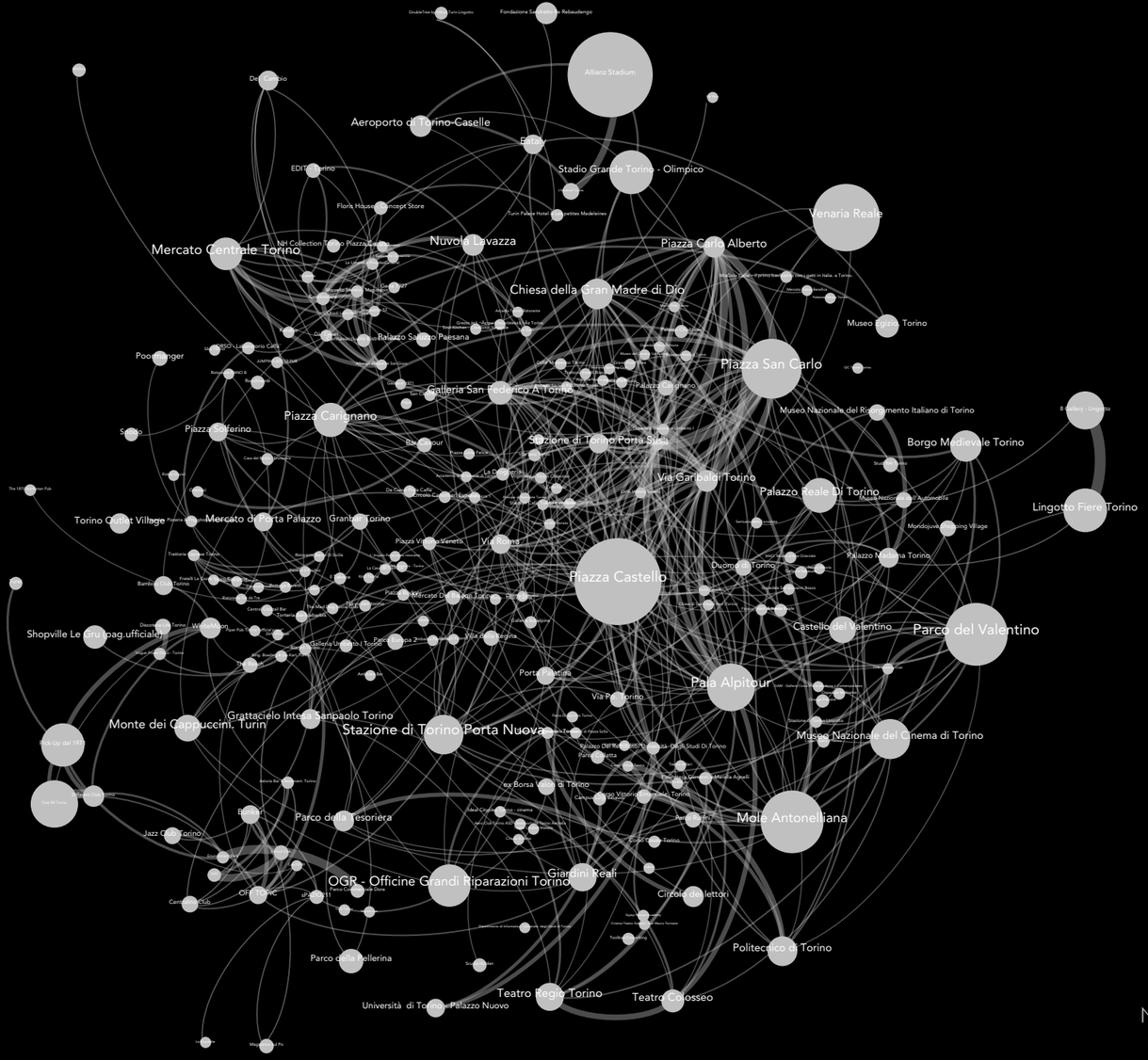


Figure 5.1_4

Network of places weighted with umber of posts per venues
(graph constructed with Gephi and Force Atlas 2 as layout)

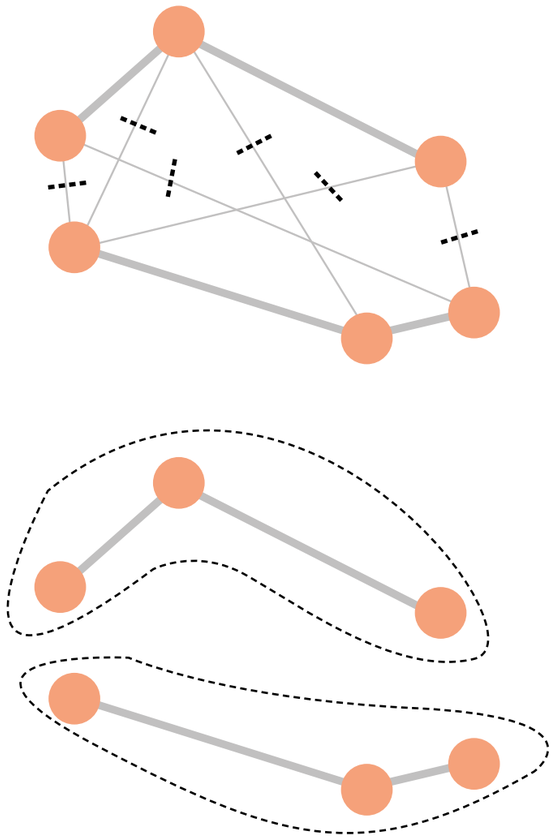


Figure 5.2_1

Graphical explanation of the expected clustering process
by "cutting" the weakest edges and grouping resultin venues

#5.2

Community detection

How can places be grouped in "neighborhoods" using the old concept of neighborhood as the expression of a community of users? Hot to cut the weakest link in the graph and group those with a higher weight?

Cranshaw's approach continued with spectral clustering, grouping locations mainly in relation to their "social affinity" combined with the spatial proximity. As this research aims at locating networks beyond physical borders, another method was used to cluster locations.

The technique called "community detection" also know as Louvain method of modularity optimization, progressively groups together nodes in network until it reaches an optimal level of clustering. (Blondel et al., 2008). Boy and Uitermark used the same methodology to map community of users in the city of Amsterdam based on the mutual recognition with likes and comments under Instagram posts. Instead, in this research I used the weighted nodes accounting only for the social affinity of places. (Boy & Uitermark, 2016).

Gephi allows to run the algorithm and the modularity was set to 0.8 detecting 12 communities empirically evaluated as consistent and with a sufficient grain of accuracy.

The clusters have been divided in three macro-groups and develloped singularly in the following paragraphs, namely "Landmarks", "Social life", "Isolated networks".

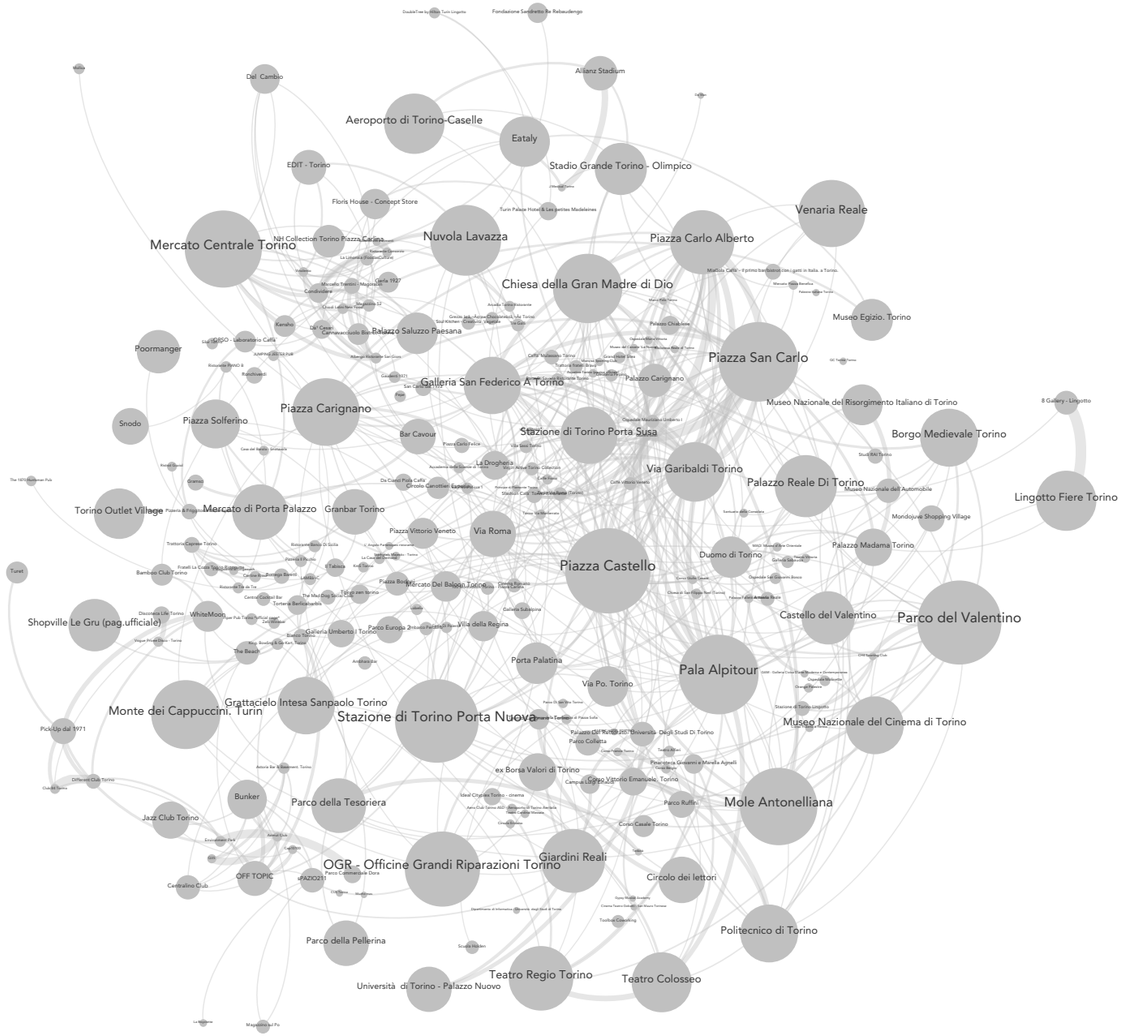


Figure 5.2_2

General network of venues
(graph constructed with Gephi and Force Atlas 2 as layout)

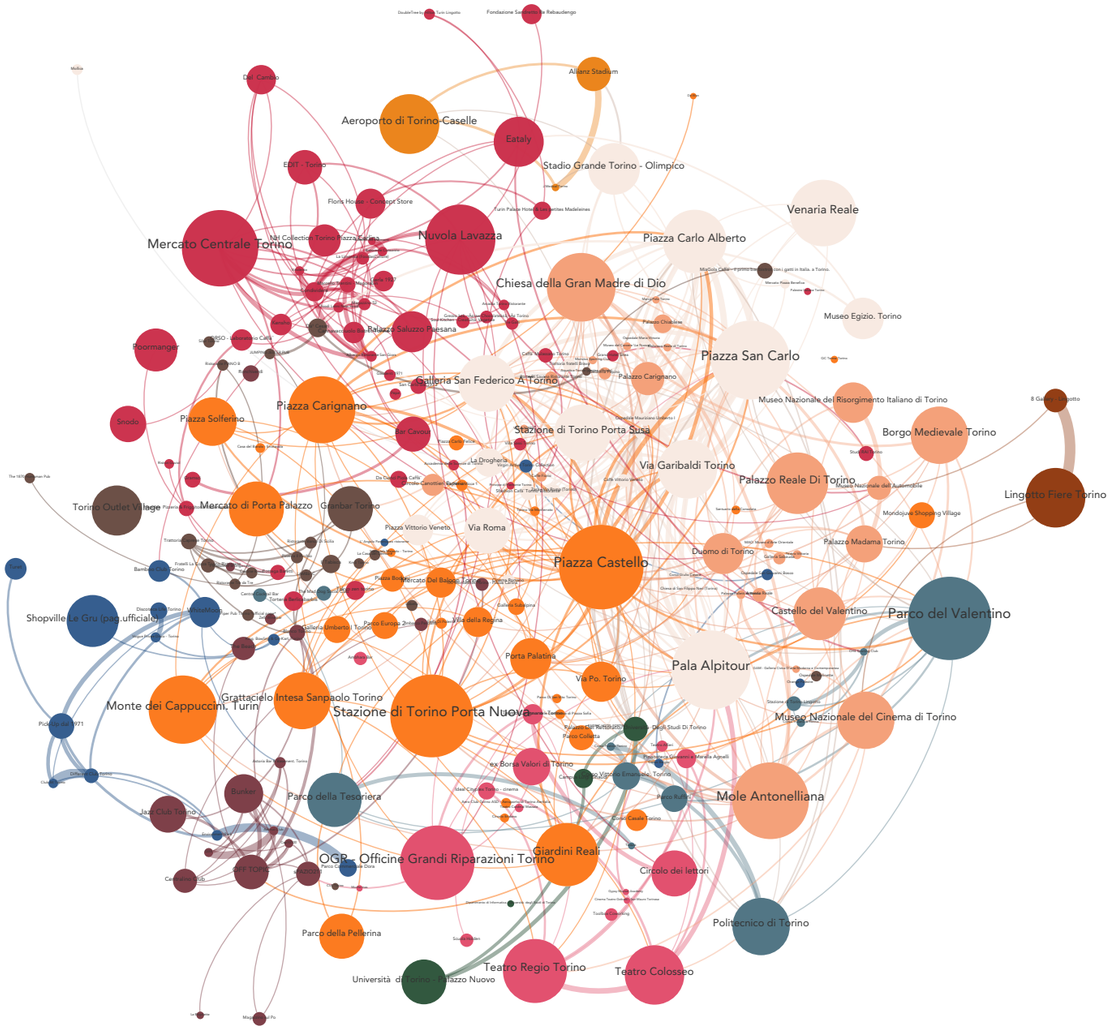


Figure 5.2_3

The twelve communities highlighted in the general network
(graph constructed with Gephi and Force Atlas 2 as layout)

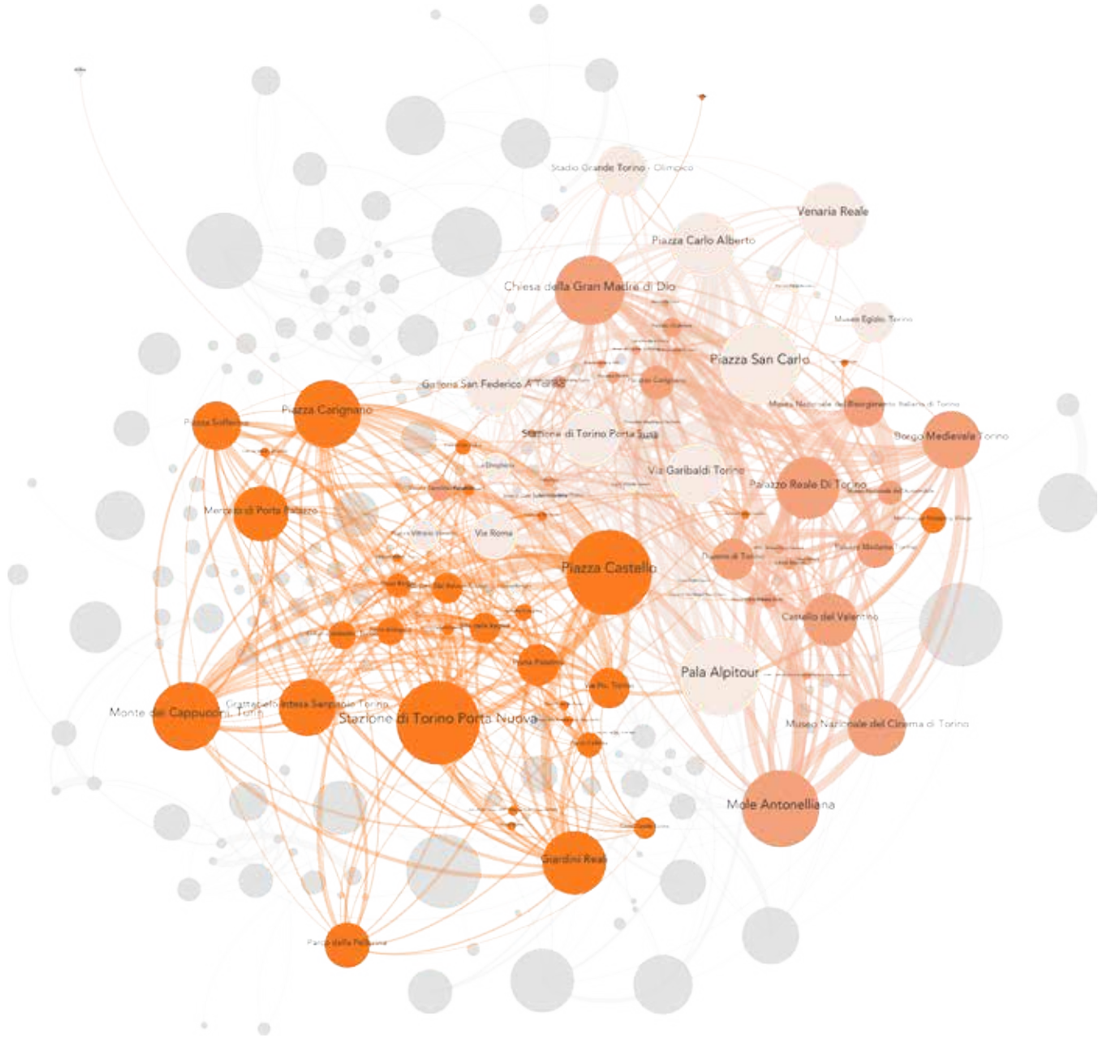


Figure 5.2.1_1

The three clusters of "Landmarks" highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

#5.2.1
LANDMARKS

The three clusters of this macro-group contain mainly venues belonging to the categories of "Public", "Leisure" and "Religion and are characterized by the presence of the most important landmarks of the city of Turin.

An empirical reading would see connections and meanings behind the clustered results. LM_1 contains venues we can link to a sort of "Turin local flavor" residents might enjoy. There are the two markets of Porta Palazzo and Baloon, the scenic spot of the Monte dei Cappuccini and other peculiar spots of the city.

LM_2, instead, groups the venue more iconic, famous and probably more touristics, like the Mole Antonelliana, the Chiesa della Gran Madre, the Castello del Valentino etc.

LM_3 is similar to LM_1 but with a subtle difference. The major nodes of the cluster are Piazza San Carlo, the palace of Venaria Reale, Pala Alpitour and the Stadio Grande Olimpico, places characterized by their meanings for locals but more oriented to specific activities. Piazza San Carlo, for example, hosts events and concerts, as well as the other venues are usually rarely visited without an organized activity which takes place there.

The other venues connected to the central hubs sometimes do not fit in the empiric definition given above, but it is interesting to notice how the community detection reconduct them to these systems and wonder about the process of identification of individuals.

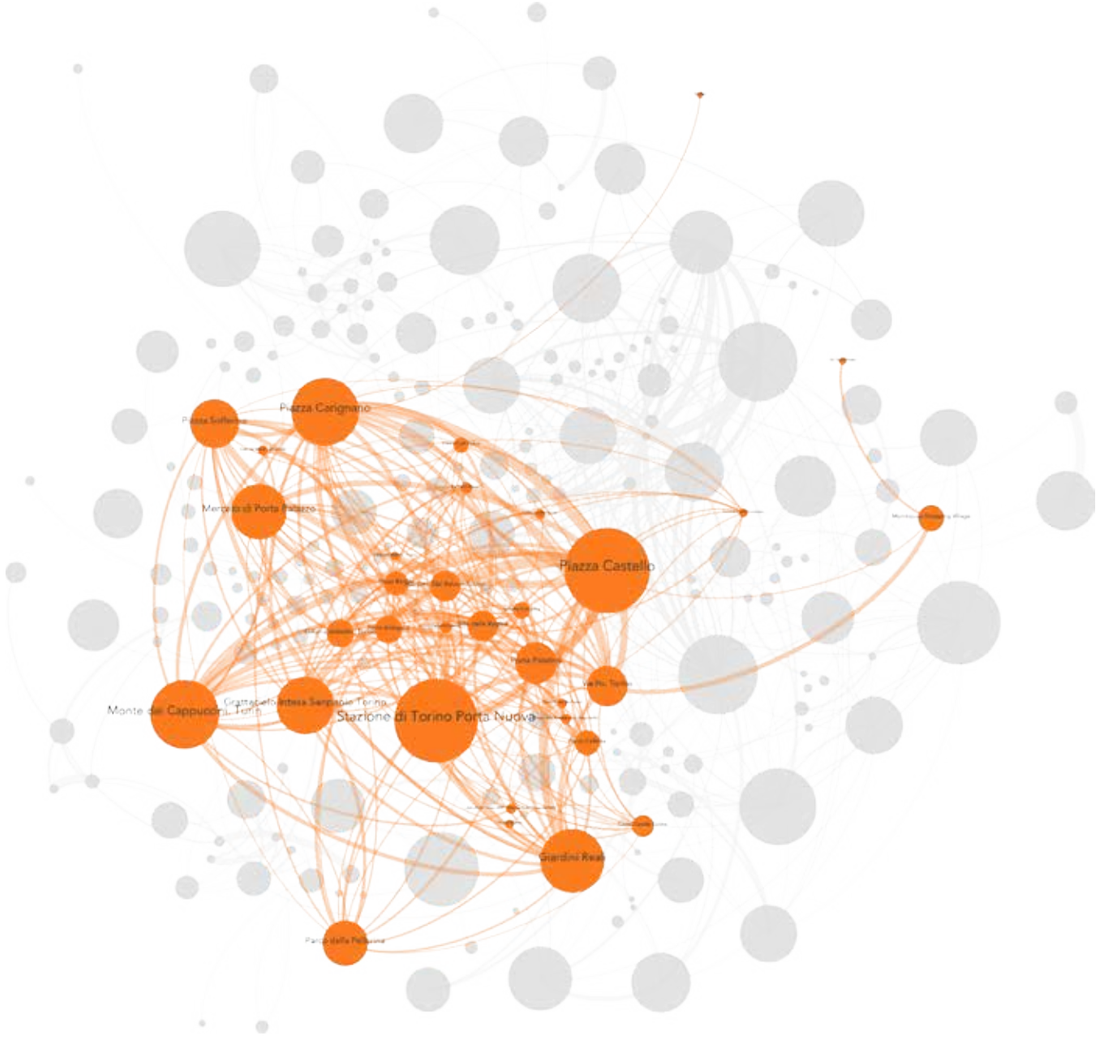


Figure 5.2.1_LM1_1

LM_1 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

LM_1

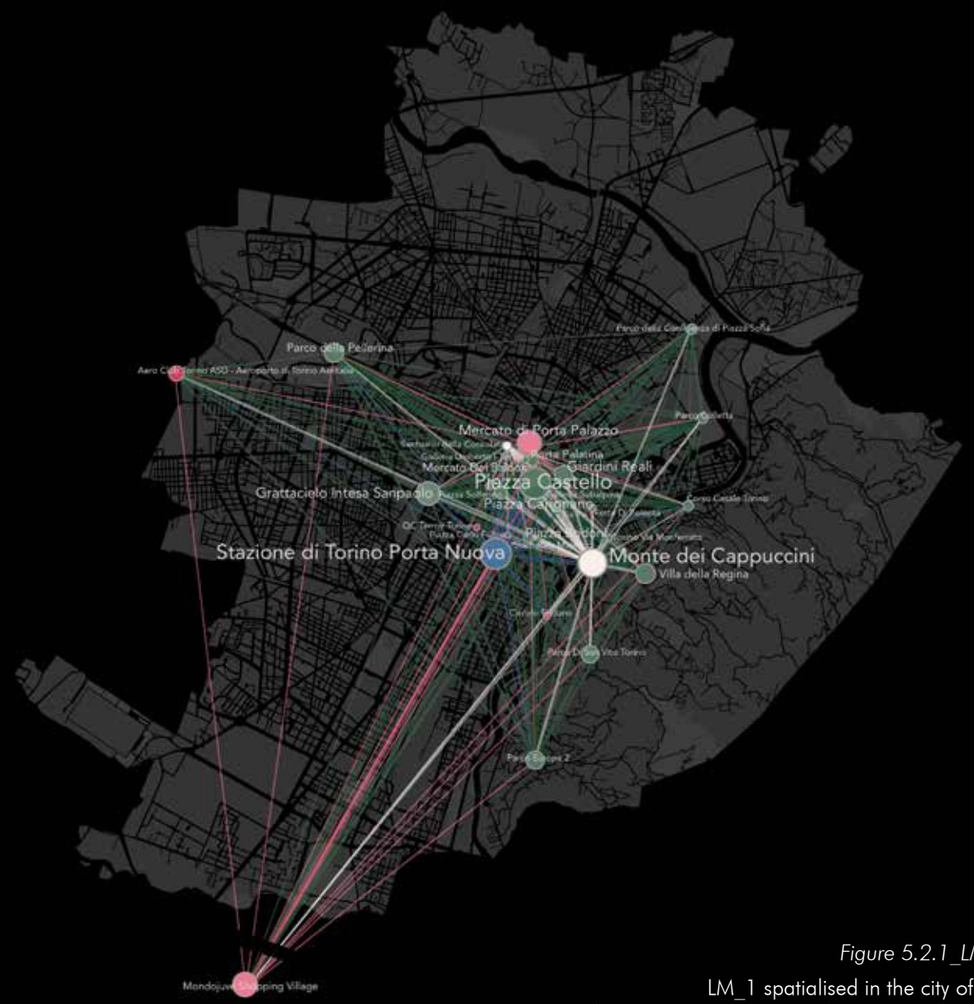


Figure 5.2.1_LM1_2

LM_1 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

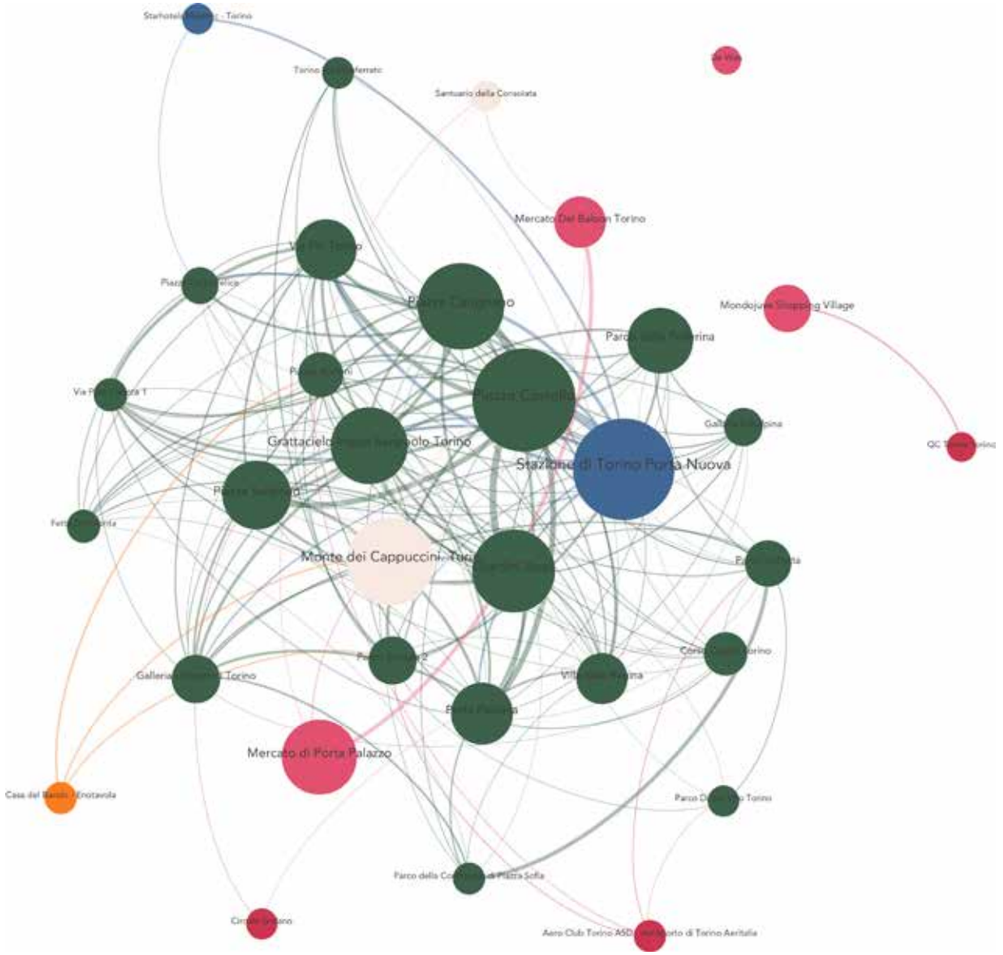
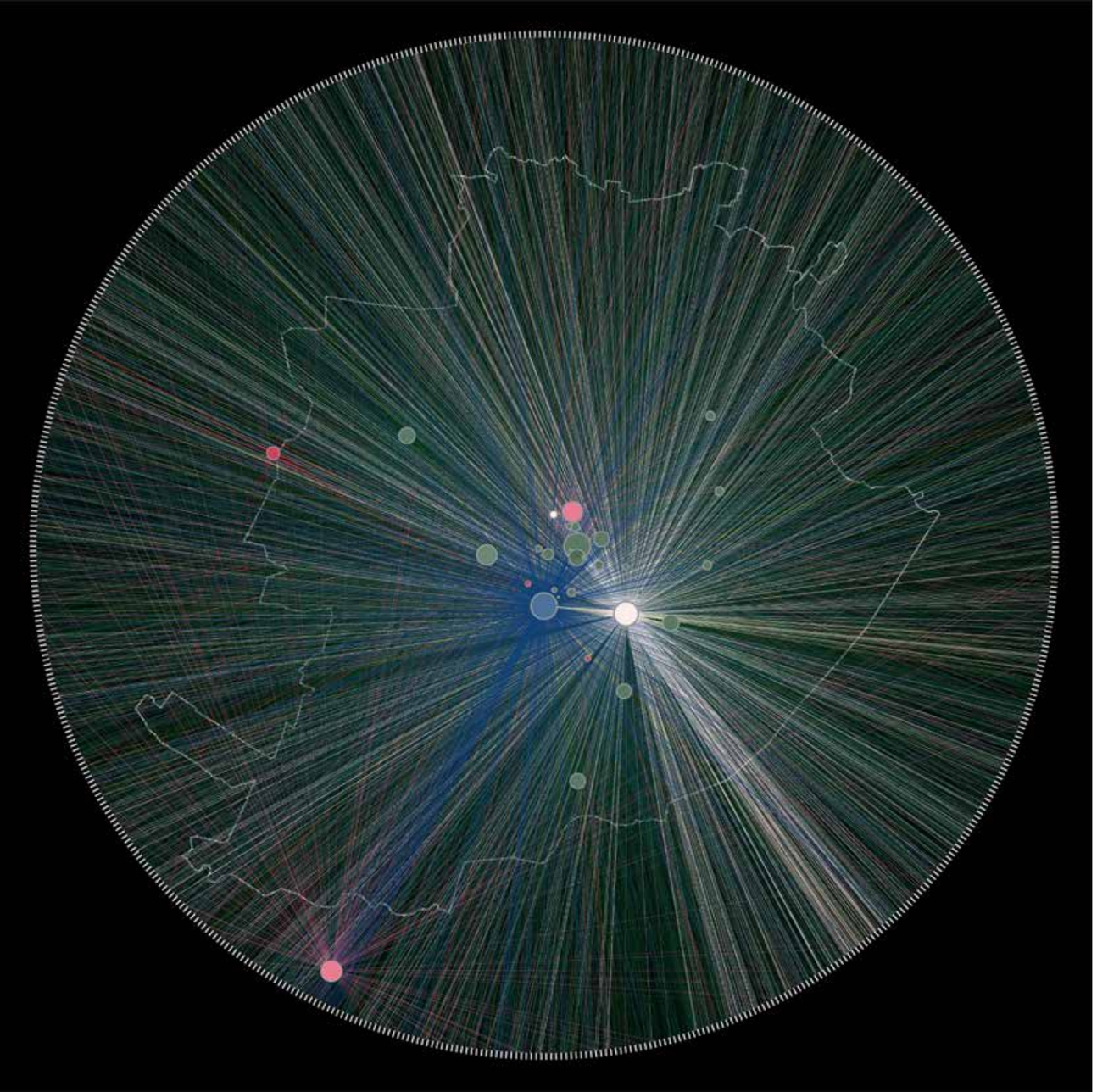


