Review

How can big data shape the field of non-religion studies? And why does it matter?

Dominik Balazka,[1,4,*] Dick Houtman,[2] and Bruno Lepri[3]
[1]University of Milan, Department of Social and Political Sciences, Milan, Lombardy 20122, Italy
[2]KU Leuven, Center for Sociological Research, Faculty of Social Sciences, Leuven, Flemish Brabant 3000, Belgium
[3]Bruno Kessler Foundation, Mobile and Social Computing Lab, Trento, Trentino-Alto Adige 38100, Italy
[4]University of Turin, Department. of Cultures, Politics and Society, Turin, Piedmont 10124, Italy
*Correspondence: dominik.balazka@unimi.it
https://doi.org/10.1016/j.patter.2021.100263

THE BIGGER PICTURE   It is becoming increasingly clear that secularity and non-religion are progressively turning into a relevant component of social life in both western and non-western countries and that they are connected with a variety of pressing issues such as civic engagement, human rights, and social integration. While traditional methodological approaches remain important to understand the varieties of non-religion, Big Data can significantly shape the way in which social scientists frame and analyze the puzzles within this emerging academic field. Nevertheless, large unstructured data collections and classification algorithms remain widely underused in sociology of religion, hindering its potential. We argue that to let Big Data in means to build an interconnected, interdisciplinary, and cooperative field situated at the intersection of non-religion studies and data science.

SUMMARY

The shift of attention from the decline of organized religion to the rise of post-Christian spiritualities, anti-religious positions, secularity, and religious indifference has coincided with the deconstruction of the binary distinction between "religion" and "non-religion"—initiated by spirituality studies throughout the 1980s and recently resumed by the emerging field of non-religion studies. The current state of cross-national surveys makes it difficult to address the new theoretical concerns due to (1) lack of theoretically relevant variables, (2) lack of longitudinal data to track historical changes in non-religious positions, and (3) difficulties in accessing small and/or hardly reachable sub-populations of religious nones. We explore how user profiling, text analytics, automatic image classification, and various research designs based on the integration of survey methods and big data can address these issues as well as shape non-religion studies, promote its institutionalization, stimulate interdisciplinary cooperation, and improve the understanding of non-religion by redefining current methodological practices.

INTRODUCTION

As Harvey Miller[1] wrote, "the data avalanche is here. Shouldn't we be digging?" At the turn of the century, big data came with a big bang providing massive amounts of underexplored or even unexplored data, opening new horizons and making possible innovative approaches to old puzzles. As a consequence, the scientific community seized the opportunity in multiple fields ranging from artificial intelligence[2] to medicine[3], from biology[4] to history,[5,6] and so on.

Despite the rise of computational social science,[7,8] several disciplines in sociology—such as cultural sociology or sociology of religion—showed a considerably lower propensity to take advantage of this new opportunity when compared with the so-called hard sciences or with more quantitatively oriented branches of

social sciences.[9,10] While certain areas of cultural studies recently displayed an increased interest in big data analytics,[11–13] computational approaches to religion remain sporadic. Recent developments in the sociology of religion, despite deeply grounded fascinations with quantitative methods,[14] have pushed multiple scholars into more qualitative directions.[15–17] The result is that only 0.06% of 139,368 papers about religion in Web of Science databases for the years 2012–2020 explicitly engage with or make use of big data. The situation is slightly better when Scopus is considered—0.31% of 138,785 papers—but either way the percentage of studies referencing big data remain considerably limited. While quantitative methods were presented as the "methodological requirements" of the requalification of cultural phenomena both inside and outside of sociology,[10] big data arguably represents the new frontier of neglected opportunities.
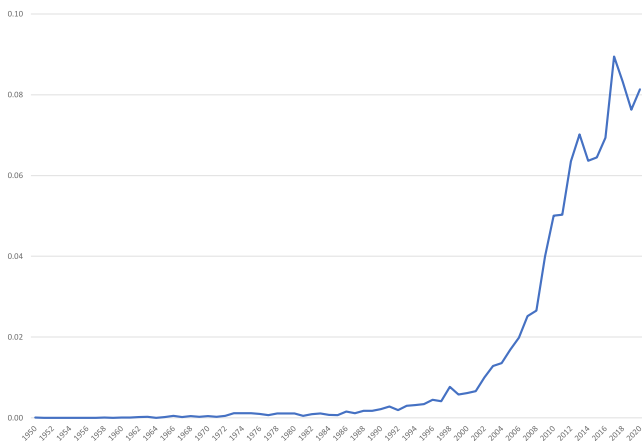
Figure 1. Number of scientific papers dealing with non-religion (expressed as proportion of total), Scopus 1950–2020 (N = 12,082)

Sociology of religion is as old as sociology itself. From Weber's[18,19] *Die Protestantische Ethik und der Geist des Kapitalismus* to Durkheim's[20] *Les Formes Élémentaires de la Vie Religieuse*, the history of sociology goes hand in hand with the study of religion. Following the increase of spirituality studies,[21,22] the binary distinction between religion and non-religion has lost much traction due to newly emerged research interests in opening up these categories to study (historical changes in) their composition empirically.[23] We argue that a better integration of big data in the research on non-religion can help to accelerate the institutionalization of non-religion studies, to promote interdisciplinarity, and to address and overcome some of the most pressing puzzles in the contemporary study of non-religion.

THE SPREAD OF NON-RELIGION STUDIES

Beyond secularization theory
Secularization theory, once the proud theoretical flagship of sociology of religion, is one of the few theories that ever attained "a truly paradigmatic status in the social sciences": "(It was) shared by all the founding fathers. Indeed, (…) everybody took it for granted."[24] This changed in the course of the 1980s, when the theory became a major target of critique,[25–27] even though various social-scientific students of religion continued to defend it.[28,29] Indeed, the hypothesis that more and more westerners are becoming less and less religious has meanwhile been strongly supported by empirical research.[30] This applies even to the United States, traditionally often invoked as a counterexample to the notion of religious decline in the West.[31]

It is as such clear that the shift away from secularization theory since the 1980s has not happened in response to its empirical refutation. It is indeed precisely the other way around. For the marked decrease in numbers of those who can plausibly be characterized as identifying with Christian religion, and the increasing numbers of non-religious westerners this has resulted in, have given rise to new research questions that go beyond the religion versus non-religion binary. What does it actually mean to identify as "non religious" nowadays,[23,32,33] not least in terms of understandings and evaluations of traditional Christian religion?[34] And what to make of apparently increased "spirituality talk"?[35]

Questions like these have meanwhile given rise to the new fields of "spirituality studies" and "non-religion studies," fields that both aim to go beyond the conceptual binary of religion versus non-religion to map "religion's other," as Smith and Cragun have put it.[33]

As to increased research interests in "spirituality," critics have identified secularization theory's notion of "religion" as in practice overly narrow, entailing basically not much more than organized religion. In doing so, they pointed out how post-Christian "self-religions" of the type that used to be called "New Age" have become more widespread in the West in precisely the same period during which Christian religion had declined,[22,36] even though without being able to compensate for the latter.[22,29] Self-religions of this type are rooted in western esotericism,[37] aim at overcoming religious and secular dualism,[38] and do as such entail a shift toward an eastern-style monistic worldview.[39] Key research questions within the new field of spirituality studies address what people actually mean when they self-identify as "spiritual but not religious,"[35,40,41] whether and how the composition of the category of religion has changed historically,[22,42–44] and whether or not spirituality entails more than a "fuzzy" step in a historical trajectory from religion to "non-religion."[45–48]

In short, widespread ambitions of moving beyond secularization theory have led to the emergence of the field of spirituality studies, which aims to open up the religion category for critical empirical scrutiny. More recently, the new field of non-religion studies has started to do the same with the other half of the conceptual binary on which secularization theory relied—i.e., the non-religion category.[34,49] The latter was traditionally studied in conjunction with faltering intergenerational transmission of churchgoing, informed by a conception of non-religion as a (indeed: one) residual category, diametrically opposed to religion and with its growth providing evidence in support of secularization theory.[50–52] The new field of non-religion studies moves beyond this by instead scrutinizing (changes in) the category's heterogeneous composition, with special attention to shifts between sub-categories, such as, for example, the spiritually inclined who dismiss Christian religion; the anti-religious who dismiss religion and spirituality alike; and the religiously indifferent who despite their own non-religiousness do not object to either religion or spirituality.[34,53] While distinctions like these are surely irrelevant in mapping the decline of Christian religion, they are absolutely vital in studying what does, and what does not, increasingly take over its former hegemony.

The new field of non-religion studies
Early pioneers in the study of non-religion started pointing out from the late 1960s onwards that "religious nones" entailed a "neglected category" with underexplored similarities with affiliated respondents[54] or a hardly homogeneous category in need of systematic empirical exploration.[55] Interestingly, these early attempts at launching non-religion as a vital research area have only very recently started to pay off.

