Social Sciences Research Methods Regarding COVID-19 Pandemic. A PRISMA Systematic Review

Ciro Clemente De Falco[*] & *Emilia Romeo*[**]

[*] University of Naples Federico II, Italy - [**] University of Salerno, Italy
E-mail: defalcociro[at]live.it - eromeo[at]unisa.it

Abstract
In the digital society, the traditional methods of social research for the study of society have been accompanied by innovative methodological proposals: digital and digitized methods , which are now applied to many themes. In 2020 the most debated topic was certainly the one concerning the Covid-19 pandemic; a "total social fact" of which not only the medical aspects have been analyzed, a substantial scientific production, indeed, has concerned the impacts that the pandemic itself and the measures that governments have taken to resist it has had on society. Starting from these considerations, the aim of this work is to offer an overview of the topics analyzed and the research methods used to investigate this disruptive event in academic research concerning the social sciences - and in particular that which focused on the Italian case - during the last year. To map the state of the art, consolidate the heterogeneous corpus of knowledge, and investigate the different methodological approaches used (tradition/digital/digitized) a methodological approach was applied based on a systematic review of the literature conducted with the PRISMA method and carried out with a third type content analysis.

Keywords: Covid-19, Digital/digitized methods, Sistematic Literature Review, SLR, Big data.

1. Introduction: new way to study the society.

In the contemporary society, data are constantly produced as direct and indirect effects of human activities. Self-tracking with wearable devices, social media posting, mobile phone interactions, are only an example of data produced every day by billions of people. This "data deluge" (Halford et al., 2013) is commonly known as "Big data" (Kitchin, 2014). It is possible to categorize big data in different ways (Kitchin, 2014; Elias, 2012) but it is known that the biggest part is produced in digital ambient (Lupton,2015). For this reason, digital data are becoming significant sources for managing crises (e.g., tool of social media monitoring) but also to study phenomena that regard the society. For researchers now it is possible to obtain a large amount of information that previously was inaccessible. Due to Big data, it is possible to carry out "non-intrusive" investigations in "natural" contexts, avoiding distortion linked to the researcher/interviewer/observer's presence (Mahrt – Scharkow, 2013). Social scientists usually use digital media data (Rice, 1990) but now it is possible to obtain tons of data from multiple tool that retrieved information from different web sources like scraping; application programming interface (API); programs and websites designed to collect trace data and custom applications and computer scripts (Hampton, 2017). It is important to point out that data are not only "big" but also "new". Rogers (2013), indeed, distinguishes from "digitized data" and "native digital data". The first refers to objects that already existed in analogic

form and that was then digitized (e.g., films, documents, books, etc.), while native digital objects are made from user activity on the web. New and big data not only requires increased computational power but also new analysis' techniques in addition to traditional techniques. Rogers (2013), in this regard, distinguishes between digital methods and digitized methods. Digital methods are an ensemble of strategy and research approaches that use data retrieved from digital environments to study socio-cultural phenomena and changes (Rogers, 2009, Caliandro and Gandini, 2016); while "digitized methods" are traditional social research tools adapted to the web (Rogers, 2009). Both methods have strengths and weaknesses (Hampton, 2017) but allows researchers to retrieve knowledge from new and big data. What is still under discussion is the extent of these data and the generalizability of them results, for example is not clear if the study of the twitter data can help researchers to understand the dynamics that characterize the social network Twitter, or the dynamics of Twitter in-social-life (Marres, 2012). Furthermore, using new and big data or new methods can be difficult for several reasons. First of all, the accessibility to all data is not guaranteed, for example, due to the growing trend towards "proprietary closure" (Manovich, 2012) is difficult to access to data produced by digital infrastructures. Companies have access to large data sets, while researchers can obtain part of these. Secondly is not always easy for social scientists to collect, store and analyze massive quantities of data. For this reason, Lupton (2015) - as many others - highlights the importance for the new generation of social scientists to improve technical background and to work in interdisciplinary groups. Finally, resistance regarding the "ontological" level could hinder the use of digital data. Another typical challenge associated with high-volume, diverse datasets is whether synthesis of data streams can be translated into actionable knowledge. The concept of translating "big data to knowledge" is important to the social sciences in several respects, big data is now being used in a variety of fields: from economics to science and politics. Governments all over the world, for example, could use big data to monitor and manage their territories through Event Detection algorithms (Mellin and Berndtsson, 2009). This kind of algorithm is becoming increasingly popular in the context of Smart cities (Borges et al., 2017) and especially in emergency management (Pohl et al., 2012). An important example of this kind of uses can be seen in the data collection from tweets during natural disasters such as Sandy (Wang et al., 2015). In addition to the purely managerial side, big data can be used as knowledge source to better understand the socio-economic dynamics and phenomena that arise from emergencies. This general shift to data-intensive science will exert an influence on all scientific disciplines, but particularly on the social sciences given the wealth of behavior and related constructs captured by big data sources. Starting from the consideration of the covid-19 pandemic as an exogenous shock similar to a natural disaster the aim of this study is highlight, in the first place, the different kinds of research conducted in social sciences field regarding Covid-19. Moreover, given the challenging period and the importance of "social distance" between people and the "lockdown" restrictions carried on most of the research was conducted online. So, from a methodological point of view, it is interesting to understand in which way online research was carried out. Have digitized methods or digital methods been used? Have the new data sources been exploited? Besides providing an idea of the current level of implementation of digital methods and the use of big data in the current social research, answering these questions will allow us to understand, from a practical point of view, how we could consider constantly monitoring the evolution of socio-economic phenomena during emergencies and whether the tools currently used can be defined as effective.

The study is organized as follows. First, an overview of the diffusion of Sars-cov2 in Italy. Then, the methodology based on the Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA, Moher et al., 2009) model is presented. In the following sections, the results of the analysis are reported and discussed by identifying: 1) the main themes discussed in social research; 2) the use of digitized, digital, or traditional methods in empirical research during the pandemic; 3) the level of exploitation of big data in empirical social research. Finally, implications, conclusion and limitations of the work are debated.

