CHAPTER 32

Geospatial Data Analysis in Russia's Geoweb

Mykola Makhortykh

32.1   Introduction

The rise of digital technologies has led to the emergence of new ways in which physical spaces are perceived, experienced, and mapped. The availability of high-quality satellite imaginary amplified by the unprecedented possibilities for crowdsourcing geospatial data (Crampton 2009) has enabled the emergence of multiple platforms dealing with geographic information. It was followed by the integration of geographically aware computing in the architecture of major social media platforms (Crampton et al. 2013) and the growing capabilities for location tracking embedded into mobile devices (Sansurooah and Keane 2015). Together, these changes have given rise to a global collection of services which use the geographic data for different domains' applications. These services are currently known as "geospatial Web" (Lake and Farley 2009) or simply "geoweb" (Crampton 2009).

The emergence of geoweb and associated "neographic" (Haklay et al. 2008) practices of publishing, sharing, and visualizing information about places and people has significant implications for academic research. In the large-scale review of studies, which use geospatial data, Stock (2018) demonstrates these data's applicability to a wide range of research fields, including recreation, crisis management, and environment studies. The reasons for the growing adoption of geospatial data vary from the emergence of geographic datasets of unprecedented size and granularity (Elwood 2010) to the transformation of citizens into geospatial subjects able to produce and employ geospatial data (Wilson 2011). Their use is amplified by innovative possibilities for identifying and

M. Makhortykh (✉)
University of Bern, Bern, Switzerland
e-mail: mykola.makhortykh@ikmb.unibe.ch

mapping spatial relationships enabled by artificial intelligence and big data (VoPham et al. 2018).

Russia is not an exception from this trend as shown by the increasing number of studies applying geospatial data to study subjects varying from electoral fraud (Kobak et al. 2016) to Silk Road tourism (Tikunov et al. 2018) to Second World War remembrance (Bernstein 2016). Yet, the use of geospatial data in the context of Digital Russian Studies has its own specifics attributed both to the general role of digital media in Russia's media ecologies and to the particular importance of geoweb in this geopolitical context. The explosive growth of Internet use in Russia in 2000s has led to profound changes in the language and communication in multiple domains, including politics (Gorham et al. 2014). The importance of the digital sphere increased even further since the beginning of the Ukraine crisis in 2014, which marked the unprecedented level of state-sponsored cynicism toward the media sphere and its growing instrumentalization for propaganda and disinformation (Roudakova 2017). In this "post-truth" (Surowiec 2017) environment, geolocation data that allow to (dis)prove the existence of specific phenomena emerge as a pivotal factor for making and refuting knowledge claims (e.g. about the presence of Russian troops in Ukraine (Shim 2018)).

To further contextualize the features of Russian geoweb and examine how recent studies address opportunities and challenges provided by it, I will start by reviewing different sources of geospatial data available in the Russian context, varying from social media platforms to crowdsourced databases. I will then move toward discussing possible ways of extracting location information; these ways vary from mapping location names provided through metadata to specific geographic coordinates to extracting location from verbal or visual texts or inferring it from users' activity on social media. Then, I will explore different ways to use geospatial data, such as mapping spatial distribution of socioeconomic phenomena and analyzing mediatization of cultural practices. Additionally, I will briefly discuss the ethical aspects of some of these uses, in particular privacy-related issues. Finally, I will conclude by recapping the main arguments of the chapter and scrutinizing possible directions for future uses of geospatial data in Digital Russian Studies.

32.2   Data Acquisition

The first question to address in research using geoweb analysis is what kind of geospatial data is to be used. As I mentioned in the introduction, the distribution of location tracking devices and geographic crowdsourcing gave rise to multiple platforms dealing with geospatial data; however, the format, scope, and quality of these data vary significantly depending on the platform. To illustrate these differences, I will review below three categories of geospatial data sources, which are of particular relevance for Digital Russian Studies: crowdsourced databases, open datasets, and social media.

32.2.1   *Crowdsourced Databases*

The availability of digital technology allowing to collect, visualize, and share geospatial data led to the emergence of multiple projects focused on crowdsourcing "volunteered geographic information" (Goodchild 2007). Unlike established sources of geographic information (e.g., open datasets produced by national mapping agencies), crowdsourced databases rely on the assumption that geospatial content produced and edited by multiple individuals will eventually converge on a consensus (Elwood et al. 2012, 575). While this assumption does not guarantee the same quality of data as in the case of sources produced by certified experts, crowdsourced projects are able to account for attributes which are usually omitted by traditional mapping agencies and capture fast-changing phenomena (e.g., natural disasters).

The scope and focus of volunteered geographic projects vary significantly. Some of them, such as Open Street Map (OSM) (https://www.openstreet-map.org), HERE Maps (https://mapcreator.here.com/), or Yandex People's Map (https://n.maps.yandex.ru/), pursue the goal of creating and sustaining free digital maps or gazetteers. Other projects have limited temporal and thematic focus. Both in Russia and in the West,[1] the latter projects often arise as part of the volunteered reporting in the context of natural disasters[2] or armed conflicts.[3]

Both categories of crowdsourced databases can be of use in the context of Digital Russian Studies. Many global initiatives provide relevant geospatial information, which can be used for Russia-centered research. For instance, Quinn and Tucker (2017) used OSM and Wikimapia (https://wikimapia.org/) to trace how crowdsourced maps are used to represent disputed areas such as Crimea and found substantial differences in the ways geopolitical disagreements were visualized and addressed. These differences were attributed to the OSM hosting more contributions from Western editors, whereas Wikimapia was more eager to transmit the Russian official discourse. Other examples include the study by Kulakov, Petrina, and Pavlova (2016), who used Wikimapia for evaluating digital smart services utilized for cultural heritage tourism planning, and the research by Karbovskii et al. (2014), who employed Wikimapia for simulating the process of decision making based on 2012 Krymsk flooding.