The following data (see Figure 1) use the Scopus database to provide some aggregated summary statistics. Compared with Web of Science, Scopus offers a better coverage of publications within the field of social sciences and within the field of arts and humanities.[56] The search is based on title, abstract, and author's keywords. The query first searched for the words "religion,"

Table 1. Average percentage per year, percentage of total and numerosity of scientific papers dealing with non-religion by period, Scopus 1950–2020 (N = 12,082)

	1950–1972	1973–1989	1990–1999	2000–2006	2007–2011	2012–2020	Total
Average percentage/year	<0.1	0.1	0.4	1.2	3.8	7.4	1.4
Percentage of total	0.3	1.8	3.8	8.6	19.2	66.2	99.9
N	41	223	465	1,038	2,321	7,994	12,082

"religious," or "religiosity" in contributions published between 1950 and 2020. Successively the results were refined looking for the following words or expressions: irreligion, unbelief, "non religion," "non religious," "no religion," "religious nones," agnosti*, or athei*.

As shown in Table 1, between 1950 and 1972 non-religion was hardly discussed by scholars within the field, with only 0.3% of the observed 12,082 papers about the topic being published during this period. Between 1973 and 1989 the debate remained substantially stagnant. The average annual production increased from 0.015% to 0.1%—i.e., an increase by 6.7 times—but remained very modest. It was during the 1990s that papers of this kind slowly started to become more common. In fact, compared with 1950–1972, the average annual production increased by more than 25 times during this period. Between 2000 and 2006 the average annual production increased even further, reaching 1.2%, and resulting in 8.6% of the total in only 7 years. This means that roughly 85% of the total scientific production about non-religion in the Scopus database was published during the past 14 years. The average annual production grew steadily during the observed period, reaching 3.8% in 2007–2011 and 7.4% in 2012–2020. A total of 66.2% of the total scientific production about non-religion, nearly two-thirds of it, was therefore published throughout the past 9 years.

Five decades after Vernon[54] and Campbell,[55] and despite the evidence of an increased attention to non-religion, testified also by the establishment of specialized research programs, such as the "Programme for the study of religion and non-religion" of the London School of Economics and Political Science or "Understanding unbelief" of the University of Kent, Nikitaki[57] reaches a similar conclusion claiming that the focus on religious nones among scholars is still "virtually nonexistent."

By looking at the sheer growth of papers about non-religion without controlling for the general trend within religious studies, there is the risk to overestimate the scope of the phenomenon. In fact, while the papers about secularity and non-religion grew over the past decades as we previously showed, the same can be said about the remaining papers with a focus on religion. As shown in Figure 2, while the prevalence of non-religion studies increased by 10 times between 1950–1972 and 2016–2020—going from 0.6% to 6.0%—the papers about non-religion still remain a minority within the broader field of religious studies. Far from being a well-established academic and scientific reality supported by a solid majority of scholars, non-religion studies thus represent an emerging field. While Web of Science shows a temporal trend that is similar to the one observed in Figure 1, compared with Scopus it tends to underestimate the share of papers dealing with non-religion.

Besides the debate about secularity and secularization discussed above, at least three other elements have arguably

contributed to the growing interest in non-religion[58]: (1) the rapid growth of the population with no religious preference in highly influential countries from Europe, Australasia, Americas, and East Asia observed over the past four decades; (2) the increased acknowledgment of consistent intra-group differences of this supposedly homogeneous residual category; (3) the establishment of a normative framework sensitive to the freedom of religious or *belief* actors accompanied by the institution and recognition of associations representing religious nones—e.g., the European Humanist Federation or the American Humanist Association.

STUDYING NON-RELIGION WITH SMALL DATA

Intellectual progress in the newly emerged research field of non-religion studies is slowed down by the lack of relevant and internationally comparative survey data. This is because the large survey programs that could in principle provide such data (e.g., the European Values Study, the World Values Survey, the European Social Survey, and the International Social Survey Programme) do still rely heavily on questionnaires informed by secularization theory. The resulting data are as such perfectly useful to map the decline of Christian religion, and to some extent changes within Christianity itself,[42,44] but they are of much less use for the study of religious change more broadly conceived,[59] especially for historical changes in understandings of Christian religion among the waxing numbers of non-religious and for the spread of spirituality. While useful questions that are repeated across multiple waves of data collection are not completely absent, to be sure, their number is limited and by far precludes the levels of detail and precision that have traditionally been attained for Christian religion. This leaves researchers interested in the dissemination of the variegated renditions of non-religion largely empty handed, facing up to a non-religion category that has considerably increased in time, yet at the same time remains much of a black box. While the resulting "small data" do surely provide a picture of various characteristics typically associated with this black box, the collection and analysis of big data offer a major promise in overcoming the current data problems.

Nones are more frequently males than females, tend to be more concentrated in densely populated areas and to be, on average, younger than affiliated subjects.[58,60,61] While education is traditionally positively correlated with non-religious preferences,[62] according to Voas[52] this trend is progressively reversing among younger cohorts in the United Kingdom. Gender, age, education, and geographical area are all typically associated with non-affiliation. Nevertheless, the predictive power of these socio-demographic characteristics was questioned,[63,64] suggesting that the degree of worldview pluralism
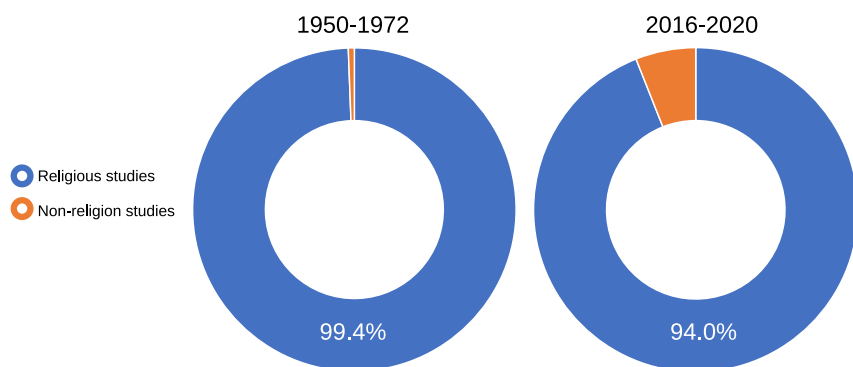
Figure 2. Prevalence of non-religion studies, a comparison of 1950–1972 (N = 7,093) and 2016–2020 (N = 81,176), Scopus

of local networks might play a fundamental role in the process of religious non-affiliation and disaffiliation.[63]

While the vast majority of social surveys do not allow respondents to elaborate further on their lack of religious affiliation, or distinguish only between atheism and agnosticism, recent research showed that generic non-religious labels hide a rich variety of internal differences ranging from new atheism to atheism plus, humanism, religious indifference, secularism, and so on.[65–68] A new comparative study of Brazil, China, Denmark, Japan, the United Kingdom, and the United States showed that only a minority of nones describe themselves as an "atheist" or "agnostic," frequently preferring other popular labels, such as "humanist," "free-thinker," "skeptic," "secular," etc.[69] The study also pointed out that widespread theoretical assumptions about atheists and non-believers in general, such as the strong dogmatic conviction of self-assessed atheists or the lack of supernatural beliefs among nones, are frequently violated at an empirical level.
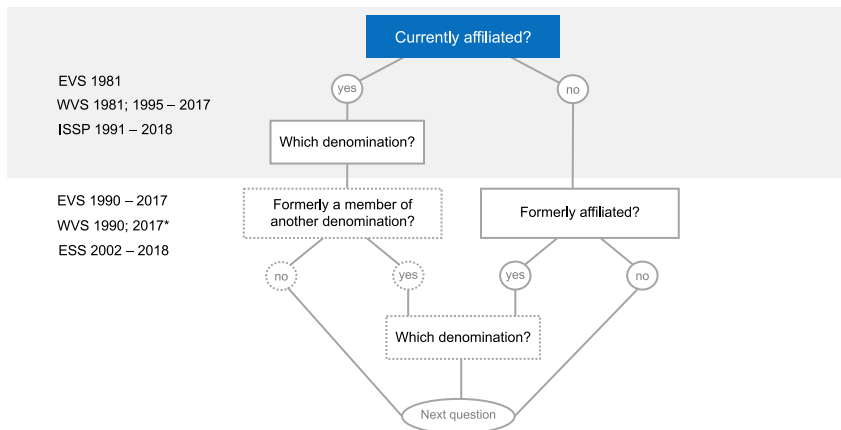
Besides the variety of strictly non- and/or anti-religious positions, it was also evidenced that non-affiliation is not necessarily a matter of non-religiosity. According to a research conducted by Lindeman and Lipsanen,[70] 25% of subjects with a high score on the Supernatural Belief Scale[71] are religious nones. Frequently addressed as "spiritual seekers" or "unchurched believers,"[35,72] these peculiar sub-categories of nones invite the research community to see religiosity beyond affiliation and to re-discuss the way secularity is both theorized and operationalized to explore the conceptual and empirical implications of the relocation of the sacred. Similar results were attained by other studies[72,73] that found significant differences in religiosity among nones using panel data about changes in the affiliation to a religious denomination as a discriminatory variable. The authors distinguished between stable religious affiliates (respondents who were affiliated at both of the considered time points), liminal nones (respondents whose affiliation changed during the observed period), and secular nones (respondents who were unaffiliated at both of the considered time points). Liminal nones tend to waver between religion and non-religion and were found to be, on average, more religious than their secular counterparts.[72,73] However, the application of this distinction in cross-sectional studies that researchers commonly use is problematic because the resulting classification strategy is not sufficiently fluid. As shown in Figure 3, the questionnaires of major international social surveys do not tell much about the meaning of non-affiliation and offer limited information about conversion