2. Defining the issue. Covid-19 pandemic: the Italian case

A total social fact according to Marcel Mauss (2002) is something capable of influencing and determining a set of phenomena involving most of the functioning mechanisms of the reference community. The total social fact, thus, would allow the interpretation of apparently distant and different "pieces" of the same society. Today, the coronavirus performs the same function. Social norms, inequalities, culture and political polarization are the main things (but not the only) that have been affected and influenced, even in a radical way. In particular, Italy represents a case of interest as it is one of the first countries to register one of the highest numbers of infected in the world and it is also among the first to respond positively to the pandemic trying to implement measures to counter the spread of the virus, or at least to limit it as much as possible. On January 31st, 2020, a state of health emergency was declared due to the new coronavirus epidemic. The decision was made immediately after the WHO declared a public health emergency of international concern. The first two Covid-19 cases were confirmed in Italy in February 2020, and, by the end of the month, the virus had spread over northern regions of the country. On February 23rd, the Government issued a decree aimed at preventing and combating further transmission of the virus but during March, the toll of new cases and deaths kept surging, rapidly over- taking the numbers registered in China and making Italy the hardest-hit country in Europe and one of the epicenters of the global pandemic. The rules decided by the government to counter the spread of the virus initially included the total schools' closure, restrictions on people's mobility and the obligation for people not to leave their home for a month; while for the commercial activities two weeks of suspension, but then an extension of the measures was deemed necessary restrictive (Briscese et al, 2020). Until March 21st, 2020, the number of cases is multiplied, reaching the peak with a very steep curve. Then, from March 21st to May 3rd, in 43 days, the curve decreased slowly. On May 4th the second phase began with the opening of some kind of activities and with the possibility for citizens to freely move according to safety conditions. Since May 18th, the number of infected people has continued to decline and the government has ordered the reopening of most activities and the possibility, for the people, to move between regions and in Schengen countries. Sadly, the perhaps hasty reopening that involved the country, mostly in order to support a gradual revival of the economy severely affected by the period of forced lockdown, has exposed Italy to the severe consequences - still in progress - of a second wave with bigger numbers of infected than the first one. The 4th of November, in fact, due to the increasing number of infections, following a new decree of Italian Prime minister the country has been divided into three areas (e.g. red, orange and yellow) according to risk profiles that take into account the evolution of the epidemic with weekly updates. The classification system according to the risk scenario of each Region in "red", "orange" and

"yellow" leads to specific restrictive measures according to the color. The red is the most serious, followed by the orange and then the yellow. So, due to Covid-19 lockdown, our lives, by now, were projected in the network, that, according to Boccia Artieri (2015), has become "a place of the society". Job, education, shopping and other different activities, in this period, have been transferred to the digital world, representing a turning point for social research. Hence, this event, from a social researcher perspective can be seen as an opportunity to exploit the advantages related to big data and computational techniques related to the digital world. According to Pitrelli (2017) automated methods of information extraction, analysis of social networks applied to the Internet, use of computational models to understand how individuals interact with each other, augmented reality, geospatial analyses have gained an increasingly significant space in multiple fields of application, therefore becomes interesting to understand how much these advances have been exploited in social research.

3. Methodology

In order to provide an overview of the covered topics and the adopted methods and sources to investigate the phenomena - directly and indirectly linked to the COVID-19 pandemic and its restrictive measures - a systematic literature review based on content analysis is performed. The review presents two main objectives: 1) identify the main topics studied; 2) analyze the use of digital data and techniques for empirical research in social science.

The systematic literature review has been carried out during a temporal range that goes from November to December 2020 by two researchers from two Italian universities - the University of Naples "Federico II" and the University of Salerno - according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) methodology (Moher et al., 2009). Moreover, thanks to a well-defined process in four steps, PRISMA ensures a strict review plan, offering methodological accuracy, transparency, and the possibility to easily replicate the research (Tranfield et al., 2003). PRISMA comprises four steps - identification, screening, eligibility, and inclusion - to perform the SLR (Moher et al., 2009), as described below and in Figure 1.

3.1 Data collection

During the first step, known as the identification phase, to properly cover the issue Scopus and Web of Science (WOS) database have been used for the paper research according to their extensiveness and relevance in social sciences. To obtain the most relevant literature in databases, the choice of keywords fell on "covid*19", "coronavirus", "Italy" and "Italian". The keywords have been connected with the Boolean operator "AND" and "OR". Thus, the following search string has been defined: (("coronavirus" OR "covid*19") AND ("Italy" OR "Italian")). The search produced 4,482 articles. To complete this step, after entering the search strings, inclusion and exclusion criteria were set to obtain the relevant literature in the chosen databases.

As shown in Table 1, works published between 2020 and 2021[1] were selected because this period indicates the arise and developing of Covid-19 emergence in Italy.

Table 1. The process of data collection

Database	*Keywords*	*Records*
Scopus	TITLE-ABS-KEY ((“coronavirus” OR “covid*19”) AND (“Italian” OR”Italy”))	3,304
WOS	TOPIC ((“coronavirus” OR “covid*19”) AND (“Italian” OR”Italy”))	1,178

Legend: ABS= Abstract; KEY = Keywords

Regarding the type of document, as reported in Table 2, reviews, editorial book chapters and working papers were excluded, while articles published in international and national peer-reviewed journals, conference papers and “in press” papers (in English or Italian languages) belonging to a specific research fields were included in the analysis. Thus, the related research areas that belong to the research domain of social science were considered for both databases, excluding disciplines as neuroscience, pharmacology, and medicine. It is important to highlight that a minimum number of citations were not included because of the novelty of the topic under study.