Figure 5.2.1_LM1_3

LM_1 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



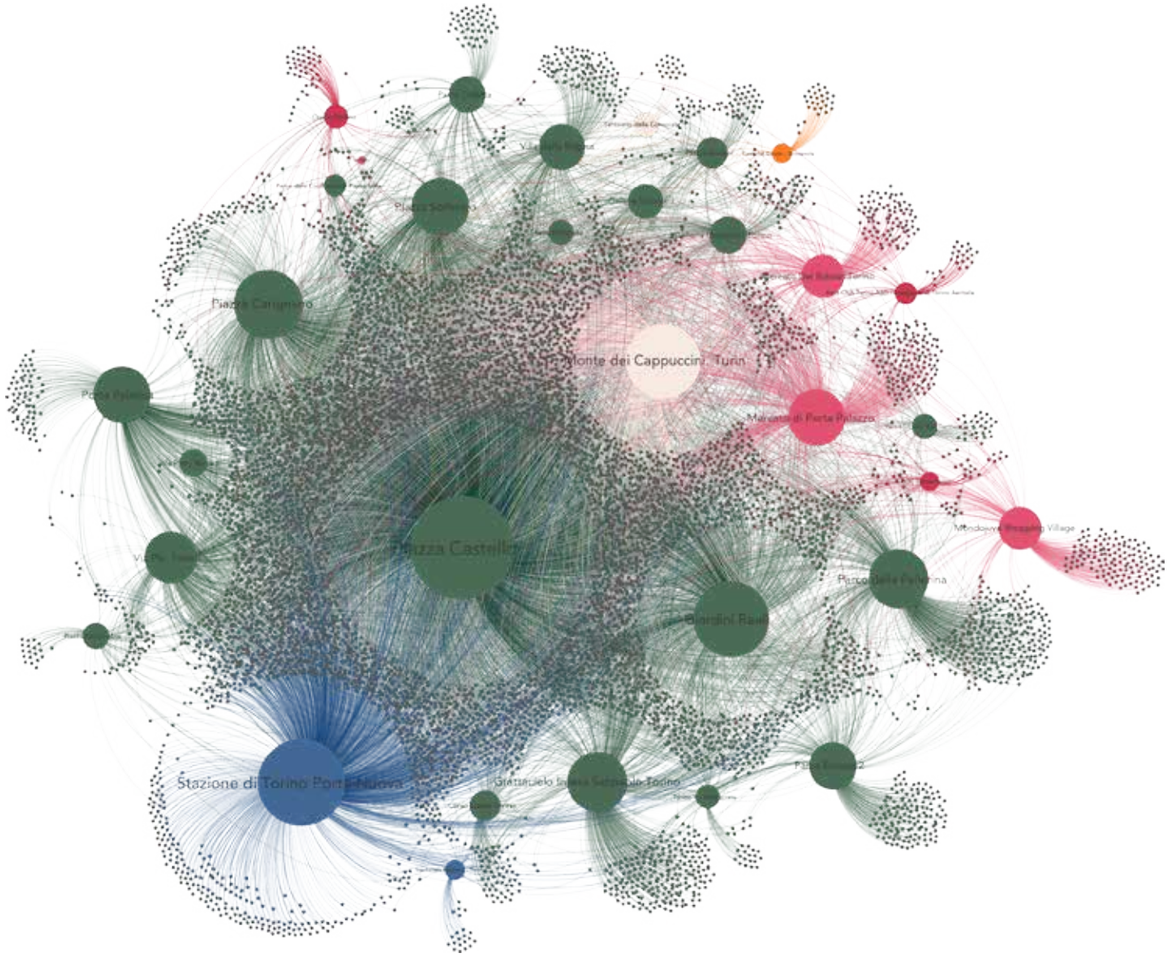


*Figure 5.2.1_LM1_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

Figure 5.2.1_LM2_1

LM_2 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

LM_2

Figure 5.2.1_LM2_2

LM_2 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

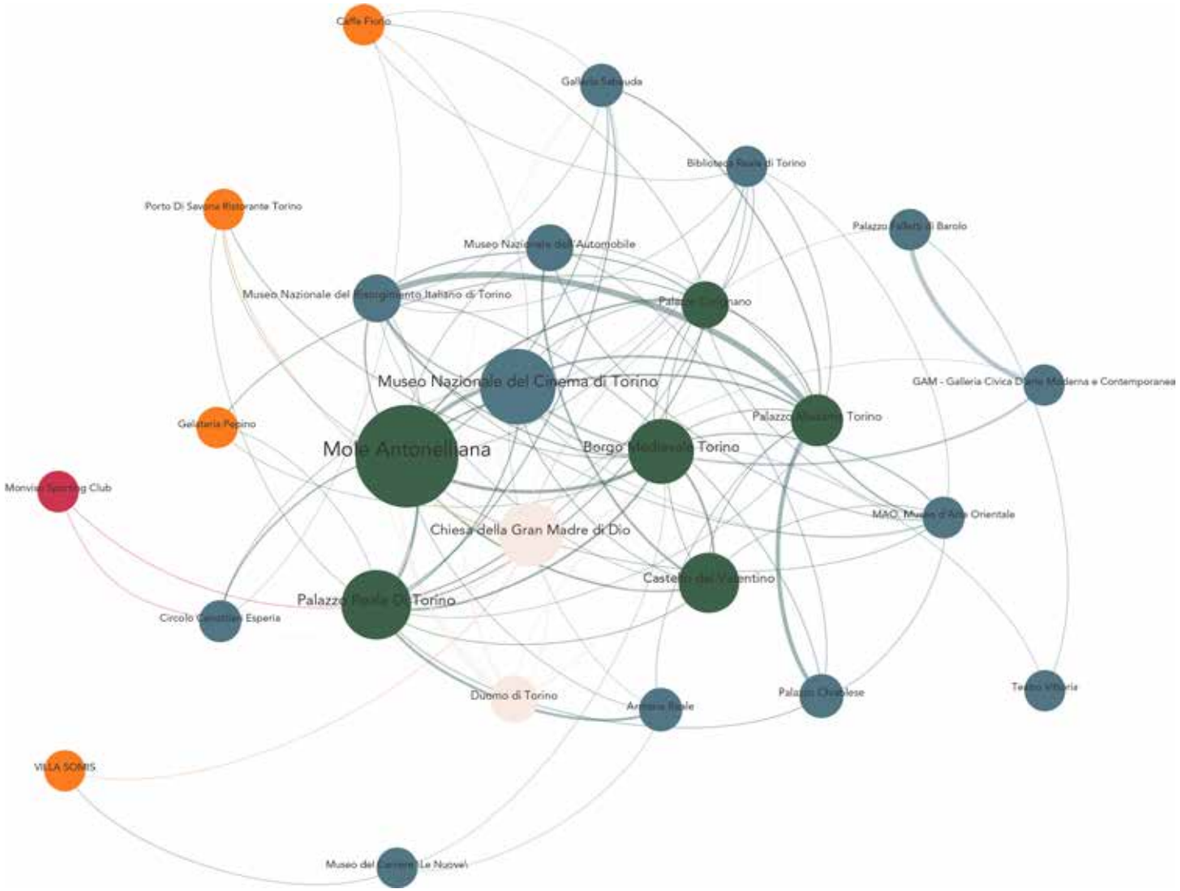
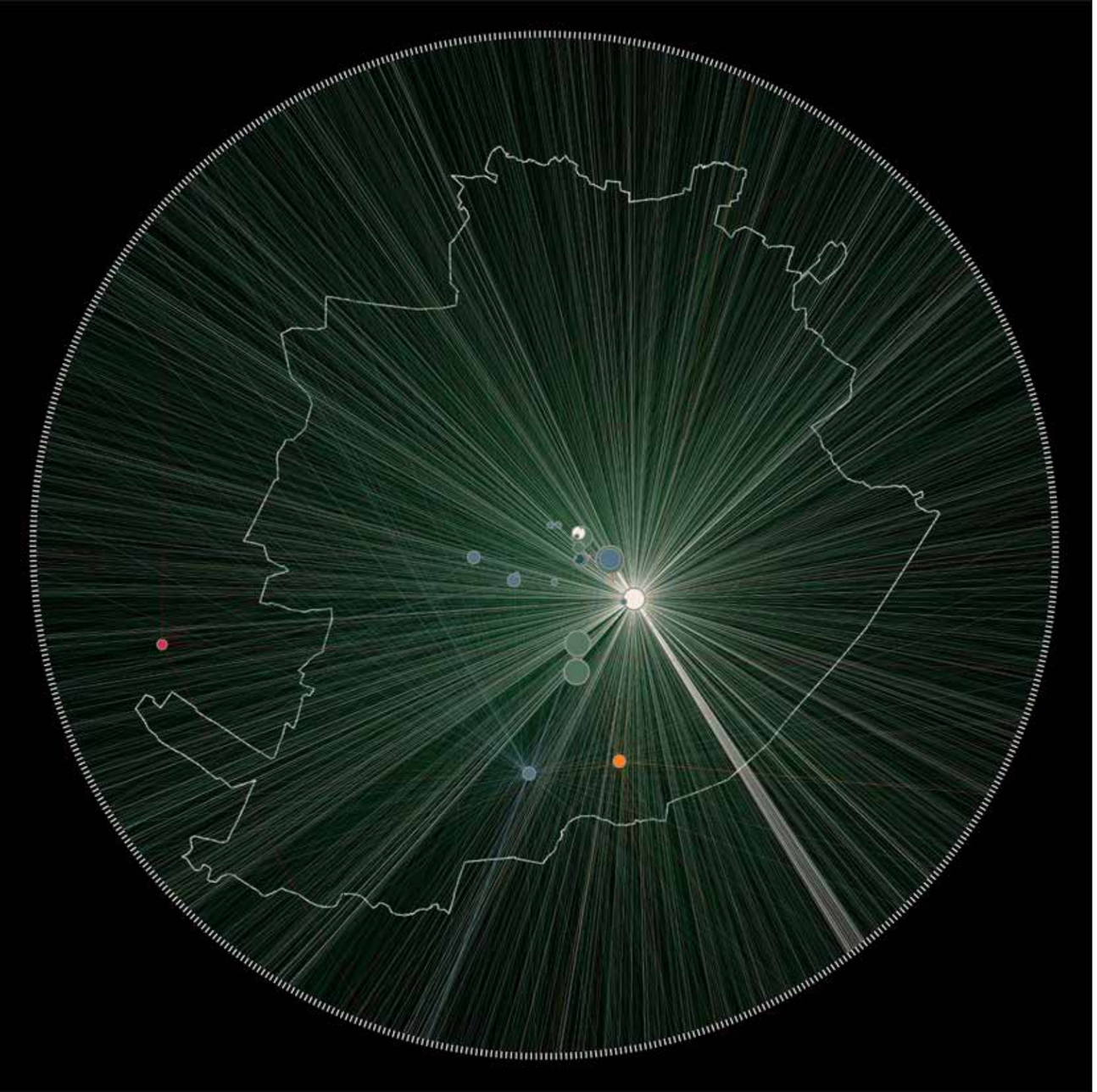


Figure 5.2.1_LM2_3

LM_2 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



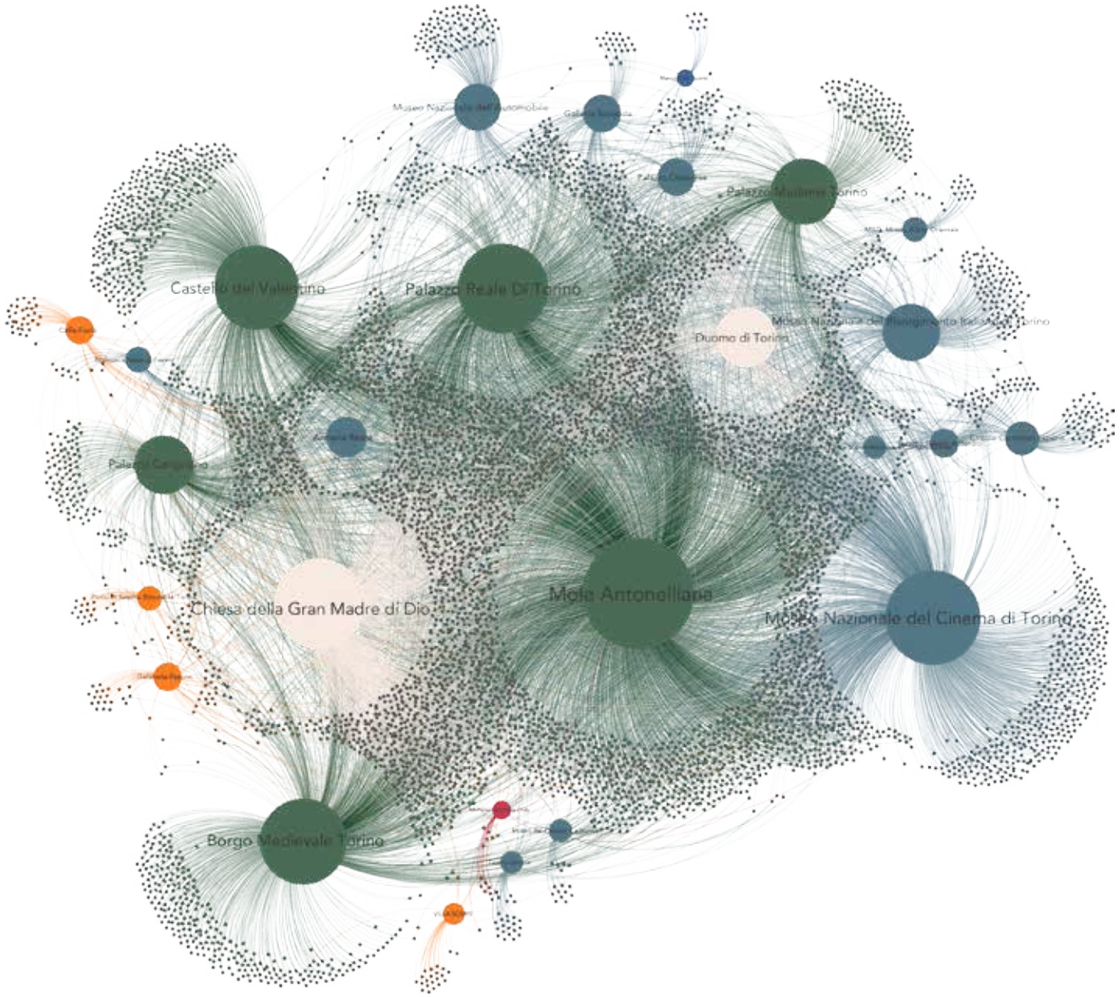


*Figure 5.2.1_LM2_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

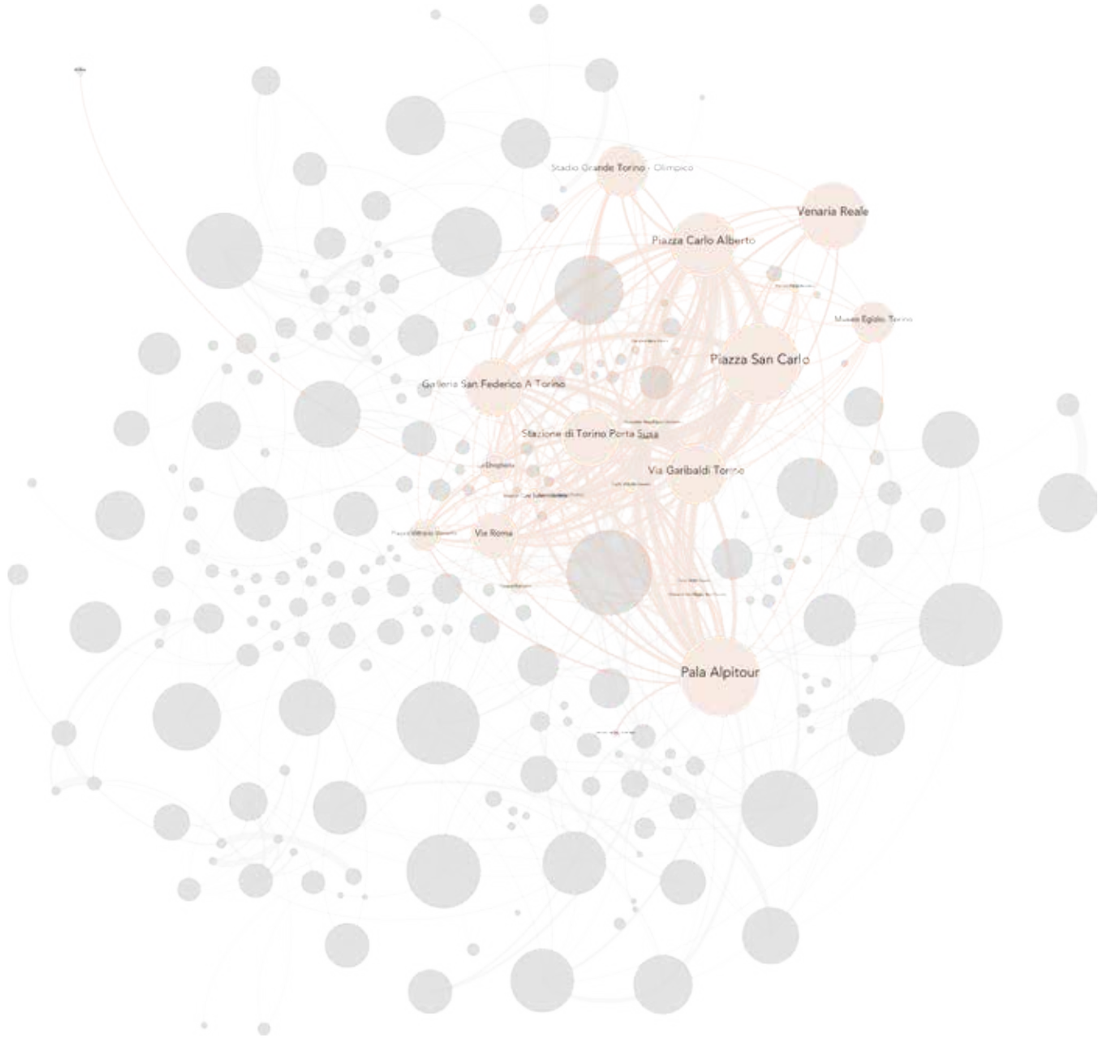


Figure 5.2.1_LM3_1

LM_3 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

LM_3

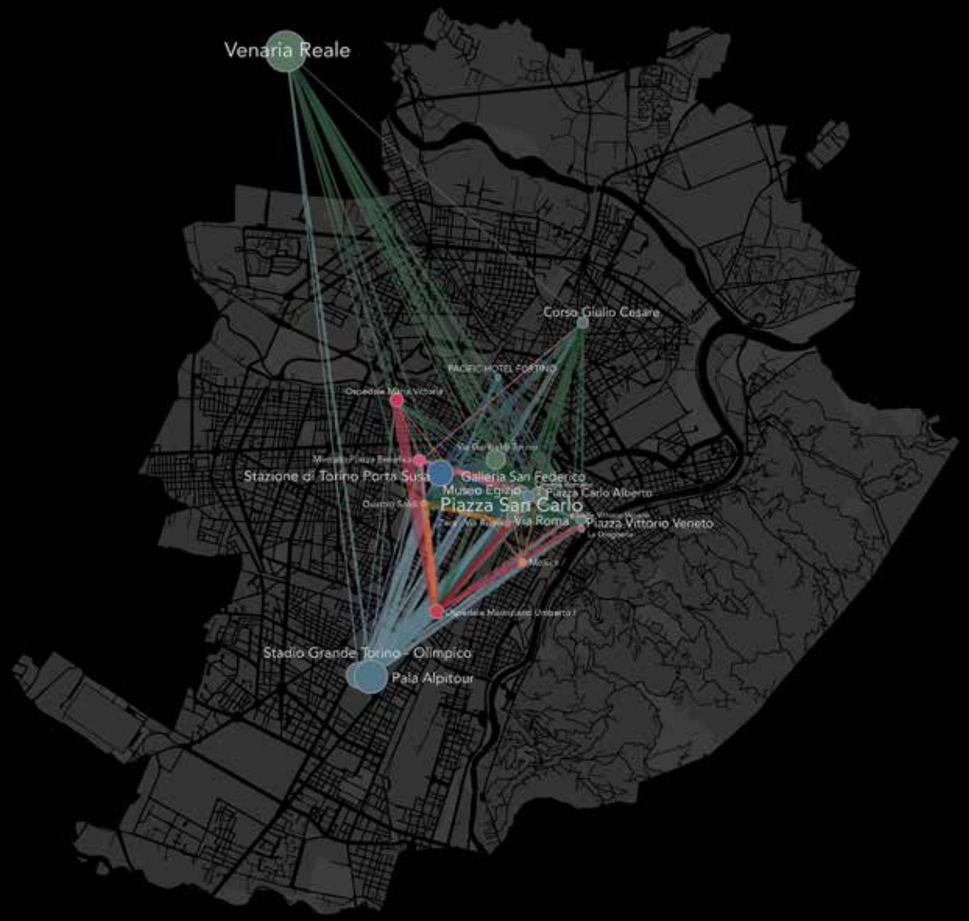


Figure 5.2.1_LM3_2

LM_3 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

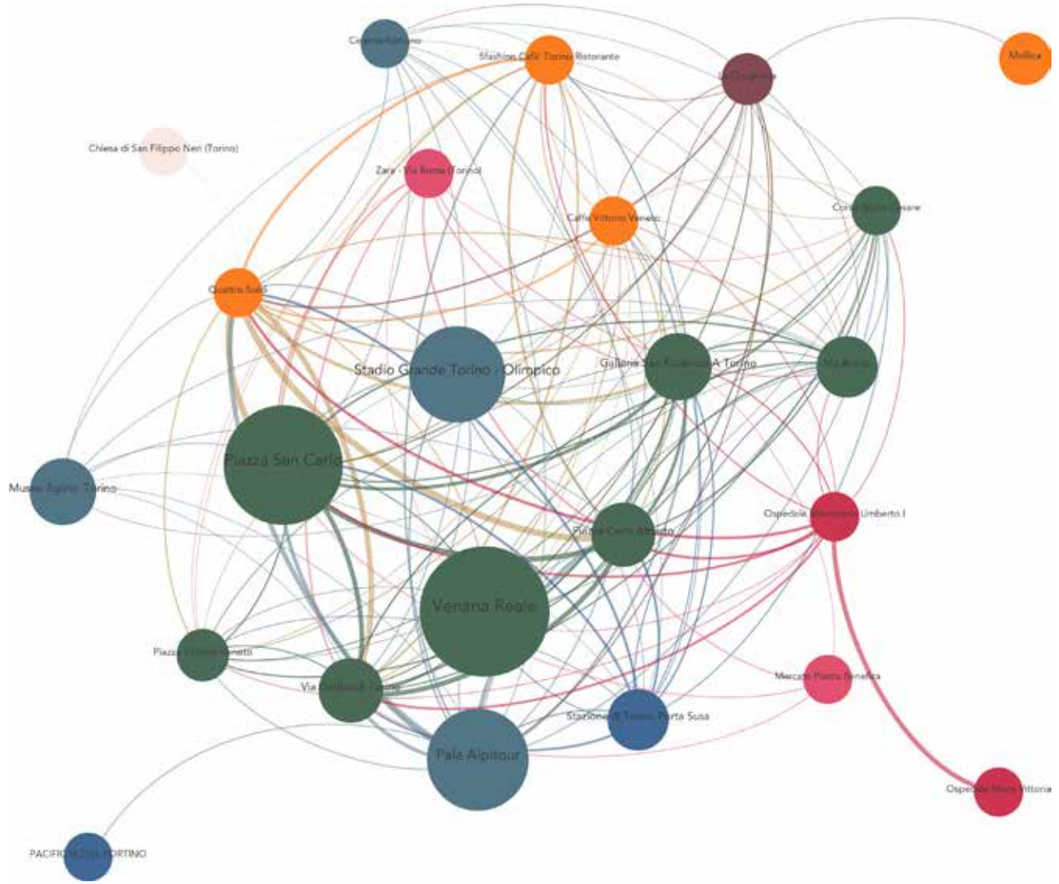
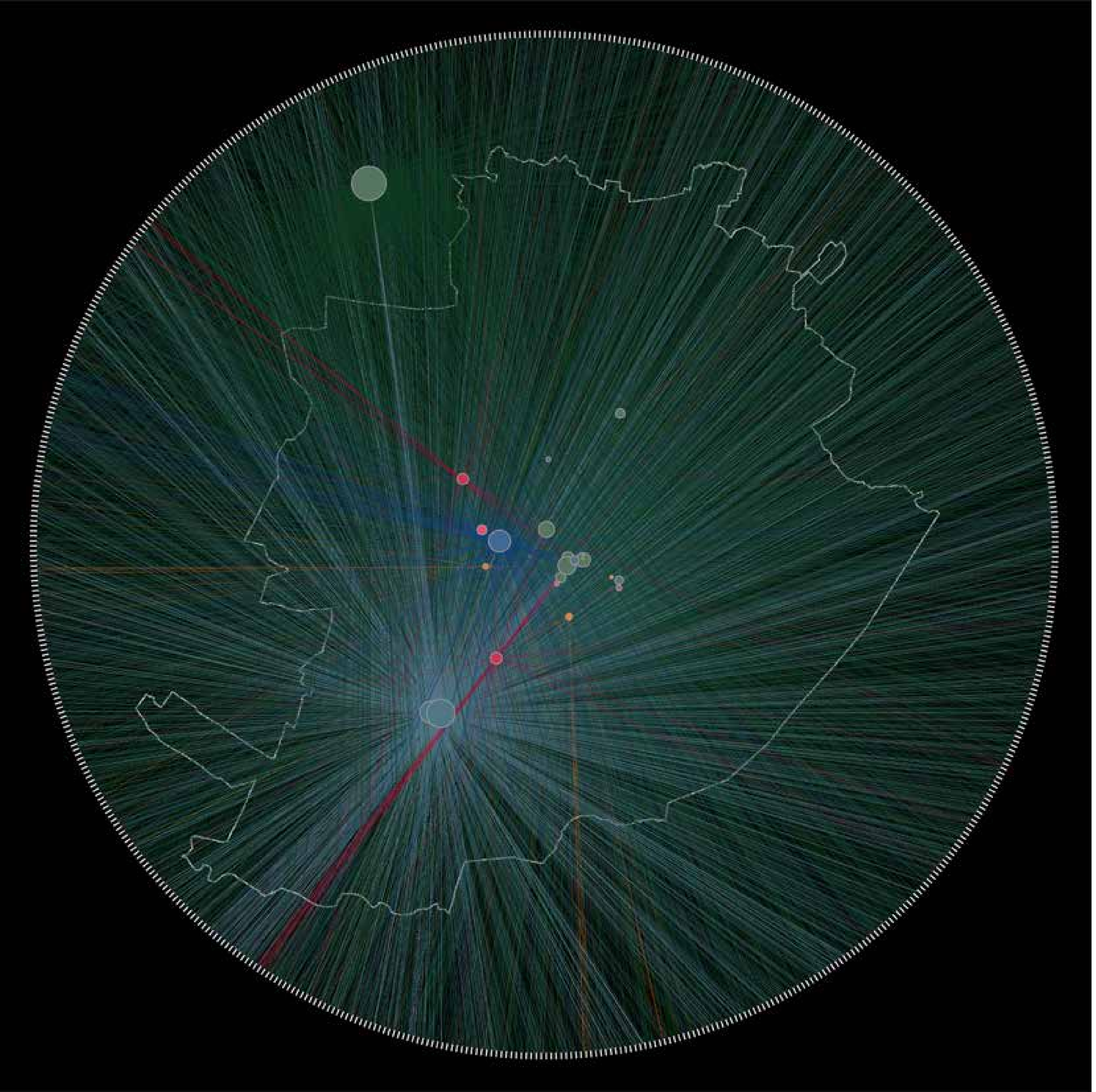


