

Andrew J. M. Smith

Emporia State University

e-mail: asmith37@emporia.edu

004:02]:929.5

https://doi.org/10.18485/ai_lik.2020.6.9.5

Оригинални научни рад

INFORMATION PRIVACY IN THE REALM OF GENEALOGICAL RESEARCH

This paper explores issues of individual privacy encountered during genealogical research. The publication of public records, including vital records and census data is discussed as well as the increasing role of genealogical DNA profiles in family history research. The role of libraries in maintaining privacy while encouraging genealogical research is also considered.

Key Words: information privacy, genealogy, vital records, genetic genealogy, information policy, consent

Interest in genealogical research has greatly increased over the past few decades, particularly in the United States, the United Kingdom and in countries with large historical immigrant populations. Family history has become the topic of successful television programs featuring celebrities discovering stories about past generations of their families and this has further expanded interest in the process by a wide range of people, eager to discover their own heritage. This thirst for knowledge has also supported the rise of large, commercial, online genealogical databases and a plethora of publishing of family history books, software and other materials in an attempt to satisfy and capitalize on this desire.

As part of this drive for knowledge, vast quantities of data supporting genealogical research have been digitized and published on the internet and elsewhere by governmental agencies, by commercial genealogical services, by libraries and archives, and by volunteer organizations and individuals. Data sets may be published in multiple forms by multiple entities, or may be exclusive to one particular entity and complex webs of

ownership and licensing agreements abound. While this free access to information may be welcomed in many spheres, it also raises many questions of individual privacy and the potential for harm that may arise from the use and dissemination of this information.

1. Privacy and the Role of Libraries

Libraries have traditionally supported genealogical research through the provision of local history resources and the preservation of everyday materials, such as local phone or business directories, that in older models of librarianship were simply added to with new editions, rather than replaced, resulting in large collections of information of great use to family historians. Libraries have continued to support genealogical research through the provision to their patrons of subscription databases of both general historical interest, such as newspapers and maps, as well as specific, commercial genealogical databases.

In their professional codes, librarians have also championed not only the idea of information access, but the right of privacy where patron usage of materials is not subject to scrutiny or retention, and no records of actual patron borrowing or access are maintained. This value of privacy is enshrined in codes of ethics, such as the “Code of Ethics” of the American Library Association (2020) or the “Code of Professional Practice” of CILIP (2012) in the United Kingdom. In more recent times, as described by Carpenter (2015), libraries have been challenged by the fact that they increasingly license material rather than having complete control over it within their own discrete collections. Since they are serving as a conduit for patron access to licensed materials, they do not have as much control over patron information and usage data, and so although the library itself may not gather and retain patron information, such information may be harvested, retained and exploited by the providers of particular licensed services.

Additionally, as libraries are now one step removed from the patron and the information sources to which the library is connecting them it is harder for the libraries to teach or encourage their ideals of privacy and ethical use of records. On the one

hand the library is making more genealogical research possible through the provision of commercial databases, but on the other they are removed from the patrons and no longer have the same opportunity for education about privacy and information ethics.

The right to privacy embraced by libraries is not universally accepted. Article 8 of the European Union “Charter of Fundamental Rights” (2012) guarantees the right to protection of personal data and to access and rectify any personal data collected. However, other jurisdictions have been less quick to grant their citizens these types of protections, and in many cases legal consideration of personal data, particularly that collected and transmitted in digital form, is rudimentary at best, if not wholly absent.¹

2. Nature of genealogical research

The nature of genealogical research, as described by the National Genealogical Society in the United States, requires not only the discovery, use and interpretation of personal data, but also the publication of that data in various forms, to support further research and discovery (National Genealogical Society 2016: “Guidelines for Sound Genealogical Research”). The Society offers extensive instruction in their “Guidelines for Sharing Information With Others on the dissemination of genealogical research, noting especially that “legal rights of privacy may limit the extent to which information from publicly available sources may be further used, disseminated, or published” (2016, 2). However, this is not always clearly understood by the public at large who operate under the misapprehension that everything is free on the internet or through misapplication of the doctrine of fair use to take whatever they find and incorporate this into their family trees with neither permission nor attribution (McClure 2003).