Table 2. Inclusion criteria

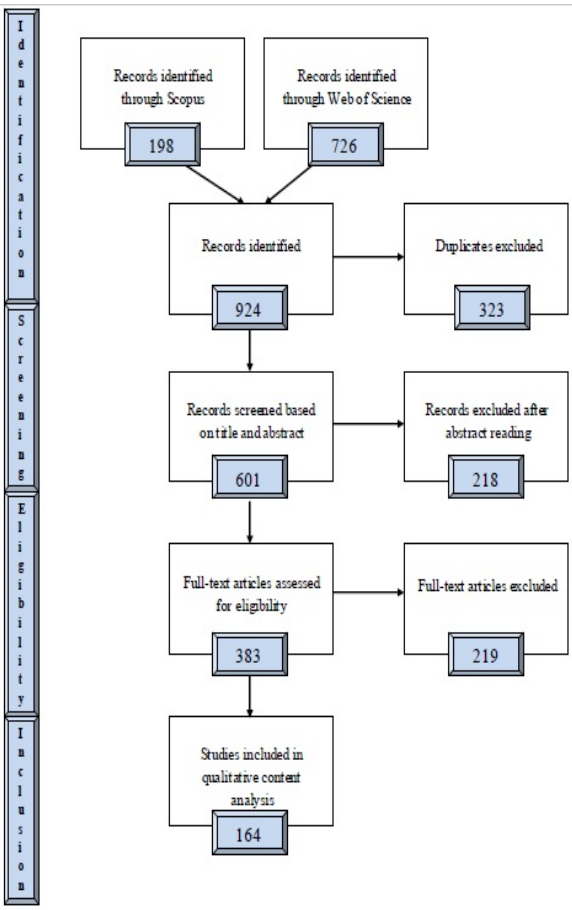
Criteria	*Description*
Document type	International and national peer-reviewed journal articles; conference paper; in press articles
Language	English or Italian
Research Domain	Social Science
Research Areas	**For Scopus**: Anthropology; Area Studies; Psychology; Behavioral Sciences; Physiology; Demography; Sociology; International Relations; Environmental Sciences Ecology; Development Studies; Family Studies; Communication; Social Work; Business Economics; Social Issues; Education Educational Research; Women Apos S Studies; Women S Studies; Ethnic Studies; Film Radio Television; Linguistics; Geography; Cultural Studies; Urban Studies; Sport Sciences; Social Sciences; Other Topics **For WOS**: Film Radio Television; Public Environmental Occupational Health; Geography; Nutrition Dietetics; Psychology; Behavioral Sciences; Education Educational Research; Demography; Environmental Sciences Ecology; Sociology; Family Studies; Cultural Studies; Women Apos S Studies; International Relations; Social Sciences Other Topics; History; Social Issues; Public Administration; Criminology Penology; Business Economics; Ethnic Studies; Linguistics; Communication; Philosophy; Anthropology; Women S Studies; Architecture; Telecommunications; Urban Studies; Area Studies
Timeline	2020-2021

As synthesized in Figure 1, in the step of identification 3,558 articles were deleted. 323 of the 924 records from Scopus and WOS resulted duplicates and thereby were rejected. On this basis, 601 records were screened. Then, in the screening phase, based on the reading of 601 abstracts, 218 papers were deleted, because of their

[1] 2021 is the year of publication of some works included in the analysis.

low pertinence to research aims and review objectives. Thus, after the reading of the full texts of all 383 remaining articles, 164 publications were included in the review process because, dealing with the empirical investigation of social impact of covid-19, they contribute to answering the review questions. Moreover, according to Robinson and Lowe (2015), who recommended the inclusion of 10-50 papers for a SLR, the number of papers included seems to be widely acceptable.

Figure 1. The assessment and selection of contributions: PRISMA flow diagram



Source: Authors' elaboration

3.2. Data Analysis

After the systematic definition of the sample, the data analysis took place through the third type content analysis technique. In explorative research, content analysis can be helpful to derive relevant knowledge from a large number of texts using both qualitative and quantitative approaches (Krippendorff, 2018). This technique, in fact, allows the extraction from textsin a non-automatized way, some content categories in order to reveal the focal points of the studies (Krippendorf, 2004). According to Losito (1996, p.76) content analysis of the third type, also known as "inquiry", is a procedure that the researcher can apply "to any type of message - verbal and non-verbal - using a semi-standardized or standardized analysis sheet to record the states in which they occur certain properties in a set of suita-

bly selected units of analysis". In this kind of procedures, a questionnaire is used to analyze a movie, a discursive response of an interviewee, a story, an advertising image, or an article in a magazine instead of an individual (Rositi 1988, 73). In this work the unit of analysis is not a sample of individuals but scientific publications, indeed. So, the data collection took place thanks to the use of an analysis sheet in which the operational definitions and classification rules of the variables have been specified and made as explicit and exhaustive as possible, in order to allow other researchers to reach the same conclusions. The analysis sheet was divided in three sections: 1) general information about the article's characteristics such as number of citations, authors, year of publication, authors' affiliation etc.; 2) title, abstract and keywords; 3) variables detected in order to understand the methodological orientation of the selected papers such as type of action to obtain information, type of technique used, type of data (e.g., digital/digitized) and number of observations. In particular, to address the first objective of this study the keywords have been used as proxy for the main themes of each article. On the other hand, for the second objective the methods, sources and dataset extent have been investigated.

4. Results

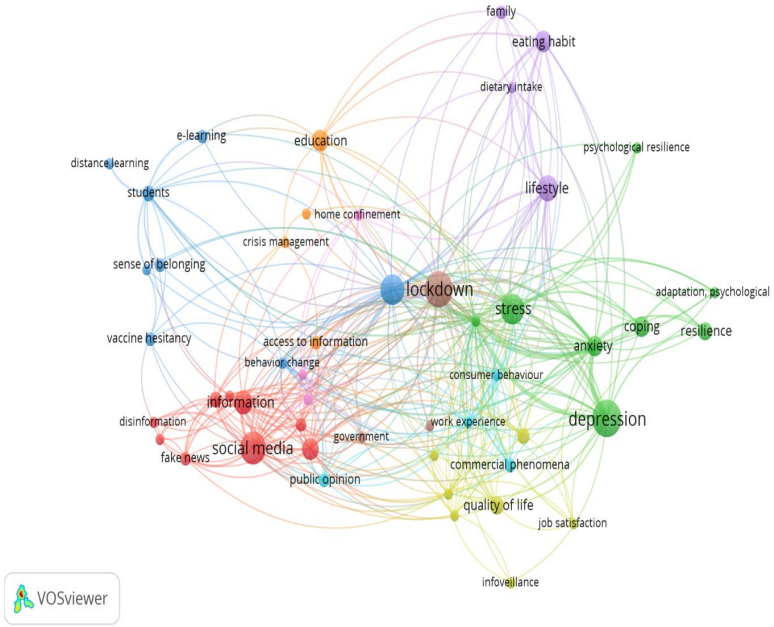
4.1. Thematic analysis

In order to address the first objective a network analysis through the analysis of co-occurrences of the articles' keywords was carried out in order to identify specific themes or research areas starting from the associations established between the keywords present (López-Fernández et al, 2016). The analysis of co-occurrences allows to understand how many times two or more keywords are simultaneously present within the analyzed texts.

The results of the analysis of co-occurrences have been organized and displayed in a reticular form (see figure 2). Moreover, thanks to cluster analysis (modularity) algorithms, it has been possible to identify keywords groups, in the whole network, that should constitute the theoretical blocks or fundamental topics of the research field under examination (Manesh et al., 2020), regarding the online research that take into account the COVID-19 pandemic impacts[2][3]. To perform the analysis, VOSviewer (Van Eck and Waltman, 2017), developed at the Center for Scientific and Technological Studies of the University of Leiden, was used. It is an extremely useful software for mapping scientific literature (Van Eck and Waltman, 2009) which allowed first to code the text units and then to analyze the keywords of the sample selected in the previous research phases.