Figure 5.2.1_LM3_3

LM_3 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



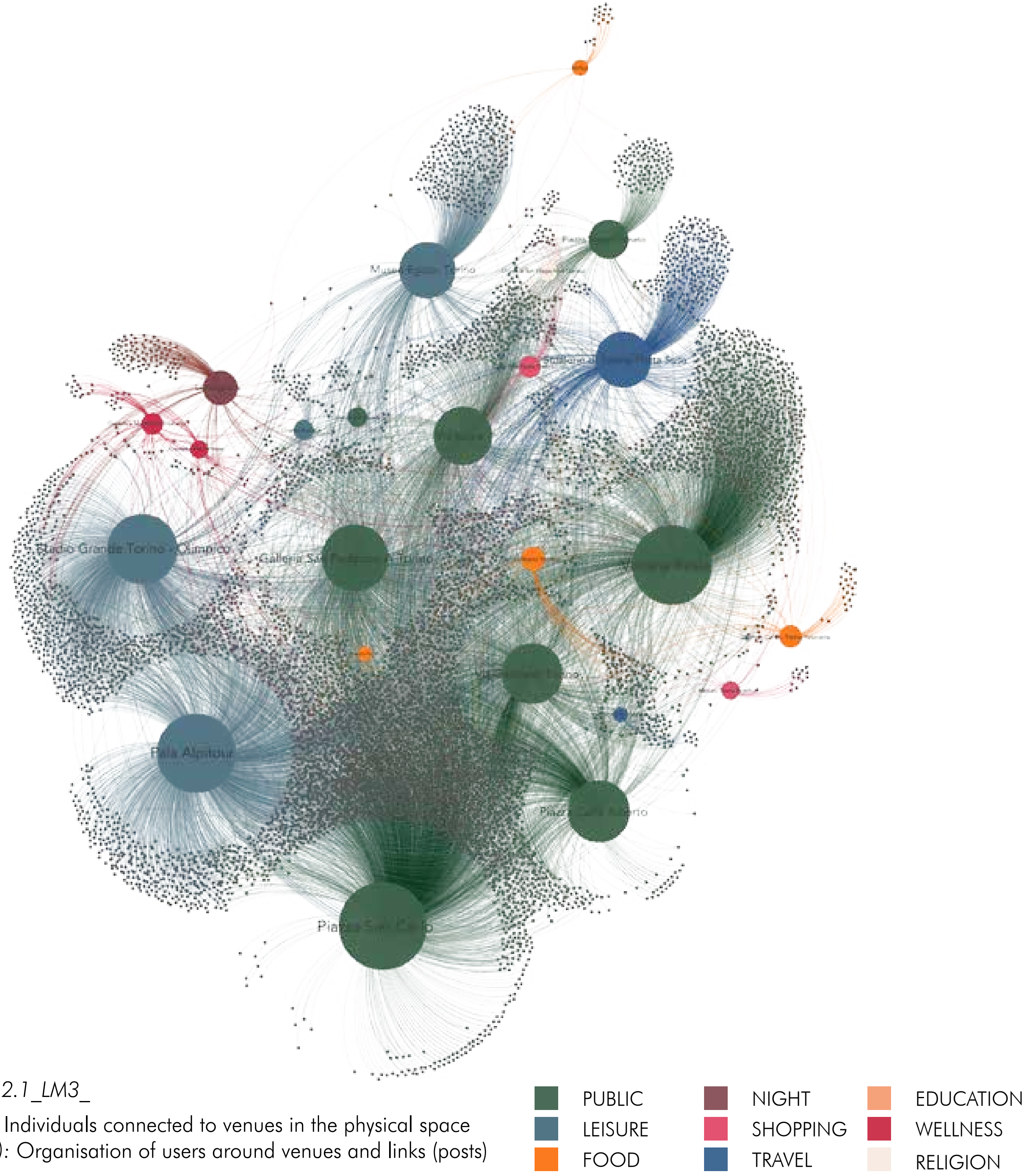


*Figure 5.2.1_LM3_*

_4 *(left)*: Individuals connected to venues in the physical space
_5 *(right)*: Organisation of users around venues and links (posts)

Figure 5.2.2_1

The five clusters of "Social life" highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

#5.2.2
SOCIAL LIFE

This macro-group contains venues mainly belonging to the category of "Food", "Leisure" and "Night", as well as some to "Shopping" less in number but
It is easier here to empirically see the connections between the venues in each cluster.

SL_1 venues of the social life connected with food and directed to an medium-upper class audience.

SL_2 and SL_4 groups venues nightclubs revolving around big shopping mall, giving the general feeling of a more commercial typology of enjoyment.

SL_3 groups venues usually visited by younger individuals with a more sophisticated attitude compared to the audience of the precedent two clusters, varying from popular clubs – the Beach, the Magazzini sul Po – to other clubs usually hosting "underground/hipster" events – Buner, OFF TOPIC, Astoria.

SL_5 is the cluster of sophistication and elegance, where the major theaters of the city are grouped together as well as post-industrial spaces used to host exhibitions and events.

It is of course an empircal reading, but it is surprising how similar venues have been grouped together in communities and interesting to see how other venues which do not fit in the definitions outlined above turn out to be part of some aesthetic systems.

Figure 5.2.2_SL1_1

SL_1 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

SL_1

Figure 5.2.2_SL1_2

SL_1 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

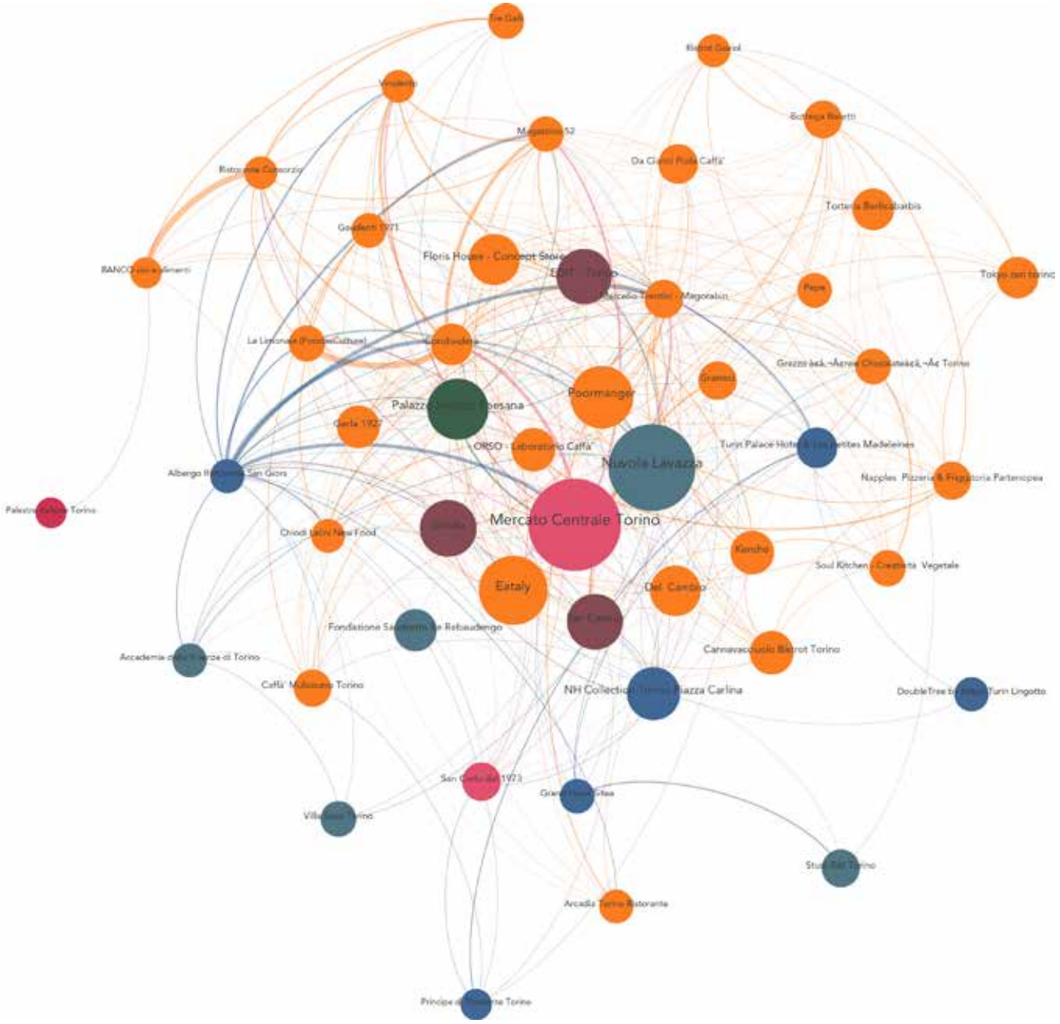
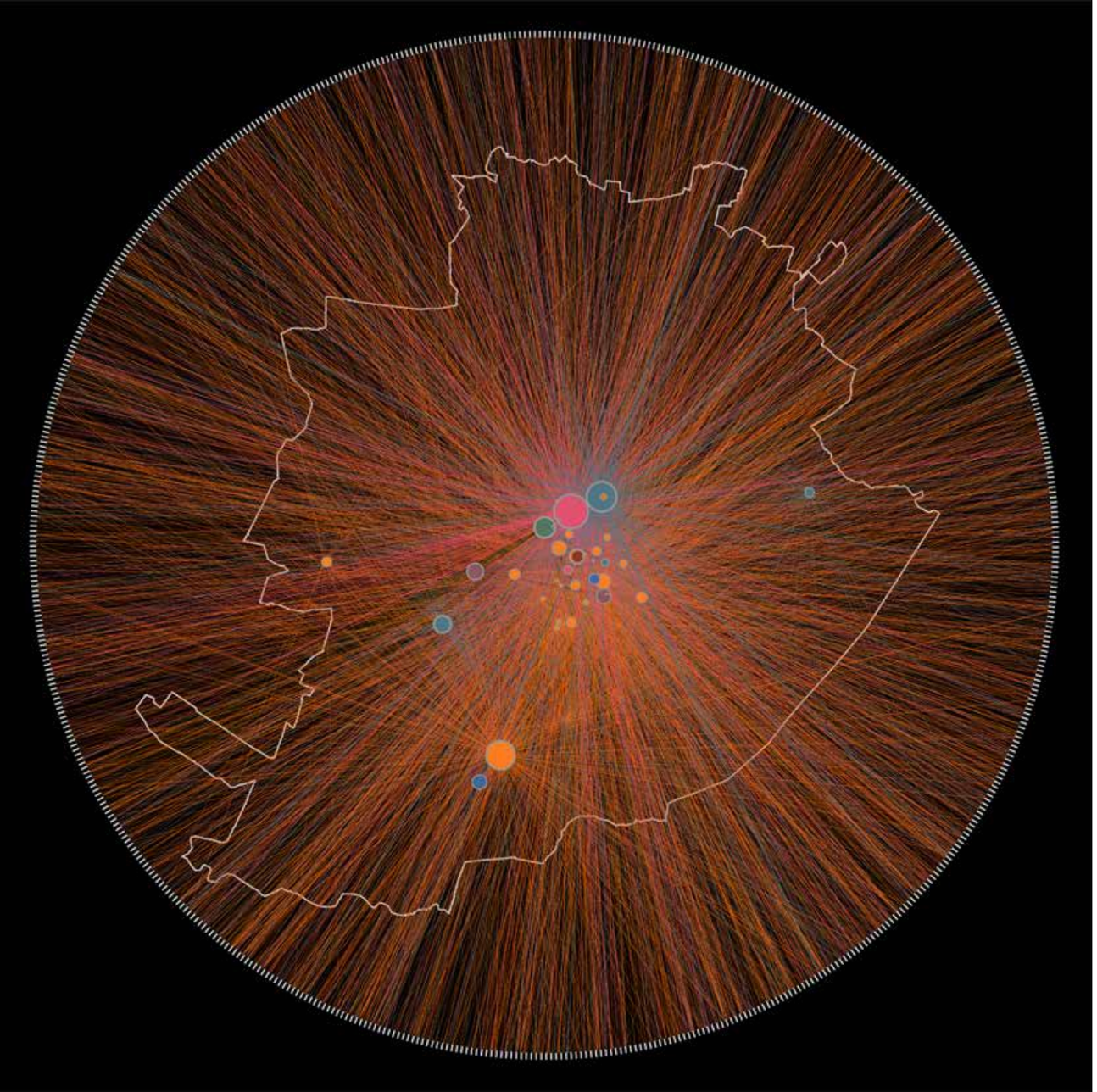


Figure 5.2.2_SL1_3

SL_1 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



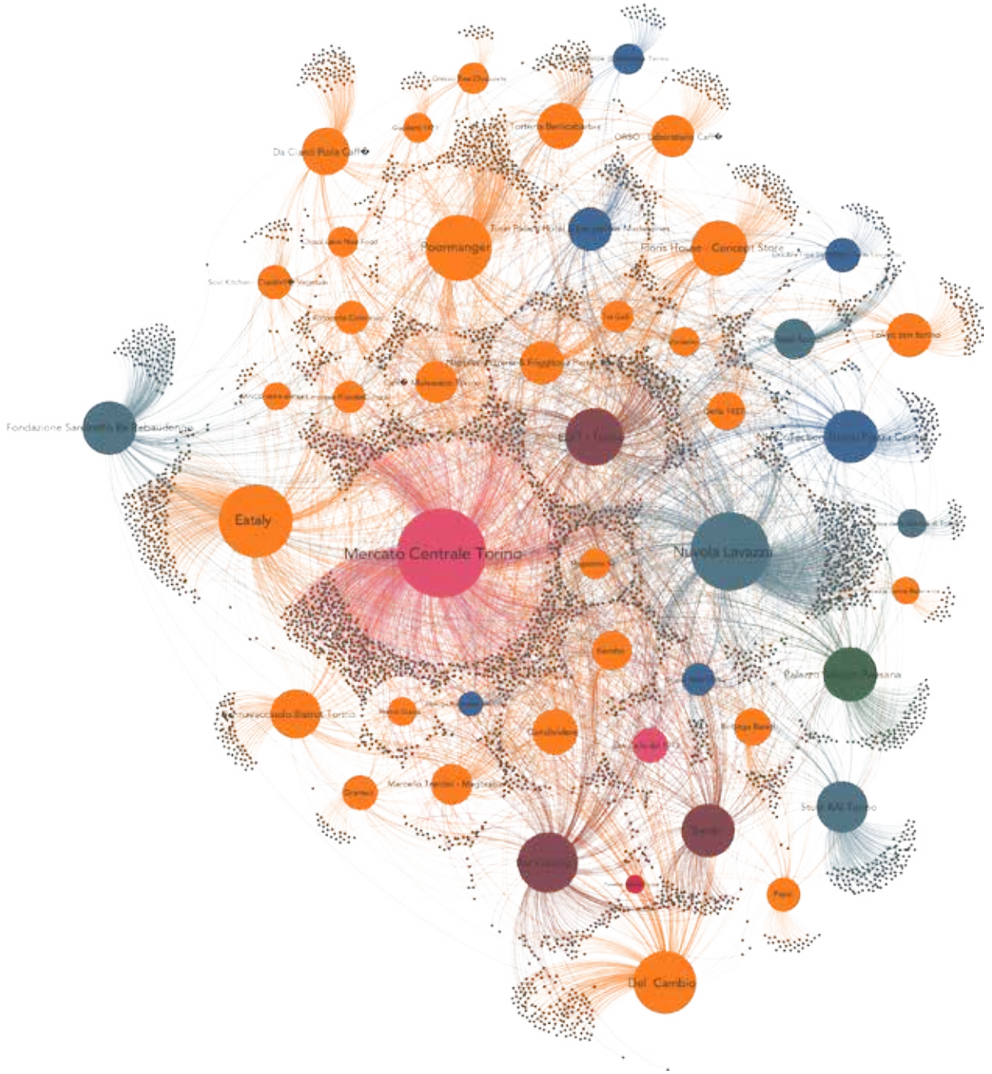
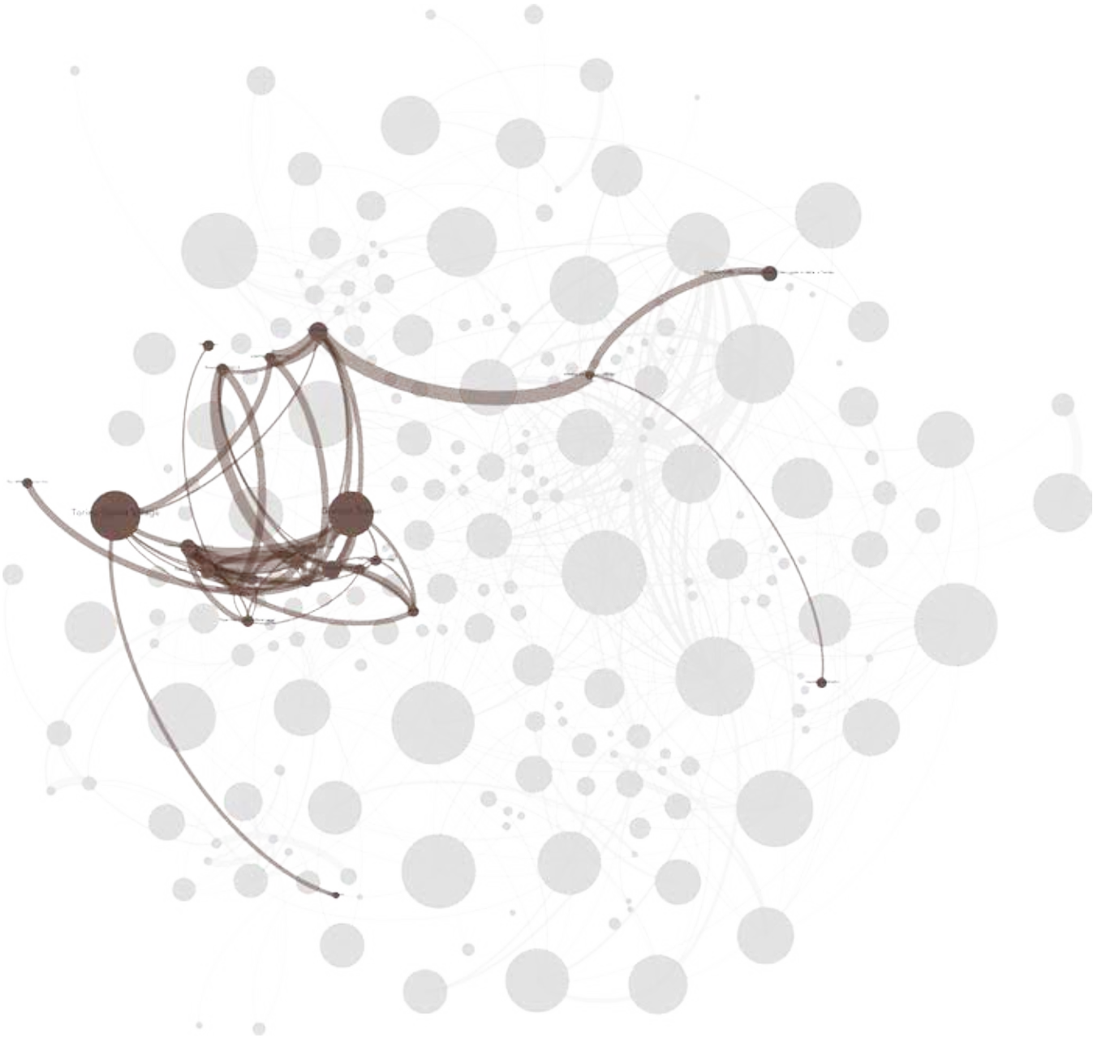


*Figure 5.2.2_SL1_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)



SL_2

Figure 5.2.2_SL2_1

SL_2 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

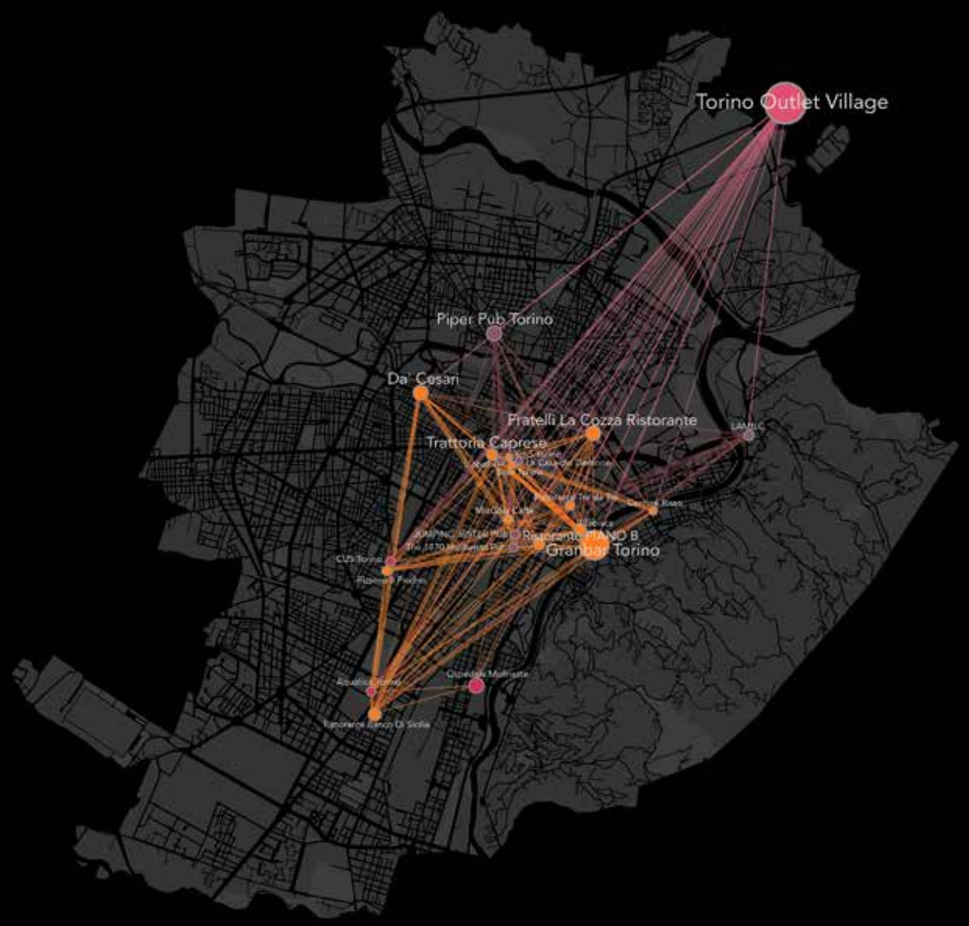


Figure 5.2.2_SL2_2

SL_2 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

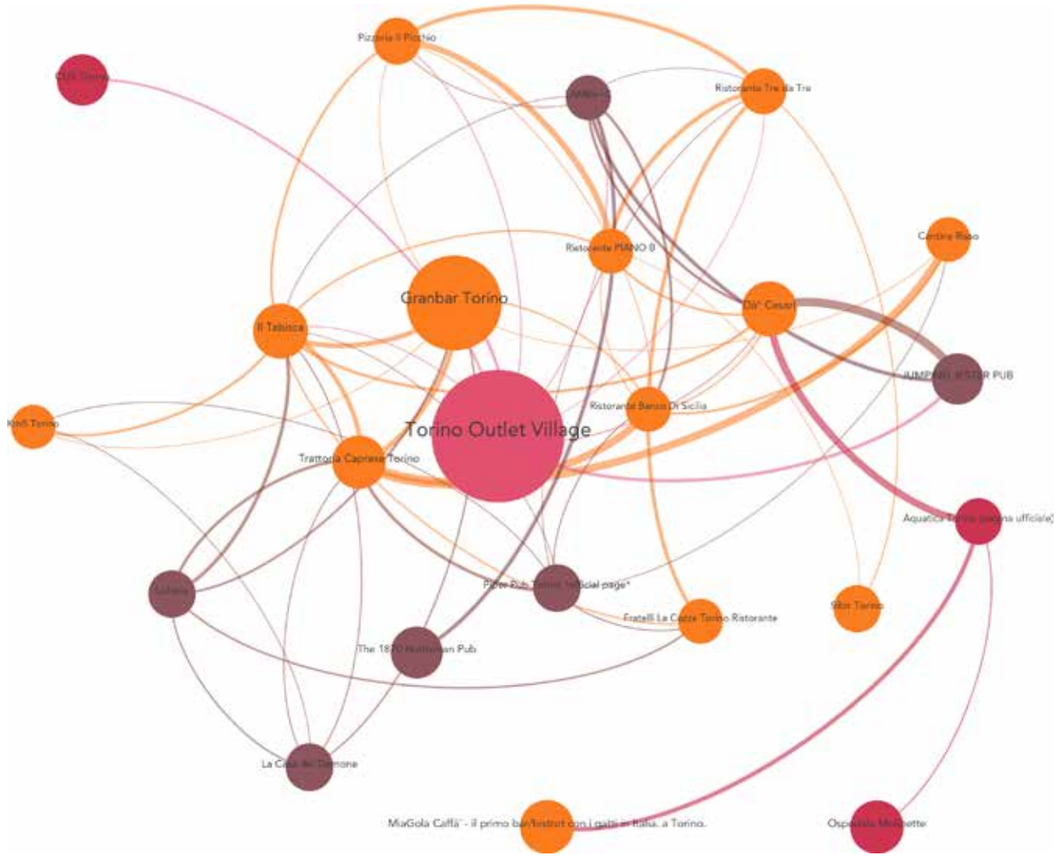
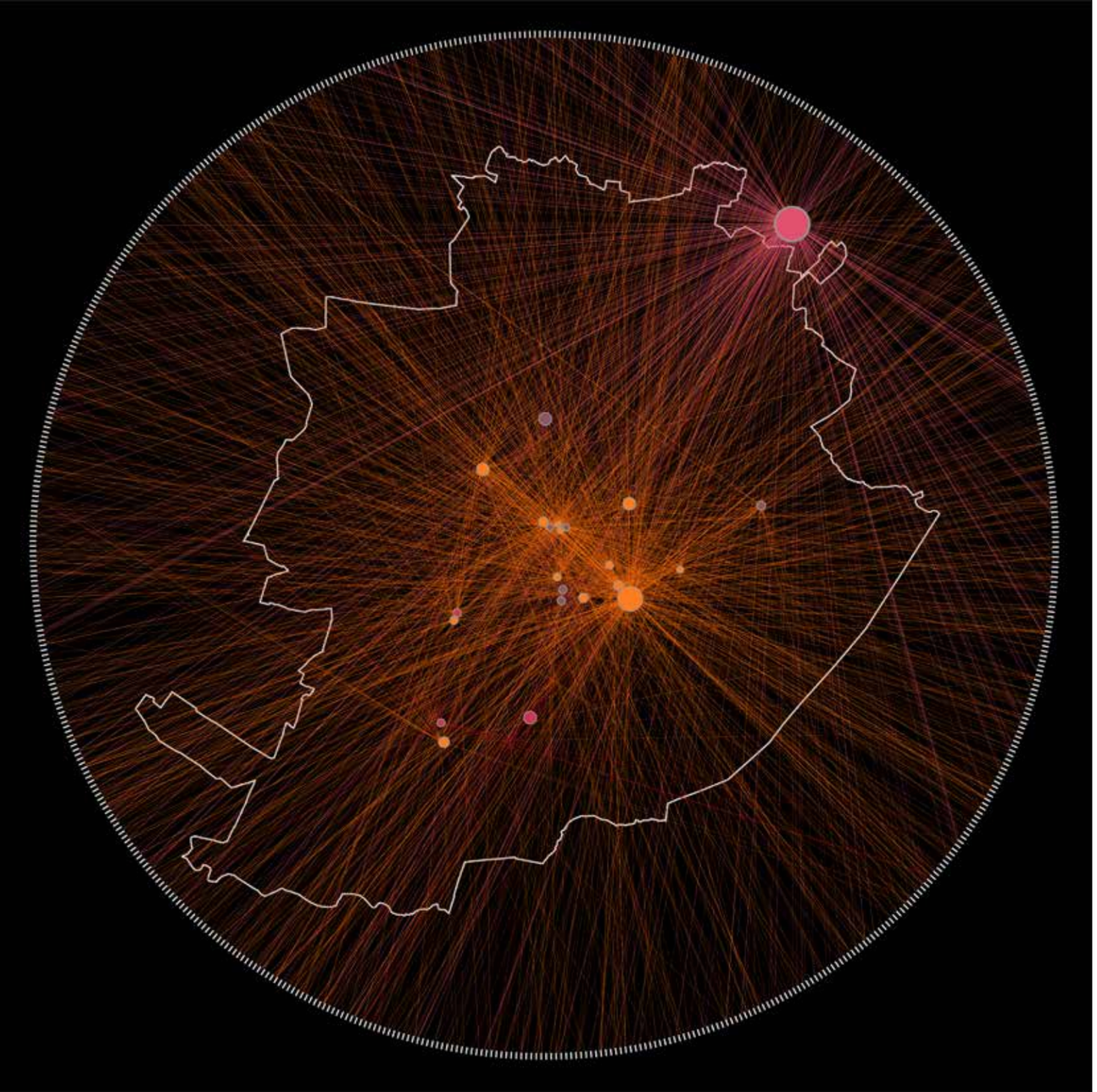