Additionally, the Russian digital landscape features a number of crowdsourced projects dealing with specific domains or topics. Despite their variety and rich data, these projects have so far received limited acknowledgement in academic scholarship. A few exceptions include, for instance, *Pomnite nas* (Remember Us) (http://www.pomnite-nas.ru/), a project devoted to collecting geospatial data about Second World War monuments devoted to Soviet soldiers (Bernstein 2016). Another example is *RosYama* (Russian Pit) (https://rosyama.ru/), a civic project initialized by Alexei Navalny, a Russian anti-systemic opposition leader and activist, who created an online crowdsourced service for reporting road potholes (Ermoshina 2014). Many of these projects are not necessarily designed as sources of geolocation data for academic research

and, instead, intended to facilitate social activities (e.g. collective remembrance of the Second World War in the case of *Pomnite nas*). Despite these non-academic goals, these projects can still be a valuable asset to the researcher who would creatively approach their data. For instance, geolocation data offered by *RosYama* can be used not only for research focused on the quality of Russian roads but also for visualizing geographic networks of activists or detecting the misappropriation of funds planned by specific regions for repairing the roads (for more projects like this, see Chap. 8).

The major challenge of using crowdsourced databases is related to the quality of data provided through them. Because of the lack of authoritative control over their content, the possibility of encountering errors or conscious distortions of geographic facts is higher than in the case of open datasets. In the larger crowdsourced databases such as Wikimapia or Yandex's People Map, such probability is lower because of the large number of contributors, which leads to faster error correction. The situation with small databases is more challenging: often, these projects are curated by small groups of users with limited time and financial resources. While the data offered by them can still be valuable (or even unavailable by other means), it is important to critically assess their quality and identify (as much as possible) who contributes to the database and for what ends.

32.2.2   *Open Datasets*

Besides the rise of volunteered geographic initiatives, the unprecedented ease of accumulating and sharing geospatial data resulted in the distribution of open datasets produced by certified actors such as state institutions and mapping agencies. Generated using authoritative geographic sources, these datasets are characterized by higher data quality when compared with crowdsourced databases. While the turn toward open data that are made available through official portals (for instance, data.gov or europeandataportal.eu) originated in the West, where these datasets are often employed in academic research on the subjects varying from earthquakes to government institutions' budgets (Ding et al. 2010; Shadbolt et al. 2012), Russia increasingly joins the open data movement.

A number of Russian official agencies make their data available through online portals, such as Russian Open Data Portal (RODP) (data.gov.ru) or Open Data Portal of Moscow City Government (data.mos.ru) (Bundin and Martynov 2015; Koznov et al. 2016; Repponen 2018). A selection of Russian portals, where open datasets are published, is provided in Table 32.1. Despite being subjected to a number of drawbacks, including often limited data pre-processing, absence of unified data standards for different organizations, and the lack of application programming interfaces (APIs) (fiftin 2017), these portals provide access to a variety of unique geospatial datasets from different domains, varying from culture (e.g., the dataset on the geospatial distribution of places related to Russian poetess Anna Akhmatova in Moscow [Data.gov

Table 32.1 Open datasets in Russian geoweb

Dataset name	Web address	Description	English interface
Russian Open Data Portal	https://data.gov.ru/	Collection of datasets related to Russian state agencies	Yes
Open Data Portal of Moscow City Government	https://data.mos.ru/	Collection of datasets produced by the Moscow City Government	Yes
Open Data Portal of Russian Ministry of Culture	https://opendata.mkrf.ru/	Collection of datasets produced by the Russian Ministry of Culture	No
Open Data Hub	https://opengovdata.ru/	Collection of datasets related to Russian state agencies, commercial organizations, and non-governmental organizations (NGOs)	Yes

2016]) to crime (e.g., data about the number of committed, resolved, and unresolved crimes by region in Russia [Data.gov 2014]; for more on government data, see Chap. 23).

Two platforms which are of particular interest in this context are Russian Open Data Portal (RODP) and Open Data Hub (ODH). Both platforms provide a large number of datasets (22,233 for RODP and 8151 for ODH) from multiple Russian organizations (1102 and 42 organizations, respectively). These organizations vary from the federal organizations (e.g., the Ministry of Justice or the Federal Statistics Service) to the local ones (e.g., Tomsk Oblast administration). Not all of these datasets deal with geospatial information, but many of them do and can serve as a valuable source of data for geospatial research.

32.2.3   Social Media

As noted in other chapters of the handbook (see Chapters 20 and 30 on social media use in the context of Digital Russian Studies), social media platforms constitute a major source of digital data. Geospatial data are not an exception as the majority of social media platforms provide in one form or another information about the location of their users and/or content produced. Stock (2018) notes that the majority of studies focus on a few Western platforms, such as Twitter and Flickr, which have accessible APIs and contain geotagged content.[4] This combination allows both identifying the location in which some content available through the platforms is produced and also searching and retrieving data for the specific geographic range (e.g., for collecting messages and images produced within recreational areas to trace visitors' numbers [Tenkanen et al. 2017] and behavior [Sessions et al. 2016]).

In addition to Western social media platforms, Russian geoweb includes several major local platforms, such as VK (also known as VKontakte),

Odnoklassniki, and Moj Mir. Among these platforms, however, only VK provides easy access to its API, which allows retrieving a wide range of geospatial data (Tikunov et al. 2018). Specifically, VK API includes a number of functions also known as methods, which can be used for data extraction (for more on social networks, see Chaps. 19 and 30).[5]

The most common type of geospatial data provided by VK is the one on the country and the city/town of residence, which constitutes part of user profile (Zamyatina and Yashunsky 2018). In the case of publicly available profiles, these data can be retrieved using users.get method. The method takes as its input user ids which are of interest for the researcher and the list of fields that have to be retrieved ("country" and "city" are a common choice). These data can be further enriched and/or verified via other profile fields available on VK such as the ones on employment and education.

Besides data available as part of user profiles, VK also provides access to check-in data, which can be retrieved via places.getCheckins method. The method takes as input latitude and longitude coordinates and returns posts made within the specified area together with ids of users who published them. Similarly, VK allows retrieving images uploaded by users together with these images' geographic coordinates using photos.get method. The method returns geographic coordinates of retrieved images if these coordinates are provided by the user. Using this method, it is possible to retrieve a sample of images from specific geographic regions in order to, for instance, examine the ways in which these regions are represented online (Tikunov et al. 2018).