[2] Works that used classic techniques (8) are not included in this analysis.
[3] To better bring out the links between the issues, the Covid-19 node has been ousted.

Figure 2. Keywords' network



Source: Authors' elaboration

Table 3 describes the occurrence of the keywords identified for each cluster in the network analysis performed through VOSviewer.

Table 3. Clusters' keywords occurrence

Cluster 1 (Red)		**Cluster 2 (Green)**		**Cluster 3 (Blue)**	
Keyword	**Occurrence**	**Keyword**	**Occurrence**	**Keyword**	**Occurrence**
communication campaign	2	adaptation, phsycology	2	behavior change	2
disinformation	2	Anxiety	6	distance learning	2
fake news	3	Coping	6	e-learning	3
health commu-nication	2	Depression	20	sense of belonging	3
information	8	mental well-being	2	social relation	14
interpersonal communication	3	phsycological resilience	2	students	4
mass media	4	Resilience	5	university	2
social behavior	7	Stress	14	vaccine hesitancy	3
social media	16				

Culture e Studi del Sociale-CuSSoc, 2021, 6(1), Special, pp. 123-141
ISSN: 2531-3975

Cluster 4 (Yellow)		Cluster 5 (Purple)		Cluster 6 (Light blue)	
Keyword	**Occurrence**	**Keyword**	**Occurrence**	**Keyword**	**Occurrence**
healthcare workers	4	dietary intake	2	commercial phenomena	3
infoveillance	2	eating habits	7	consumer behavior	2
job satisfaction	2	family	3	public opinion	3
loneliness	2	lifestyle	10	work experience	2
quality of life	5				
social network	2				
socioeconomics	2				

Cluster 7 (Orange)		Cluster 8 (Brown)		Cluster 9 (Pink)	
Keyword	**Occurrence**	**Keyword**	**Occurrence**	**Keyword**	**Occurrence**
access to information	3	government	2	home confinement	2
crisis management	2	lockdown	19	mobility	2
education	7	tourism	2	social distancing	2
school	2				

In the first cluster "information and disinformation" (red), there are studies that focus on how during the Covid-19 pandemic, risk communication has often been ineffective, and how from this perspective "fake news" has found fertile ground, both as a cause and a consequence of its ineffectiveness (Moscadelli et al, 2020). Many scholars investigate, indeed, the role of information sources in influencing public opinion and the choice of the Italian people to trust and rely on institutions, accepting such necessary measures and sustain behavioral changes (Falcone and Sapienza, 2020; Graffigna et al, 2020). Therefore, considering the fact of being the first full-scale global pandemic of the digital age, Covid-19 has presented novel challenges to government, above all the spread of news and misinformation, and the trade-off between the accessibility of science and the premature public use of unproven medical interventions (Badell-Grau et al, 2020).

The second cluster "psychological risk" (green) collects studies aimed at identifying psychological changes (e.g., changes in depression, stress, and anxiety levels) (Roma et al, 2020; Germani et al, 2020, Lenzo et al, 2020) among the Italian public during the lockdown period, in addition to factors associated with these changes. According to some scholars (Cusinato et al, 2020), the novel coronavirus outbreak has forced parents and children to adopt significant changes in their daily routine, a big challenge for families with important implications for family stress. But also

emerging adults have manifested risks for their psychological adjustment showing concerns about their role as a possible asymptomatic carrier of Covid-19 as is possible to see in Lenzo et al (2020).

In the third cluster "e-learning issue" (blue), it is possible to find works that analyze new challenges for university students (Capone et al, 2020; Procentese et al, 2020; Gaggi et al, 2020) examining the associations between mental health and academic stress or with other factors like self-efficacy, satisfaction for degree course, Covid-19 risk perception and so on. Conversely, other works address the role of individual resources and community dimensions into facing Covid-19-related lockdowns and the changes that these have implied, always with reference to academic communities. Some studies investigate the most adopted technologies, didactic methodologies, as well as the impact on school's population.

The fourth cluster "pandemic consequences" (yellow) collects the studies that have focused on the analysis of how social network sites have been used during this lockdown period and how they influenced the quality of life both by allowing spreading of information and by allowing the maintenance of long-distance relationships. The surveys mostly concerned how the social networking sites have became the prevailing information medium, where the main topics, during the lockdown phase, have been discussed (De Santis et al, 2020). Some scholars have collected information about the use of social network sites by the elderly, and about the impact on self-perceived loneliness and social commitment with family and friends (Rolandi et al, 2020).

The fifth cluster "food issue" (purple) shows works that investigate how the confinement due to the Covid-19 pandemic influenced dietary profiles (Ruiz-Roso et al, 2020; Pellegrini et al, 2020), evaluating the changes in weight and dietary habits, especially those of adolescents, who are highly susceptible to acquiring bad eating habits.

The studies in the sixth cluster "consumer behaviour" (light blue) contribute to the debate on consumer behaviour, evaluating impact of the coronavirus pandemic on the purchasing willingness, anxiety, trust, and concern of the Italian people (Aquilanti et al, 2020; Troise et al, 2020).

The orange cluster "education management" looks at studies that analyse how schools, one of the most impacted services in Italy, dealt with lockdown and social distancing measures, highlighting how, during the planning of an uncertain future, the public institutions cannot just shut down but must increase their activities in this field (Iacuzzi et al, 2020).

The eighth cluster "effect on tourism" (brown) presents the contributions that analyze the impact of lockdown measures introduced in response to the spread of novel coronavirus disease 2019 on socioeconomic conditions of Italian citizens, in particular in tourism sector (Aiello et al, 2020; Bonaccorsi et al, 2020). Some scholars explore, indeed, the touristic intentions and the role of public finance that could sustain the demand of tourist services in hotels and B&B, which is lowering because of coronavirus.

In the last cluster "impact of restrictions" (pink), there are studies relating to the ways through which mobility restrictions affected Italian citizens, and studies that looking into socioeconomic and institutional factors of urbanity and modernity that have significant effects on pandemic severity (Bonaccorsi et al, 2020; Rahman et al, 2020).