Figure 5.2.2_SL2_3

SL_2 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



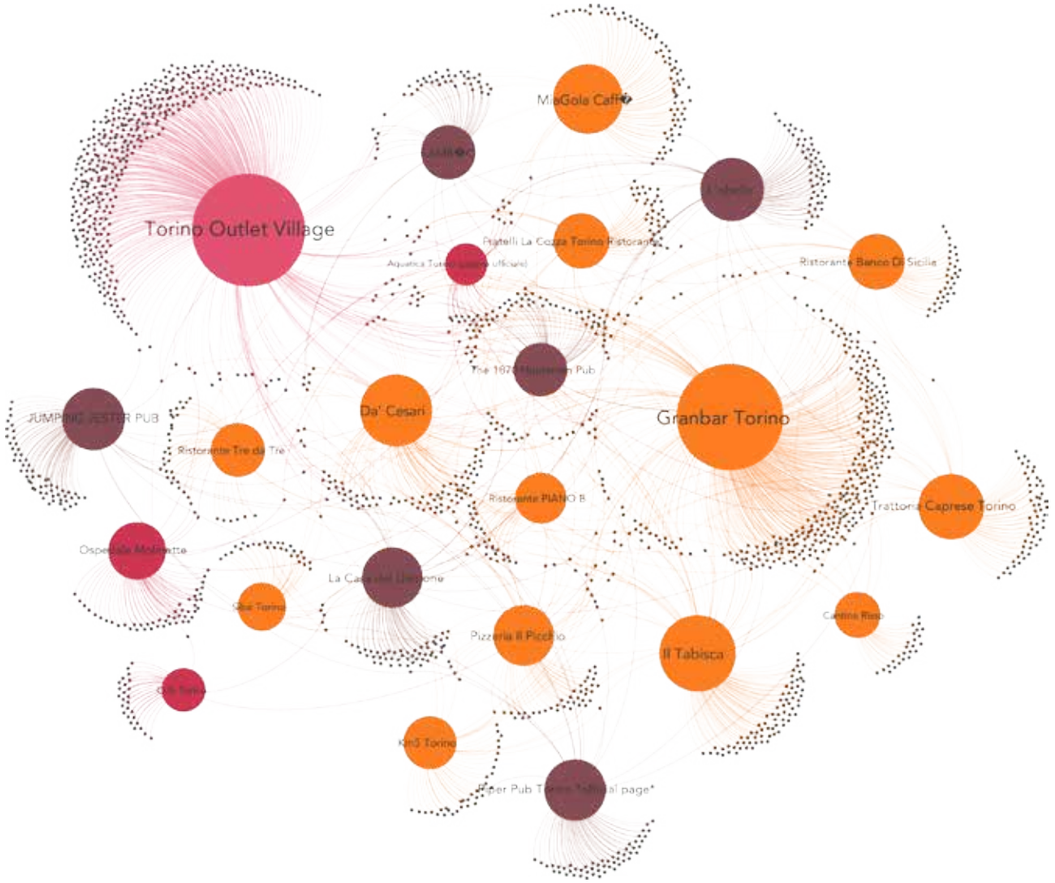


*Figure 5.2.2_SL2_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

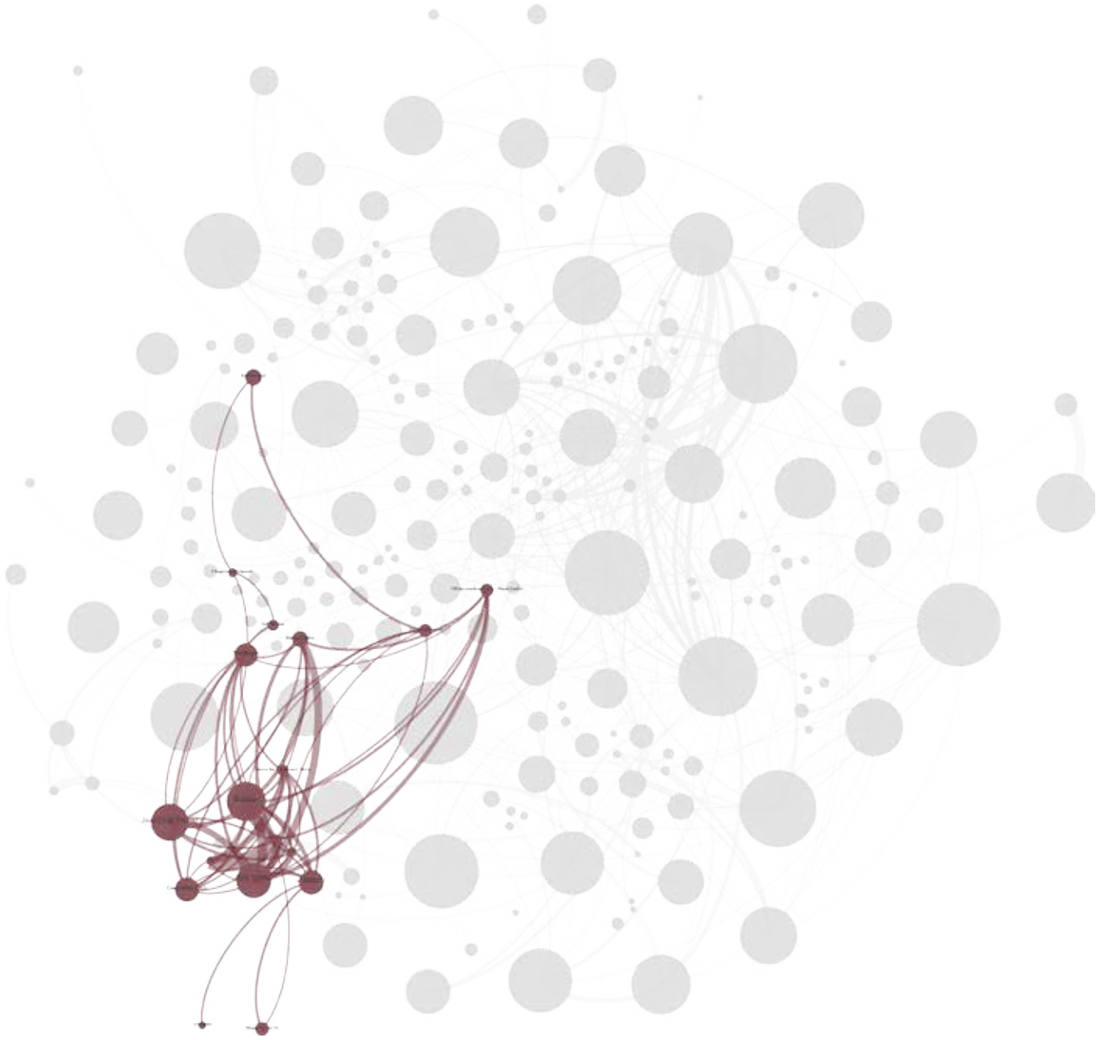


Figure 5.2.2_SL3_1

SL_3 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

SL_3

Figure 5.2.2_SL3_2

SL_3 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

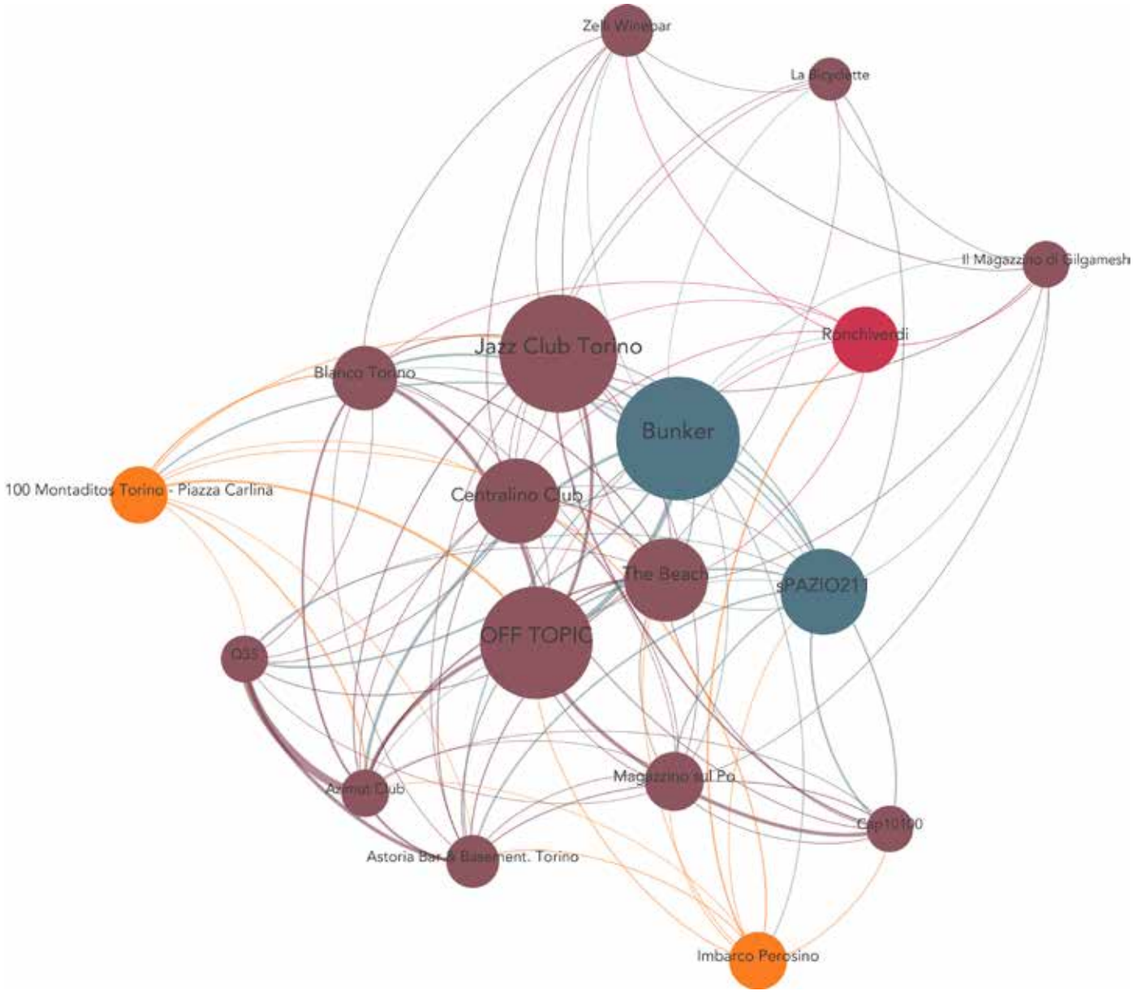
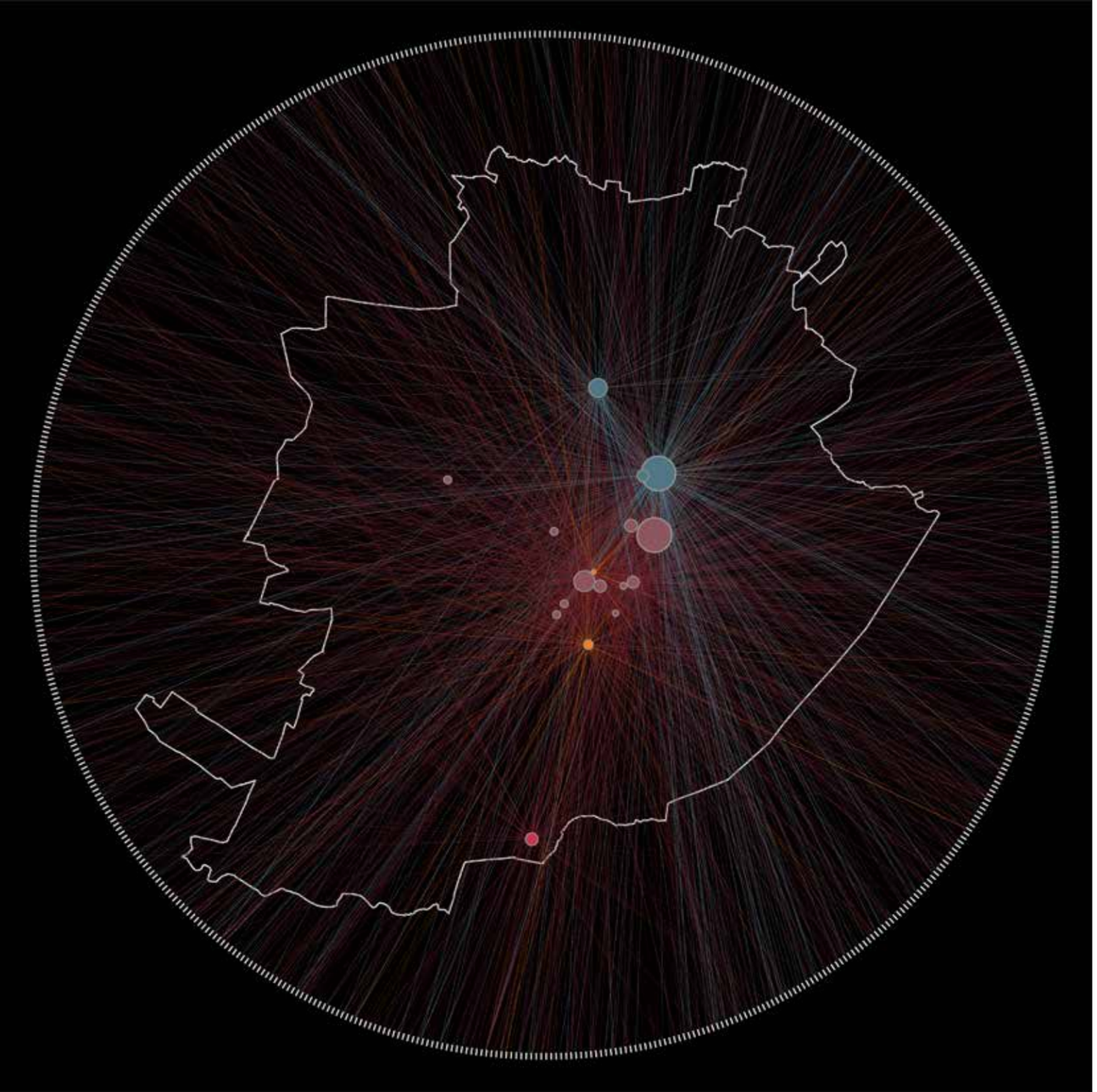