32.3   LOCATION EXTRACTION

After choosing the specific data source(s) and acquiring actual data, the next step is to process these data. In the case of geospatial data, the major purpose of processing involves the extraction of specific location(s) to which the data refer to or represent. Depending on the data format and available metadata, the process of location extraction can be as simple as retrieving exact geotags present in the metadata or mapping the location name to data from a geographic information system. In other cases, it can be more complex and involve the use of machine learning techniques to recognize the names of geographic entities in visual or verbal texts or to infer the location based on online user activity.

Geographic coordinates extraction from document metadata. The easiest—and most common (Stock 2018)—way of detecting location is by using geographic coordinates included in the document (meta)data. Such an approach is particularly applicable for data available from open datasets as well as crowd-sourced databases, which often include specific geographic coordinates. Additionally, some platforms such as Twitter and VK provide geographic coordinates for some types of their content.[6] The question of validity of these data, however, is an open one: especially in the case of geotagged content from social

media platforms, there is also a need to differentiate between the place in which the content was published and the place to which it actually refers.

Location name extraction from document metadata. In the cases when geographic coordinates are not provided, one of the alternatives is to extract place names from the metadata. This process usually consists of two steps: (1) toponym recognition: that is, identification of the toponym in the body of the metadata (Sagcan and Karagoz 2015), and (2) toponym resolution: that is, assigning of geographic coordinates to the recognized toponyms (Lieberman and Samet 2012). An example of the platform for which this approach can be highly beneficial is VK, which allows users to report their place of residence in their profiles. While the platform itself does not connect these data to a geographic information system, the location names can be retrieved via VK API and then connected to a geocoding service (e.g., Google Maps) to generate geographic coordinates (Lee et al. 2013; Baucom et al. 2013).

The most popular approach to location name extraction from the metadata is the gazetteer-based one, where the extracted location names are matched with the list of geographic named entities such as the ones provided by GeoNames (https://www.geonames.org/). Because of the limited number of gazetteers for the Russian language, such lists are often taken from Wikipedia or from a few training datasets such as FactRuEval (Starostin et al. 2016). At the same time, this approach suffers from a number of issues, including, for instance, intended or unintended mispronunciation (such as *Maskva* instead of *Moskva*) or instances of double naming (e.g. *Sankt-Peterburg* and *Piter*). To address these limitations, more complex approaches were proposed (for reviews, see Leidner 2007; Leidner and Lieberman 2011); a recent study comparing different approaches to the task indicates that approaches using lexical context of toponyms and their importance (e.g., by solving typonym-related ambiguity by always preferring options with the largest population) perform particularly well (Weissenbacher et al. 2019).

*Location name extraction from raw tex*t. This approach is similar to the location name extraction from document metadata and involves the same two steps: toponym recognition and toponym resolution. However, unlike the former approach which relies on the document's metadata, the latter one takes as input raw text data. Stock (2018, 219) notes that a major benefit of this approach is that it can be used for any text-based message (e.g. photo/video descriptions or blog posts). This approach tends to be less accurate than the one relying on supplied geotags, especially as geographic names are often ambiguous. However, it is often the only way to extract location in the cases when geographic coordinates are not provided.

The usual way of extracting location from raw texts employs the named entity recognition approach: that is, automatic detection of the words which refer to certain geographic locations. The process of detection is based on named entity recognition tools, such as Stanford or GATE, which combine machine learning techniques with pre-made geographic gazetteers, such as

GeoNames or OpenStreetMap (Stock 2018, 220; for practical examples see Jaiswal et al. 2013; Inkpen 2016; Bassi et al. 2016).

While most of the research on named entity recognition approach is tailored to the English language, in recent years the growing number of works employs this technique for the Russian context.[7] Because of the limited number of pre-made Russian gazetteers, a number of studies (see, for instance, Sysoev and Andrianov 2016) employ Wikipedia as a source of information. Additionally, there are several training datasets which include geographic data. An example of such a dataset is FactRuEval, an open annotated corpus of Russian texts.[8] The paper by Ivanitskiy et al. (2016) discusses in more details how FactRuEval can be used for geographic named entity retrieval from Russian sources.

Location inference from user activity. In some cases, the documents in question do not provide explicit references to the geographic entity; however, even under these conditions, it is still possible to infer the location based on earlier user activity. Jurgens et al. (2015) summarizes several approaches based on user networks which can be applied for dealing with this task. The majority of these approaches involve identification of users sharing the closest connections with the user in question and then using data from them to infer the user's location.

Another approach is based on content produced by the user online. A number of studies (Cheng et al. 2010; Chang et al., 2012; Han et al. 2014) discuss the possibility of inferring geographic location from local terms also known as location indicative words (LIWs) (Han et al. 2014). LIWs are terms which are particularly representative for specific places, either because of being indicative of certain locations (e.g. "rockets" for Houston) or language practices (e.g. "howdy" for Texas). Consequently, LIWs can be used to predict the location of a user who uses these terms through machine learning techniques.

Several studies (Han et al. 2014; Mourad et al. 2017) apply the latter approach to detect location based on Russian LIWs. The main idea behind it is to acquire textual data produced by users at certain geographic locations (Twitter was used in the above-mentioned studies, but the same principle can be employed for Instagram or VK) and then create separate text corpora for each location in question. Then, for each location LIWs are extracted and the model is trained. Han et al. (2014) offer a detailed discussion of different approaches toward LIWs extraction and show that information gain ratio approach provides the best performance.

Location name extraction from image. While location extraction from images is more challenging than from textual data, several techniques allow addressing this task. The first of them is based on the use of geographic information, in particular geotags, embedded in the image metadata. Usually provided in EXIF format (Stock 2018, 222), these metadata are created by the camera and include data about the image creation date, camera settings, and geolocation. Some platforms, such as Flickr, provide API access to these metadata, thus allowing to search these platforms' contents for images from specific areas and specific time span (McDougall and Temple-Watts 2012).

The second technique can be employed in the cases where no metadata is provided and involves the comparison of image similarity. Stock (2018, 222) identifies a number of approaches used to address this task, varying from the use of scale-invariant feature transformation (SIFT) for comparing selected image features (Crandall et al. 2009) to color and texton histograms employed in the domain of computer vision (Gallagher et al. 2009). After identifying these features for the image in question, they can then be compared with large image datasets (e.g., coming from Flickr) to identify similarities.