4.2. Analysis of methods, source, and dataset extension

As previously seen, the new context in which researchers operate is characterized by the possibility of using large datasets (e.g., big data) and digital/digitized data and methods which have been added to the traditional techniques. As seen in the previous paragraph, there are many ways in which the pandemic and its impact on individuals and the social system as a whole was studied. In this paragraph the analysis focuses on evaluating the extent to which the tools and data of the digital society have been used (Lupton, 2015). In order to address appropriately to the study's research objectives, and, therefore, to evaluate both the type of techniques and the type of data analyzed, it was decided to combine the classification criterion just mentioned (e.g., relationship with the digital world) with another classification criterion of the techniques, that, as stated in Amaturo (2012), are divided according to the type of action taken to obtain the information. There are three types: interrogate; observe and read. In the first case the information is directly requested from the subjects involved (questionnaire, interviews, focus groups, etc.), in the second case the researcher can obtain information observing the interactions between subjects (participant observation, periscope observation, etc.). Finally, in the third case, the products of the actions, that can be read and analyzed (statistical sources, documentary analyzes), are taken into consideration.

As Table 4 shows, the classification of the strategies for the information collection shows a clear prevalence of the "questioning" category (71%) followed by the "reading" category (28%) and the "observing" category (<1%), that as the data show, except one time, it has never been used. It is important to highlight that in a small percentage of contributes (7%) more than one strategy for data collection has been used.

Table 4. Type of action taken to obtain the information.

Type	n	%	% of article
questioning	125	71	76
reading	50	28	30
observing	1	< 1	< 1
	176	100	

Regarding the kind of techniques, as reported in Table 5, within the "questioning" category there is a clear prevalence of the use of digitized techniques. In particular, the online focus group was used just once, the online interview twice and the online questionnaire in all remaining cases. In general, the online questionnaire has been used in the 70% of the analysed works. No digital techniques were detected for the collection of information through observation, while the traditional techniques are detected eight times and mostly correspond to "face to face" and "telephone" questionnaires. In general, classic techniques are present only in the 5% of articles.

Table 5. Type of technique in questioning category

Classic		**Digitized**		**Digital**
type	*n*	*type*	*N*	
face to face questionnaire	5 (3%)	online questionnaire	114 (70%)	
Telephone questionnaire	3 (2%)	online focus group	1 (<1%)	
		interview	2 (1%)	

Concerning the "reading" category (see Table 6), a good use of digital native data is detected. In fifty times in which "reading" occurred as adopted technique to retrieve information, 31 times digital native data were taken into account, while 19 times digitized data mostly coming from classical statistical sources (e.g., Istat, ministries, world bank, etc.). It is important to highlight, referring to digital data, that 15 times over 31 the digital data were "read" thanks to native digital techniques (e.g., Google Trends, or Youtube analytics) while in the remaining 15 the data were extracted from the web, and mainly from social platforms such as Facebook and Twitter, thanks to scraping tools.

Table 6. Type of data in "reading category"

Digitized data	**Digital Data**	
	direct	*undirect*
19 (12%)	16 (10%)	15 (9%)

Regarding the "observation" category, the technique used is an online one and it is the netnography. Thus, overall, it can be possible to observe that in the works analysed so far there is a clear prevalence of the use of digitized techniques (117) instead of digital ones (16) while as relating to the used sources, these are mostly digital (32) than digitized.

With reference to the big data issue, as can be seen from the table 7, most of the works (90%) is based on datasets containing a maximum of one thousand observations, only 1% of the works, indeed, is based on datasets that contain more than one million "lines". Therefore, most of the contributes drew its conclusions on information contained in matrixes with a not excessive number of rows. A large amount of data similar to what in the literature is defined as "big data" maybe has been processed by those works that involved digital methods such as Google Trends, but from those works it was possible only to identify the time range of observation and the number of keywords used, not the precise number of data analyzed. In the end it can be possible to assert that most of the works has not used large data sets.

Table 7. Number of observations used

	N.	*%*	*% cumul*
1-100	20	14	14
101-500	35	25	39
501-1000	49	34	73
1 K -10 K	25	18	91
10 K - 100 K	9	6	97
100 K -1000 K	3	2	98
> 1000 K	2	1	100
	143	100	

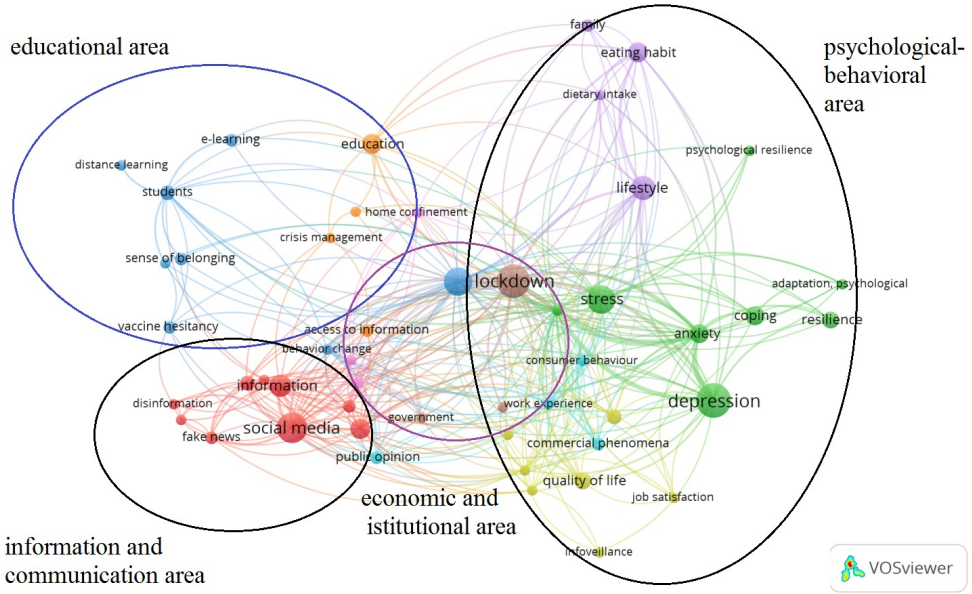
5. Discussion and research agenda

The SLR highlighted what are the different fields of society on which the Covid-19 pandemic impacted and how the social research investigated these phenomena and their consequences. Covid-19 pandemic, in fact, affected countries and people all over the world – and in particular Italy - in all their core activities, above all relational, economical, and educative. The content analysis performed suggested a series of theoretical-conceptual results through both the keywords' co-occurrence analysis (see Figure 2 and Table 3) and the collected data on research methods used. From the thematic analysis, nine clusters in which social research has been most focused could be identified. These clusters deal with different themes such as communication flows and fake news; pandemic's psychological effects and changes in the population; the impact on the educational world; social media both as the main discussion forum about pandemic-related topics and as a social infrastructure; lifestyle changes during lockdown (e.g., eating habits); consumer behavior, the role of institutions during the pandemic and the socio-economic effects this has had on the population. Considering these clusters as a starting point, four macro-areas had been derived then (see Figure 3): 1) information and communication area; 2) psychological and behavioural area; 3) economical-institutional area; 4) educational area.