Figure 5.2.2_SL3_3

SL_3 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



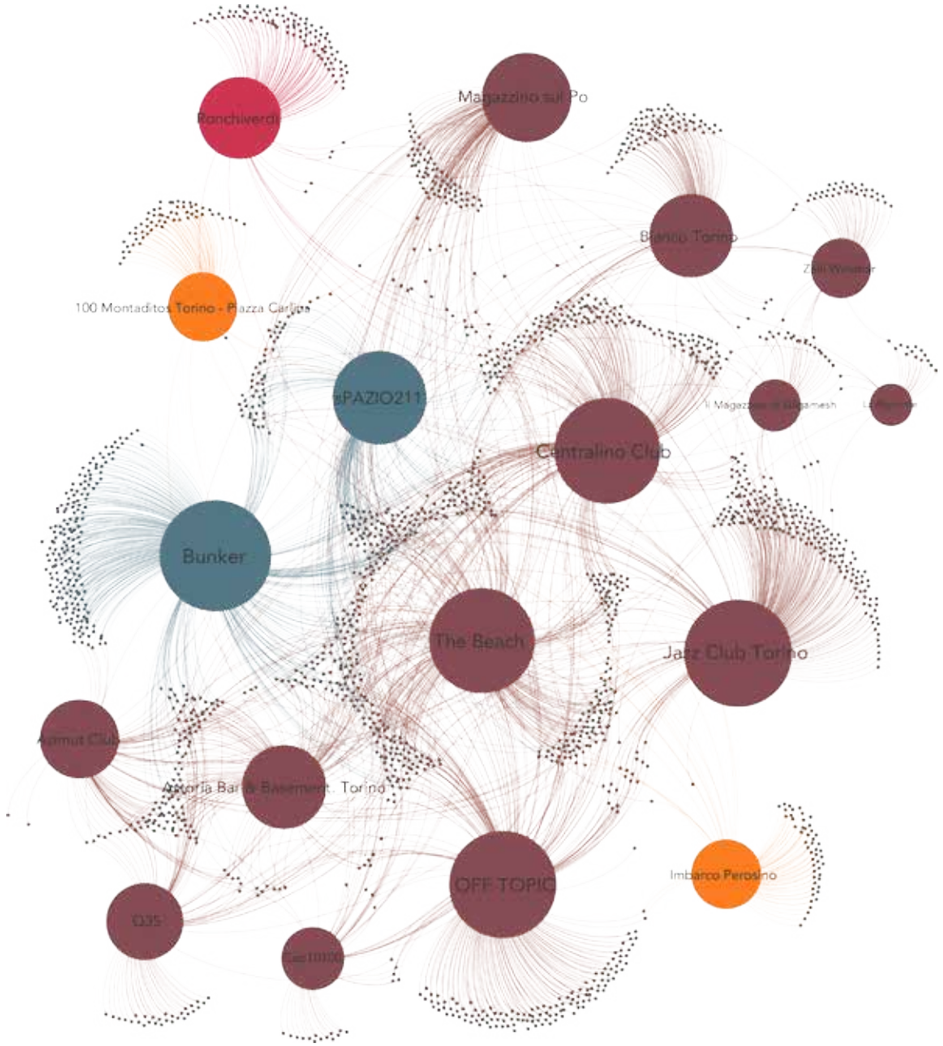


*Figure 5.2.2_SL3_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

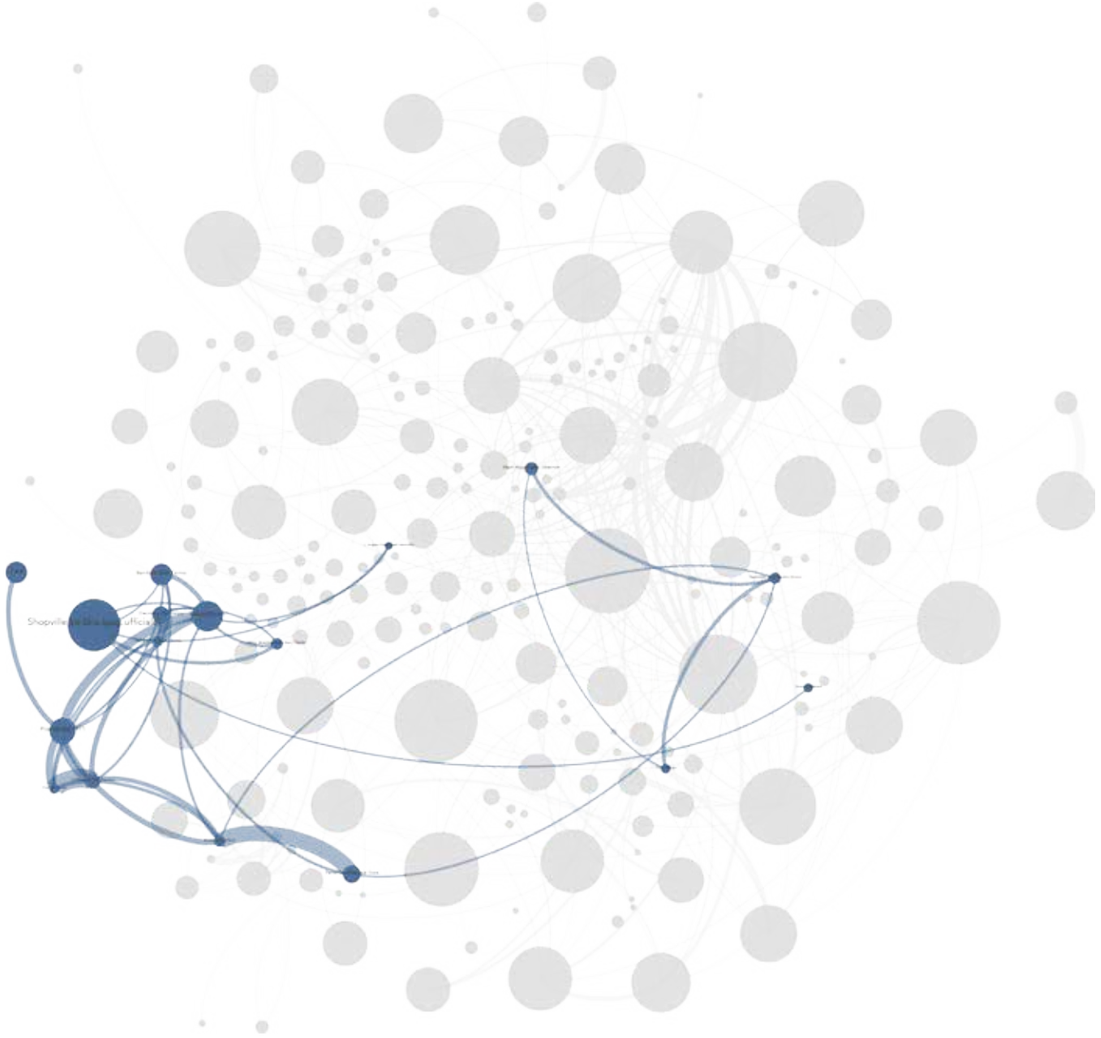


Figure 5.2.2_SL4_1

SL_4 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

SL_4

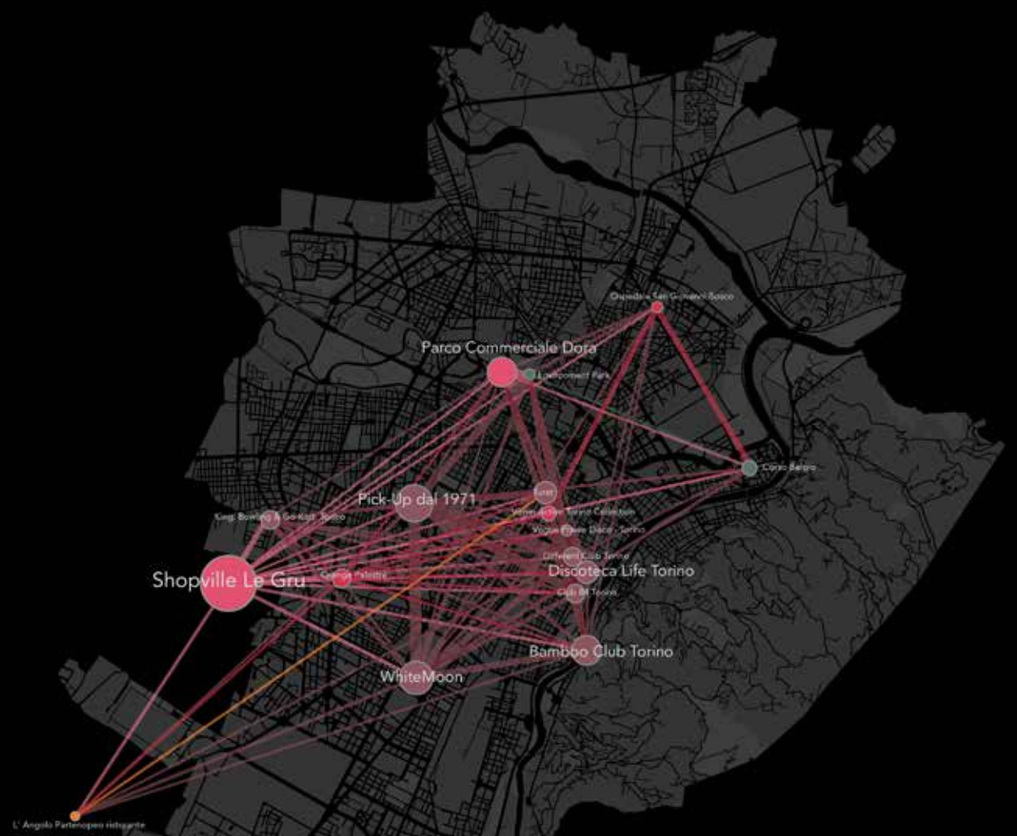


Figure 5.2.2_SL4_2

SL_4 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

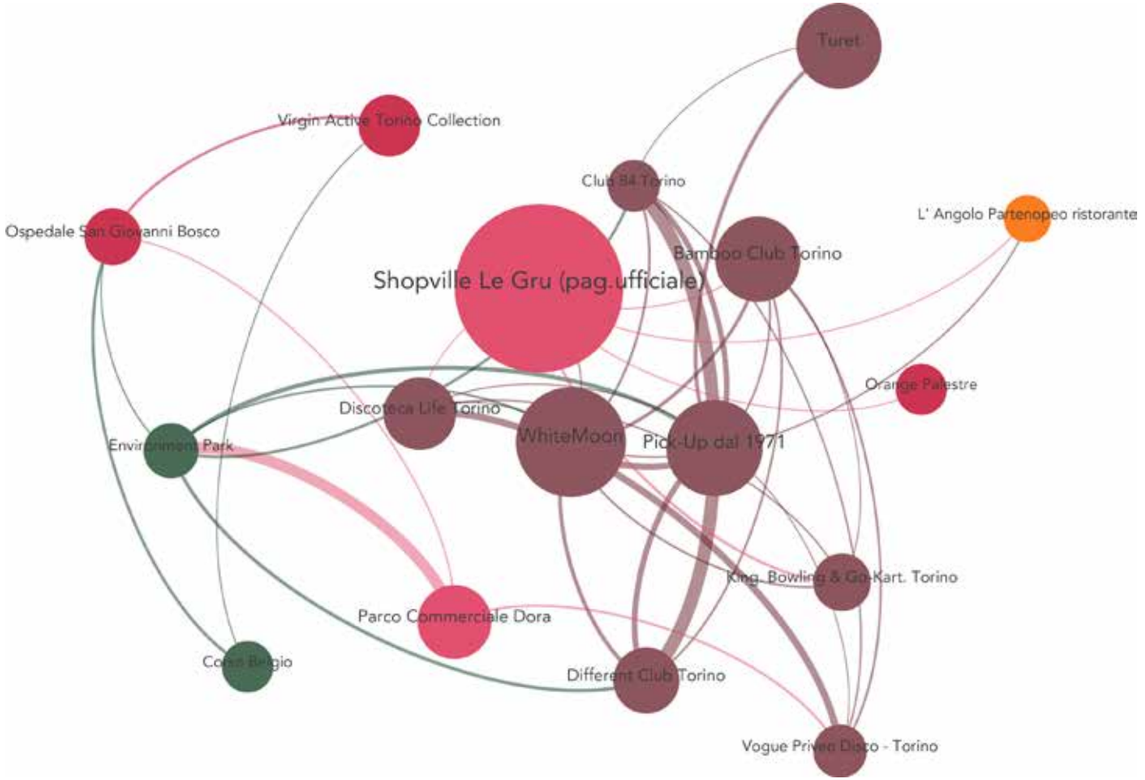
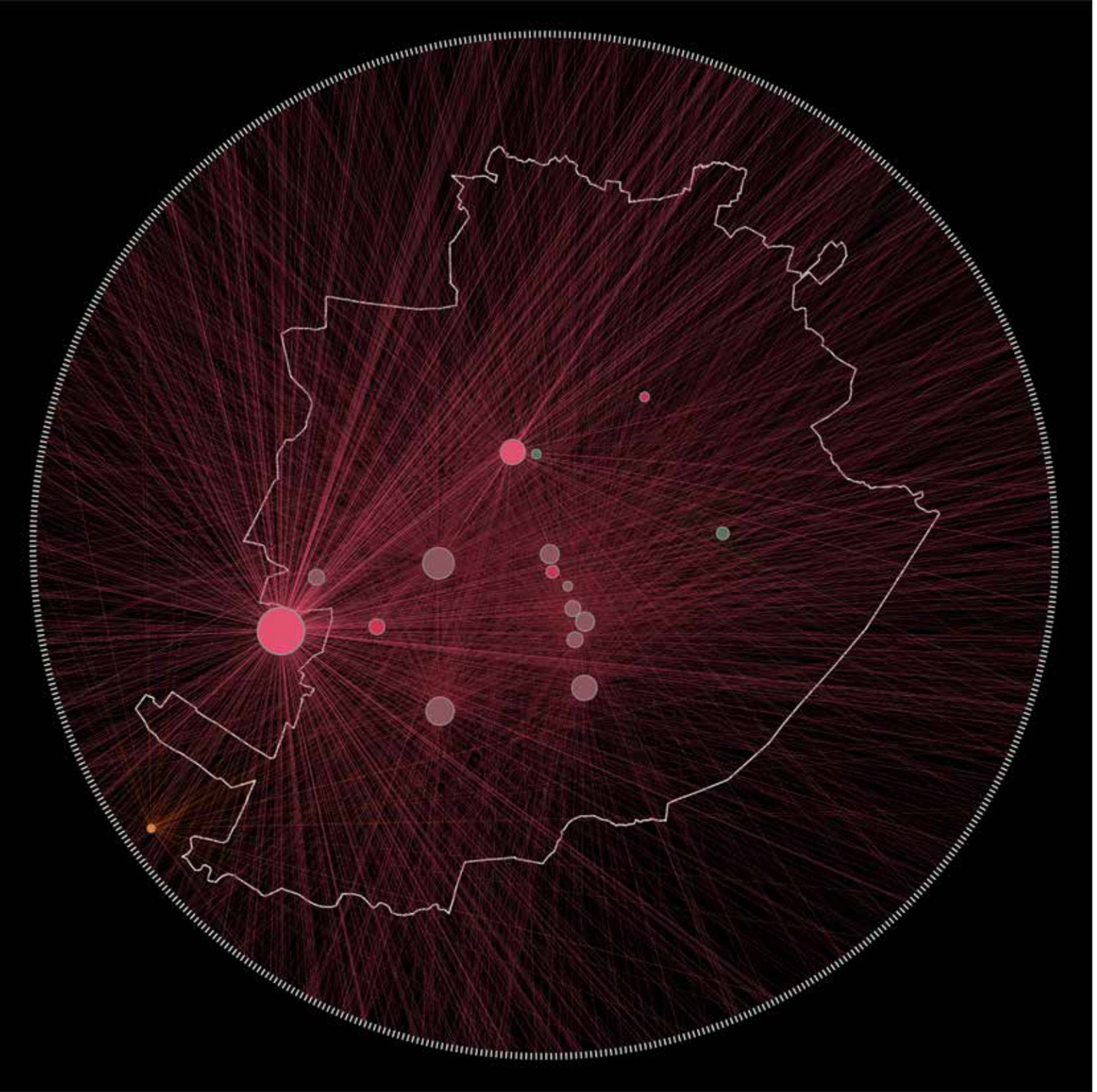


Figure 5.2.2_SL4_3

SL_4 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



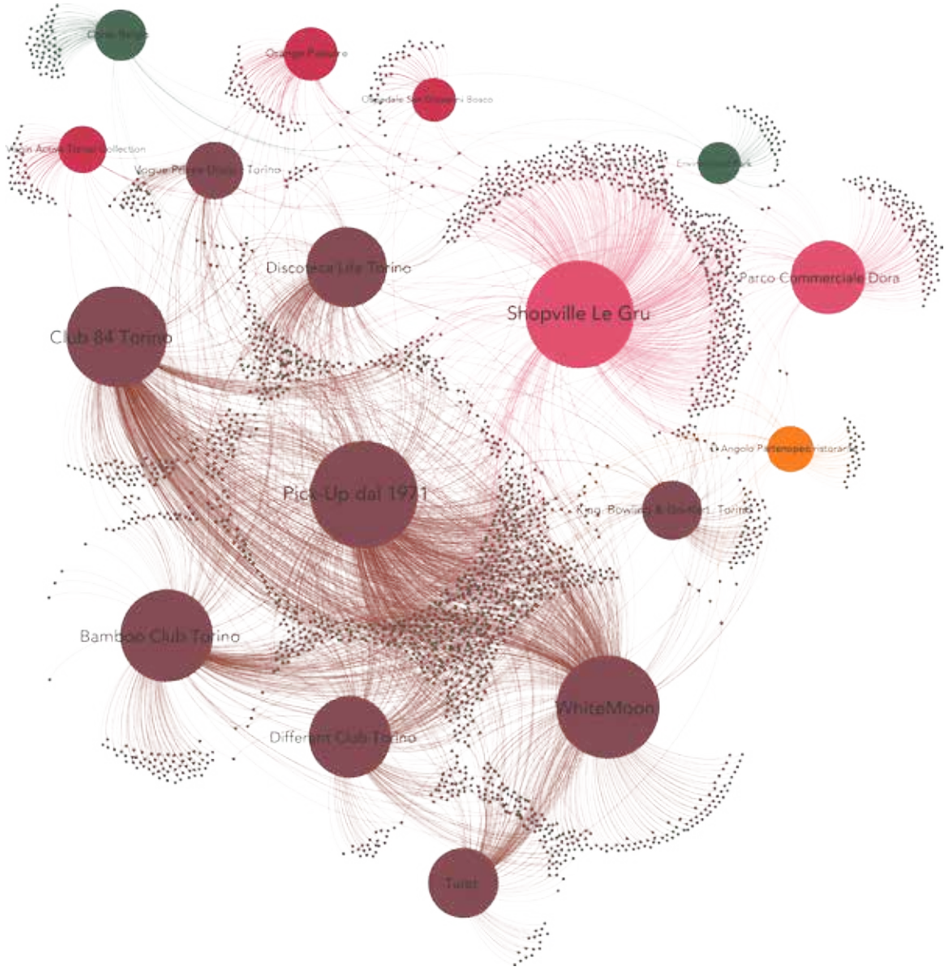


*Figure 5.2.2_SL4_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

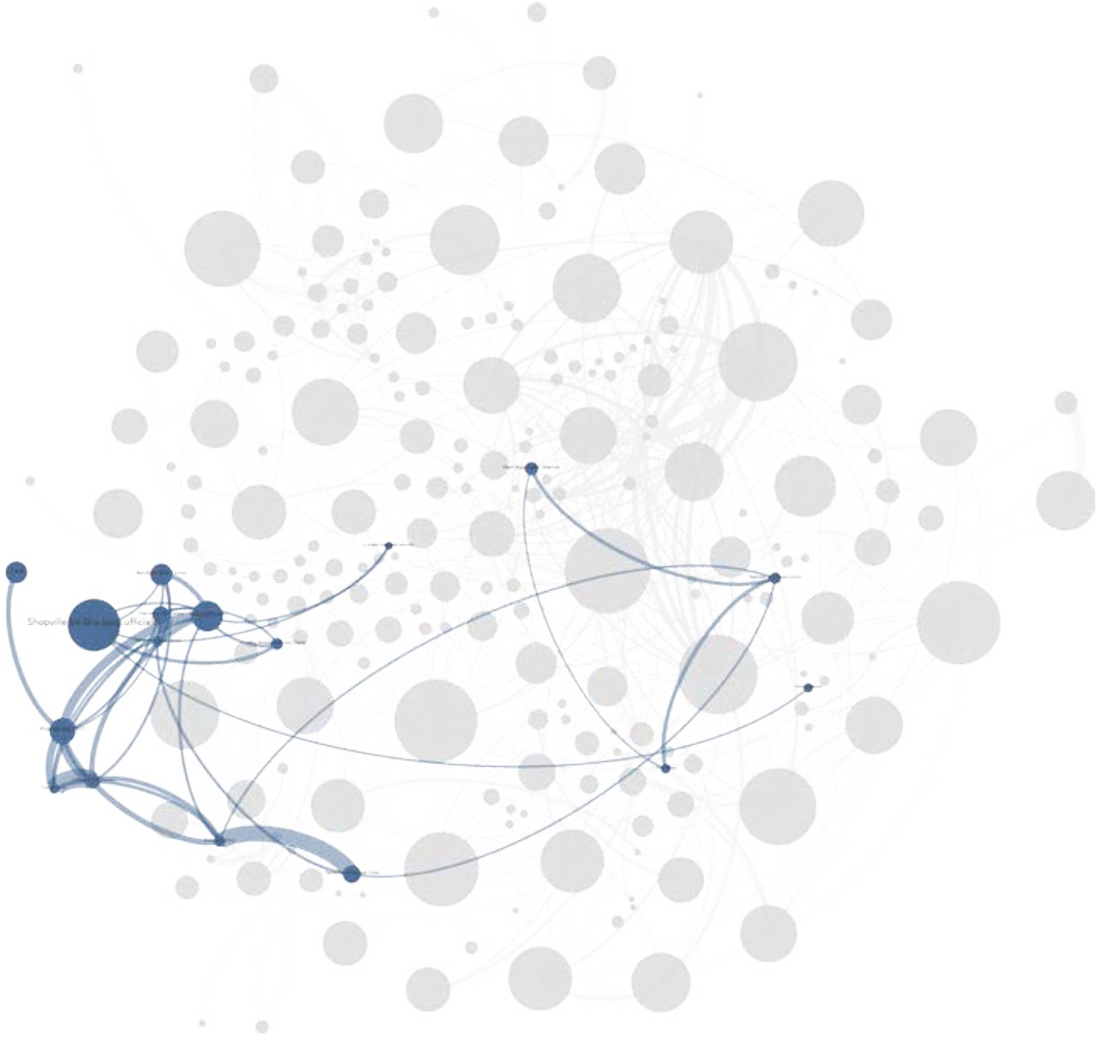


Figure 5.2.2_SL5_1

SL_5 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

SL_5

Figure 5.2.2_SL5_2

SL_5 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

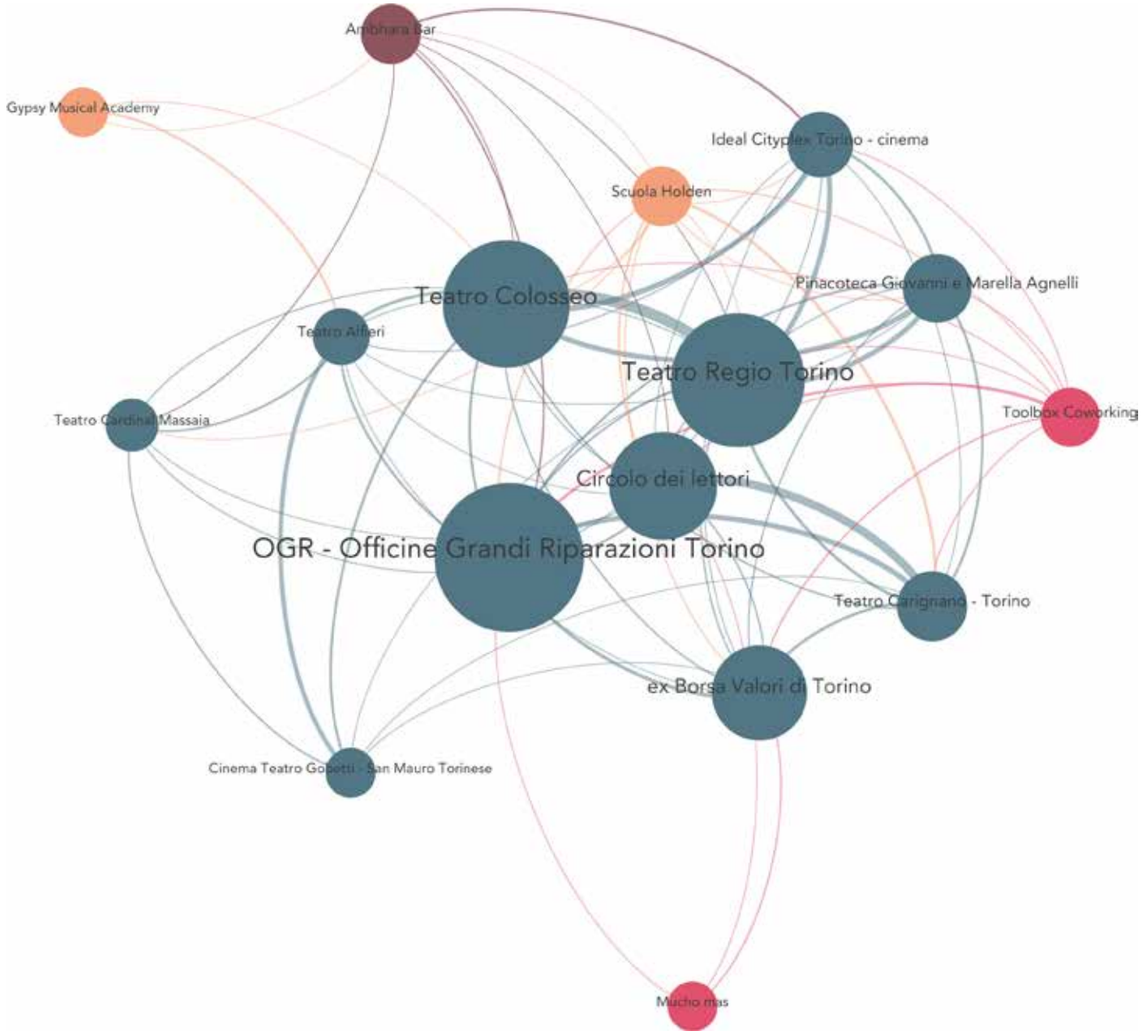
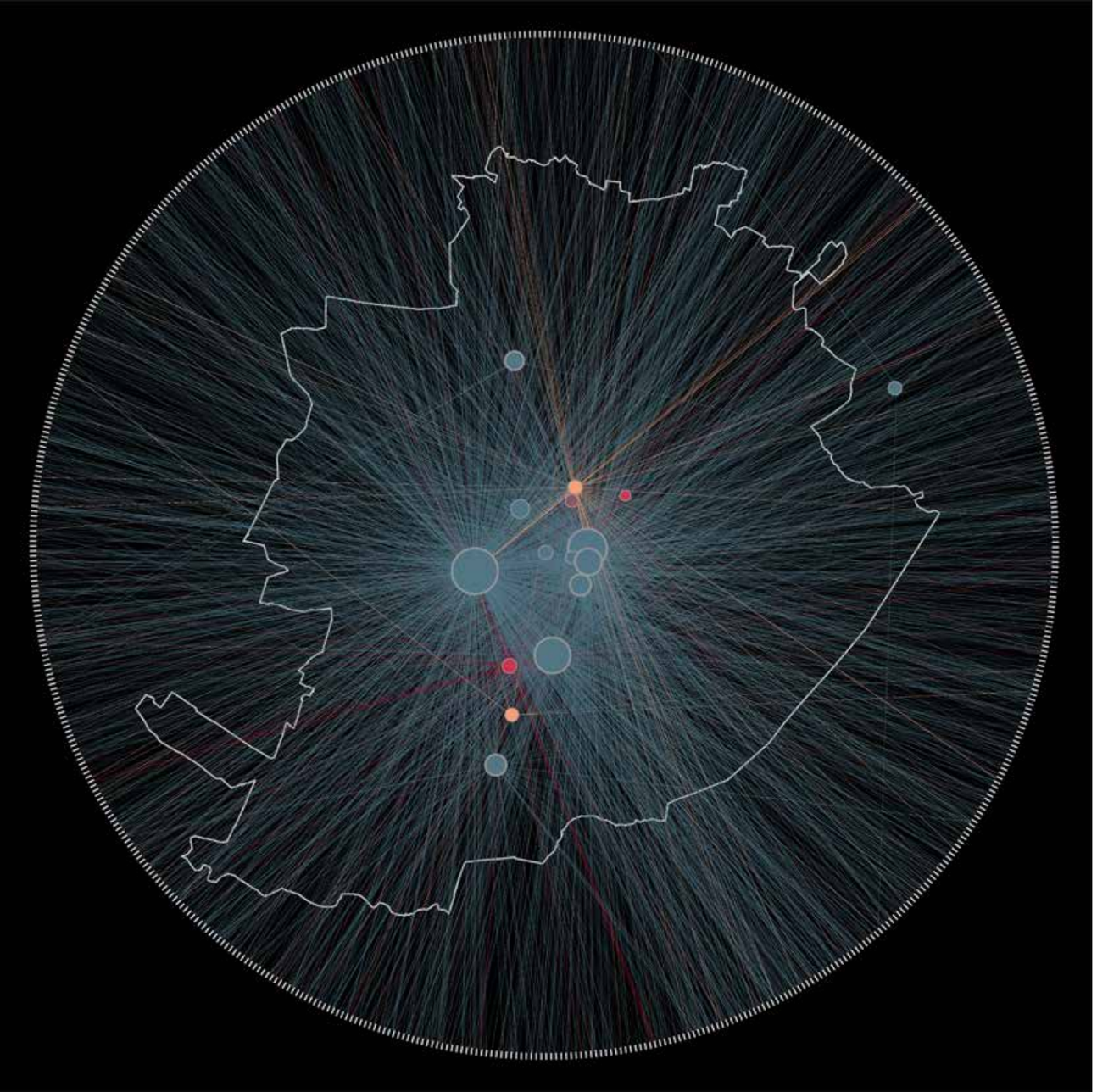


Figure 5.2.2_SL5_3

SL_5 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)



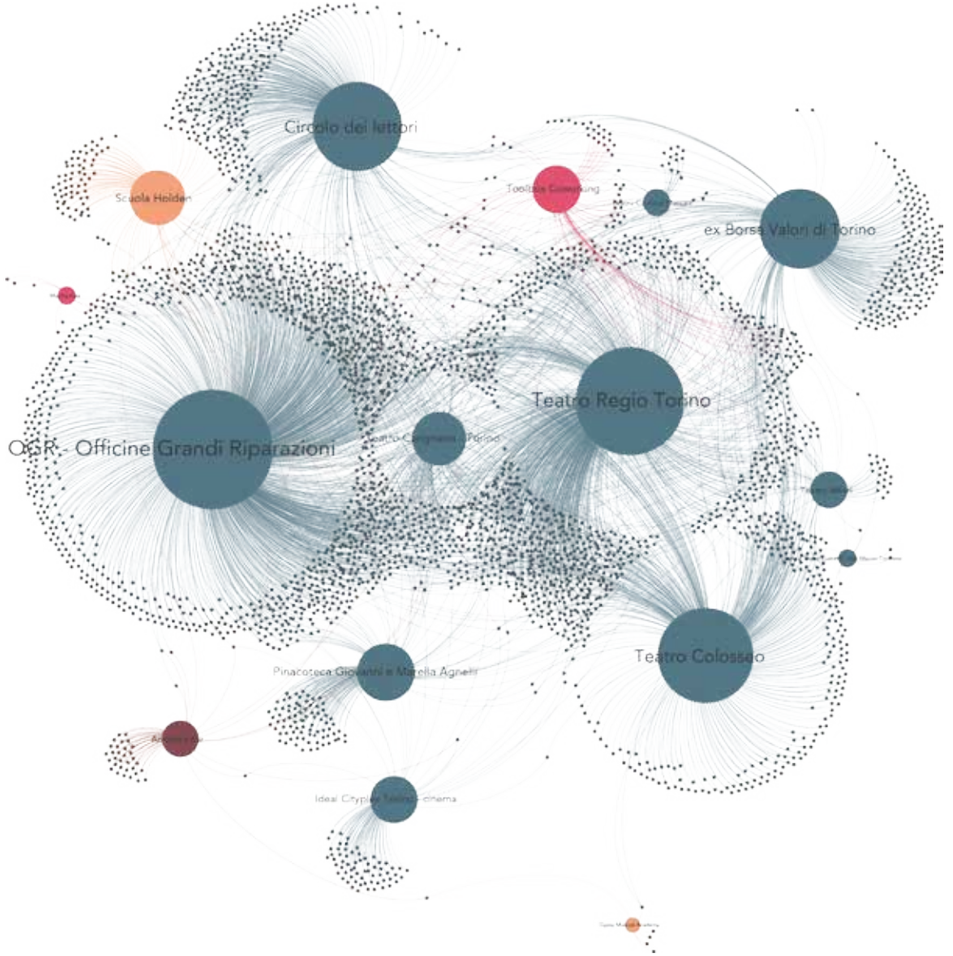


*Figure 5.2.2_SL5_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

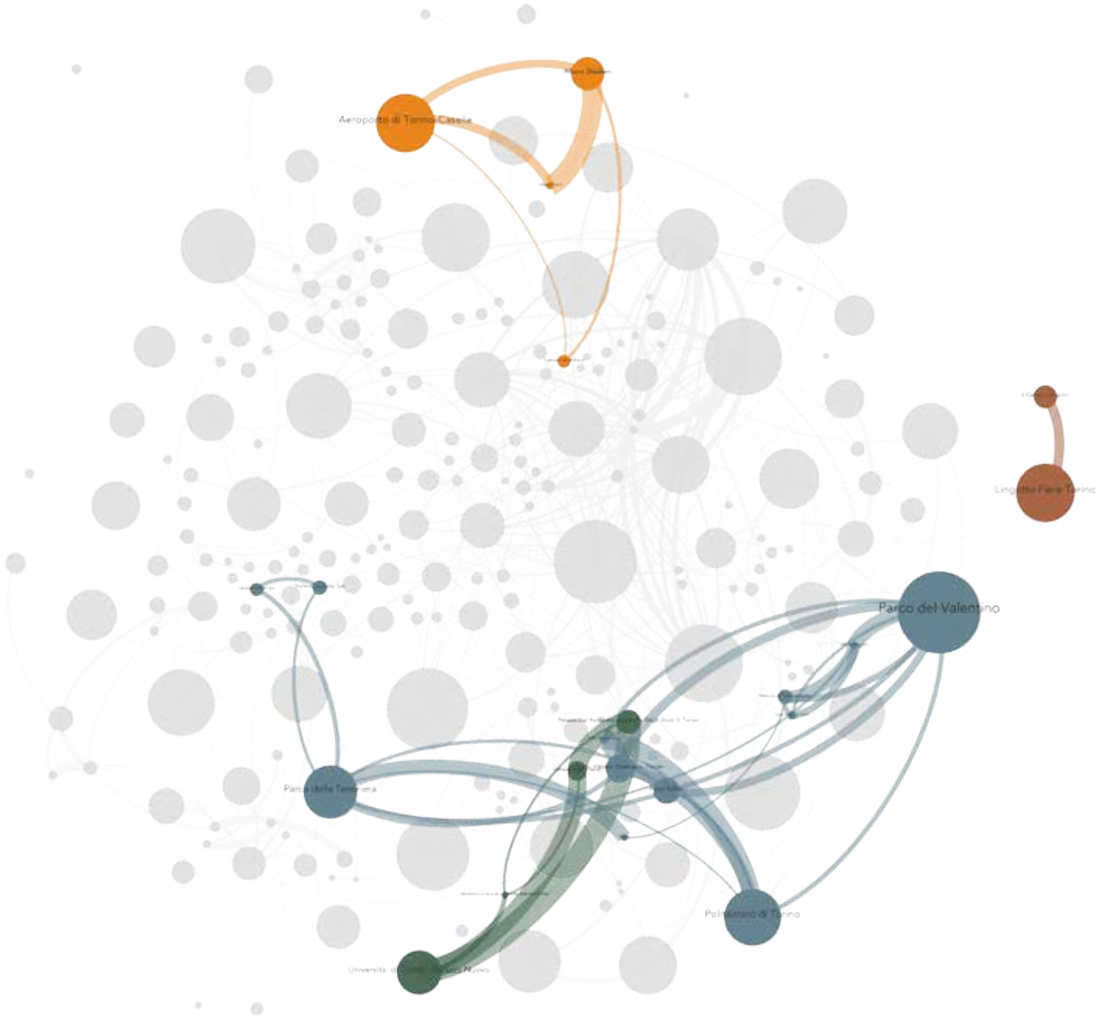


Figure 5.2.3_1

The four clusters of "Isolated Networks" highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

#5.2.1

ISOLATED NETWORKS

Although the remaining four clusters grouped in the "Isolated networks" macro-group seems to be nothing but what was left over during the running of the algorithm of community detection, they still allow to some further readings.

IN_1 for example groups together some major parks of the city, highly connected between each other, with the Polytechnic of Turin bearing some weaker connections and bringing with it in the cluster to night bar, as can be seen in [Figure 5.2.3_1.5].

IN_2 grouped surprisingly together all the venues in the filtered list part of the system of the University of Turin.

IN_3 and IN_4 are clearly what is left in the process of clustering. IN_3 groups the Allianz Stadium with a neighboring medical centre, the airport and a restaurant on the other side of the city. [Figure 5.2.3_3.5] shows how weak the ties are between the last two and the first. The neighboring Lingotto Fiere and 8 Gallery seems to create a synergy in the process of identification between the two venues, but it is clear from [Figure 5.2.3_4.5] how they have a small community strongly geotagging/identifying in the two venues.