and deconversion trajectories. A considerable number of these surveys (top half of the figure) register only the current state—usually relying on general labels, such as "none," "no religion," or "none of the above"—without follow-up questions other than affiliation at the age of 12. The European Social Survey and more recent waves of the European Values Study (bottom half of the figure) represent an exception in this sense. Despite minor variations, both surveys ask nones whether they formerly belonged to a religious denomination. In 2008, affiliated European Values Study respondents were also asked whether they previously belonged to a different denomination. Starting from 2017, the European Values Study no longer asks which religious denomination disaffiliated nones formerly belonged to. In Figure 3, the minor coverage of these questions is represented by dashed contours. Affiliates are typically not asked whether they used to be non-religious at some point of their life, a characteristic that would make them liminal nones rather than stable affiliates. Moreover, stable non-affiliation does not necessarily exclude spirituality, whether inspired by religious traditions[42] or secular alternatives.[74,75]

Currently, the field of non-religion studies is facing at least three major obstacles. First, there is the well-known overabundance of labels employed to address nones and non-religion in the scientific literature.[76,77] It is important to contain the creative impulse of social researchers in favor of a systematic reorganization of the existing corpus of knowledge. Such an effort is crucial for a sustainable interdisciplinary cooperation and constitutes the prerequisite of clear, explicit, and testable theoretical statements. Nevertheless, the terminological issue is not just a matter of theory. Wording matters, which leads to the second issue faced by social researchers. As testified by the case of the UK census of 2011—which replaced the term "none" with "no religion"—different labels can significantly affect survey results.[78,79] To develop alternative classification strategies of religious non-affiliation it is thus important to go beyond binary choices resting on a misleading understanding of the no religious preference[35,80] and to study the effect of different labels on respondents' (non-)religious self-identification. The third obstacle concerns the availability of suitable data to test emerging hypotheses about non-religiosity, individualized religion or New Age spiritualities. Not only the questionnaires of international surveys are considerably biased toward traditional western religiosity,[81] frequently neglecting the variety of eastern beliefs and their implications for individualized religion, post-Christian spirituality and religious bricolage,[82] but they rarely cover supernatural, non-religious, or anti-religious preferences.

Survey methods can be improved and new variables included to cover neglected aspects of the contemporary religious and non-religious landscape. It will be costly and time consuming, but several scholars are heading in this direction, as testified

EVS 1981
WVS 1981; 1995 – 2017
ISSP 1991 – 2018

EVS 1990 – 2017
WVS 1990; 2017*
ESS 2002 – 2018

Figure 3. The logic of religious (non-) affiliation in the master questionnaires of the European Social Survey (ESS), the European Values Study (EVS), the International Social Survey Programme (ISSP), and the World Values Survey (WVS)
*Between 2017 and 2020 EVS and WVS jointly collected data in several European countries following the Memorandum of Understandings.

by the widespread critique of the predominant survey model,[59,67,78,79,80,83,84] by the arrival of dedicated surveys,[69,85,86] by the development of new measurement methods,[47,87] or by the progressive institutionalization of non-religion studies.[88] This will significantly improve researchers' capacity to analyze the multifaceted reality of non-affiliation today, but it will leave their ability to explore past historical trends—which is a crucial aspect in the debate about secularization—largely unaffected.

Another potential limit is represented by the numerosity of these groups. Religious nones grow in numbers and constitute a stable majority in several countries, such as the Czech Republic, France, Great Britain, Hungary, Japan, or the Netherlands, but once the scientific community opens the Pandora's box of religious non-affiliation to have a closer look, this solid majority will be fragmented into a multitude of smaller realities. Some of them will be big enough to avoid large confidence intervals, but other groups will be considerably smaller or just difficult to reach with small N studies.

In 1981, when the European Values Study and the World Values Survey started their very first fieldwork, the percentage of nones in participating countries was on average 11.3%. Today the most recent data of the same two surveys reveal that the overall percentage of nones has doubled, and that it tripled in Europe. As of 2020, nones represent the new majority in the Czech Republic, Estonia, France, Hong Kong, Hungary, Japan, Macau, the Netherlands, Great Britain, South Korea, and Vietnam, ranging between 54.1% and 86.8%. Australia, Belgium, Colombia, New Zealand, and the United States are already approaching the 50% threshold, while several other countries are heading in the same direction.[58] In the meanwhile, researchers became increasingly aware of relevant within-group differences in terms of, for example, beliefs,[69,72,73] civic engagement,[89] or political orientation,[90] recognizing secularity as a socially relevant issue. We argue that, by letting big data in, non-religion studies scholars can exploit a variety of novel sources of data to improve coverage (in both time and space), granularity, and predictive power while containing costs. Furthermore, a major opening to quantitative methods and big data analytics will positively affect also the interdisciplinarity of this emerging field and speed up its institutionalization.

STUDYING NON-RELIGION WITH BIG DATA

In recent years, the discussion about big data kept rapidly growing, not without a considerable level of hype among both scholars and business professionals. According to a popular—and strongly criticized[91–93]—narrative, big data came to revolutionize everything making previous science obsolete.[94,95] More recently, rather than suggesting replacing small data with big ones, researchers argued that integration between these two methodologies or "cultures of modelling"[96] is both possible and necessary.

In the following, the promises, limitations, and potential pitfalls for non-religion studies of user profiling, text analytics, automatic image classification, and various research methods mixing small and big data will be considered to discuss some of the most promising routes made available to scholars by recent developments in data analytics.

User profiling

User profiling leverages predictive analytics and machine learning to infer a variety of individual-level information, ranging from personality traits to political preferences, religion, and behavioral habits. Whether based on explicit and/or implicit techniques, user profiling is a grounding element of modern service personalization.[97] Following the large-scale diffusion of smartphone devices, a rich variety of user-generated data were employed for this purpose: call logs, messaging logs, lists of installed apps, app usage patterns, and other mobile phone data.[98–101] The usage of social media (e.g., Facebook, Twitter, Instagram, etc.), the browsing activity and website navigation, credit card transactions, etc., constitute other examples of data frequently used in user profiling.[102–104]

For example, Nguyen and Lim[105] have employed Support-Vector Machine classifiers to predict religion labels of users based on microblogging data collected from Twitter users in Singapore. Given the small numbers of self-declared Buddhist, Taoist, and Hinduist users, the authors focused on Christian and Muslim respondents, showing that the proposed classification strategy can accurately predict religious affiliation by combining textual features and social links.

Gathering labeled data to create training samples is not an easy task, but gathering labeled data about users' religious non-affiliation is arguably even trickier. An approach called "co-labeling" was developed to deal with ambiguous problems where labels of users in the training sample are uncertain.[106,107] Co-labeling is basically a multi-view learning method that combines classifiers trained on different views to improve the

effectiveness of unlabeled data by allowing different classifiers to inform each other.

Privacy concerns[108] and the effective availability of usable data[91] to the researchers are potential obstacles to a research design relying on co-labeling to develop an efficient classification strategy of religious nones. Regarding this, further limitations, at least from a sociological standpoint, are certainly represented by potential sampling/selection bias due to the access to limited segments of the population (e.g., the ones present and active on social media platforms, the owners of smartphones) as well as the relatively limited time span accessible through this particular approach. User profiling relies on large-scale diffusion of big data and does not provide means to reconstruct detailed historical trends for the past century.[109] Nevertheless, co-labeling constitutes a powerful technique that uses massive amounts of data generated by users on a daily basis that could be used to train and employ dynamic classification strategies of users' latent attributes, such as religious non-affiliation. This can greatly improve the knowledge of the varieties of non-religion—effectively assisting researchers in their effort to deconstruct the black box of non-religion—as well as the access to outliers and minorities,[110] minimizing uncertainty and avoiding large confidence intervals. While the latter categories can be studied through a variety of ethnographic and qualitative techniques integrating the current theoretical framework, an approach based on big data can significantly accelerate the process at the international level. Furthermore, the results can be used in correlational studies of institutional trust or political participation, and to refine classification strategies employed by international social surveys.