As seen from the thematic analysis, many topics were analyzed using online survey as one of the main methods of study. The analysis of the techniques and tools showed that a substantial amount of these works used one of the classic tools of social research in his online version: the questionnaire. On one hand, this is a confirmation of its usefulness and flexibility, on the other hand, it is possible to identify at least two critical elements related to the use of this tool for online research. In the first place the type of information collected by the questionnaire - the main tool of standard social research - is already established upstream, deleting the possibility of identifying emerging or, however, not taken into consideration aspects of an analyzed phenomenon. Studies conducted on Twitter, for example, allow the researcher to broadly explore the semantic field of a given phenomenon. In second place, more generally, the default use of the questionnaire for online research precludes the possibility of exploiting the potential of the web, such as the possibility of studying subjects' behavior by collecting traces of their activities. Notwithstanding that the questionnaire is not an outdated or useless tool, because it often allows to obtain detailed information on a given phenomenon. However, it seems clear that in some cases it became possible to obtain information on a phe-

nomenon without resorting to intrusive techniques such as the questionnaire, but simply collecting data from the traces left by subjects during their activities on the web.

Figure 3. Clusters grouped into four macro-areas



This process would also allow the collection of a more consistent amount of data, while, as seen in the sample of articles analyzed, most of the works presented fairly limited datasets. Clearly big does not mean necessarily better, considering that one of the limitations of big data is related to the fact that their use does not allow any form of generalization, however it should be emphasized that also results deriving from data collected through an online questionnaire cannot allow any generalization. Large datasets are not the only potential that web research offers, for example the online experiment, as demonstrated in research, can be a successfully tool to study some phenomena. Combining different strategies for detecting information, indeed, could also be useful, as widely emphasized in the research that use mixed methods. In the study, just 7% of articles analyzed used multiple tools and sources. From the results of the analysis emerges that most research were conducted through the web and social networks, but considering the period, this was an expected result. An interesting result to discuss is the fact that between "digital" and "digitized" research there has been a strong propensity for the second type, taking into the account the fact that the analysis' results are certainly partial, since many papers that will be published shortly not have been considered. It can also be assumed that what emerged from the analysis is due to the fact that research conducted with digital tools needs more time to be concluded. On the other hand, it should be emphasized that, considering the unpredictability of the phenomenon, the importance of setting up research in a short time has led the research groups to apply and use known and well-tested tools. From this point of view, although used, the "digital" tools would not yet be part of the standard instrumentation of the research groups, not solely due to theoretical resistances towards digital but for the costs that the implementation of these particular ways of doing research might have. The new "degree", "masters" and "Ph.D" courses have trained social scien-

Culture e Studi del Sociale-CuSSoc, 2021, 6(1), Special, pp. 123-141
ISSN: 2531-3975

tists with remarkable technical skills, but like any other human enterprise (Lakatos, 1970) it can be difficult, at the moment, to implement new research procedures within groups. So, beyond the theoretical questions, doing deeply research on digital big data requires new expertise and it has a cost both in an economical terms and time. Cost that not every research group can or want to sustain.

Thus, both the thematic analysis and the analysis of the techniques and tools of the social research literature that focuses on Covid-19 pandemic led to the identification of some gaps and future research directions.

Research direction 1: Computational fact-checking
Being the first full-scale global pandemic of the digital age, Covid-19 has presented novel challenges to governments such as the spread of news and misinformation. The review shed light on the necessity of accurate information to avoid the spread of misinformation. Thus, the future research should focus on developing mechanisms of computational fact-checking for identifying the fake news topics that can arise during times of crisis, such as the Covid-19 pandemic. In a period of uncertainty and almost constantly changing news, in fact, fact-checking provides a service by clarifying the assessment of a claim and also delivering context and background information.

Research direction 2: Increase of multi- and inter-disciplinarity
The social research on Sars-Cov-2 through big data is dynamic and has developed along fragmented disciplinary lines. In particular, the review, through the analysis of existing works, has shown a need of interconnections among different fields of knowledge (Emani, Cullot & Nicolle, 2015), for better manage the topic by looking at the totality of the phenomenon.

Research direction 3: Increase use of geo-located data
Big data analytics definitely lead to valuable knowledge in many domains. The disaster management, for example, could benefit from it as there are many real-time information, which is available from official and non-official sources (e.g., wireless sensor networks, Internet-based systems, etc.). Thus, through the review emerges the need for more studies that investigate the potential of geo-tagged data in the disaster management (Chen et al., 2016; Chatfield and Brajawidagda, 2013).

Conclusion

Theoretical and Practical Implications

This paper provides a synthesis of the current conceptual and empirical literature on the disruptive phenomenon of Covid-19 pandemic in Italy. A first theoretical contribution of this study pertains to the identification of the current state of the art of the Covid-19 related topics, and the way these issues have been investigated by social scientists. Another implication of the review lies in the harmonisation of the existing knowledge in four macro-areas that underlines which are the main stressed topics in literature. Then, drawing the attention to one unique and valuable source like big data is, this paper tries to demonstrate the opportunities that they provide to social scientists, highlighting how these data and methods can be used to benefit social science research. However, the final conceptual contribution of the

paper rests in the identification of specific gaps in the Covid-19 pandemic literature that led to the description of three directions for future research in social sciences.

Regarding the practical implications, this study can contribute to a general understanding of the phenomenon for policymakers who wants to address better future emergencies, especially suggesting through the future research directions, what kind of methodologies should put in practise to avoid mistakes. For example, the review identified a major concern relates to Covid-19 pandemic into misinformation that can have drastic health care consequences, particularly the misinformation surrounding vaccines, is a significant obstacle in overcoming Covid-19 (Carrieri et al., 2019). So, an important role for public health organizations and governments in providing accessible online information and refutation of misinformation is highlighted. The presentation of accurate information is a moral imperative for scientists, public health authorities, and governments to safeguard the truth and maintain an accessible discourse with the public opinion to limit fear and to safeguard public safety. Therefore, in line with the future research directions highlighted above, policymakers should invest more in the development and integration of digited methods to exploit most of the cognitive power of big data in every aspect: from quality public information to disaster management and clearly to social research.