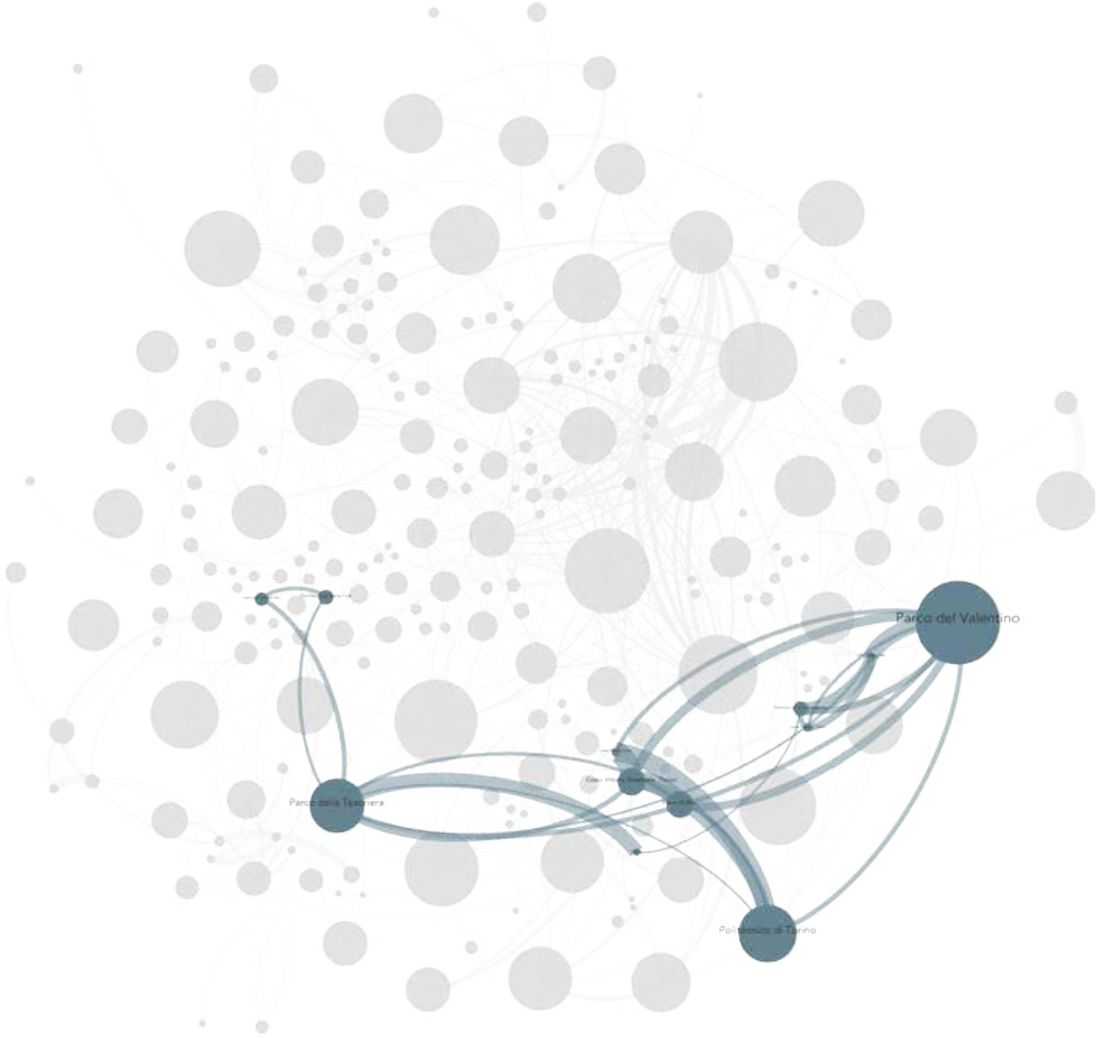


Figure 5.2.3_IN1_1

IN_1 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

IN_1

Figure 5.2.3_IN1_2

IN_1 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

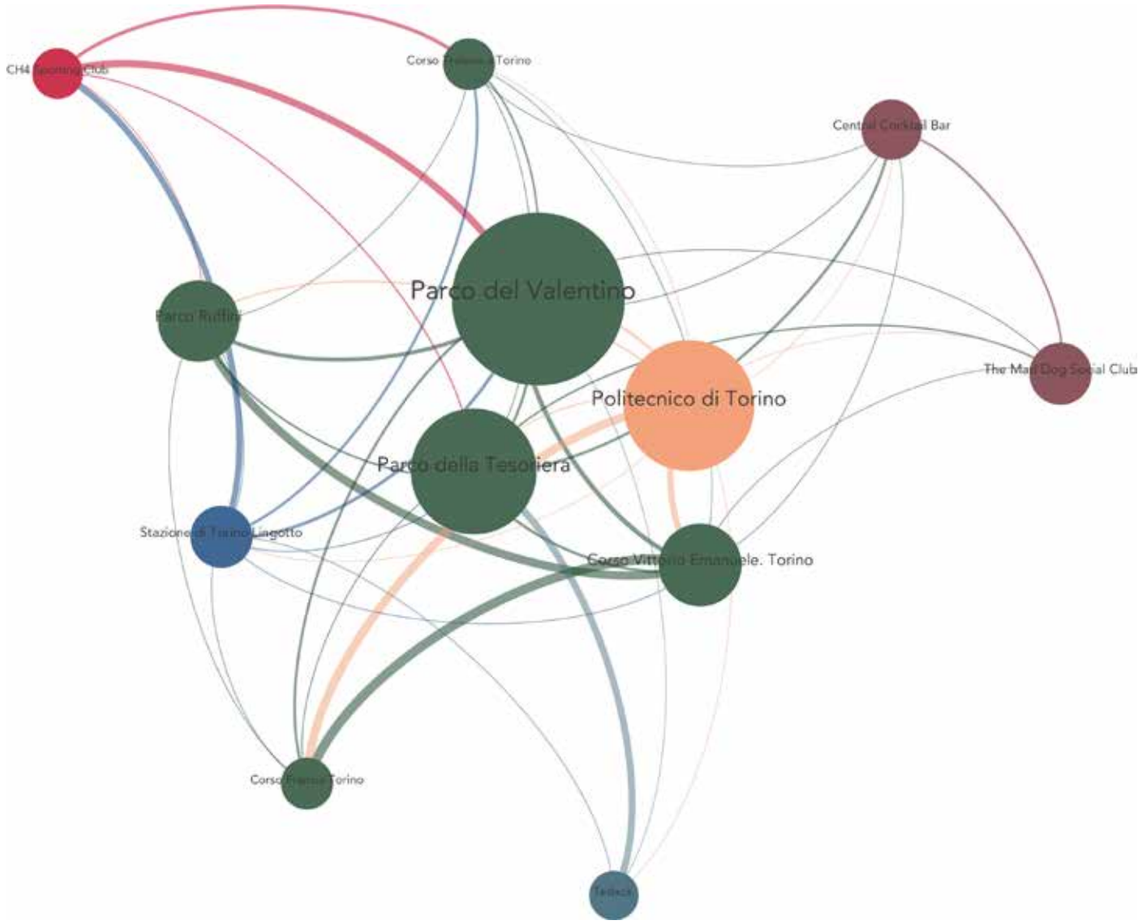
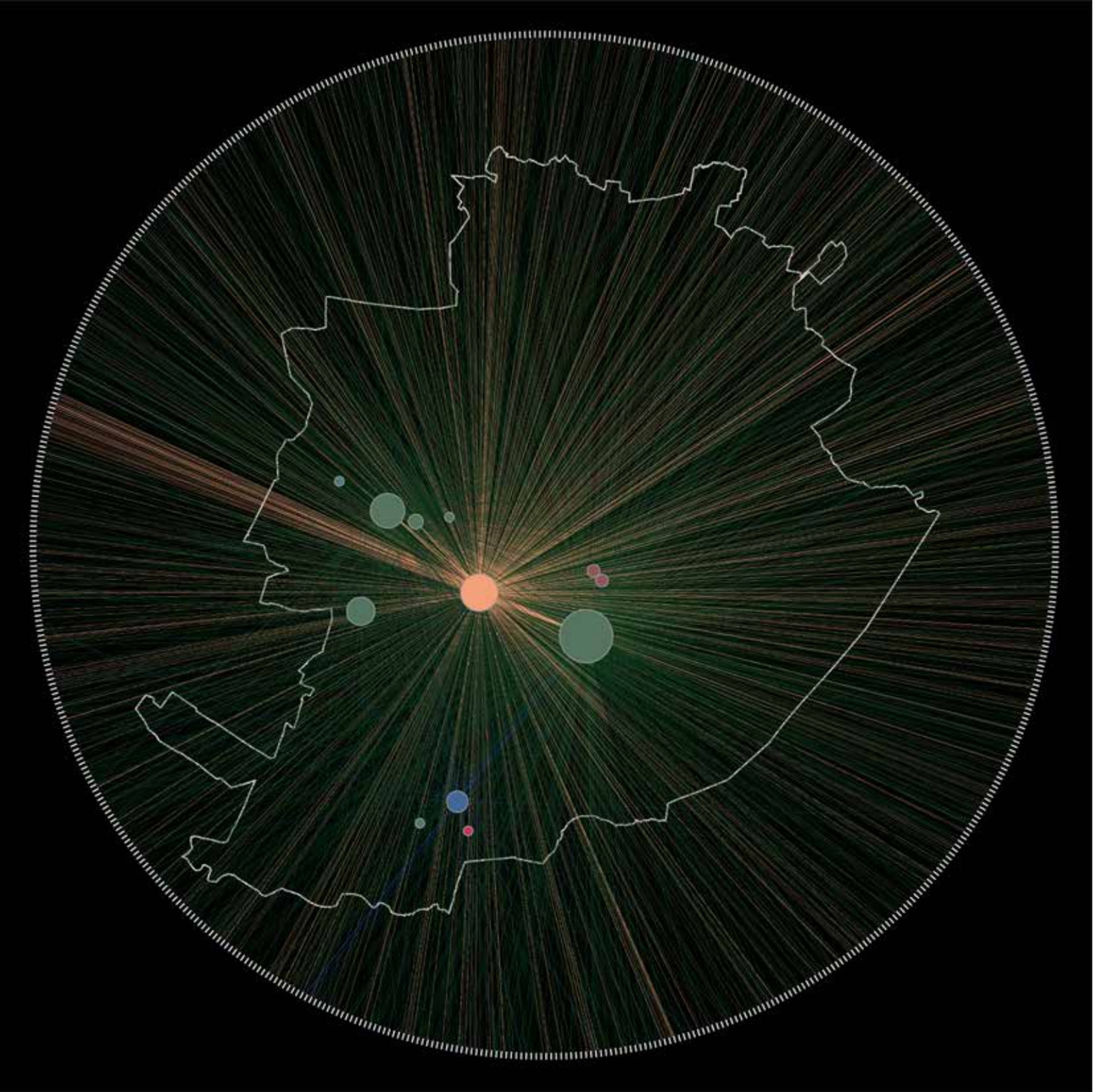


Figure 5.2.3_IN1_3

IN_1 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



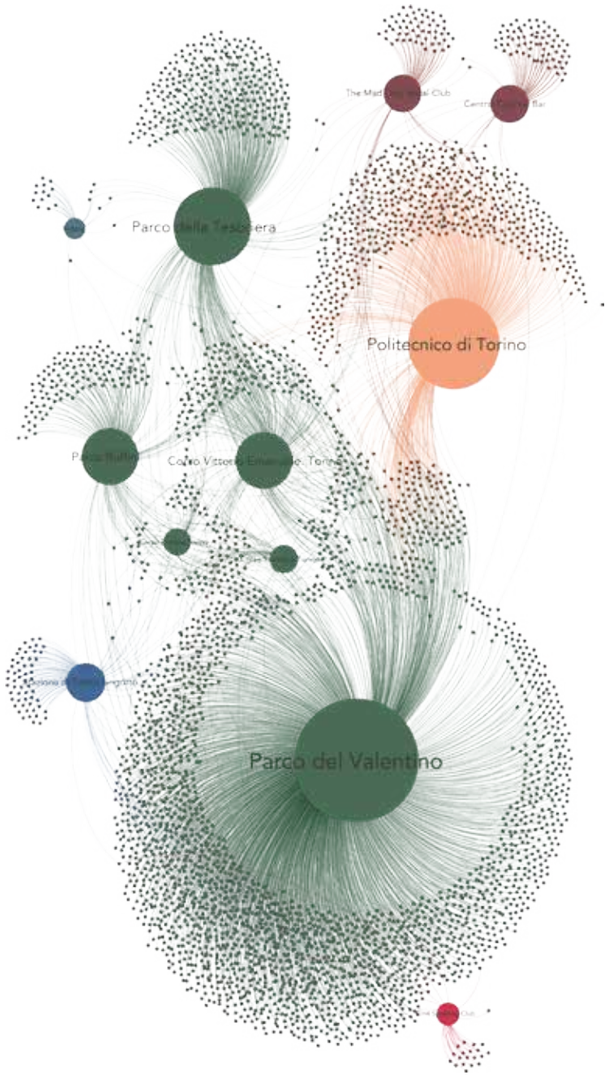


*Figure 5.2.3_IN1_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

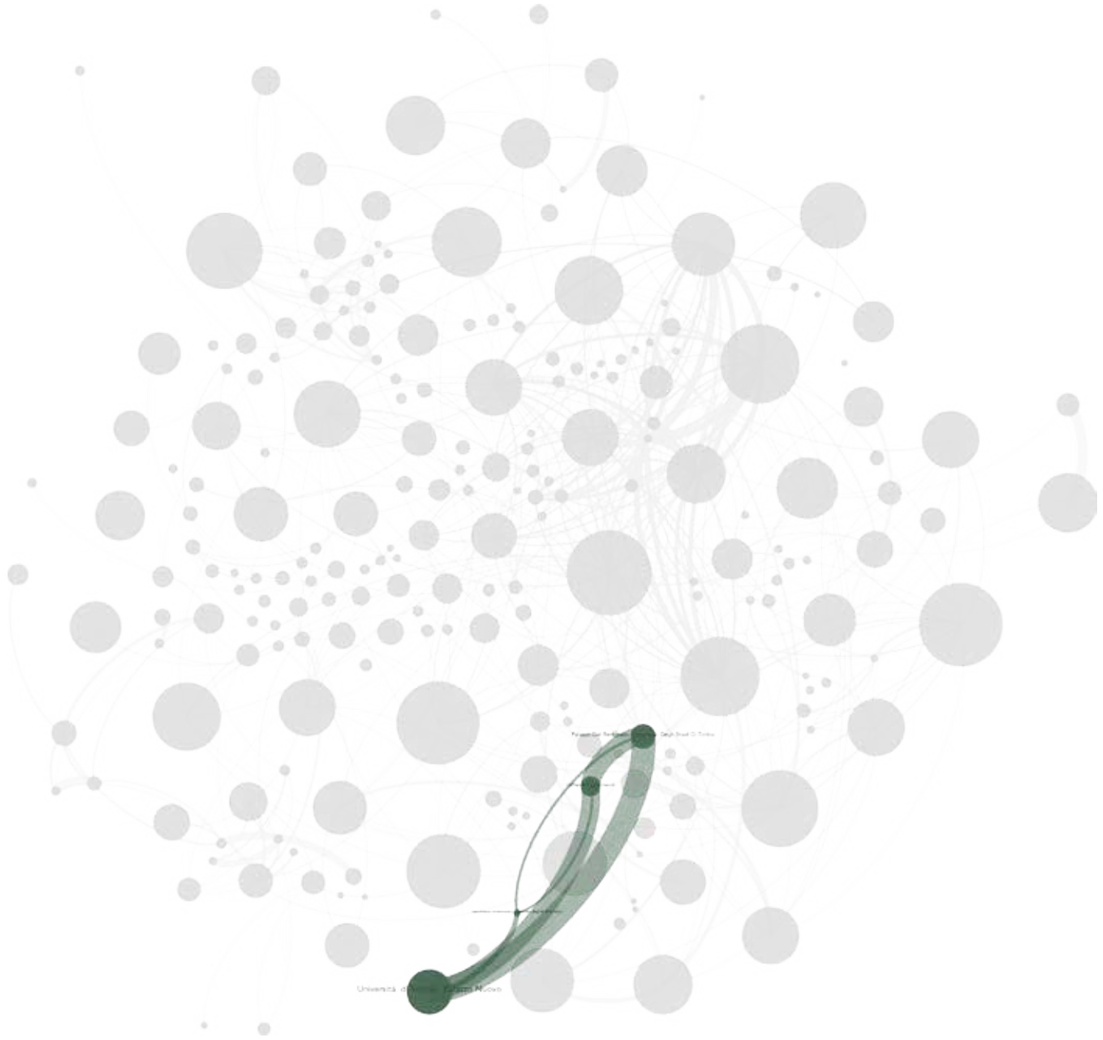


Figure 5.2.3_IN2_1

IN_2 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

IN_2

Figure 5.2.3_IN2_2
IN_2 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

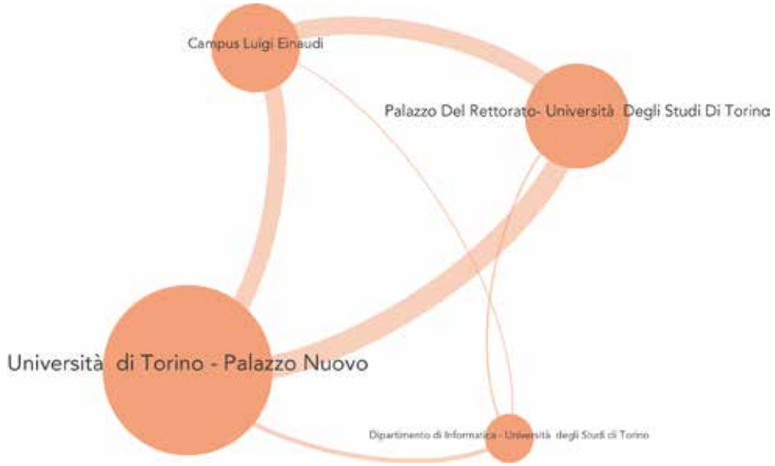
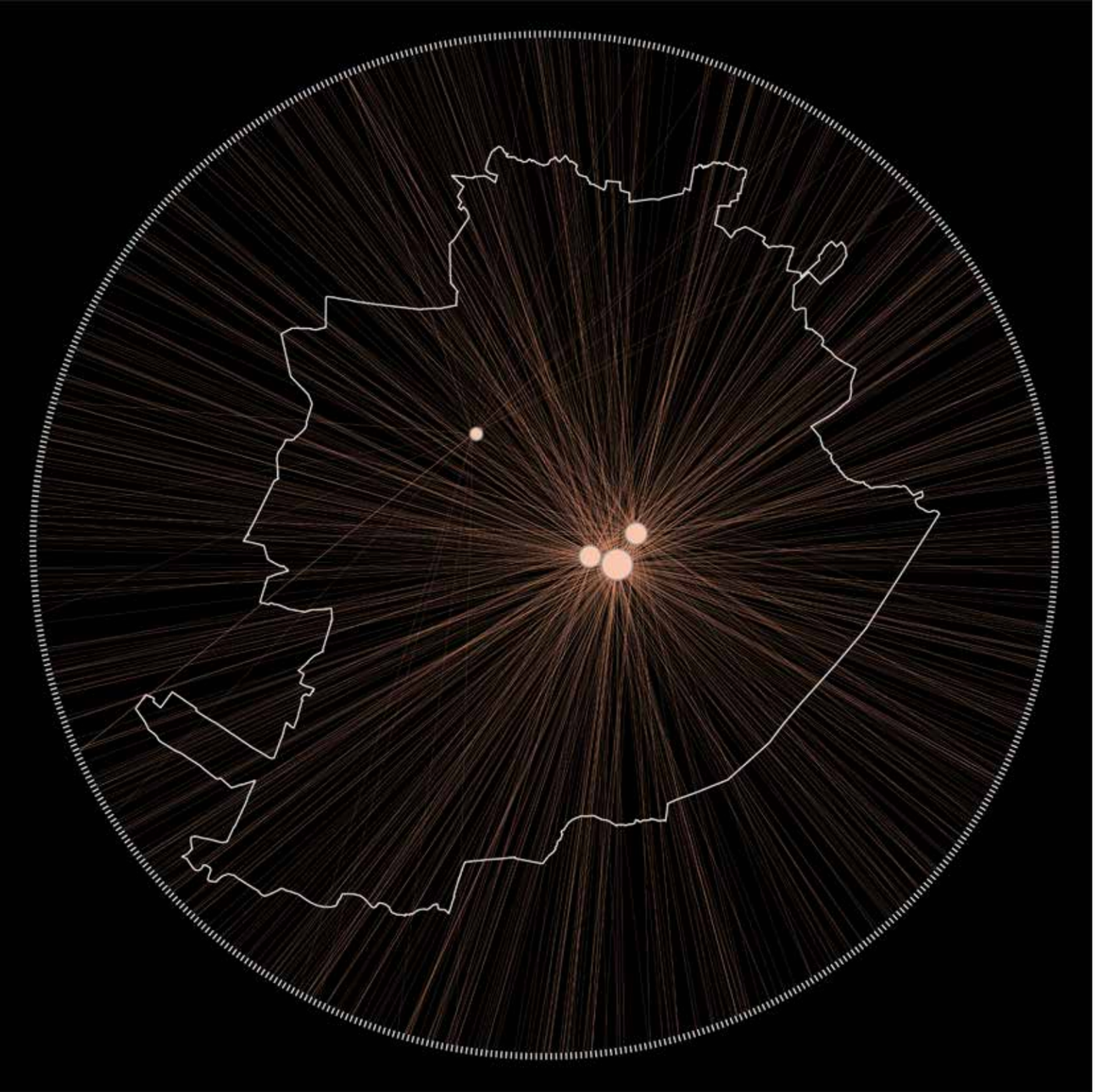