Text analytics

At least since the late 1970s, researchers' toolkits contain rudimentary text-mining techniques to parse unstructured textual data and identify patterns.[111] Nevertheless, with the improvement of natural language processing and machine learning techniques, and with the large-scale diffusion of unstructured data, text mining underwent a rapid development in computational sociology and digital humanities. Specifically, topic modeling and sentiment analysis are of particular interest for non-religion studies scholars.

Topic modeling typically uses Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA), or Probabilistic Latent Semantic Analysis (PLSA) to analyze large and unstructured collections of textual documents in a longitudinal and/or comparative perspective to identify recurrent topics and latent semantic structures.[112,113] While a classical clustering algorithm returns one label to characterize a document (e.g., k-means clustering), with topic modeling it is possible to analyze a probabilistic composition of the text. Sentiment analysis, on the other side, is constituted by a series of techniques—based on natural language processing—used to study affective states, opinions, and other subjective information conveyed by a text.[114,115] While automated approaches to text still struggle to adequately frame sarcasm and irony,[116] posing relevant methodological concerns, topic modeling and sentiment analysis can be combined[117] to compare and contrast the characteristics of non-religious discourses in time and space.

While big data may be young, meaning that it is hard to provide the analysis with a long-term longitudinal dimension, "big data of

the past"[118] are potentially as old as humanity itself and offer a precious source of under-analyzed information. For example, magazines published periodically and explicitly intended for an atheist, humanist, and/or secular audience cover over a century of modern and contemporary history. Among others: *New Humanist Magazine* (United Kingdom, 1885–current), *The Humanist Magazine* (United States, 1941–current), *The Skeptic* (Australia, 1981–current), *Les Cahiers Rationalistes* (France, 1930–current), *Espace de libertés* (Belgium, 1980–current), *L'ateo* (Italy, 1996–current—renamed as *Nessun Dogma:Agire laico per un mondo più umano* in 2020).

The main advantage of this approach is the availability of rich and long-term historical data that non-religion studies are currently missing: it combines nominally available data with recent advancements in text analytics to extrapolate new value from old information by analyzing the changes over time in non-religious discourses, but also national specificities and international differences. In this sense, an integration of big data and modern computational techniques in current research practices might be the only way for social scientists to historically trace the development of varieties of non-religiosity.

However, the application of text analytics to secular magazines may imply careful and time-consuming archival work to digitize older issues and prepare a dataset for the analysis. The necessary pre-processing can be reduced by reorienting the focus on readily available material (e.g., tweets, Reddit comments, dedicated forums), but this will also considerably decrease the time window of the analysis undermining one of the main advantages of this particular approach for non-religion studies. Books on atheism, secularity, or non-religion represent another alternative source of information. Nevertheless, compared with atheist magazines, books constitute a very heterogeneous group of materials—ranging from philosophical treatises to essays, science fiction, commentaries, and more—which are not always easily comparable. For this reason, and because of the temporal regularity of publications, magazines arguably represent a better proxy of non-religious discourses in time.

A second limitation is imposed by the very nature of books and magazines, which implies a form of institutionalization of the non-religious position through the editorial board that might or might not reflect accurately the social reality, thus making marginal, indifferent, and/or "liminal" positions[73] particularly difficult to observe with this approach. Once again, shifting attention to other sources, such as Twitter or Reddit, represents a valid alternative that solves this issue but poses its own problems—namely, the risk of exchanging highly active and vocal minorities for something much bigger.

Multipurpose international social surveys started to appear in the early 1980s, but were initially confined prevalently to Northern America and Western Europe. It was only a decade later that these survey programs started to improve their reach in other geographical areas, but even leaving aside persisting gaps in Asian and African data, two problems remain: (1) harmonization of different waves of the same survey or of different surveys is a problematic process.[119] In this sense, increasing the time horizon or extending the observation to a larger group of countries typically comes at the cost of a considerable decrease of comparable variables; (2) and, more importantly,

questionnaires of large survey programs remain theoretically informed by secularization thesis and keep struggling with contemporary religious and non-religious landscapes.[59] While the field of data harmonization is a rapidly expanding sector significantly improving the quality of newly collected data, new surveys, such as Understanding Unbelief,[69] the Secular Voices Survey,[85] or the Secular Communities Survey[86] tackle the latter issue. Nevertheless, future waves of dedicated surveys remain uncertain, coverage of countries is still considerably limited and data are currently not public. Most importantly, in both cases the quality of existing longitudinal data is suboptimal for the study of non-religion. Once magazines and other historical sources are fully recognized as data, it becomes possible to considerably extend the accessible time horizon without necessarily sacrificing the depth of the analysis. Potential applications of text analytics ranges from longitudinal designs of the kind described above to comparative studies of local groups or meta analyses of the scientific community studying non-religion.[120] Due to the massive increase of availability of unstructured textual data observed throughout the past two decades, text mining represents an important tool in the future study of non-religion.

Automatic image classification

While text clearly constitutes a relevant part of the unstructured data available today, images are another precious source of information. The use of visual methods in social sciences is hardly new. In 1998, Homer presented the "visual culture" as a new paradigm[121] in research. Over the years multiple studies applied visual methods, for example: Peraica studied contemporary forms of self-representation analyzing selfies,[122] Arbulla and Bucchi focused on warning messages posted in public spaces,[123] Vindrola-Padros used drawings to study the daily life of children in Argentina,[124] and Eide analyzed preschool Christmas holiday specials to understand the depiction of religious holiday practices in the United States.[125] What all these very different studies have in common is a qualitative approach to the analysis of visual data. While qualitative approaches to visual culture are still more appealing to social scientists, the increased availability of large collections of visual data and the development of innovative artificial intelligence approaches and tools force the scientific community to consider possible strategies to implement automated methods in the research process.

So how can computational science help researchers interested in the analysis of images? Arguably, at least two general strategies can be defined. The first strategy is to use automatic image classification to reduce the complexity of large data collections by extracting a small but highly consistent cluster of visual objects with similar characteristics to analyze with qualitative techniques. The researcher might want to distinguish, for example, between photographs and graphics,[126] between scenes occurring indoor and outdoor,[127] between portraits of men and women,[128] between images with and without people, or between photos of individuals and groups.[129] In addition, following the contribution of computer vision to the analysis of artworks, a weakly supervised object detection approach could be implemented to identify, distinguish, and classify various iconographic elements[130,131] observable in the considered images. In principle, a researcher can apply multiple criteria to progressively reduce the number of cases into a sub-set of theoretically relevant, conceptually consistent, and qualitatively manageable objects. Alternatively, if the resulting selection is still too big, sampling methods can be used to further reduce the number of observations. The second general strategy uses similar clustering techniques and the available metadata but, instead of extracting a specific cluster of images to analyze qualitatively, it focuses on the totality of the resulting clusters—or on a selection of relevant clusters—from a quantitative standpoint. Projects of this kind can analyze visual signatures of geographical areas exploring spatial and temporal visualizations of images,[132] study the physicality and visual content of cultural production,[133] or explore identity and self-representations of users on social media.[134,135]

The choice of the strategy will depend on available resources and on the research question, but the major limitation of this approach is arguably represented by the limited availability of large thematic collections about non-religion. So what kind of visual data might non-religion studies' scholars be interested in? In the previous section we considered the use of textual data from a variety of sources, such as secular popular press, books, or dedicated forums. However, these sources are usually not limited to textual data. The covers of books and magazines, as well as the visual content that they convey, can become a relevant part of the analysis of visual culture in organized non-religious groups. Communication strategies of non-religious communities could be analyzed using the visual content related to campaigns of various secular associations around the globe—such as the recent campaign of the *UK Humanist* asking non-religious people to answer "no religion" in the upcoming British census or the campaign of the Italian Union of Rationalist Agnostics and Atheists for the legalization of debaptism. Discussions generated on dedicated forums, such as "atheist forums," "think humanism," or "the agnostic forum" are frequently accompanied by the production and/or sharing of graphical objects, for example, memes, that can be used to study secular representations of religious others. Conversely, religious representations of atheism or secularity could be approached in a similar way. A final example of potentially interesting sources of visual data about non-religion is represented by social media. In particular, Twitter and Instagram could be used to analyze co-occurrences of popular non-religious hashtags and characteristics of visual objects associated with them. In principle, this could be extended to religious communities in a comparative analysis of religious, non-religious, and anti-religious self-representations.

Mixing small and big data

Mixing small and big data, rather than a specific line of investigation, constitutes an umbrella of possible approaches based on the manipulation of the observed situation and/or on the possibility of integrating big data with survey methods or various qualitative techniques directly.[136]

The relevance of mobile applications for user profiling was already discussed above. While mobile data pose considerable difficulties in terms of access, scientists around the world have started to design applications to combine mobile data with survey methods to address research topics ranging from time allocation[137] to academic performance,[138,139] behavioral

habits,[98,101,140] or mental well-being.[139,141] The possibility of designing mobile apps to collect relevant information about non-religion and to cross-reference it with survey data to analyze non-religious preferences, behavioral patterns, and more, offers multiple advantages and possibilities.