1 See Tashea (2019/2020) and Zeevi (2019) for extensive discussion of the legal situation of data privacy laws in the United States, particularly in reference to genetic information.

3. Privacy and government records

The paradox of privacy is best illustrated in the case of government vital records, or records of births, deaths and marriages. On the one hand, these are public information, and subject to open information regulations, while on the other they also constitute personal information and should enjoy some measure of protection to protect the privacy of the subject of the each record.

An examination of information practices within the United Kingdom can illuminate some of the difficulties. The Guides to the Statutory Registers on the Scotlands People website, the official website of the Scottish Government for providing access to records and archives, maintains that digital images of original items in the statutory registers are restricted for internet research to items older than 100 years for birth entries, older than 75 years for marriage entries, and older than 50 years for death entries. Items available on the internet may be viewed, saved or downloaded for a fee, on a pay-per-view basis.

Indexes of the registers, however, that show name(s), year and location of event (where registered), as well as mother's maiden name (death entries after 1974) are available up to and including the current year. Access to more recent records is available either through the purchase of copies of the online-embargoed records or through an in-person visit (subject to a search fee) to one of several research centres where all records up until a few months before may be viewed. Mechanical copies may not be made or printed, although notes may be taken. So all vital record information is available for a price, although not always through the internet searching portal.

For England and Wales the situation is slightly different, as the General Register Office "Research Your Family History..." guide reveals that online indexes are only available for non-embargoed records (100+ year births, 75+ year marriages, and 50+ year deaths). Digital images are not available online for any records, although copies may be purchased for both historical and current records. So access to records is much more restricted in England and Wales, and copies of public records are only available at a greater cost than comparable records in Scotland.

3.1 1939 Register

An exactly opposite situation exists in the United Kingdom for a record set known as the 1939 register. Originally compiled at the outbreak of World War II, the 1939 Register formed the basis for all wartime activities involving the general population, such as issuing identity cards and rationing books for food and clothing (The National Archives, 2020). It also formed the basis of records upon the creation of the National Health Service in 1948 (National Records of Scotland 2020) and was in use for several decades until the advent of computer indexes. In England and Wales, 1939 Register entries are indexed and available through commercial partners as part of their regular subscription service for all deceased persons listed in the Register, regardless of when they died i.e. the records are not subject to the same 50-year after-death rule as vital records. However, in Scotland, although the 1939 Register records of deceased people are open, they are not indexed, nor are they made available online, so the only method of obtaining them is to request a search and to purchase copies of the register entries for each particular individual or family.

3.2 Census Records

Census records are a valuable source for genealogical research, allowing individuals and families to be tracked across their lifetimes in five or ten-year increments. Differing privacy laws and confidentiality concerns result in varying degrees of accessibility for genealogists, regardless of the original assurances or guarantees made to census informants at the time of data collection. The United States Census Bureau releases census records 72 years following the official census day, which makes it possible for the personal information of anyone over the age of 72 to be publicly displayed. Canada provides a longer embargo period of 95 years for all census data up to and including 1926, although all census data gathered after this date will no longer be made public (Libraries and Archives Canada 2019). All component jurisdictions of the United Kingdom follow a 100-year rule for the publication of census data (Scotlands People, Guides: Census

returns”) although the release of the 1921 census in 2021 will be last for some time as the data of the 1931 census for England and Wales was destroyed by fire during the Second World War, and no census was taken in 1941, because of the war. The 1931 census in Scotland was stored separately and so survived, although equally no data was gathered in 1941. Although technically not an official census, the 1939 register provides much the same data, and its release only 73 years after the data was gathered reveals an inconsistency in the application of personal information privacy rules across the different but similar data sets in the United Kingdom.

There are contrasting situations as well in Australia and New Zealand. New Zealand also follows a 100 year privacy rule. Although earlier legislation forbade the release of individualized data from a census, the Public Records Act of 2005 (quoted in StatsNZ, “A History of Census-taking in New Zealand.”) did allow for the publication of individual returns after a 100 year embargo period. This contrasts with the approach of the Australian government, where not only is individual information not released at any time, but all individual information was destroyed after compilation of the official census statistics for all censuses from 1901 to 1996 (State Library Victoria 2020). Provision now exists for individuals to give permission for their individual information to be retained and published, although the minimum embargo period will be 100 years from the date of data collection.