Figure 5.2.3_IN2_3
IN_2 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



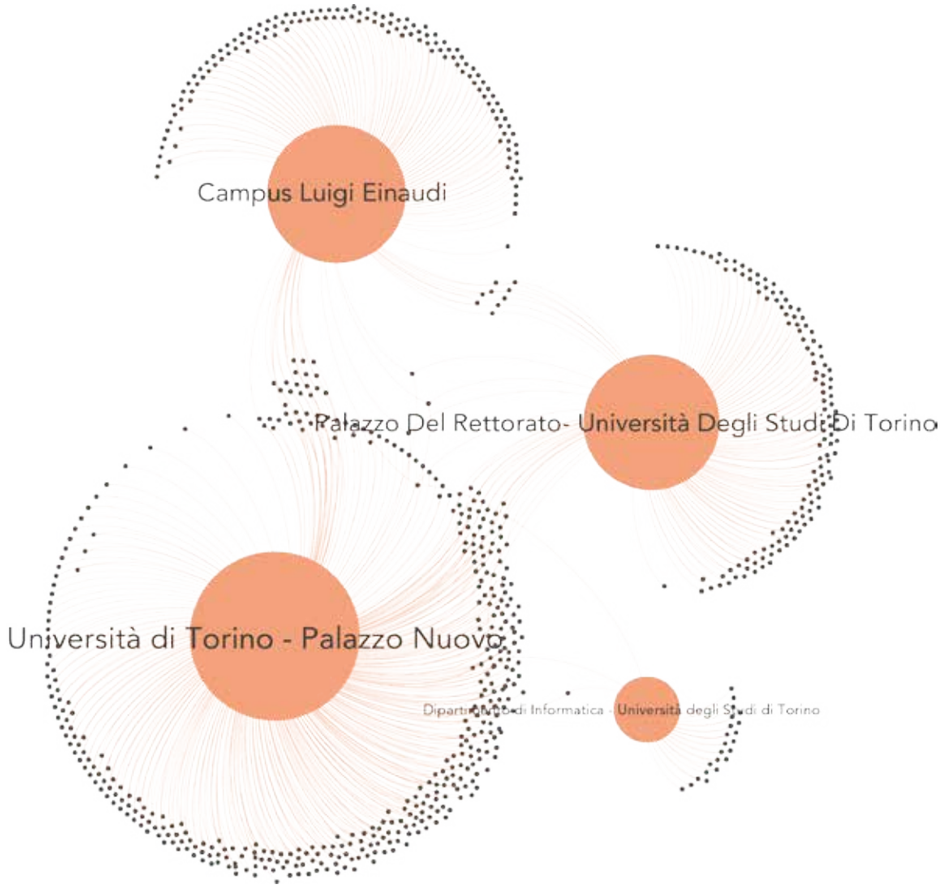


*Figure 5.2.3_IN2_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

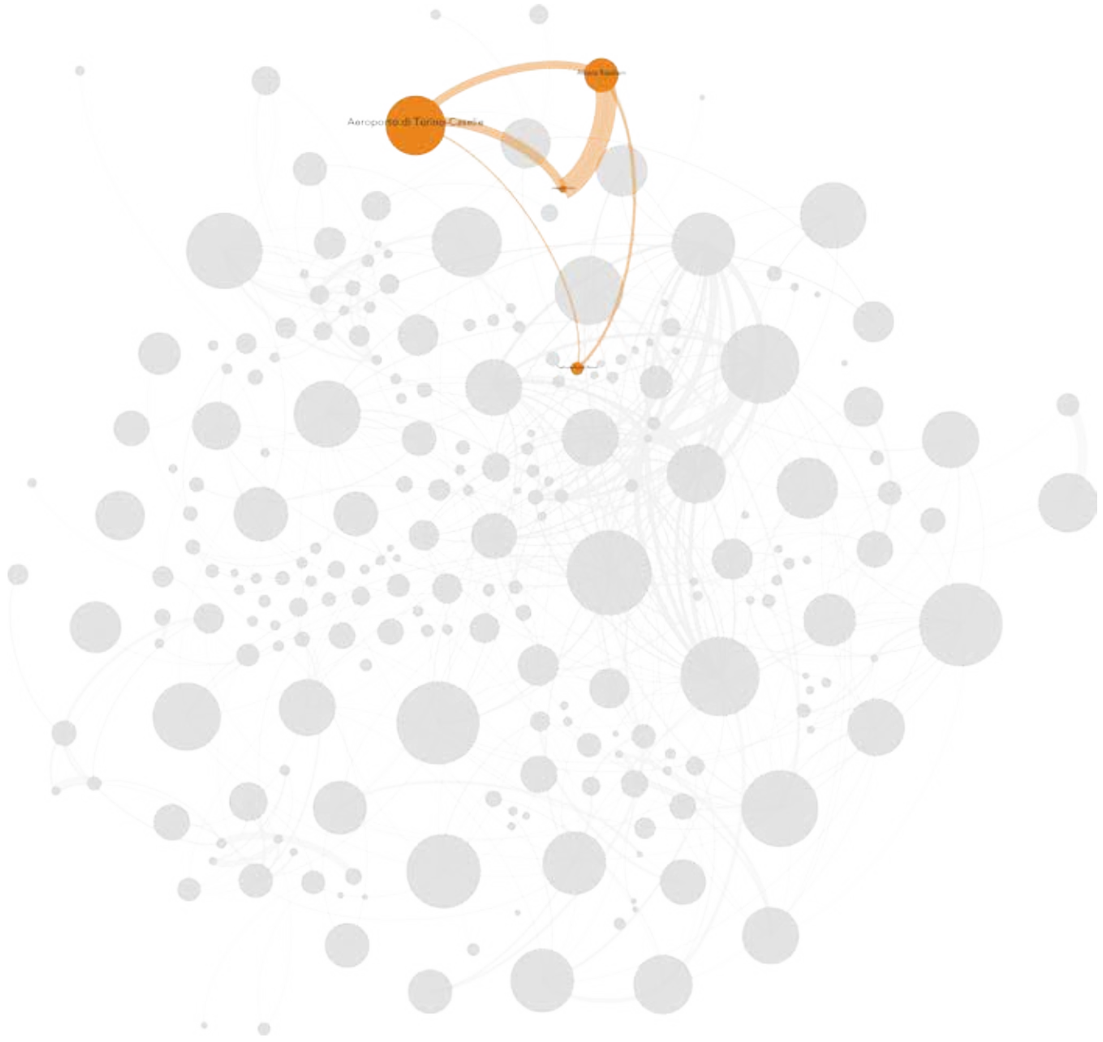


Figure 5.2.3_IN3_1

IN_3 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

IN_3

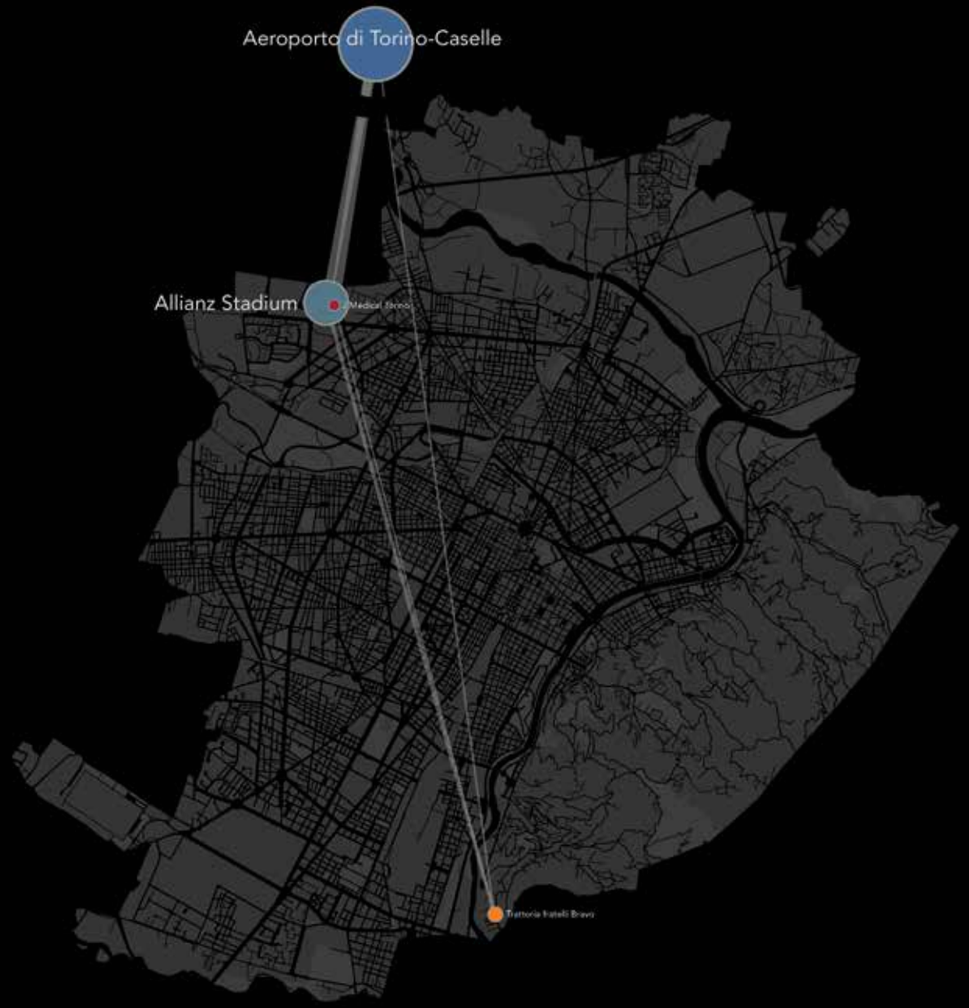


Figure 5.2.3_IN3_2

IN_3 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

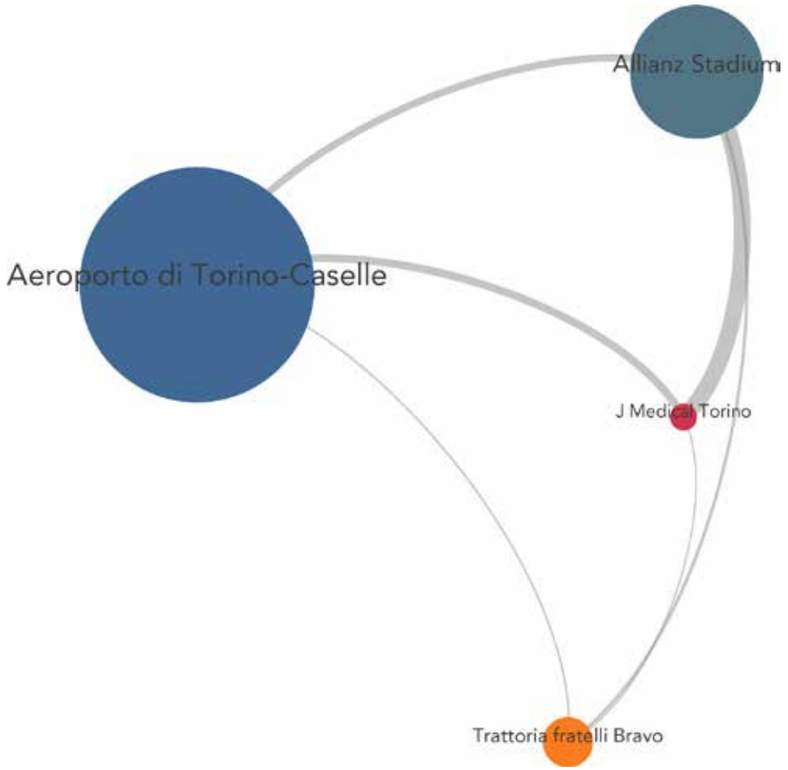
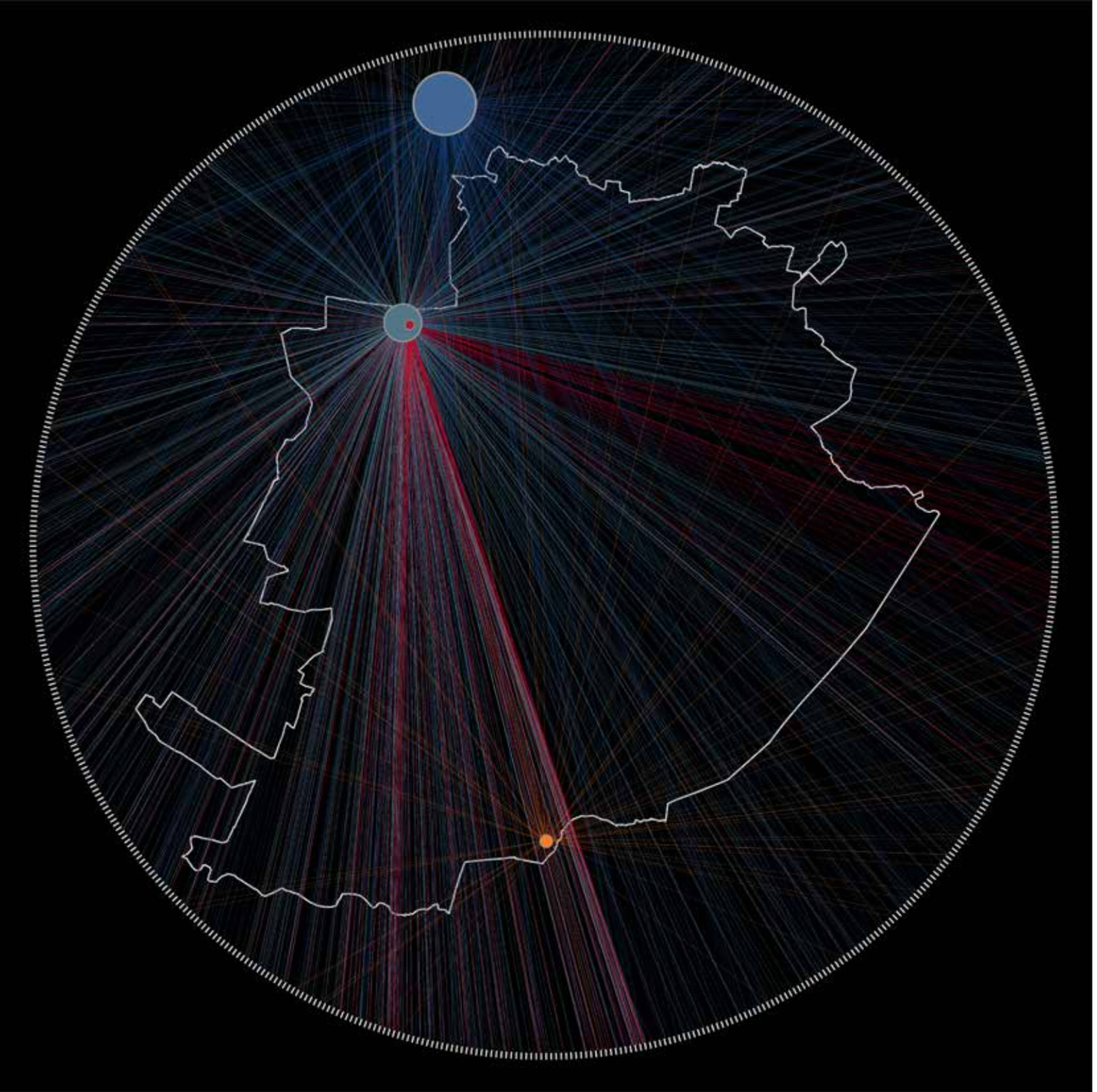


Figure 5.2.3_IN3_3

IN_3 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



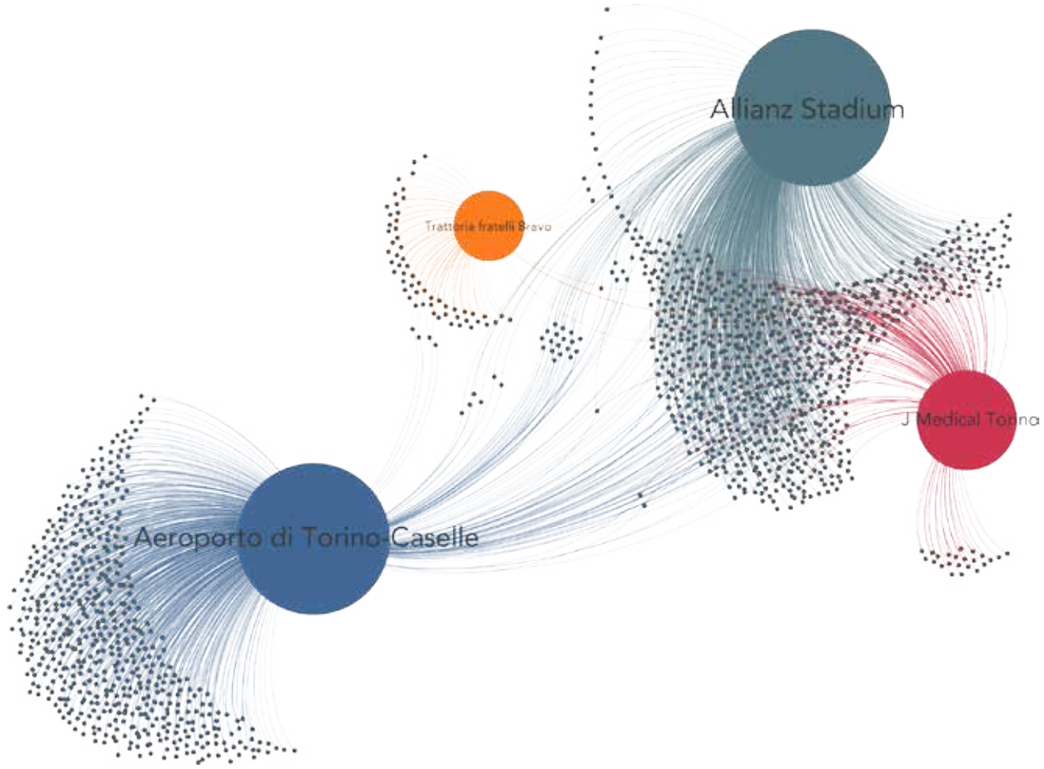
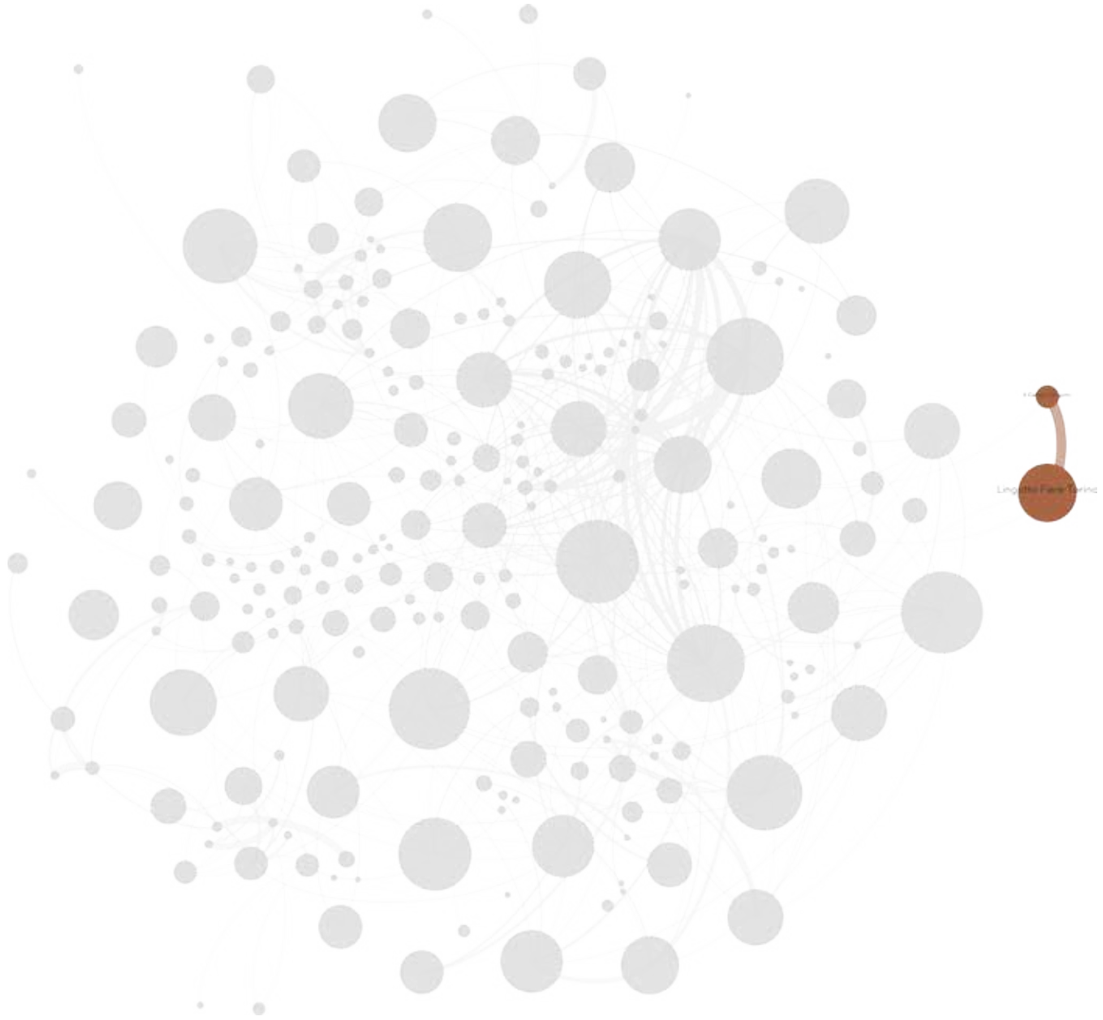


*Figure 5.2.3_IN3_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



IN_4

Figure 5.2.3_IN4_1

IN_4 highlighted in the general graph
(graph constructed with Gephi and Force Atlas 2 as layout)

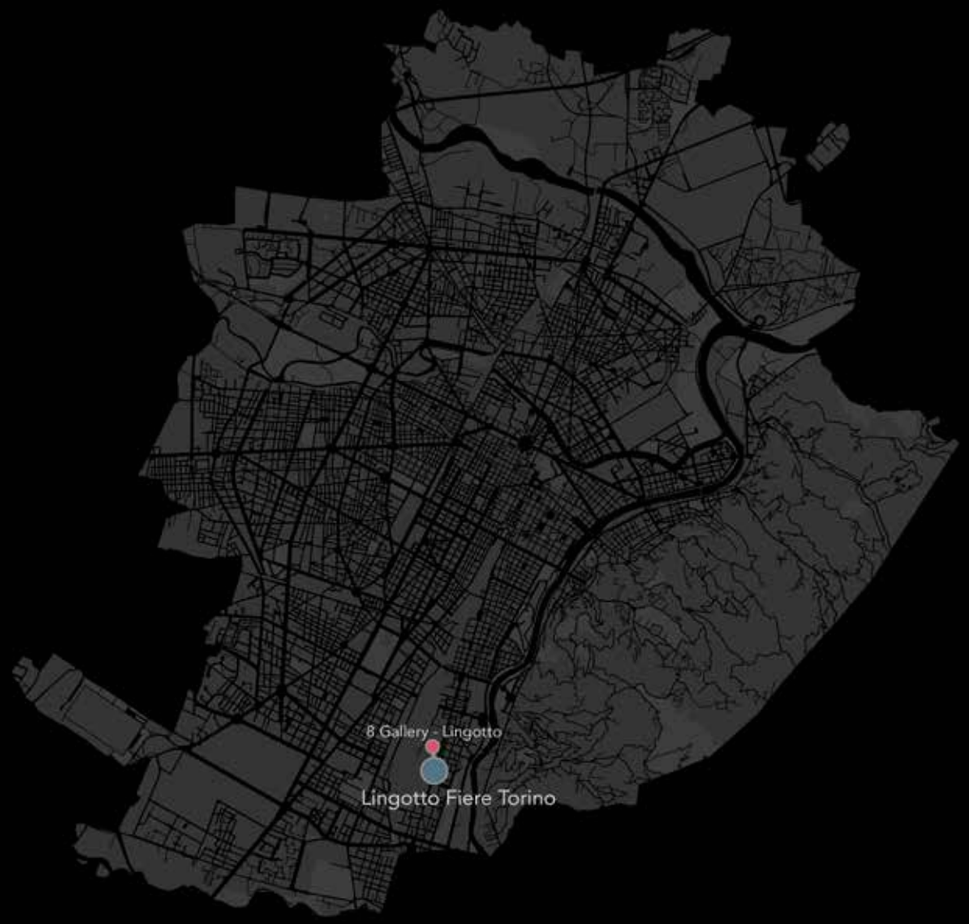


Figure 5.2.3_IN4_2
IN_4 spatialised in the city of Turin
(radius of nodes is proportional to Betweeness Centrality)

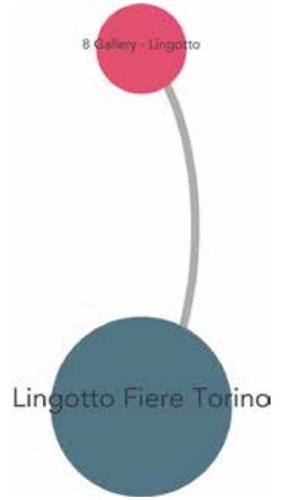
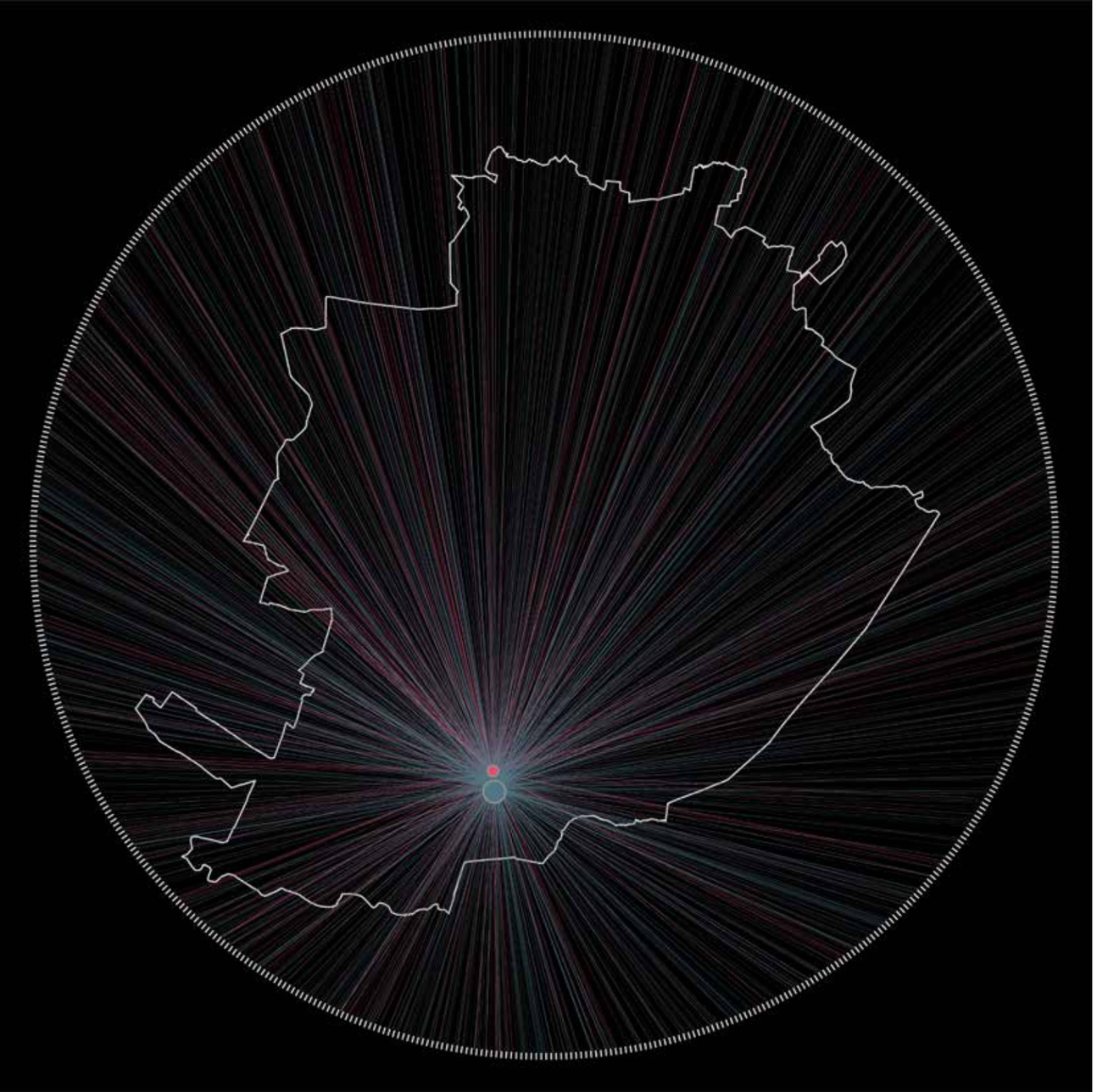


Figure 5.2.3_IN4_3
IN_4 nodes organization
(graph constructed with Gephi and Force Atlas 2 as layout)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION



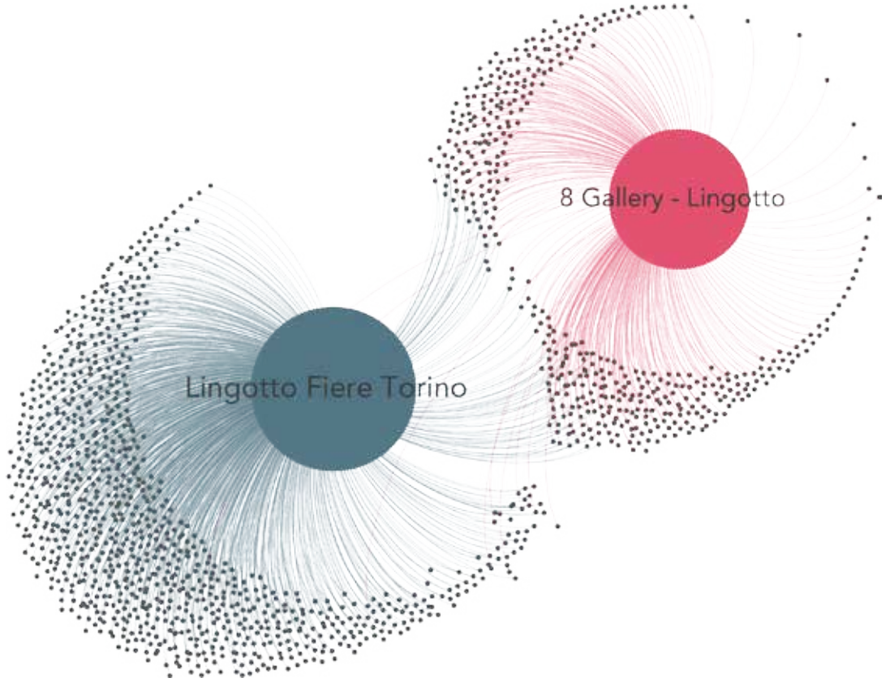


*Figure 5.2.3_IN4_*

_4 (left): Individuals connected to venues in the physical space
_5 (right): Organisation of users around venues and links (posts)

PUBLIC
LEISURE
FOOD
NIGHT
SHOPPING
TRAVEL
EDUCATION
WELLNESS
RELIGION

BIBLIOGRAPHY

Boy, J. D., & Uitermark, J. (2016). How to Study the City on Instagram. *PLoS ONE 11(6): e0158161*. doi:10.1371/journal.pone.0158161

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 58–65.

Blondel, V. D., Guillaume, JL, & Lambiotte, R. (2008). Fast Unfolding Communities in Large Networks. *Journal of Statistical Mechanics; Theory and Experiment, April 2008*.

#6
CONCLUSIONS

A space of flow and mixture, promiscuous 'meshworks' and hierarchies of different relations, [...] hybrids involving almost continuous improvisation in which the in-between' of interaction is crucial.

Amin & Thrift

The communities of venues were detected weighting the social affinity among them, how much two places were similar in the process of individual identification of users in them. What are this communities of venues? Are they "spread neighborhoods" referred to defined communities of individual?

The fourth chapter explained how Instagram data can show existing urban dynamics taking place in time and space in the city, but what about the social dimension introduced in the last chapter? How the clustering of venues should be read?

We saw in the introductory chapters how the static idea of community is not suitable to the describe the way individuals construct their identity in the society of the Information age and how the swarm is the best metaphor to harness this complexity. The swarm, an evolving mass of particles, that turn the urban space into a space of «flow and mixture, promiscuous 'meshworks' and hierarchies of different relations, [...] hybrids involving almost continuous improvisation in which the in-between' of interaction is crucial» (Amin & Thrift, 2002, p.81).

The metaphor of the swam can be used to understand the real nature of the clustering emerged in the previous chapter. The clusters are formed by venues which derive their similarity by the mass of individuals who metaphorically gather around them in their identification process, but who do not establish a univocal relation with them. The identified clusters are a way to read connections and systems otherwise unattainable.

The conclusions of this dissertations and its outcomes will be further developed continuing in the exploration of the metaphor of the swarm and, particularly, of two of its defining aspects: the shape and the dynamism.

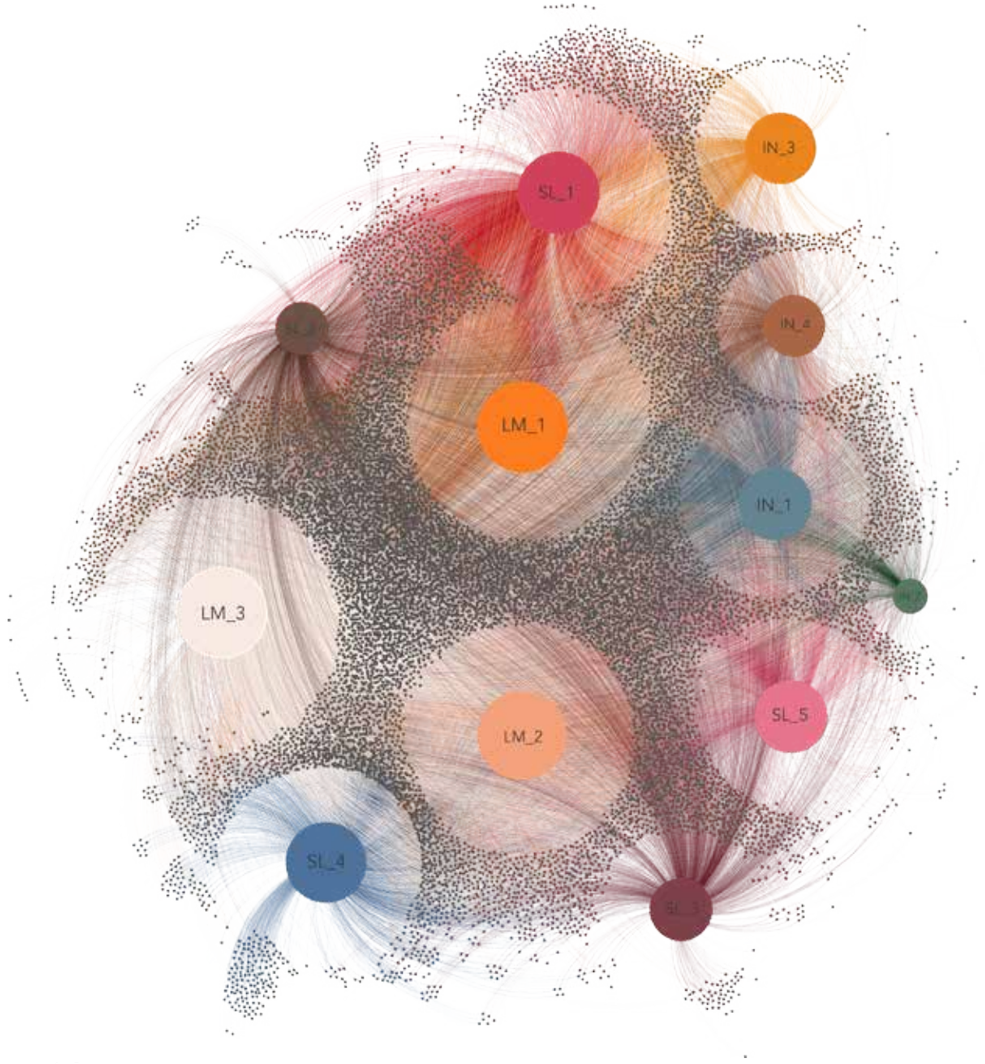


Figure 6_1

Organization of individuals around clusters
(graph constructed with Gephi and Force Atlas 2 as layout)