Limitations and Further Research

Despite the value of the findings presented here, the paper has some limitations. First, the results of this work, as mentioned before, need to be updated with works that will be published in the near future and it will be interesting to understand if the trend highlighted in this article could radically changes or not. Furthermore, as soon as a more consistent amount of works will be reached, it will be equally interesting to understand the ways in which the different areas within the social sciences approach to digital methods and big data and what kind of orientation will emerge.

References

Aiello F., Bonanno G. & Foglia F. (2020). On the choice of accommodation type at the time of Covid-19. Some evidence from the Italian tourism sector, *Current Issues in Tourism*. https://doi.org/10.1080/13683500.2020.1846504

Amaturo E. (2012). Metodologia della ricerca sociale, Utet

Aquilanti, L., Gallegati, S., Temperini, V., Ferrante, L., Skrami, E., Procaccini, M., & Rappelli, G. (2020). Italian Response to Coronavirus Pandemic in Dental Care Access: The DeCADEStudy.*International journal of environmental research and public health*, 17(19), 6977.https://doi.org/10.3390/ijerph17196977

Badell-Grau, R. A., Cuff, J. P., Kelly, B. P., Waller-Evans, H., & Lloyd-Evans, E. (2020). Investigating the Prevalence of Reactive Online Searching in the COVID-19 Pandemic: Infoveillance Study. *Journal of medical Internet research*, *22*(10), e19791.https://doi.org/10.2196/19791

Boccia Artieri, G.(2015). *Gli effetti sociali del web. Forme della comunicazione e metodologie della ricerca online*. Milano, Italia: Franco Angeli.

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F.,Pammolli, F. (2020). Economic and social consequences of human mobility restrictions under COVID-19. *Proceedings of the National Academy of Sciences*, *117*(27), 15530-15535. https://doi.org/10.1073/pnas.2007658117

Borges J., Hain H., Sudrich S., & Beigl M. (2017). "Event detection for smarter cities," *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Com-*

puting, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pp. 1-8. https://doi.org/10.1109/UIC-ATC.2017.8397590.

Briscese G., Lacetera N., Macis M., & Tonin M. (2020). Compliance with COVID-19 Social-Distancing Measures in Italy: The Role of Expectations and Duration. *Iza Discussion Paper No.13092,* available at SSRN: https://ssm.com/abstract=3568295

Caliandro, A., & Gandini, A. (2016). *Qualitative research in digital environments: A research toolkit*. Routledge.

Capone, V., Caso, D., Donizzetti, A. R., & Procentese, F. (2020). University student mental well-being during COVID-19 outbreak: What are the relationships between information seeking, perceived risk and personal resources related to the academic context? *Sustainability*, *12*(17), 7039. https://doi.org/10.3390/su12177039

Carrieri V, Madio L, & Principe F. (2019). Vaccine hesitancy and (fake) news: quasi-experimental evidence from Italy. Health Econ 2019 Nov; 28(11):1377-1382. https://doi.org/10.1002/hec.3937

Chatfield, A. T., & Brajawidagda, U. (2013). Twitter early tsunami warning system: A case study in Indonesia's natural disaster management. In 2013 46th Hawaii international conference on system sciences (pp. 2050-2060). IEEE. https://doi.org/ 10.1109/HICSS.2013.579

Chen, X., Elmes, G., Ye, X., & Chang, J. (2016). Implementing a real-time Twitter-based system for resource dispatch in disaster management. GeoJournal, 81(6), 863-873. https://doi.org/10.1007/s10708-016-9745-8

Cusinato, M., Iannattone, S., Spoto, A., Poli, M., Moretti, C., Gatta, M., &Miscioscia, M. (2020). Stress, Resilience, and Well-Being in Italian Children and Their Parents during the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health*, *17*(22). 8297.https://doi.org/10.3390/ijerph17228297

De Santis, E., Martino, A., & Rizzi, A. (2020). An Infoveillance System for Detecting and Tracking Relevant Topics From Italian Tweets During the COVID-19 Event. *IEEE Access*, *8*, 132527-132538. 10.1109/ACCESS.2020.3010033

Elias, P. (2012*)*. Big data and the social sciences: a perspective from the ESRC, presentation at the conference Shaping society.

Emani, C. K., Cullot, N., & Nicolle, C. (2015). Understandable big data: a survey. *Computer science review*, *17*, 70-81. https://doi.org/10.1016/j.cosrev.2015.05.002

Falcone, R., & Sapienza, A. (2020). How COVID-19 Changed the Information Needs of Italian Citizens. *International Journal of Environmental Research and Public Health*, *17*(19), 6988. https://doi.org/10.3390/ijerph17196988

Gaggi, O., Kolasinska, A. B., Mirri, S., &Prandi, C. (2020). The new classmate: an exploration of how CoVid-19 affected primary schools activities in Italy. In *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, 36-41, https://doi.org/10.1145/3411170.3411239

Germani, A.; Buratta, L.; Delvecchio, E.; Mazzeschi, C. (2020). Emerging Adults and COVID-19: The Role of Individualism-Collectivism on Perceived Risks and Psychological Maladjustment. *Int. J. Environ. Res. Public Health* 2020, *17*, 3497. https://doi.org/10.3390/ijerph17103497

Graffigna, G., Bosio, C., Savarese, M., Barello, M., & Barello, S. (2020). "# I-Am-Engaged": Conceptualization and First Implementation of a Multi-Actor Participatory, Co-designed Social Media Campaign to Raise Italians Citizens' Engagement in Preventing the Spread of COVID-19 Virus. *Frontiers in psychology, 11*, 2428. https://doi.org/10.3389/fpsyg.2020.567101

Halford, S., Pope, C., & Weal, M. (2013). Digital futures? Sociological challenges and opportunities in the emergent semantic web. *Sociology*, *47*(1), 173-189. https://doi.org/10.1177/0038038512453798

Hampton, K. N. (2017). Studying the digital: Directions and challenges for digital methods. Annual Review of Sociology, 43, 167-188. https://doi.org/10.1177/0038038512453798

Iacuzzi, S., Fedele, P., & Garlatti, A. (2020). Beyond Coronavirus: the role for knowledge management in schools responses to crisis. *Knowledge Management Research & Practice*, 1-6. https://doi.org/10.1080/14778238.2020.1838963

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. New York: Sage

Krippendorff, K. (2004). Reliability in content analysis. *Human communication research, 30*(3), 411-433.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. New York: Sage publications.