The current attempt to redefine survey methods in non-religion studies is accompanied by severe coverage issues, with relevant data being collected for only a minority of countries at a single time point and being still mostly unavailable to the wider scientific community. An increased usage of big data offers the possibility to fill the existing gap of relevant information considerably faster and at a lower cost. While the belief that big data is replacing rather than integrating survey methods still permeates computational sciences,[95,142] it is increasingly more and more common to encounter examples of research relying on a combination of big and small data.[98,143] The resulting reflections on commonalities and potential integrations[96,144] of heterogeneous data sources—whether structured, semi-structured, or unstructured—arguably represent one of the major challenges in the contemporary field of data harmonization.[145,146]

Merging public, private, and governmental sources of data to produce a multidimensional dataset is also another way to approach otherwise hardly accessible clusters of population, such as, for example, people with a vegetarian lifestyle.[143] The decision tree learning algorithm employed by Lusk[143] offers additional insights and can outperform logit and regression models by fully exploiting the size and complexity of big data. This research design thus overcomes the problem of small or hardly accessible groups[110] frequently experienced in non-religion studies and might be useful to develop a classification tree of religious nones and to monitor the impact of (non-)religious preferences on their decision making.

FINAL CONSIDERATIONS

Religious nones and non-religion play a fundamental role in sociology of religion and in the emerging field of non-religion studies. A big data-based approach to the topic can greatly improve our understanding of the phenomenon in contemporary societies, of its evolution over time, and of relevant cross-national specificities and differences. This is certainly a major point of concern for sociology of religion and non-religion studies—which can in turn accelerate its institutionalization within sociology of religion and further support its emerging theoretical interests through an increased big data usage. Nevertheless, the implications of an improvement in the scientific study of non-religion go beyond sector-specific theoretical interests: social and institutional trust, political participation, political orientation, civic engagement, business strategies, palliative care, general patient care, and social integration[89,90,147–154] are, in fact, some of the research areas frequently associated with religious (non)affiliation. These fields can thus greatly benefit from an improvement in the parallel field of non-religion studies by further exploring empirical implications of a better understanding of religious nones.

The proposed approaches tackle some of the more pressing issues currently faced by scholars in this emerging field: development and implementation of alternative classification strategies of non-religion; lack of coverage of relevant variables in international surveys; lack of detailed longitudinal data; and

access to small and/or hardly reachable sub-populations of religious nones.

The prefigured outcome, as well as the proposed methodological strategies, presupposes analytical approaches that go well beyond the currently dominating techniques in sociology of religion. As such, the success of the operation might require the establishment of an interconnected, cooperative, and interdisciplinary sociology of non-religion integrated not just within the broader context of sociology of religion, but also in the rapidly expanding fields of computational social sciences, digital humanities, data visualization, and data harmonization.

REFERENCES

1. Miller, H.J. (2010). The data avalanche is here. Shouldn't we be digging? J. Reg. Sci. *50*, 181–201. https://doi.org/10.1111/j.1467-9787.2009.00641.x.

2. Duan, Y., Edwards, J.S., and Dwivedi, Y.K. (2019). Artificial intelligence for decision making in the era of big data—evolution, challenges and research agenda. Int. J. Inf. Manag. *48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

3. Krumholz, H.M. (2014). Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. Health Aff. *33*, 1163–1170. https://doi.org/10.1377/hlthaff.2014.0053.

4. Leonelli, S. (2014). What difference does quantity make? On the epistemology of big data in biology. Big Data Soc. *1*, 1–11. https://doi.org/10.1177/2F2053951714534395.

5. Barron, A.T.J., Huang, J., Spang, R.L., and DeDeo, S. (2018). Individuals, institutions, and innovation in the debates of the French revolution. Proc. Natl. Acad. Sci. U S A *115*, 4607–4612. https://doi.org/10.1073/pnas.1717729115.

6. Whitehouse, H., Francois, P., Savage, P.E., Currie, T.E., Feeney, K.C., Cioni, E., Purcell, R., Ross, R.M., Larson, J., Baines, J., et al. (2019). Complex societies precede moralizing gods throughout world history. Nature *568*, 226–229. https://doi.org/10.1038/s41586-019-1043-4.

7. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barbási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., et al. (2009). Computational social science. Science *323*, 721–722. https://doi.org/10.1126/science.1167742.

8. Lazer, D., Pentland, A., Watts, D.J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., et al. (2020). Computational social science: obstacles and opportunities. Science *369*, 1060–1062. https://doi.org/10.1126/science.aaz8170.

9. McFarland, D.A., Lewis, K., and Goldberg, A. (2016). Sociology in the era of big data: the ascent of forensic social science. Am. Sociolog. *47*, 12–35. https://doi.org/10.1007/s12108-015-9291-8.

10. Houtman, D., and Achterberg, P. (2016). Quantitative analysis in cultural sociology: why it should be done, how it can be done. In Sage Handbook of Cultural Sociology, D. Inglis and A.-M. Almia, eds. (Sage), pp. 225–236.

11. Hall, G. (2013). Toward a postdigital humanities: cultural analytics and the computational turn to data-driven scholarship. Am. Lit. *85*, 781–809. https://doi.org/10.1215/00029831-2367337.

12. Manovich, L. (2016). The science of culture? Social computing, digital humanities and cultural analytics. J. Cult. Analyt. https://doi.org/10.22148/16.004.

13. Piccialli, F., Benedusi, P., Carratore, L., and Colecchia, G. (2020). An IoT data analytics approach for cultural heritage. Personal. Ubiquitous Comput. *24*, 429–436. https://doi.org/10.1007/s00779-019-01323-z.

14. Buehler, C., Hesser, G., and Weigert, A. (1972). A study of articles on religion in major sociology journals: some preliminary findings. J. Sci. Study Religion *11*, 165–170. https://doi.org/10.2307/1384929.

15. Davidsen, M.A. (2012). Future directions in the sociology of non-institutional religion. Implicit Religion *15*, 553–570. https://doi.org/10.1558/imre.v15i4.553.

16. Tomlins, S. (2018). Atheism and religious nones: an introduction to the study of nonreligion in Canada. In Exploring Religion and Diversity in Canada: People, Practice and Possibility, C. Holtmann, ed. (Springer International), pp. 237–254.

17. Travers, M., and Ezzy, D. (2019). Interpretative issues in researching law and religion. In Research Handbook on Interdisciplinary Approaches to Law and Religion, R. Sandberg, N. Doe, B. Kane, and C. Roberts, eds. (Edward Elgar Publishing), pp. 207–220.

18. Weber, M. (1904). Die Protestantische Ethik und der Geist des Kapitalismus. Archiv für Sozialwissenschaft und Sozialpolitik *20*, 1–54, n2:0174-819X.

19. Weber, M. (1905). Die Protestantische Ethik und der Geist des Kapitalismus. Archiv für Sozialwissenschaft und Sozialpolitik *21*, 1–110, n2:0174-819X.

20. Durkheim, É. (1912). Les Formes Élémentaires de la Vie Religieuse (Alcan).

21. Heelas, P., and Woodhead, L. (2005). The Spiritual Revolution: Why Religion Is Giving Way to Spirituality (Blackwell).

22. Houtman, D., and Mascini, P. (2002). Why do churches become empty, while new age grows? Secularization and religious change in The Netherlands. J. Scientific Study Religion *41*, 455–473. https://doi.org/10.1111/1468-5906.00130.

23. Bullivant, S., and Lee, L. (2012). Interdisciplinary studies of non-religion and secularity: the state of the union. J. Contemp. Religion *27*, 19–27. https://doi.org/10.1080/13537903.2012.642707.

24. Casanova, J. (1994). Public Religions in the Modern World (University of Chicago Press).

25. Hadden, J.K. (1987). Towards desacralizing secularization theory. Social Forces *65*, 587–610. https://doi.org/10.2307/2578520.

26. Stark, R., and Bainbridge, W.S. (1986). The Future of Religion. Secularization, Revival and Cult Formation (University of California Press).

27. Swatos, W.H., Jr., and Christiano, K.J. (1999). Secularization theory: the course of a concept. Sociol. Religion *60*, 209–228. https://doi.org/10.2307/3711934.

28. Bruce, S. (2002). God Is Dead: Secularization in the West (Blackwell).

29. Bruce, S. (2017). Secular Beats Spiritual: The Westernization of the Easternization of the West (Oxford University Press).

30. Norris, P., and Inglehart, R. (2004). Sacred and Secular: Religion and Politics Worldwide (Cambridge University Press).