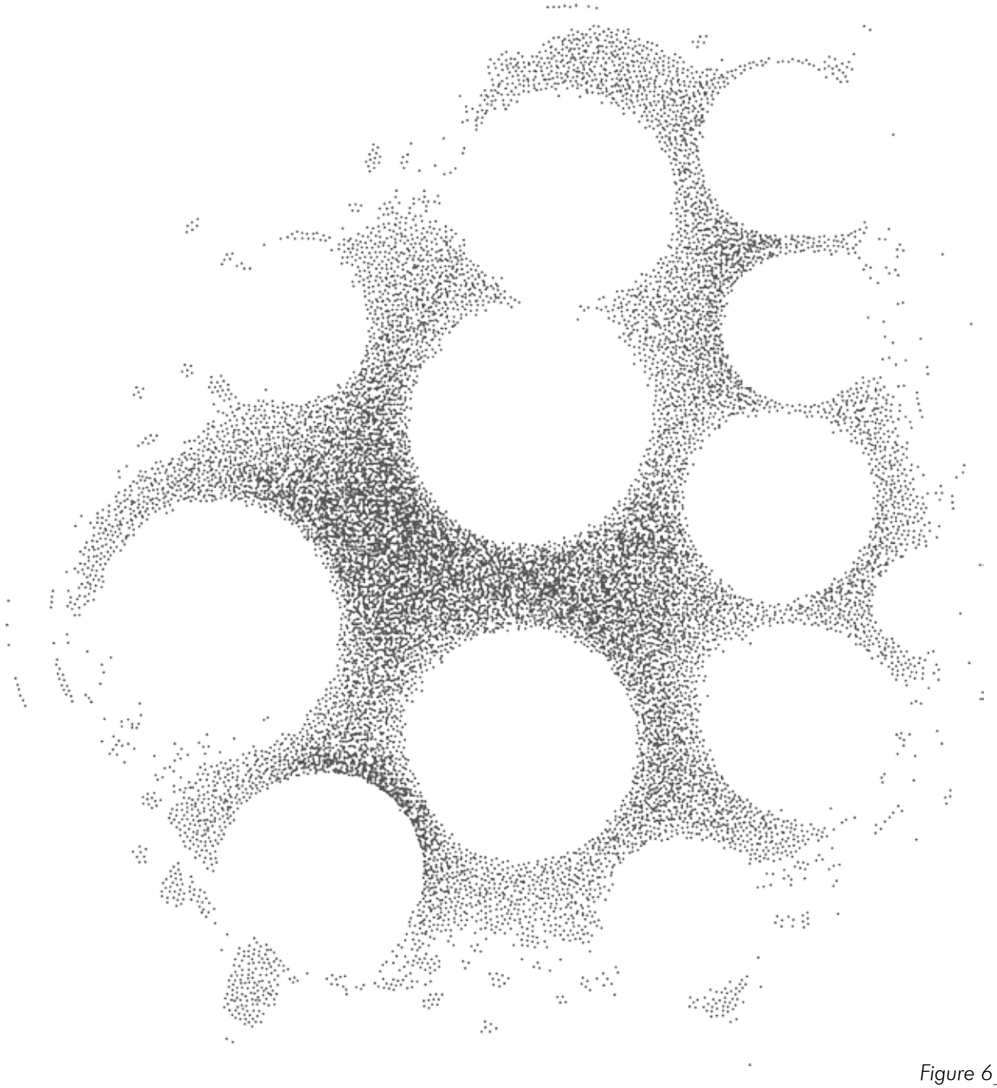


Figure 6_2

The individuals isolated as a swarm
(graph constructed with Gephi and Force Atlas 2 as layout)

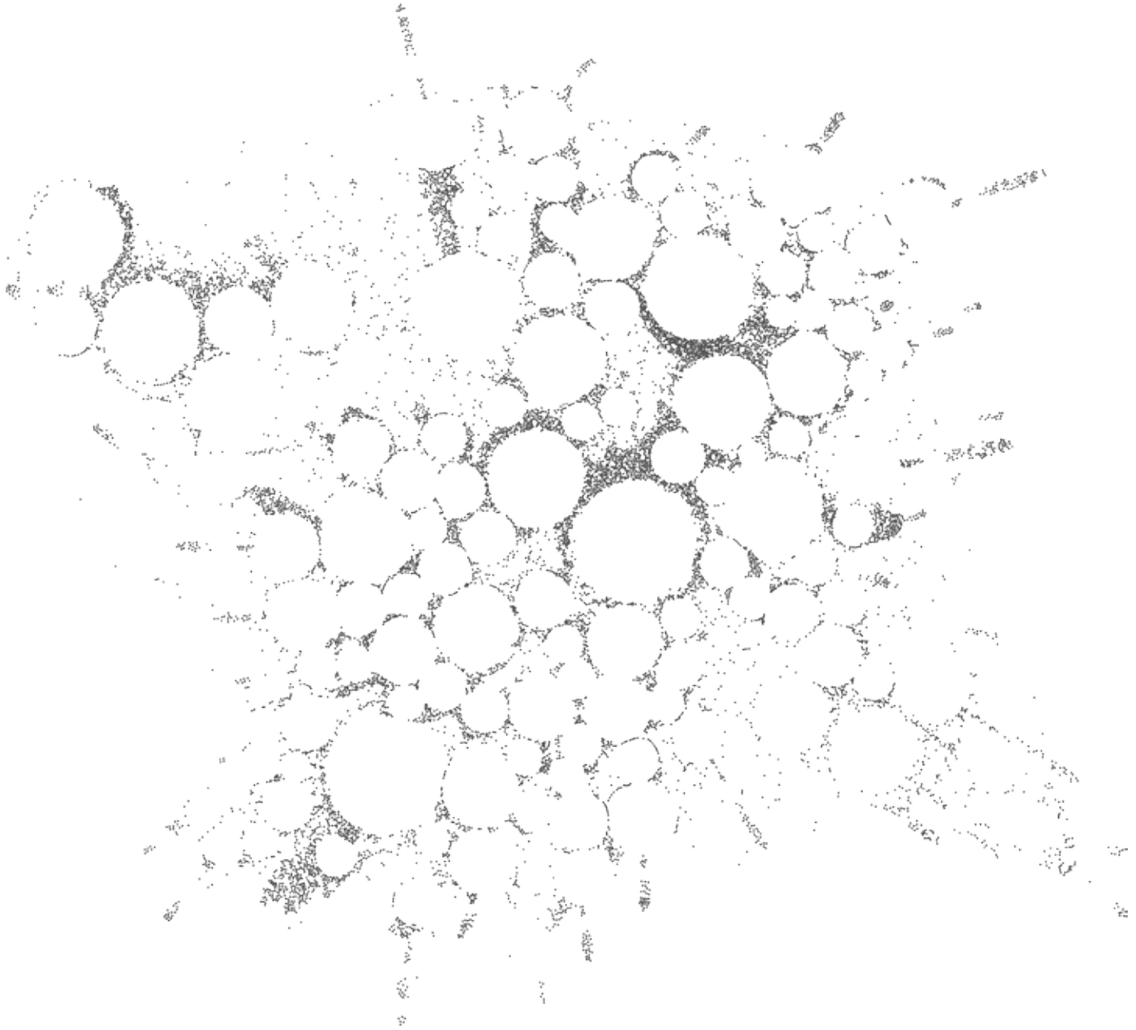


Figure 6.1_1

The swarm-like organization of individuals around the venues of “Social life”
(graph constructed with Gephi and Force Atlas 2 as layout)

#6.1

THE SHAPE: Centrality and Periphery

The swarm’s shape is a continuous process of redefinition, but it would be wrong to deny the fact that it has actually a shape, a center – or multiples - and a periphery.

What it was already readable in the analysis of the individual clusters in the previous chapter, it is even more evident when grouping venues belonging to different clusters within the same macro-group.

In [Figure 6.1_3] the venues of the five clusters referred to social life are shown in relation to the users who geotagged post in them. The division into clusters is still evident, but what it is striking is the position venues acquire in the graph.

The graph in itself has peripheries, pushed towards the boundaries, and central hubs where the overlapping of lines makes it difficult to decode the subtleties. But what about this hierarchy transferred into the real space?

The social mapping of the centrality and periphery of city venues shows how “distance” is an evanescent concept. Some spatial proximities are confirmed in the process of social identification, showing how some venues located in the city center keep their status in the graph and how they are relatively “close” to actual neighbors, while venues located in marginal areas of the city preserve their position. However, the hierarchy is subverted for other venues which acquire a centrality unknown in the real space.

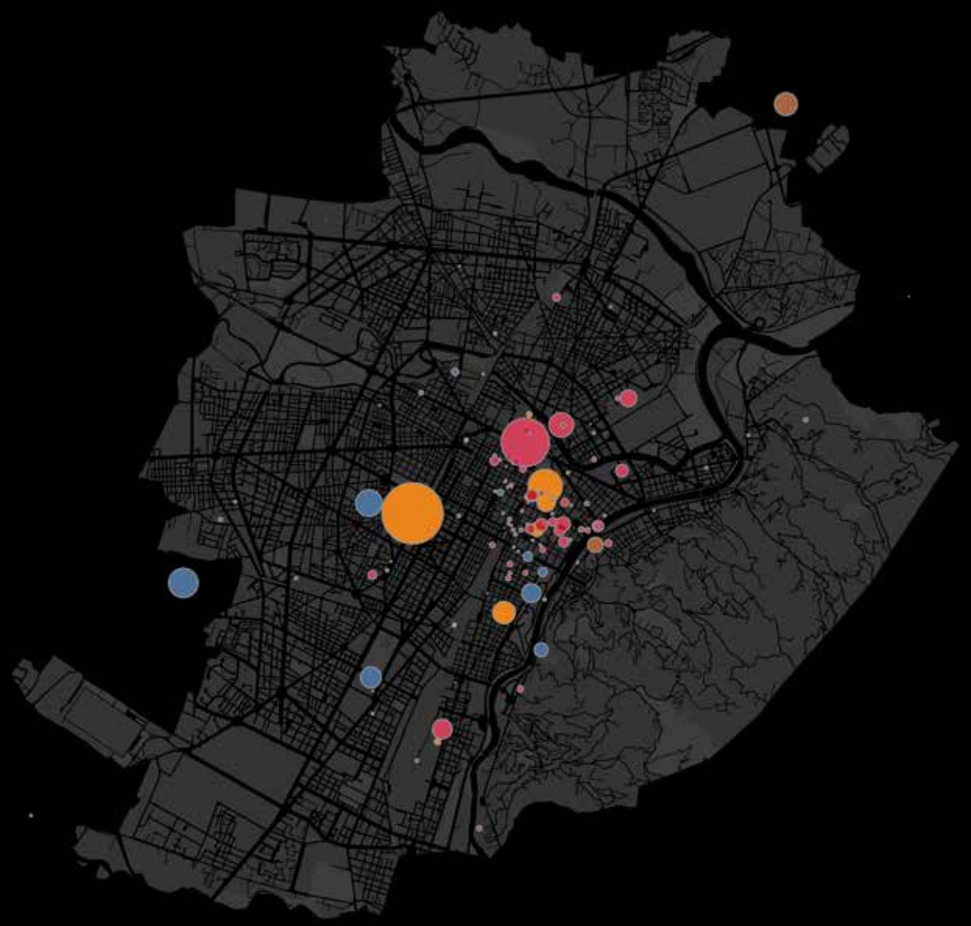


Figure 6.1_2
Venues of "Social life" spatialised in the city
(radius of nodes is proportional to In-Degree in the graph)

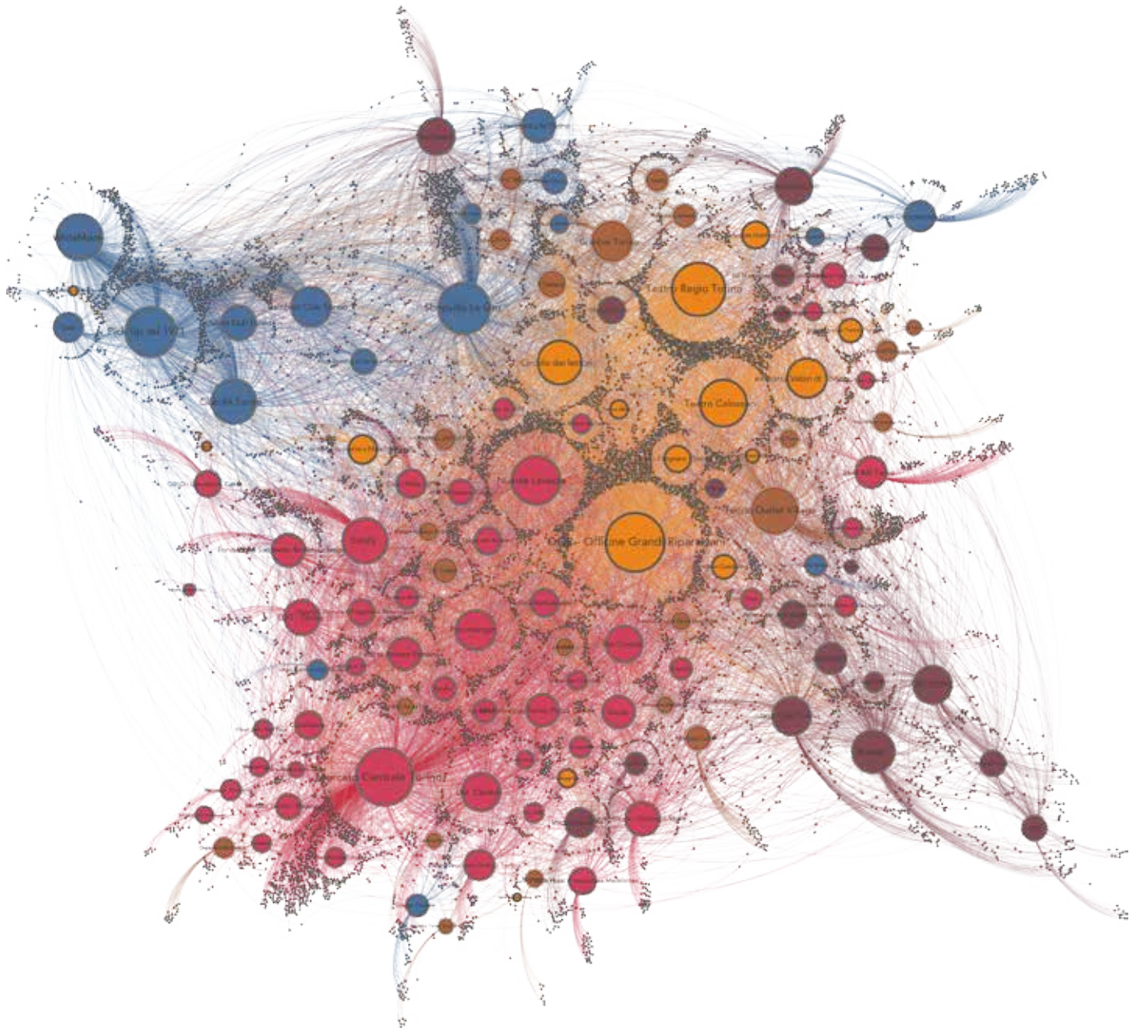


Figure 6.1_3
Organization of individuals around the venues of "Social life"
(graph constructed with Gephi and Force Atlas 2 as layout)

SL_1
SL_2
SL_3
SL_4
SL_5

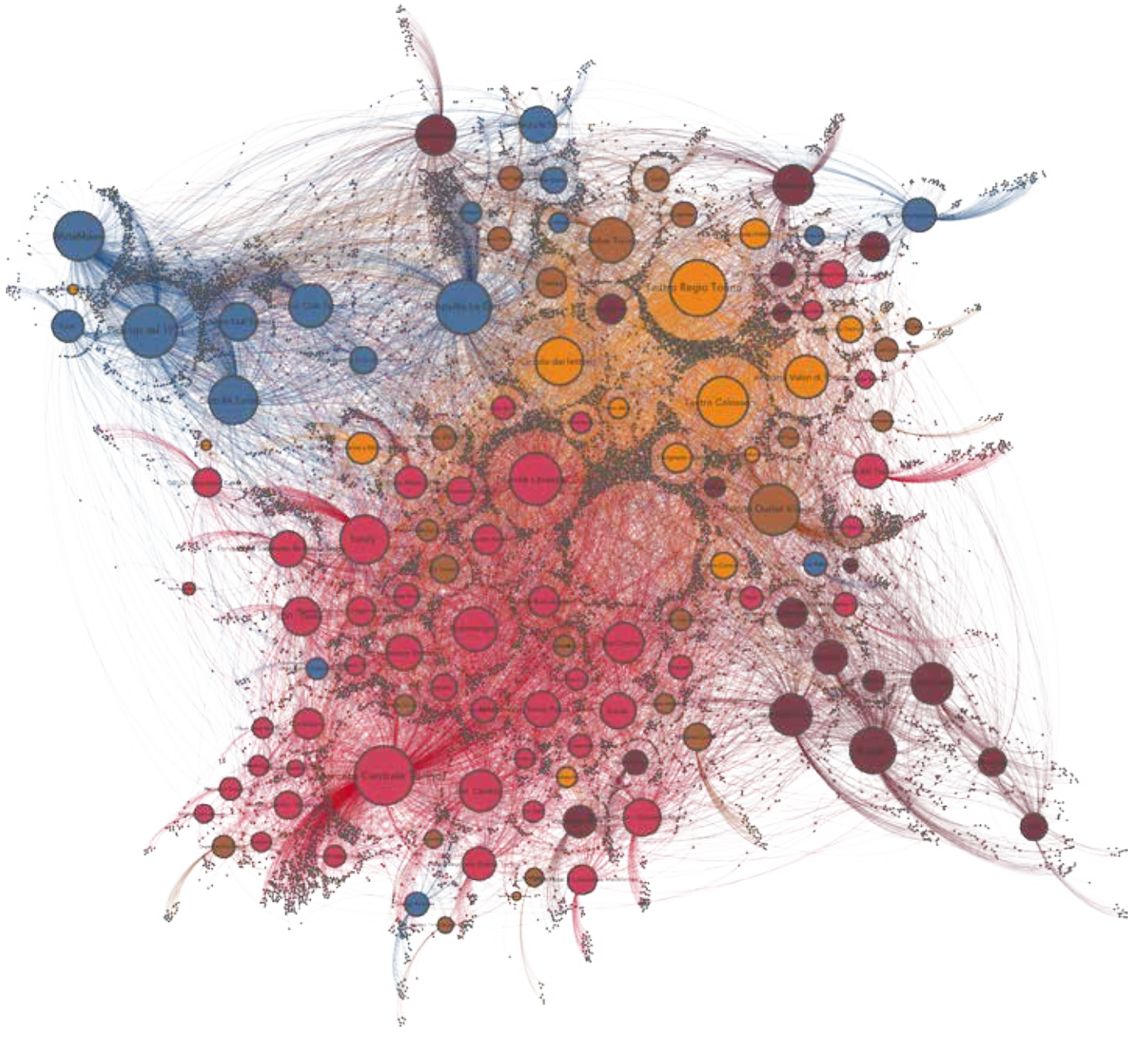
#6.2
THE DYNAMISM

The evolutionary process is continuous, and this notion of dynamism can be transferred in the understanding of the organization of venues around clusters. Although, thus, it would be wrong to state the swarm has no shape, it would be equally misleading to say that it has only one.

To continue with the already isolated group of clusters referred to social proximity, how this division and hierarchy work? How much the are contingent?
The database is referred to the solar year (from December 2018 to December 2019), and we already see how some central venues are both “historical” places – theaters, clubs, shopping mall – and “recent” ones – like the OGR, the Nuvola Lavazza and the Mercato Centrale, opened in 2017, 2018 and 2019.

What would it happen on the long period if one of them – for example the OGR – suddenly closed or went out of fashion?

As we can see in [Figure 6.2_1] and in [Figure 6.2_2], the shape and the structure, the disposition of elements would change while preserving some similarities.



SL_1
SL_2
SL_3
SL_4
SL_5

Figure 6.2_1

Figure 6.1_3 without the OGR at the center
(graph constructed with Gephi and Force Atlas 2 as layout)

Figure 6.2_2

Rerganization of individuals around the venues of "Social life"
(graph constructed with Gephi and Force Atlas 2 as layout)

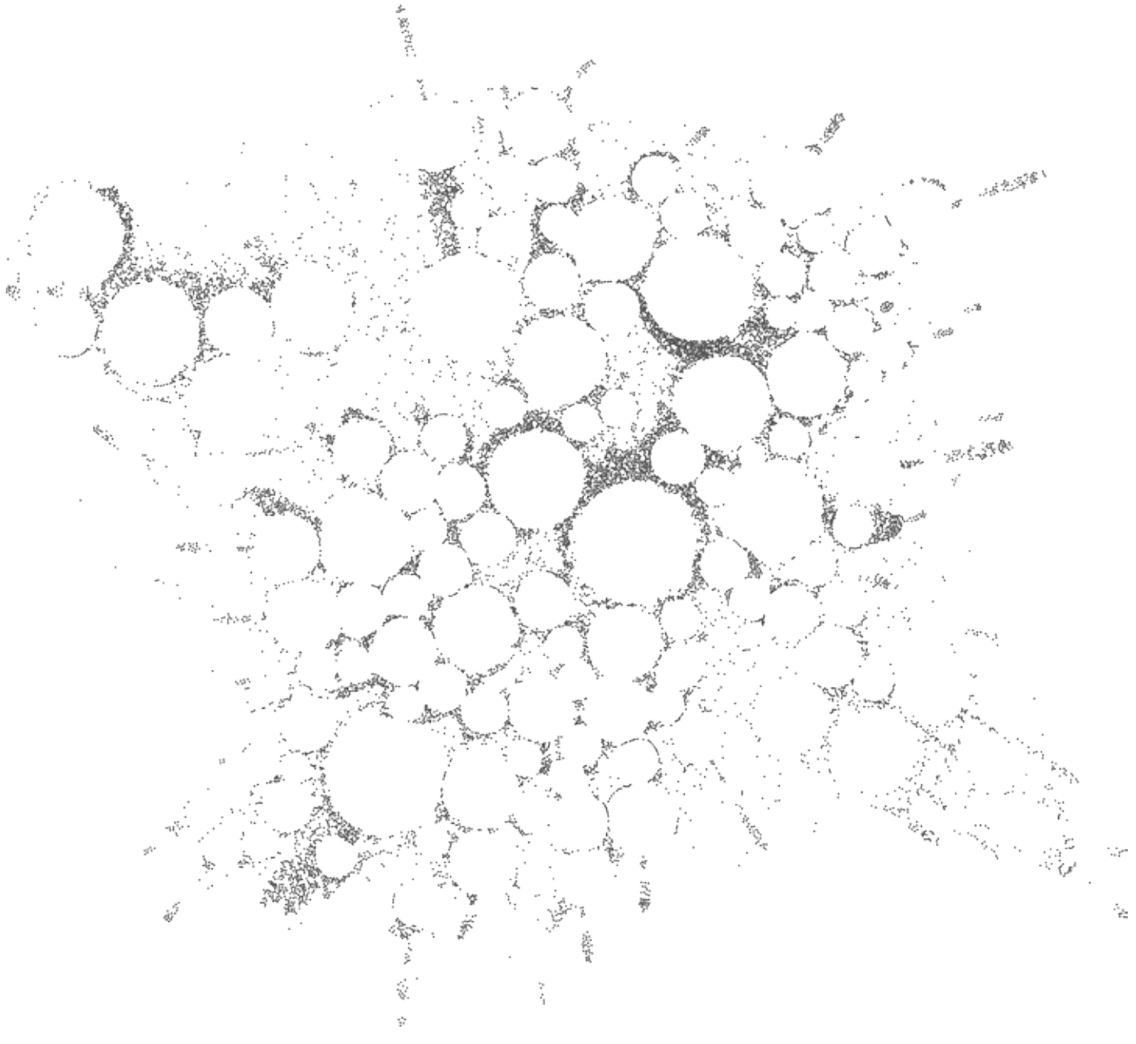


Figure 6.2_3

The swarm-like organization of individuals around the venues of "Social life" without the OGR
(graph constructed with Gephi and Force Atlas 2 as layout)

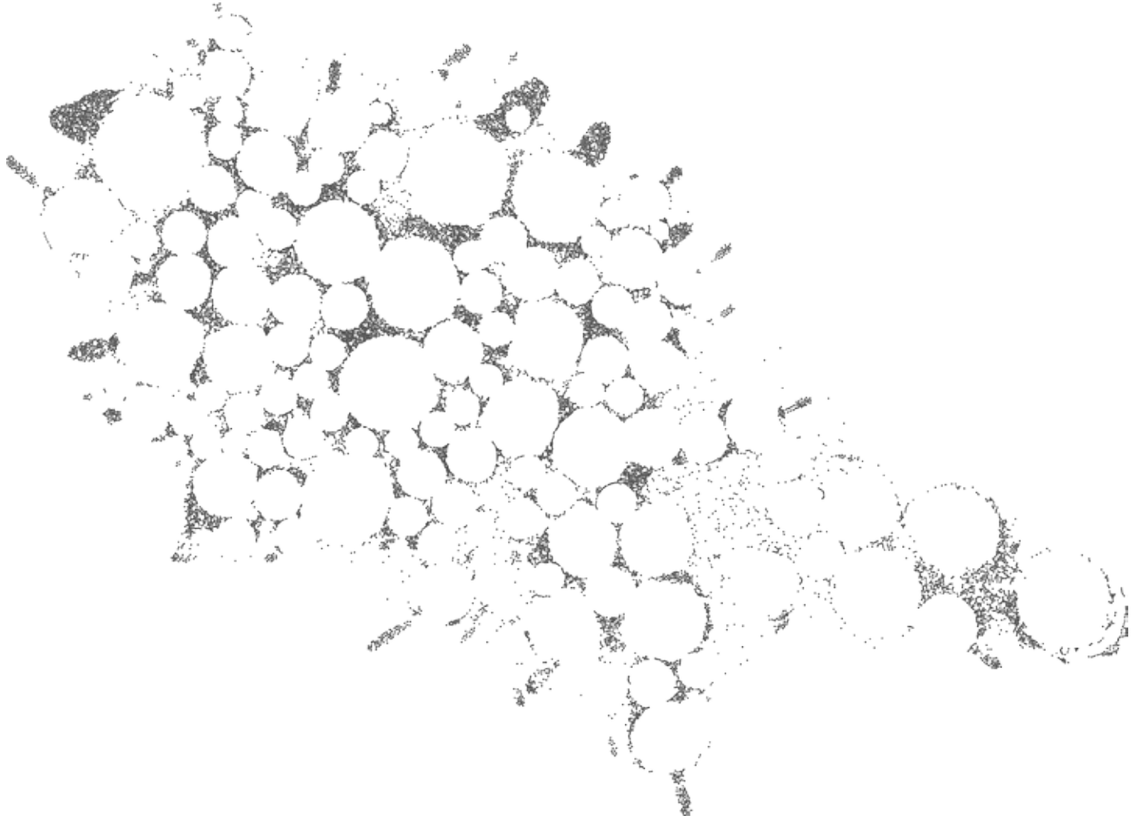


Figure 6.2_4

The "swarm" resulting from the reorganization
(graph constructed with Gephi and Force Atlas 2 as layout)

#6.3
FURTHER DEVELOPMENTS

This work proved the validity of social urban data and particularly from Instagram to understand and harness the complex dynamics occurring in the complex system of individuals, venues and networks.

The image city that emerges is dynamic, heterogeneous but not unstructured. Our role as planners is to be aware of this everchanging organism, to overcome the superstructures with which we have been reading the texture and the connections, to rethink proximity and to go establish connections beyond physical borders. A structure still exists, but it is more fluid, dynamic.

The failure of zoning and utopian city is fully known, but we still design for functions, with borders, for imagined communities in defined perimeters of space. Will urban planning be able to overcome these structures? Will it be able to zoom to a point - a square, a building, an area - and then dezoom to a larger scale and seek hidden networks buried within the surface? Will urban planning be able to see that, in the last analysis, the city is not merely the space we see, but it is rather the one we, as individuals, create?

The validity of looking at social urban data coming from Instagram can be useful for further works focusing on the analysis of urban phenomena and, even, the prediction of them, starting from the process of identification in venues of users.

The work proposed in this thesis can be furtherly expanded in the future. One possible development could consist in the evaluation of clusters with the population through surveys and interviews and lower degrees of modularity could be deployed to detect higher number of communities with a finer grain. The same process of data mining can be also extended to other typologies of search, using hashtag or user profiles, and even extend it to other social media like Facebook and Twitter.

The most important conclusion, however, that we can derive from this kind of work is the invite to change the mindset with which we look at the city and involve the tools the Digital Age has provided us with to read, understand and plan the space we live in.

ACKNOWLEDGEMENTS

When you arrive at the end of something, there is always somebody to whom be thankful. And maybe to say sorry to.

First of all I would like to thank who helped me in the development of this work and who helped and academically supported me.

Professor Matteo Robiglio, who believed in a strange student and his crazy idea a few months ago and supported me all the way till here, giving me strength and confidence

Professor Marco Brambilla, who decided to become the coadvisor of an architecture student from Turin who knew nothing in the field of informatics, but still he had confidence and accepted.

Elenga Guidetti, whom I thank for all the support not only in the academic field, who was always there next to me when I needed her the most.

The whole team of the Future Urban Legacy Lab who accompanied and helped me along these months of work.

My family needs a special mention, for you a thank and a sorry, even if one of each is not enough to express my feelings.

To my mother.
Thank you for always trusting and believing in me, for giving me the unconditional love you are only capable of. And sorry for not being sometimes as you would like me to be, when I fail you.

To my father.
Thank you for making me who I am today, and I am sorry if am not able to fully appreciate the one half of me that comes to you.

To my brother.
Thank you for having become the best brother I could ever dream of. Sorry for letting you down sometimes, for not understanding your needs, your desires and focus only on me.