Lakatos, I. (1970). History of science and its rational reconstructions. In *PSA: Proceedings of the biennial meeting of the philosophy of science association* (Vol. 1970, pp. 91-136). D. Reidel Publishing.

Lenzo, V., Quattropani, M. C., Musetti, A., Zenesini, C., Freda, M. F., Lemmo, D., ... &Cattivelli, R. (2020). Resilience contributes to low emotional impact of the COVID-19 outbreak among the general population in Italy. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.576485

López Fernández, M. C., Serrano Bedia, A. M., & Pérez Pérez, M. (2016). Entrepreneurship and family firm research: A bibliometric analysis of an emerging field. *Journal of Small Business Management, 54*(2), 622-639. https://doi.org/10.1111/jsbm.12161

Losito, G. (1996). L'analisi del contenuto nella ricerca sociale (Vol. 1). Milano: FrancoAngeli.

Lupton D. (2015), *Digital sociology*, London, Routledge.

Mahrt, M., & Scharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, *57*(1), 20-33. https://doi.org/10.1080/08838151.2012.761700

Manesh, M. F., Pellegrini, M. M., Marzi, G., & Dabic, M. (2020). Knowledge management in the fourth industrial revolution: Mapping the literature and scoping future avenues. *IEEE Transactions on Engineering Management, 68*(1), 289-300. https://doi.org/10.1109/TEM.2019.2963489

Manovich, L. (2012). Trending: The Promises and the Challenges of Big Social Data. In Gold M. (Ed.), *Debates in the Digital Humanities* (pp. 460-475). Minneapolis; London: University of Minnesota Press.

Marres, N. (2012). The redistribution of methods: On intervention in digital social research, broadly conceived. *The Sociological Review*, *60*, 139165.https://doi.org/10.1111/j.1467-954X.2012.02121.x

Mauss M. (2002). Saggio sul dono. Torino: Einaudi.

Mellin J., & Berndtsson M. (2009). Event Detection. In: LIU L., ÖZSU M.T. (eds) Encyclopedia of Database Systems. *Springer*, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_506

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269. https://doi.org/10.1371/journal.pmed.1000097

Moscadelli, A., Albora, G., Biamonte, M. A., Giorgetti, D., Innocenzio, M., Paoli, S., ... &Bonaccorsi, G. (2020). Fake news and Covid-19 in Italy: Results of a quantitative observational study. *International Journal of Environmental Research and Public Health*, *17*(16), 5850. https://doi.org/10.3390/ijerph17165850

Pellegrini, M., Ponzo, V., Rosato, R., Scumaci, E., Goitre, I., Benso, A., ... &Broglio, F. (2020). Changes in weight and nutritional habits in adults with obesity during the "lockdown" period caused by the COVID-19 virus emergency. *Nutrients*, *12*(7), 2016. https://doi.org/10.3390/nu12072016

Pitrelli, N. (2017). Big data e metodi digitali per la ricerca in comunicazione della scienza: opportunità, sfide e limiti. *JCOM*, 16 (02), C01 it. https://doi.org/10.22323/2.16020301

Pohl, D., Bouchachia, A., & Hellwagner, H. (2012). Automatic sub-event detection in emergency management using social media. In *Proceedings of the 21st international conference on world wide web* (pp. 683-686). https://doi.org/10.1145/2187980.2188180

Procentese, F., Capone, V., Caso, D., Donizzetti, A. R., & Gatti, F. (2020). Academic Community in the Face of Emergency Situations: Sense of Responsible Togetherness and Sense of Belonging as Protective Factors against Academic Stress during COVID-19 Outbreak. *Sustainability*, *12*(22), 9718. https://doi.org/10.3390/su12229718

Rahman, M., Thill, J. C., & Paul, K. C. (2020). COVID-19 pandemic severity, lockdown regimes, and people's mobility: Early evidence from 88 countries. *Sustainability*, *12*(21), 9101. https://doi.org/10.3390/su12219101

Rice R. (1990). Computer-mediated communication system network data. *Int. J. Man-Mach. Stud.*, 32 (6), 627–647. https://doi.org/10.1016/S0020-7373(05)80104-X

Robinson, P., & Lowe, J. (2015). Literature reviews vs systematic reviews. *Australian and New Zealand journal of public health*, *39*(2), 103-103. https://doi.org/10.1111/1753-6405.12393

Rogers, R. (2009). *The end of the virtual: Digital methods* (Vol. 339). Amsterdam University Press.

Rogers R (2013). *Metodi digitali: fare ricerca sociale con il web.* Bologna: Il Mulino.

Rolandi, E., Vaccaro, R., Abbondanza, S., Casanova, G., Pettinato, L., Colombo, M., & Guaita, A. (2020). Loneliness and social engagement in older adults based in Lombardy during the COVID-19 lockdown: The long-term effects of a course on social networking sites use. *International journal of environmental research and public health*, *17*(21), 7912. https://doi.org/10.3390/ijerph17217912

Roma, P., Monaro, M., Colasanti, M., Ricci, E., Biondi, S., Di Domenico, A., ... & Mazza, C. (2020). A 2-Month Follow-Up Study of Psychological Distress among Italian People during the COVID-19 Lockdown. *International Journal of Environmental Research and Public Health*, *17*(21), 8180. https://doi.org/10.3390/ijerph17218180

Rositi, F. (1988). Analisi del contenuto. La ricerca sull'industria culturale, 59-94.

Ruiz-Roso, M. B., de CarvalhoPadilha, P., Mantilla-Escalante, D. C., Ulloa, N., Brun, P., Acevedo-Correa, D., ... & Carrasco-Marín, F. (2020). Covid-19 confinement and changes of adolescent's dietary trends in Italy, Spain, Chile, Colombia and Brazil. *Nutrients*, *12*(6), 1807.https://doi.org/10.3390/nu12061807

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management*, 14(3), 207-222.https://doi.org/10.1111/1467-8551.00375

Troise, C., O'Driscoll, A., Tani, M., & Prisco, A. (2020). Online food delivery services and behavioural intention–a test of an integrated TAM and TPB framework. *British Food Journal*.https://doi.org/10.1108/BFJ-05-2020-0418

Van Eck, N. J. V., &Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American society for information science and technology*, 60(8), 1635-1651. https://doi.org/10.1002/asi.21075

Van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*(2), 1053-1070.

Wang, H., Hovy, E., & Dredze, M. (2015). The hurricane sandy twitter corpus. In *Workshops at the twenty-ninth AAAI conference on artificial intelligence*.