31. Voas, D., and Chaves, M. (2016). Is the United States a counterexample to the secularization thesis? Am. J. Sociol. *121*, 1517–1556. https://doi.org/10.1086/684202.

32. Thiessen, J., and Wilkins-Laflamme, S. (2020). None of the above: Nonreligious Identity in the US and Canada (New York University Press).

33. Smith, J.M., and Cragun, R.T. (2019). Mapping religion's other: a review of the study of nonreligion and secularity. J. Scientific Study Religion *58*, 319–335. https://doi.org/10.1111/jssr.12597.

34. Ribberink, E., Achterberg, P., and Houtman, D. (2013). Deprivatization of disbelief? Non-religiosity and anti-religiosity in 14 western European countries. Polit. Religion *6*, 101–120. https://doi.org/10.1017/S1755048312000740.

35. Ammerman, N.T. (2013). Spiritual but not religious? Beyond binary choices in the study of religion. J. Sci. Study Religion *52*, 258–278. https://doi.org/10.1111/jssr.12024.

36. Heelas, P. (1996). The New Age Movement: The Celebration of the Self and the Sacralization of Modernity (Blackwell).

37. Hanegraaff, W.J. (1996). New Age Religion and Western Culture: Esotericism in the Mirror of Secular Thought (Brill).

38. Huss, B. (2014). Spirituality: the emergence of a new cultural category and its challenge to the religious and the secular. J. Contemp. Religion *29*, 47–60. https://doi.org/10.1080/13537903.2014.864803.

39. Campbell, C. (2007). The Easternization of the West: A Thematic Account of Cultural Change in the Modern Era (Paradigm).

40. Fuller, R.C. (2001). Spiritual, but Not Religious: Understanding Unchurched America (Oxford University Press).

41. Palmisano, S. (2010). Spirituality and catholicism: the Italian experience. J. Contemp. Religion *25*, 221–241. https://doi.org/10.1080/13537901003750910.

42. Davie, G. (1994). Religion in Britain since 1945: Believing without Belonging (Wiley-Blackwell).

43. Voas, D., and Crockett, A. (2005). Religion in Britain: neither believing nor belonging. Sociology *39*, 11–28. https://doi.org/10.1177/0038038505048998.

44. Tromp, P., Kulkova, A., and Houtman, D. (2020). 'Believing without belonging' in twenty European countries (1981–2008): de-institutionalization of Christianity or spiritualization of religion? Rev. Religious Res. *62*, 509–531. https://doi.org/10.1007/s13644-020-00432-z.

45. Storm, I. (2009). Halfway to heaven: four types of fuzzy fidelity in Europe. J. Scientific Study Religion *48*, 702–718. https://doi.org/10.1111/j.1468-5906.2009.01474.x.

46. Woodhead, L. (2010). Real religion and fuzzy spirituality? Taking sides in the sociology of religion. In Religions of Modernity: Relocating the Sacred to the Self and the Digital, S. Aupers and D. Houtman, eds. (Brill), pp. 31–48.

47. Houtman, D., and Tromp, P. (2020). The Post-Christian Spirituality Scale (PCSS): misconceptions, obstacles, prospects. In Assessing Spirituality in a Diverse World, P.L. Ai, P. Wink, R.F. Paloutzian, and K.A. Harris, eds. (Springer International), pp. 35–57.

48. Marshall, J., and Olson, D.V.A. (2018). Is 'spiritual but not religious' a replacement for religion or just one step on the path between religion

and non-religion? Rev. Religious Res. *60*, 503–518. https://doi.org/10.1007/s13644-018-0342-9.

49. Lee, L. (2015). Recognizing the Non-religious: Reimagining the Secular (Oxford University Press).

50. Bruce, S. (2006). Secularization and the impotence of individualized religion. Hedgehog Rev. *8*, 34–45.

51. Thiessen, J., and Wilkins-Laflamme, S. (2017). Becoming a religious none: irreligious socialization and disaffiliation. J. Scientific Study Religion *56*, 64–82. https://doi.org/10.1111/jssr.12319.

52. Voas, D. (2015). The normalization of non-religion: a reply to James Lewis. J. Contemp. Religion *30*, 505–508. https://doi.org/10.1080/13537903.2015.1081354.

53. Zuckerman, P. (2020). Society without God: What the Least Religious Nations Can Tell Us about Contentment, Second Edition (New York University Press).

54. Vernon, G.M. (1968). The religious "nones": a neglected category. J. Scientific Study Religion *7*, 219–229. https://doi.org/10.2307/1384629.

55. Campbell, C. (1971). Toward a Sociology of Irreligion (Macmillan).

56. Mongeon, P. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics *106*, 213–228. https://doi.org/10.1007/s11192-015-1765-5.

57. Nikitaki, S. (2018). The place of nonbelief studies in the academic study of religion: issues and insights. In Formatting Nonreligion in Late Modern Societies—Institutional and Legal Perspectives (University of Oslo).

58. Balazka, D. (2020). Mapping Religious Nones in 112 Countries: An Overview of European Values Study and World Values Survey Data (1981-2020), Technical report, Prot.8/07-2020 (Bruno Kessler Foundation).

59. Houtman, D., Heelas, P., and Achterberg, P. (2012). Counting spirituality? Survey methodology after the spiritual turn. Annu. Rev. Sociol. Religion *3*, 25–44. https://doi.org/10.1163/9789047429470_003.

60. Taira, T. (2018). Who are the nones in Finland? In Formatting Nonreligion in Late Modern Societies – Institutional and Legal Perspectives (University of Oslo).

61. Voas, D., and McAndrew, S. (2012). Three puzzles of non-religion in Britain. J. Contemp. Religion *27*, 29–48. https://doi.org/10.1080/13537903.2012.642725.

62. Lewis, J.R. (2015). Education, irreligion, and non-religion: evidence from select anglophone census data. J. Contemp. Religion *30*, 265–272. https://doi.org/10.1080/13537903.2015.1025556.

63. Cragun, R., McCffree, K., Puga-Gonzalez, I., Wildman, W., and Shults, F.L. (2021). Religious exiting and social networks: computer simulations of religious/secular pluralism. Secularism & Nonreligion *10*, 1–20. https://doi.org/10.5334/snr.129.

64. Strawn, K.D. (2019). What's behind the "nones-sense"? Change over time in factors predicting likelihood of religious nonaffiliation in the United States. J. Scientific Study Religion *58*, 707–724. https://doi.org/10.1111/jssr.12609.

65. Franzmann, M. (2016). From "atheism" to "religious indifference": suggestions for future research on secularization. Annu. Rev. Sociol. Religion *7*, 1–16. https://doi.org/10.1163/9789004319301_002.

66. Keller, B., Hood, R.W., Jr., and Streib, H. (2016). Mapping the varieties of "spiritual" biographies. In Semantics and Psychology of Spirituality. A Cross-Cultural Analysis, H. Streib and R.W. Hood, Jr., eds. (Springer), pp. 275–280.

67. Lee, L. (2014). Secular or nonreligious? Investigating and interpreting generic 'not religious' categories and populations. Religion *44*, 466–482. https://doi.org/10.1080/0048721X.2014.904035.

68. Lee, L., and Bullivant, S. (2016). Oxford Dictionary of Atheism (Oxford University Press).

69. Bullivant, S., Farias, M., Lanman, J., and Lee, L. (2019). Understanding Unbelief. Atheists and Agnostics Around the World, Technical report (UK, May: St Mary's University).

70. Lindeman, M., and Lipsanen, J. (2016). Diverse cognitive profiles of religious believers and nonbelievers. Int. J. Psychol. Religion *26*, 185–192. https://doi.org/10.1080/10508619.2015.1091695.

71. Jong, J., Bluemke, M., and Halberstadt, J. (2013). Fear of death and supernatural beliefs: developing a new supernatural belief scale to test the relationship. Eur. J. Personal. *27*, 495–506. https://doi.org/10.1002/per.1898.

72. Lim, C., MacGregor, C.A., and Putnam, R.D. (2010). Secular and liminal: discovering heterogeneity among religious nones. J. Scientific Study Religion *49*, 596–618. https://doi.org/10.1111/j.1468-5906.2010.01533.x.

73. Hout, M. (2017). Religious ambivalence, liminality, and the increase of no religious preference in the United States, 2006-2014. J. Scientific Study Religion *56*, 52–63. https://doi.org/10.1111/jssr.12314.

74. Bullock, J. (2016). Reimagining the Secular: The Sunday Assembly—Belonging Without Believing. In "The Diversity of Nonreligion" Closing Conference & NSRN Annual Conference, Zurich, Switzerland, 07-09 July 2016.

75. Chastain, D. (2017). Gifts without givers: secular spirituality and metaphorical cognition. Sophia *56*, 631–647. https://doi.org/10.1007/s11841-016-0554-9.

76. Conrad, N.G. (2018). An argument for unbelief: a discussion about terminology. Secularism Nonreligion *7*, 1–8. https://doi.org/10.5334/snr.110.