Location name extraction from video. Similar to location extraction from image, several other major approaches for location extraction can be identified. The first of them involves the use of video metadata (e.g., geographic coordinates produced by Global Positioning System [GPS] and compass sensors, which are embedded into video descriptions). This information can be used to identify the region in which the video was produced. Then geoinformation services (e.g., OSM) can be used to extract data about visible objects in the region (e.g., monuments or office buildings) in 2D or 3D.[9] Using OSM data, the descriptive tags can be generated for different objects in the area (e.g., their addresses and names), and then the object models can be compared with objects from the videos. Then, the relevance of each tag for specific video frame is calculated (i.e., to detect if a specific tag is present or absent on the frame) (Shen et al. 2011). While currently there are no papers applying this approach to the Russian context, such an approach is language-agnostic and can be implemented for any video independently of the language in which it is produced, until there is some metadata available.

The second approach can also be employed in the cases where no video metadata is present and combines audio and visual features of videos for identifying the location shown in them. For this purpose, a geotagged collection of videos is required; this collection is then used for calculating the audiovisual similarity with non-geotagged content. Specifically, visual frames and soundtrack are extracted from the videos, and then visual and acoustic features are computed for each one of them. Following the extraction, k-nearest neighbor algorithm (a classification algorithm, which classifies the unknown objects according to the classes of k closest neighbors) is employed to identify geotagged videos which look and sound more similar to the non-geotagged content (Sevillano et al. 2015).

32.4   LOCATION USE

After the location is extracted and identified, it can be used for actual analysis. As I noted earlier, the advantage of geospatial data is their versatility and applicability for addressing a wide range of research questions. In this section, I scrutinize some of the uses of geoweb in the context of Digital Russian Studies, from mapping the spatial distribution of phenomena and specifying actors' identities and relationships to scrutinizing the role of location in online cultural practices.

Mapping the spatial distribution of phenomena. An important feature of using geospatial data is its rich potential for mapping socioeconomic and (geo) political phenomena. These phenomena vary from tourist mobility (e.g., spatial and temporal dimensions of tourist flows [Lu and Stepchenkova 2015; Kirilenko and Stepchenkova 2017]) to electoral fraud during Russia's federal elections (Kobak et al. 2016) and migration patterns (Zamyatina and Piliasov 2013). Geotag data can be also used for mapping contested phenomena, when official reports are often subjected to censorship or disinformation, such as the involvement of Russian troops in the conflict in Eastern Ukraine based on Instagram data (Czuperski et al. 2015). While the use of geospatial data for studying such contested cases often raises multiple concerns (e.g., concerning the reproducibility and the quality of available data), it can still provide valuable insights for researchers.

Specifying actor identities and relationships. Another common use of geospatial data is for identifying specific actors and tracking connections between them. Such tasks are particularly common for studies in political communication and/or disinformation online: for instance, Zelenkauskaite and Balduccini (2017) used geospatial data to specify the origins of users commenting on Russian language news portals in Lithuania, whereas Helmus et al. (2018) employed geoweb to track the identities of users involved in Russian propaganda and counter-propaganda efforts on Twitter. Disinformation, however, is not the only subject which can be investigated in this context as shown by Smirnov et al. (2016), who used geospatial data for identifying friendship networks between youngsters on VK.

Scrutinizing digitization of cultural practices. The use of geospatial data increasingly becomes part of the mediatization of cultural practices, varying from war remembrance to tourism. Bernstein (2016) in his research on Second World War memory in Russia showed how the formation of a geotagged database of Soviet monuments enriches existing memory practices by producing virtual embodiments of existing memorials and re-iterating the mainstream Soviet narrative of the war. Another example is the use of geotagged images as part of sharing—and shaping—travel experiences as shown by several studies focused on the use of geospatial information to examine vacation culture in Russia (Kirilenko and Stepchenkova 2017; Tikunov et al. 2018).

Exploring identity narration. Besides extensive possibilities for tracking phenomena, digital platforms also enable new ways of (re)-imagining individual and collective identities. A number of studies (Stefanidis et al. 2013; Croitoru et al. 2015) suggest that geospatial data can serve as a strong identifier of group belonging and individual self-expression. Examples of such identifications are, for instance, elements of individual user profiles on Wikipedia, where userboxes are employed for declaring individuals' interests, preferences, and personal details (Neff et al. 2013). In the context of Digital Russian Studies, these means of self-expression often deal with geospatial data (e.g., place of residence [Dounaevsky 2014]) or geopolitical aspects of territoriality (e.g., belonging of the Southern Ossetia to Georgia). Another example is the use of geolocation

data for producing digital maps of the conflict in Eastern Ukraine (e.g., MilitaryMaps or Liveuamap), which are used to visualize the borders of imagined communities (e.g., of the self-declared confederation of Novorossiya [Makhortykh 2018]).

32.5   Geospatial Data and Research Ethics

The advent of big data research opens unprecedented possibilities for studying different phenomena, but it also raises multiple ethical concerns. Some of these concerns are related to the general considerations of using big data for research purposes (e.g., acquiring proper permissions for data use [Richards and King 2014]), but some are rather specific for geospatial data, in particular in the Russian context. In this pre-final section, I will briefly discuss three of these concerns: validity, privacy, and reliability.

Privacy. Security and privacy are two key concerns of using geospatial data for research purposes (Li et al. 2016). The use of portable GPS receivers in mobile devices together with the enrichment of social media data with geospatial information raise concerns about the use of these data for tracking individuals' actions and movements (Loebel 2012). While such data can be beneficial for many types of research, their use also requires the researcher to recognize the potential consequences for the privacy of users. Such consequences are particularly important in cases dealing with highly sensitive and/or polarizing subjects, where the use of geotag data can cause material or immaterial harm for research participants.

The privacy risks are even greater when geotag data is used for studying phenomena occurring in authoritarian states. An example of a highly privacy-sensitive subject is research on anti-government protests, where geospatial data can be (ab)used to identify the location of individual protesters and expose their involvement in the protests, thus bringing legal repercussions by the state. To address this concern, the use of personal data should be minimized and (pseudo)anonymization techniques should be used. On the official level, however, Russian legislation is still catching up with the notion of big data and their uses for research purposes (for an overview, see Zharova and Elin 2017). Consequently, the protection of the data rights of individuals in Russia is still significantly less strict than in the European Union (EU) countries, where it is regulated by the EU General Data Protection Regulation (GDPR).