77. Lee, L. (2012). Research note: talking about a revolution: terminology for the new field of non-religious studies. J. Contemp. Religion *27*, 129–139. https://doi.org/10.1080/13537903.2012.642742.

78. Day, A. (2013). Yes, but not in the North: nuances in religion and language cultures. Stud. Ethn. Nationalism *13*, 105–108. https://doi.org/10.1111/sena.12013.

79. Field, C.D. (2013). Measuring religious affiliation in Great Britain: the 2011 census in historical and methodological context. Religion *44*, 357–382. https://doi.org/10.1080/0048721X.2014.903643.

80. Pasquale, F.L. (2007). Empirical study and neglect of unbelief and irreligion. In The New Encyclopedia of Unbelief, T. Flynn, ed. (Prometheus Books), pp. 760–766.

81. Wiktor-Mach, D. (2012). Measuring Muslims: the problems of religiosity and intra-religious diversity. Annu. Rev. Sociol. Religion *3*, 207–227. https://doi.org/10.1163/9789047429470_011.

82. Houtman, D., and Aupers, S. (2007). The spiritual turn and the decline of tradition: the spread of post-Christian spirituality in 14 western countries, 1981–2000. J. Scientific Study Religion *46*, 201–222. https://doi.org/10.1111/j.1468-5906.2007.00360.x.

83. Cragun, R.T. (2019). Questions you should never ask an atheist: toward better measures of nonreligion and secularity. Secularism Nonreligion *8*, 1–6. https://doi.org/10.5334/snr.122.

84. Järnefelt, E. (2020). Beneath the surface: a critique of the common survey model in the study of nonreligion. Secularism Nonreligion *9*, 1–9. https://doi.org/10.5334/snr.106.

85. Navarro-Rivera, J. (2020). Reshaping American Politics. The 2020 Secular Voices Survey Summary Report, Technical report (Hyattsville, Maryland: Socioanalitica Research).

86. Blankholm, J., Navarro-Rivera, J., and Hoesly, D. (2021). Take the first-ever secular communities survey. In The Humanist https://thehumanist.com/commentary/announcement-secular-communities-survey-launches-today.

87. Cragun, R.T., Hammer, J.H., and Nielsen, M. (2015). The nonreligious-nonspiritual scale (NRNSS): measuring everyone from atheists to zionists. Sci. Religion Cult. *2*, 36–53. https://doi.org/10.17582/journal.src/2015/2.3.36.53.

88. Bullivant, S. (2020). Explaining the rise of 'nonreligion studies': subfield formation and institutionalization within the sociology of religion. Social Compass *67*, 86–102. https://doi.org/10.1177/2F0037768619894815.

89. Frost, J., and Edgell, P. (2018). Rescuing nones from the reference category: civic engagement among the nonreligious in America. Nonprofit Voluntar. Sect. Q. *47*, 417–438. https://doi.org/10.1177/2F0899764 017746251.

90. Schwadel, P. (2020). The politics of religious nones. J. Scientific Study Religion *59*, 180–189. https://doi.org/10.1111/jssr.12640.

91. Balazka, D., and Rodighiero, D. (2020). Big data and the little big bang: an epistemological (R)evolution. Front. Big Data *3*, 1–13. https://doi.org/10.3389/fdata.2020.00031.

92. Lazer, D., and Radford, J. (2017). Data ex machina: introduction to big data. Annu. Rev. Sociol. *43*, 19–39. https://doi.org/10.1146/annurev-soc-060116-053457.

93. Resnyansky, L. (2019). Conceptual frameworks for social and cultural big data analytics: answering the epistemological challenge. Big Data Soc. *6*, 1–12. https://doi.org/10.1177/2053951718823815.

94. Anderson, C. (2008). The end of theory: the data deluge makes the scientific method obsolete. In WIRED http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

95. Mayer-Schönberger, V., and Cukier, K. (2013). Big Data. A Revolution that Will Transform How We Live, Work, and Think (Eamon Dolan Book/Houghton Mifflin Harcourt).

96. Veltri, G.A. (2017). Big data is not only about data: the two cultures of modelling. Big Data Soc. *4*, 1–6. https://doi.org/10.1177/2F2053951717703997.

97. Cufoglu, A. (2014). User profiling—a short review. Int. J. Comput. Appl. *108*, 1–9. https://doi.org/10.5120/18888-0179.

98. Centellegher, S., López, E., Saramäki, J., and Lepri, B. (2017). Personality traits and ego-network dynamics. PLoS One *12*, 1–17. https://doi.org/10.1371/journal.pone.0173110.

99. Mønsted, B., Mollgaard, A., and Mathiesen, J. (2018). Phone-based metric as a predictor for basic personality traits. J. Res. Personal. *74*, 16–22. https://doi.org/10.1016/j.jrp.2017.12.004.

100. Stachl, C., Au, Q., Schoedel, R., Gosling, S.D., Harari, G.M., Buschek, D., Völkel, S.T., Schuwerk, T., Oldemeier, M., Ullmann, T., et al. (2020). Predicting personality from patterns of behavior collected with smartphones. Proc. Natl. Acad. Sci. *117*, 17680–17687. https://doi.org/10.1073/pnas.1920484117.

101. Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., My Madsen, M., Larsen, J.E., and Lehmann, S. (2014). Measuring large-scale social networks with high resolution. PLoS One *9*, 1–24. https://doi.org/10.1371/journal.pone.0095978.

102. Gladstone, J.J., Matz, S.C., and Lemaire, A. (2019). Can psychological traits be inferred from spending? Evidence from transaction data. Psychol. Sci. *30*, 1087–1096. https://doi.org/10.1177/2F095679761984 9435.

103. Kosinski, M., Matz, S.C., Gosling, S.D., Popov, V., and Stillwell, D. (2015). Facebook as a research tool for social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am. Psychol. *70*, 543–556. https://doi.org/10.1037/a0039210.

104. Kosinski, M., Stillwell, D., and Graepel, D. (2013). Private traits and attributes are predictable from digital records of human behavior. Proc. Natl. Acad. Sci. *110*, 5802–5805. https://doi.org/10.1073/pnas.1218772110.

105. Nguyen, M.-T., and Lim, E.-P. (2014). On predicting religion labels in microblogging networks. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information, Gold Coast, Australia, pp. 1211–1214. https://doi.org/10.1145/2600428.2609547.

106. Li, W., Duan, L., Tsang, I.W.-H., and Xu, D. (2012). Co-labeling: a new multi-view learning approach for ambiguous problems. In 2012 IEEE 12th International Conference on Data Mining, pp. 419–428. https://hdl.handle.net/10356/99661.

107. Xu, X., Li, W., Xu, D., and Tsang, I.W. (2016). Co-labeling for multi-view weakly labeled learning. IEEE Trans. Pattern Anal. Machine Intelligence *38*, 1113–1125. https://doi.org/10.1109/TPAMI.2015.2476813.

108. Mai, J.-E. (2016). Big data privacy: the datafication of personal information. Inf. Soc. *32*, 192–199. https://doi.org/10.1080/01972243.2016.1153010.

109. Barnes, T.J. (2013). Big data, little history. Dialogues Hum. Geogr. *3*, 297–302. https://doi.org/10.1177/2F2043820613514323.

110. Welles, B.F. (2014). On minorities and outliers: the case for making big data small. Big Data Soc. *1*, 1–2. https://doi.org/10.1177/2F2053951714540613.

111. Sager, N. (1978). Natural language information formatting: the automatic conversion of texts to a structured data base. Adv. Comput. *17*, 89–162. https://doi.org/10.1016/S0065-2458(08)60391-5.

112. Asmussen, C.B., and Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. J. Big Data *6*, 1–18. https://doi.org/10.1186/s40537-019-0255-7.

113. Stokes, D.C., Andy, A., Guntuku, S.C., Ungar, L.H., and Merchant, R.M. (2020). Public priorities and concerns regarding COVID-19 in an online discussion forum: longitudinal topic modeling. J. Gen. Intern. Med. *35*, 2244–2247. https://doi.org/10.1007/s11606-020-05889-w.

114. Bose, R., Dey, R.K., Roy, S., and Sarddar, D. (2019). Analyzing political sentiment using Twitter data. In Information and Communication Technology for Intelligent Systems. Smart Innovation, Systems and Technologies, S. Satapathy, A. Joshi, ed. (Springer), pp. 427–436.

115. Zubiaga, A., Procter, R.N., and Maple, C. (2018). A longitudinal analysis of the public perception of the opportunities and challenges of the internet of things. PLoS ONE *13*, 1–18. https://doi.org/10.1371/journal.pone.0209472.

116. Sykora, M., Elayan, S., and Jackson, T.W. (2020). A qualitative analysis of sarcasm, irony and related #hashtags on Twitter. Big Data Soc. *7*, 1–15. https://doi.org/10.1177/2F2053951720972735.

117. Dahal, B., Kumar, S.A.P., and Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. Social Netw. Anal. Mining *9*, 1–20. https://doi.org/10.1007/s13278-019-0568-8.

118. Kaplan, F., and di Lenardo, I. (2017). Big data of the past. Front. Digital Human. *4*, 1–12. https://doi.org/10.3389/fdigh.2017.00012.

119. Dubrow, J.K., and Tomescu-Dubrow, I. (2016). The rise of cross-national survey data harmonization in the social sciences: emergence of an interdisciplinary methodological field. Qual. Quant. *50*, 1449–1467. https://doi.org/10.1007/s11135-015-0215-z.

120. Balazka, D., and Rodighiero, D. (2021). Cartography of Nonreligion and Secularity Research Network (Zenodo). https://doi.org/10.5281/zenodo.4495456.

121. Homer, W.I. (1998). Visual culture: a new paradigm. Am. Art *12*, 6–9.

122. Peraica, A. (2017). Culture of the Selfie: Self-Representation in Contemporary Visual Culture (Institute of Network Cultures).

123. Arbulla, B., and Bucchi, M. (2015). "Queue up, you stupid!": communicating about technology problems. An exploratory study of warning messages posted on machines in public places. J. Sci. Commun. *14*, 1–12.

124. Vindrola-Padros, C. (2011). The everyday lives of children with cancer in Argentina: going beyond the disease and treatment. Child. Soc. *26*, 430–442. https://doi.org/10.1111/j.1099-0860.2011.00369.x.

125. Eide, M. (2020). Religion in children's visual media: a qualitative content analysis of preschool holiday specials. J. Media Religion *19*, 108–126. https://doi.org/10.1080/15348423.2020.1812339.

126. Chen, Y., Li, Z., Li, M., and Ma W.-Y. (2006). Automatic Classification of Photographs and Graphics. 2006 IEEE International Conference on Multimedia and Expo, Toronto, Canada, 9–12 July, pp. 973-976. https://doi.org/10.1109/ICME.2006.262695.

127. Kumari, S., Jha, R.R., and Nigam, A. (2019). Indoor-outdoor scene classification with residual convolutional neural work. In Proceedings of 3rd International Conference on Computer Vision and Image Processing,

B. Chaudhuri, M. Nakagawa, P. Khanna, and S. Kumar, eds. (Springer), pp. 325–337.

128. Kumar, S., Singh, S., and Kumar, J. (2019). Gender Classification Using Machine Learning with Multi-Feature Method. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). https://doi.org/10.1109/CCWC.2019.8666601.

129. Sacchi, C., Gera, G., Marcenaro, L., and Ragazzoni, C.S. (2001). Advanced image-processing tools for counting people in tourist site-monitoring applications. Signal Process. 81, 1017–1040. https://doi.org/10.1016/S0165-1684(00)00280-2.

130. Gonthier, N., Gouseau, Y., Ladjal, S., and Bonfait, O. (2018). Weakly supervised object detection in artworks. In Computer Vision—ECCV 2018 Workshops. Lecture Notes in Computer Science, L. Leal-Taixé and S. Roth, eds. (Springer), pp. 692–709.

131. Lang, S., and Ommer, B. (2018). Attesting similarity: supporting the organization and study of art image collections with computer vision. Digital Scholarship in the Humanities 33, 845–856. https://doi.org/10.1093/llc/fqy006.

132. Hochman, N., and Manovich, L. (2013). Zooming into an Instagram city: reading the local through social media. First Monday 18. https://doi.org/10.5210/fm.v18i7.4711.

133. Hochman, N., and Manovich, L. (2014). A view from above: exploratory visualizations of the Thomas Walther collection. In OBJECT:PHOTO. Modern Photographs from the Thomas Walther Collection 1909-1949. An Online Project of the Museum of Modern Art, M. Abbaspour, L.A. Daffner, and M.M. Hambourg, eds. (The Museum of Modern Art), pp. 1–6.

134. Bruno, N., Bertamini, M., and Protti, F. (2015). Selfie and the city: a worldwide, large, and ecologically valid database reveals a two-pronged side bias in naïve self-portraits. PlosOne 10, 1–6. https://doi.org/10.1371/journal.pone.0124999.

135. Tifentale, A., and Manovich, L. (2015). Selfiecity: exploring photography and self-fashioning in social media. In Postdigital Aesthetics, D.M. Berry and M. Dieter, eds. (Palgrave Macmillan), pp. 109–122.

136. Blok, A., and Pedersen, M.A. (2014). Complementary social science? Quali-quantitative experiments in a big data world. Big Data Soc. 1, 1–6. https://doi.org/10.1177/2F2053951714543908.

137. Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., and Gershuny, J. (2018). Using new technologies for time diary data collection: instrument design and data quality findings from a mixed-mode pilot survey. Soc. Indicators Res. 137, 379–390. https://doi.org/10.1007/s11205-017-1569-5.

138. Giunchiglia, F., Zeni, M., Gobbi, E., Bignotti, E., and Bison, I. (2018). Mobile social media usage and academic performance. Comput. Hum. Behav. 82, 177–185. https://doi.org/10.1016/j.chb.2017.12.041.

139. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., and Campbell, A.T. (2014). StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 3–14. https://doi.org/10.1145/2632048.2632054.

140. Aharony, N., Pan, W., Ip, C., Khayal, I., and Pentland, A. (2011). Social fMRI: investigating and shaping social mechanisms in the real world.

Pervasive Mobile Comput. 7, 643–659. https://doi.org/10.1016/j.pmcj.2011.09.004.

141. Harari, G.M., Lane, N.D., Wang, R., Crosier, B.S., Campbell, A.T., and Gosling, S.D. (2016). Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges. Perspect. Psychol. Sci. 11, 838–854. https://doi.org/10.1177/1745691616650285.

142. Raveilhac, M., Steinmetz, S., and Morselli, D. (2020). Can social media data complement traditional survey data? A reflection matrix to evaluate their relevance for the study of public opinion. In BigSurv20—Big Data Meets Survey Science, Worldwide, 6/11–4/12 2020.

143. Lusk, J.L. (2017). Consumer research with big data: applications from the food demand survey (FooDS). Am. J. Agric. Econ. 99, 303–320. https://doi.org/10.1093/ajae/aaw110.

144. Christian, L., and Ricci, K. (2020). Integrating organic data and designed data for higher quality measurment: overcoming coverage limitations of big data. In BigSurv20—Big Data Meets Survey Science, Worldwide, 6/11–4/12 2020.

145. Avazpour, I., Grundy, J., and Zhu, L. (2019). Engineering complex data integration, harmonization and visualization systems. J. Ind. Inf. Integration 16, 1–13. https://doi.org/10.1016/j.jii.2019.08.001.

146. Kumar, G., Basri, S., Imam, A.A., and Balogun, A.O. (2020). Data harmonization for heterogeneous datasets in big data – a conceptual model. In Software Engineering Perspectives in Intelligent Systems,, R. Silhavy, P. Silhavy, and Z. Prokopova, eds. (Springer), pp. 723–734.

147. Biolcati, F., Luijkx, R., and Vezzoni, C. (2017). Introduction. Religion and social integration in Europe. Rassegna Italiana di Sociolog. 28, 727–740. https://doi.org/10.1423/88792.

148. FBK-ISR. (2020). Engaging Religious and Belief Actors in the European Approach to AI. Report (Bruno Kessler Foundation).

149. Hsiung, B.O., and Djupe, P.A. (2019). Religion and the extension of trust. Polit. Behav. 41, 609–631. https://doi.org/10.1007/s11109-018-9466-4.

150. Koenig, H.G. (2012). Religion, spirituality, and health: the research and clinical implications. Int. Scholarly Res. Netw. Psychiatry 2012, 1–33. https://doi.org/10.5402/2F2012/2F278730.

151. Omelicheva, M.Y., and Ahmed, R. (2018). Religion and politics: examining the impact of faith on political participation. Religion, State Soc. 46, 4–25. https://doi.org/10.1080/09637494.2017.1363345.

152. Rochmawati, E., Wiechula, R., and Cameron, K. (2018). Centrality of spirituality/religion in the culture of palliative care service in Indonesia: an ethnographic study. Nurs. Health Sci. 20, 231–237. https://doi.org/10.1111/nhs.12407.

153. Van Buren, H.J., III, Syed, J., and Mir, R. (2020). Religion as a macro social force affecting business: concepts, questions, and future research. Business Soc. 59, 799–822. https://doi.org/10.1177/2F0007650319845097.

154. Welch, M.R., Sikkink, D., Sartain, E., and Bond, C. (2004). Trust in god and trust in man: the ambivalent role of religion in shaping dimensions of social trust. J. Sci. Study Religion 43, 317–343. https://doi.org/10.1111/j.1468-5906.2004.00238.x.