Validity. Sheppard (2005, 74) defines validity as the degree to which the use of a specific instrument or finding is sound, defensible, and well-grounded for the issue at hand. The question of validity is of particular relevance for the use of geospatial data, because of their significant potential for being used for manipulation: both through the data and their visualization (Sheppard and Cizek 2009). In some cases, the use of data can be invalidated by their wrong interpretation (i.e., when geospatial information is used to prove a point which is incorrect), whereas in other cases obscure visualizations of data can mislead the public.

An example of the invalid use of geographic data is the contrasting reporting of the 2018 clashes near Chigari village in Eastern Ukraine. Both the Ukrainian authorities and pro-Russian insurgents produced video records showing them controlling certain landmarks, which were claimed to be related to the village in question. Despite these claims, not all of the shown landmarks were related to Chigari and eventually it was proven that the village was controlled by the Ukrainian army, but not before causing significant confusion. A possible way of increasing validity according to Sheppard and Cizek (2009, 2112) is to use more flexible and interactive approaches for geospatial data analysis, thus allowing end users more control over results' reporting.

Reliability. Sheppard (2005) argues that reliability is another major concern of using geospatial data. Unlike validity, which focuses on the possible (ab)uses of geospatial data for drawing invalid conclusions, reliability concerns the internal consistency of analysis and the possibility to produce the same results under similar conditions. The issue of reliability is of particular importance for analyses produced via crowdsourced databases and social media as both data sources are subjected to frequent changes and often provide limited possibilities for consistent data access.

An example of reliability issues which accompany the use of geospatial data is MilitaryMaps mentioned earlier. This crowdsourced database aggregates updates from conflicts in the post-Soviet space as well as in the Middle East and provides geotags indicating the movement of troops and outbursts of violence. From September 2018, however, the previously open project switched toward paid subscription, which made it harder to recreate analyses based on MilitaryMaps data. Another reliability-related limitation of the project is its reliance on the GoogleMaps framework, which stores markers that are added to the map only for a one-year period. Sheppard and Cizek (2009) suggest that the main way to amend these and other reliability issues is the use of more prescriptive approaches to data analysis and presentation based on recognized quality standards.

32.6   Conclusions

In this chapter, I discussed the possible uses of data available through geoweb, the integrated and discoverable collection of geographically related web services and data (Lake and Farley 2009), in the context of the Digital Russian Studies. Increasingly employed for academic studies worldwide, geoweb data are of particular importance for Russia-centered digital research, serving both as a pivotal factor for making and verifying knowledge claims by regional actors and an integral means of producing individual and collective narratives on subjects varying from international conflicts (Shim 2018) to presidential elections (Kobak et al. 2016).

The use of geoweb for Digital Russian Studies is facilitated by the large volume of geospatial data available today. As I discussed above, these data can be divided into three broad categories according to their source: (a) crowdsourced databases, (b) open datasets, and (c) social media. Out of these three, social

media data are the hardest to get and often require extensive pre-processing; however, they are also applicable to a wide range of research questions, in particular the ones related to inter-user interactions. Furthermore, the largest Russian social media platform, VK, provides public access to multiple forms of geospatial data (e.g. users' self-declared place of residence/work and check-in data), thus enabling more possibilities for data collection than many Western platforms.

The research possibilities provided by geospatial data are amplified by the quickly developing toolkit of analytical techniques used to extract geographic location from different data formats. The complexity of techniques varies depending on the data format. In the simplest scenarios, geographic coordinates or the location's administrative address are provided in the metadata and only has to be matched with data from existing geographic information systems. In the more difficult scenarios, the location has to be extracted from the content or inferred from the user's earlier activity using a combination of machine learning and geographic gazetteers. Much still can be done to better adapt these techniques to the Russophone context, in particular in terms of improving named entity recognition techniques and developing better gazetteers. Yet, even in the current state of research, there are plenty of possibilities for using the mentioned techniques for different types of Russia-centered studies.

The importance of location extraction techniques is exemplified by the wide range of research questions to which Russian geospatial data are applicable. These research questions vary from the spatial distribution of socioeconomic and political phenomena, such as migration and electoral fraud, to the verification of knowledge claims about the presence of Russian troops in Eastern Ukraine to the analysis of mediatization of cultural practices of war remembrance and the exploration of narrative uses of geospatial data for communicating individual and collective identities.

Despite their significant potential for Digital Russian Studies, the future of geospatial data is not fully clear. The existing concerns about complex interrelations between privacy and geospatial data are amplified by the current calls for tightening the government's control over the Internet in Russia, leading to increasing restrictions on data retrieval from Russian platforms' APIs, including VK. These limitations might curb the amount of geospatial data available from social media; however, the growing number of open datasets and crowdsourced databases suggests that Russia's geoweb will remain a valuable research venue for Digital Russian Studies for years to come.

Notes

1. For Western examples see Goodchild and Glennon (2010) and Cavelty and Giroux (2011).
2. See, for instance, Virtual Bell (http://russian-fires.ru/) project, which was used from 2010 till 2013 to collect reports of wildfires in Russia.

3. An example of such a project is MilitaryMaps (https://militarymaps.info/), which crowdsources the recent developments in conflict zones where Russian geopolitical interests are involved.
4. See, for instance, Cheng, Caverlee, and Lee (2010), Chang et al. (2012), Van Canneyt, Schockaert, and Dhoedt (2016), Tenkanen et al. (2017).
5. For more information see the description provided by VK documentation (https://vk.com/dev/methods; also available in English).
6. In the case of social media, however, geotagged content is often subjected to numerous limitations, which have to be recognized: in the case of Twitter, only a small number of users (i.e., less than one percent [Cheng, Caverlee, and Lee 2010]) enable geolocation in their profiles, whereas in VKontakte the use of the check-in option is used by the limited number of users and often confined to specific social situations (e.g., tourist visits).
7. See, for instance, Gareev et al. (2013), Sysoev and Andrianov (2016), Anh, Arkhipov, and Burtsev (2017).
8. For the description, see Starostin et al. (2016); the dataset is available in open access at https://github.com/dialogue-evaluation/factRuEval-2016.
9. For an overview of tools which can use OSM data to make 3D models see https://wiki.openstreetmap.org/wiki/3D.

References

Anh, Le T., Mikhail Y. Arkhipov, and Mikhail S. Burtsev. 2017. Application of a Hybrid Bi-LSTM-CRF Model to the Task of Russian Named Entity Recognition. In *AINL 2017: Artificial Intelligence and Natural Language*, 91–103. Cham: Springer.

Bassi, Jonathan, Sukanya Manna, and Sun Yu. 2016. Construction of a Geo-Location Service Utilizing Microblogging Platforms. In *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, 162–165. Piscataway, NJ: IEEE.

Baucom, Eric, Azade Sanjari, Xiaozhong Liu, and Miao Chen. 2013. Mirroring the Real World in Social Media: Twitter, Geolocation, and Sentiment Analysis. In *Proceedings of the 2013 International Workshop on Mining Unstructured Big Data Using Natural Language Processing*, 61–68. New York: ACM.

Bernstein, Seth. 2016. Remembering War, Remaining Soviet: Digital Commemoration of World War II in Putin's Russia. *Memory Studies* 9 (4): 422–436.

Bundin, Mikhail, and Aleksei Martynov. 2015. Russia on the Way to Open Data: Current Governmental Initiatives and Policy. In *Proceedings of the 16th Annual International Conference on Digital Government Research*, 320–322. New York: ACM.

Chang, Hau-wen, Dongwon Lee, Mohammed Eltaher, and Jeongkyu Lee. 2012. @Phillies Tweeting from Philly? Predicting Twitter User Locations with Spatial Word Usage. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining*, 111–118. Piscataway, NJ: IEEE.

Cheng, Zhiyuan, James Caverlee, and Kyumin Lee. 2010. You are Where You Tweet: A Content-Based Approach to Geo-Locating Twitter Users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, 759–768. New York: ACM.

Crampton, Jeremy. 2009. Cartography: Maps 2.0. *Progress in Human Geography* 33 (1): 91–100.

Crampton, Jeremy W., Mark Graham, Ate Poorthuis, Taylor Shelton, Monica Stephens, Matthew W. Wilson, and Matthew Zook. 2013. Beyond the Geotag: Situating 'Big Data' and Leveraging the Potential of the Geoweb. *Cartography and Geographic Information Science* 40 (2): 130–139.

Crandall, David, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. 2009. Mapping the World's Photos. In *Proceedings of the 18th International Conference on World Wide Web*, 761–770. New York: ACM.

Croitoru, Arie, N. Wayant, A. Crooks, Jacek Radzikowski, and Anthony Stefanidis. 2015. Linking Cyber and Physical Spaces Through Community Detection and Clustering in Social Media Feeds. *Computers, Environment and Urban Systems* 53: 47–64.

Czuperski, Maksymilian, John Herbst, Eliot Higgins, Alina Polyakova, and Damon Wilson. 2015. *Hiding in plain sight: Putin's war in Ukraine*. Washington, DC: Atlantic Council of the United States.

Data.gov. 2014. *Statističeskaâ informaciâ 'Informaciâ o zaregistrirovannyh, raskrytyh i neraskrytyh prestupleniâh' [Statistical data 'Information about Registered, Resolved and Unresolved Crimes']*. https://data.gov.ru/opendata/7727739372mvdgiac38. Accessed 17 Aug 2019.

———. 2016. *Ahmatovskie mesta v Moskve [Ahmatova's Places in Moscow]*. https://data.gov.ru/opendata/7708660670-ahmatova. Accessed 17 Aug 2019.

Ding, Li, Dominic DiFranzo, Alvaro Graves, James R. Michaelis, Xian Li, Deborah L. McGuinness, and Jim Hendler. 2010. Data-Gov Wiki: Towards Linking Government Data. In *2010 AAAI Spring Symposium: Linked Data Meets Artificial Intelligence*, 38–43. Menlo Park, CA: AAAI.

Dounaevsky, Helene. 2014. Building Wiki-History: Between Consensus and Edit Warring. In *Memory, Conflict and New Media: Web Wars in Post-Socialist States*, ed. Ellen Rutten, Julie Fedor, and Vera Zvereva, 130–143. London and New York: Routledge.

Dunn Cavelty, Myriam, and Jennifer Giroux. 2011. Crisis Mapping: A Phenomenon and Tool in Emergencies. *CSS Analyses in Security Policy* 103: 1–4.

Elwood, Sarah. 2010. Geographic Information Science: Emerging Research on the Societal Implications of the Geospatial Web. *Progress in Human Geography* 34 (3): 349–357.

Elwood, Sarah, Michael F. Goodchild, and Daniel Z. Sui. 2012. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers* 102 (3): 571–590.

Ermoshina, Ksenia. 2014. Democracy as Pothole Repair: Civic Applications and Cyber-Empowerment in Russia. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 8(3). https://cyberpsychology.eu/article/view/4315/3365. Accessed 17 Aug 2019.

fiftin. 2017. Počemu otkrytye dannye nikomu *ne nužny [Why Nobody* Uses Open Data]. July 17. *Habr*. https://habr.com/post/331036/. Accessed 17 Aug 2019.

Gallagher, Andrew, Dhiraj Joshi, Jie Yu, and Jiebo Luo. 2009. Geo-Location Inference from Image Content and User Tags. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 55–62. Piscataway, NJ: IEEE.

Gareev, Rinat, Maksim Tkachenko, Valery Solovyev, Andrey Simanovsky, and Vladimir Ivanov. 2013. Introducing Baselines for Russian Named Entity Recognition. In *Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics*, 329–342. Berlin: Springer.

Goodchild, Michael. 2007. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* 69 (4): 211–221.

Goodchild, Michael, and Alan Glennon. 2010. Crowdsourcing Geographic Information for Disaster Response: A Research Frontier. *International Journal of Digital Earth* 3 (3): 231–241.

Gorham, Michael, Ingunn Lunde, and Martin Paulsen. 2014. Introduction. In *Digital Russia: The Language, Culture and Politics of New Media Communication*, ed. Michael Gorham, Ingunn Lunde, and Martin Paulsen, 1–11. New York: Routledge.

Haklay, Muki, Alex Singleton, and Chris Parker. 2008. Web Mapping 2.0: The Neogeography of the GeoWeb. *Geography Compass* 2 (6): 2011–2039.

Han, Bo, Paul Cook, and Timothy Baldwin. 2014. Text-Based Twitter User Geolocation Prediction. *Journal of Artificial Intelligence Research* 49: 451–500.

Helmus, Todd, Elizabeth Bodine-Baron, Andrew Radin, Madeline Magnuson, Joshua Mendelsohn, William Marcellino, Andriy Bega, and Zev Winkelman. 2018. *Russian Social Media Influence: Understanding Russian Propaganda in Eastern Europe*. Santa Monica: RAND Corporation.

Inkpen, Diana. 2016. Text Mining in Social Media for Security Threats. In *Recent Advances in Computational Intelligence in Defense and Security*, 491–517. Cham: Springer.

Ivanitskiy, Roman, Alexander Shipilo, and Liubov Kovriguina. 2016. Russian Named Entities Recognition and Classification Using Distributed Word and Phrase Representations. In *Proceedings of the 3rd Annual International Symposium on Information Management and Big Data*, 150–156. Aachen: CEUR.

Jaiswal, Anuj, Wei Peng, and Tong Sun. 2013. Predicting Time-Sensitive User Locations from Social Media. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 870–877. New York: ACM.

Jurgens, David, Tyler Finethy, James McCorriston, Yi Tian Xu, and Derek Ruths. 2015. Geolocation Prediction in Twitter Using Social Networks: A Critical Analysis and Review of Current Practice. In *Proceedings of the Ninth International AAAI Conference on Web and Social Media*, 188–197. Menlo Park, CA: AAAI.

Karbovskii, Vladislav, Daniil Voloshin, Kseniia Puzyreva, and Aleksandr Zagarskikh. 2014. Personal Decision Support Mobile Service for Extreme Situations. *Procedia Computer Science* 29: 1646–1655.

Kirilenko, Andrei, and Svetlana Stepchenkova. 2017. Sochi 2014 Olympics on Twitter: Perspectives of Hosts and Guests. *Tourism Management* 63: 54–65.

Kobak, Dmitry, Sergey Shpilkin, and Maxim Pshenichnikov. 2016. Integer Percentages as Electoral Falsification Fingerprints. *The Annals of Applied Statistics* 10 (1): 54–73.

Koznov, Dmitrij, Olga Andreeva, Uolevi Nikula, Andrey Maglyas, Dmitry Muromtsev, and Irina Radchenko. 2016. A Survey of Open Government Data in Russian Federation. In *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 173–180. Porto: SCITEPRESS.

Kulakov, Kirill, Oksana Petrina, and Anastasiya Pavlova. 2016. Smart Service Efficiency: Evaluation of Cultural Trip Planning Service. In *Proceedings of the 19th Conference of Open Innovations Association*, 106–112. Piscataway, NJ: IEEE.

Lake, Ron, and Jim Farley. 2009. Infrastructure for the Geospatial Web. In *The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, Scharl, Arno, and Klaus Tochtermann, eds*, 15–27. London: Springer Science & Business Media.

Lee, Ryong, Shoko Wakamiya, and Kazutoshi Sumiya. 2013. Urban Area Characterization Based on Crowd Behavioral Lifelogs Over Twitter. *Personal and Ubiquitous Computing* 17 (4): 605–620.

Leidner, Jochen. 2007. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names.* Irvine, CA: Universal-Publishers.

Leidner, Jochen, and Michael Lieberman. 2011. Detecting Geographical References in the Form of Place Names and Associated Spatial Natural Language. *SIGSPATIAL Special* 3 (2): 5–11.

Li, Songnian, Suzana Dragicevic, Francesc Antón Castro, Monika Sester, Stephan Winter, Arzu Coltekin, Christopher Pettit, et al. 2016. Geospatial Big Data Handling Theory and Methods: A Review and Research Challenges. *ISPRS Journal of Photogrammetry and Remote Sensing* 115: 119–133.

Lieberman, Michael, and Hanan Samet. 2012. Adaptive Context Features for Toponym Resolution in Streaming News. In *Proceedings of the 35th international ACM SIGIR Conference on Research and Development in Information Retrieval*, 731–740. New York: ACM.

Loebel, Jens-Martin. 2012. Is Privacy Dead?—An Inquiry into GPS-Based Geolocation and Facial Recognition Systems. In *IFIP International Conference on Human Choice and Computers*, 338–348. Berlin: Springer, Berlin.

Lu, Weilin, and Svetlana Stepchenkova. 2015. User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *Journal of Hospitality Marketing & Management* 24 (2): 119–154.

Makhortykh, Mykola. 2018. *Charting the Conflicted Borders: Narrating the Conflict in Eastern Ukraine Through Digital Maps.* Paper presented at The Power of Borderland(s): In Media's Res. June 28–29, Greifswald.

McDougall, Kevin, and P. Temple-Watts. 2012. The Use of LIDAR and Volunteered Geographic Information to Map Flood Extents and Inundation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 1: 251–256.

Mourad, Ahmed, Falk Scholer, and Mark Sanderson. 2017. Language Influences on Tweeter Geolocation. In *Proceedings of the European Conference on Information Retrieval*, 331–342. Cham: Springer.

Neff, Jessica J., David Laniado, Karolin E. Kappler, Yana Volkovich, Pablo Aragón, and Andreas Kaltenbrunner. 2013. Jointly They Edit: Examining the Impact of Community Identification on Political Interaction in Wikipedia. *PLOS One* 8 (4). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0060584. Accessed 17 Aug 2019.

Quinn, Sterling, and Doran A. Tucker. 2017. How Geopolitical Conflict Shapes the Mass-Produced Online Map. *First Monday* 22 (11). https://journals.uic.edu/ojs/index.php/fm/article/view/7922/6558. Accessed 17 Aug 2019.

Repponen, Ilona. 2018. Linked Ontology on Metadata on Russian Open Governmental Data. http://urn.fi/urn:nbn:fi:csc-kata20181221153713379920. Accessed 17 Aug 2019.

Richards, Neil, and Jonathan King. 2014. Big Data Ethics. *Wake Forest Law Review* 49: 393–432.

Roudakova, Natalia. 2017. *Losing Pravda: Ethics and the Press in Post-Truth Russia.* Cambridge: Cambridge University Press.

Sagcan, Meryem, and Pinar Karagoz. 2015. Toponym Recognition in Social Media for Estimating the Location of Events. In *2015 IEEE International Conference on Data Mining Workshop*, 33–39. Piscataway, NJ: IEEE.

Sansurooah, Krishnun, and Bradley Keane. 2015. The Spy in Your Pocket: Smartphones and Geo-Location Data. In *Proceedings of the 13th Australian Digital Forensics Conference*, 95–103. https://ro.ecu.edu.au/adf/154/. Accessed 17 Aug 2019.

Sessions, Carrie, Spencer A. Wood, Sergey Rabotyagov, and David M. Fisher. 2016. Measuring Recreational Visitation at US National Parks with Crowd-Sourced Photographs. *Journal of Environmental Management* 183: 703–711.

Sevillano, Xavier, Xavier Valero, and Francesc Alías. 2015. Look, Listen and Find: A Purely Audiovisual Approach to Online Videos Geotagging. *Information Sciences* 295: 558–572.

Shadbolt, Nigel, Kieron O'Hara, Tim Berners-Lee, Nicholas Gibbins, Hugh Glaser, and Wendy Hall. 2012. Linked Open Government Data: Lessons from data.gov.uk. *IEEE Intelligent Systems* 27 (3): 16–24.

Shen, Zhijie, Sakire Arslan Ay, Seon Ho Kim, and Roger Zimmermann. 2011. Automatic Tag Generation and Ranking for Sensor-Rich Outdoor Videos. In *Proceedings of the 19th ACM International Conference on Multimedia*, 93–102. New York: ACM.

Sheppard, Stephen. 2005. Validity, Reliability, and Ethics in Visualization. In *Visualization in Landscape and Environmental Planning: Technology and Applications*, ed. Ian Bishop and Eckart Lange, 79–97. London: Taylor and Francis.

Sheppard, Stephen, and Petr Cizek. 2009. The Ethics of Google Earth: Crossing Thresholds from Spatial Data to LandscapeVisualisation. *Journal of Environmental Management* 90 (6): 2102–2117.

Shim, David. 2018. Satellites. In *Visual Global Politics*, ed. Roland Bleiker, 265–271. New York: Routledge.

Smirnov, Ivan, Elizaveta Sivak, and Yana Kozmina. 2016. In Search of Lost Profiles: The Reliability of VKontakte Data and its Importance in Educational Research. *Voprosy Obrazovaniya* 4: 106–122.

Starostin, Anatoli, Dmitry Granovsky, Svetlana Alexeeva, Anastasia Bodrova, Alexander Chuchunkov, et al. 2016. FactRuEval 2016: Evaluation of Named Entity Recognition and Fact Extraction Systems for Russian. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016"*, 1–19. Moscow: RSUH.

Stefanidis, Anthony, Andrew Crooks, and Jacek Radzikowski. 2013. Harvesting Ambient Geospatial Information from Social Media Feeds. *GeoJournal* 78 (2): 319–338.

Stock, Kristin. 2018. Mining Location from Social Media: A Systematic Review. *Computers, Environment and Urban Systems* 71: 209–240.

Surowiec, Pawel. 2017. Post-Truth Soft Power; Changing Facets of Propaganda, Kompromat, and Democracy. *Georgetown Journal of International Affairs* 18 (3): 21–27.

Sysoev, Andrey, and Ivan Andrianov. 2016. Named Entity Recognition in Russian: the Power of Wiki-Based Approach. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016"*, 1–10. Moscow: RSUH.

Tenkanen, Henrikki, Enrico Di Minin, Vuokko Heikinheimo, Anna Hausmann, Marna Herbst, Liisa Kajala, and Tuuli Toivonen. 2017. Instagram, Flickr, or Twitter:

Assessing the Usability of Social Media Data for Visitor Monitoring in Protected Areas. *Scientific reports* 7 (1): 1–11.

Tikunov, Vladimir S., Vitaly S. Belozerov, Aleksander N. Panin, and Stanislav Antipov. 2018. Geoinformation Monitoring of Key Queries of Search Engines, and Geotagging Photos in the North-Caucasian Segment of the Tourist Route 'Great Silk Road'. *Annals of GIS* 24 (4): 255–260.

Van Canneyt, Steven, Steven Schockaert, and Bart Dhoedt. 2016. Categorizing Events Using Spatio-Temporal and User Features from Flickr. *Information Sciences* 328: 76–96.

VoPham, Trang, Jaime E. Hart, Francine Laden, and Yao-Yi Chiang. 2018. Emerging Trends in Geospatial Artificial Intelligence (GeoAI): Potential Applications for Environmental Epidemiology. *Environmental Health* 17 (1): 1–6.

Weissenbacher, Davy, Arjun Magge, Karen O'Connor, Matthew Scotch, and Graciela Gonzalez. 2019. Semeval-2019 task 12: Toponym Resolution in Scientific Papers. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, 907–916. Stroudsburg, PA: ACL.

Wilson, Matthew. 2011. 'Training the Eye': Formation of the Geocoding Subject. *Social & Cultural Geography* 12 (4): 357–376.

Zamyatina, Nadezhda, and Aleksandr Piliasov. 2013. Sever, Social'nye Seti i 'Diaspora Naoborot'. *Demoscop* 547–548. http://www.demoscope.ru/weekly/2013/0547/analit07.php. Accessed 17 Aug 2019.

Zamyatina, Nadezhda, and Alexey Yashunsky. 2018. Virtual Geography of Virtual Population. *Monitoring of Public Opinion: Economic and Social Changes* 143 (1): 117–137.

Zelenkauskaite, Asta, and Marcello Balduccini. 2017. 'Information Warfare' and Online News Commenting: Analyzing Forces of Social Influence Through Location-Based Commenting User Typology. *Social Media + Society* 3 (3): 1–13.

Zharova, Anna, and Vladimir Mikhailovich Elin. 2017. The Use of Big Data: A Russian Perspective of Personal Data Security. *Computer Law & Security Review* 33 (4): 482–501.