4. Genetic genealogy

Another area of concern in genealogy for privacy issues is that of DNA (deoxyribonucleic acid) testing or genetic genealogy. Many commercial enterprises now offer DNA testing kits with the offer of providing a quick way to identify ethnicity, probable place of origin, and matches to other family members based on comparisons within each company’s database. Options also exist for downloading this data and uploading it to different sites that allow comparison across other contributed samples, regardless of which company provided the original DNA analysis.

While this may provide benefit for the volunteer participants, law enforcement agencies, particularly in the United States have taken advantage of these public DNA databases to search for familial matches to DNA taken from crime scenes, many of them long since relegated to the cold case file. While there have been some notable successes in identifying violent offenders, Zeevi (2019) notes that there are currently no legal restrictions in employing public DNA databases for this type of criminal investigation. The absence of legislation means that such exploration may not be confined to major crimes, but may be expanded to other types or levels of crime and with no recourse from the owners of the DNA samples or their relatives to the use of their data, or its misuse through error or the dissemination of embarrassing or unwanted disclosures that often occur in genealogical DNA testing.

As Greytak et al (2019) note, the high profile cases of law enforcement agencies solving cases by matching DNA results of crime-scene evidence or suspects' DNA with DNA profiles created for genealogical purposes, has so far only been done by using DNA results that were voluntarily uploaded to a specific genealogical DNA site. The complete testing results were not made available to law enforcement by the testing companies, but rather the individuals concerned chose to download private test results from one website and publish them to another, public site. This does not preclude attempts in the future to compel private DNA testing companies to turn over DNA data by legal means, but this is not currently the case, and so the privacy issues concerned are different, being wholly within the control of the DNA sample subject. Davidowitz (2019) reveals that the DNA sharing site in question has now changed its terms of service to require new users to opt in if they wished their records available for any law enforcement use, and automatically set all users already in the system to opt out of law enforcement searches.

However there have been cases where DNA testing companies have allowed direct access to their DNA databases by law enforcement agencies in the United States, although not all companies have permitted this use of their databases. Tashea (2019/2020) describes how the two companies that have allowed access by law enforcement agencies in the past have changed their terms of service to restrict the types of crimes where they

would allow this type of investigation, following a backlash from users. There have also been disagreements in interpretation as to whether certain crimes were serious enough to fit within the definitions that permitted searches.

5. Contradictory privacy behaviours

Another of the difficulties encountered in the consideration of privacy and genealogy is the issue of contradictory behaviours. While people may assert the wish to privacy of their family connections in the realm of genealogy, they may at the same time make detailed genealogical information available in other formats. In the United States, for example, it is quite common for obituaries in newspapers to contain extremely detailed information on family relationships. Such obituaries may contain not just birth and death dates of the deceased, but the names of the parents (including the mother's maiden name); grandparents; names and current location of siblings, including their spouses; the names, spouses and locations of children, grandchildren, and other relations of both older and younger generations. All countries are moving toward online databases of obituaries and other family notices, where family information that formerly may have been confined to the person's local community is now available instantly from any location.

Grave markers are another source of family information. While these have always contained the name and often the dates of all those buried in the marked grave, modern practice often includes names of living family members on the grave markers as well. With the advent of websites such as BillionGraves.com and FindAGrave.com transcribing, collating, indexing and publishing this information online, it is straightforward to search for and locate this information, as it has been made public.

5.1 Naming conventions

The conventions used to describe women following their marriage can also lead to information disclosure. While some

family researchers always choose to refer to females by their birth name, others may refer to them by their married name, including their birth name as a middle name or in parenthesis before or after their married surname. Various software programs also default to variations on these practices. The results can be that even when genealogy software is hiding people's names because they have been designated in the family tree as living people or private entries i.e. not to be published, the fact that a designation of private followed by a last name and then another name in parenthesis allows the reader to identify immediately that there was a female child in the family who is now married. The presence of both names makes it a straightforward search for a marriage, resulting in identification of the child and her spouse.

6. Online family trees

All the commercial genealogy database companies encourage the creation of online family trees. The reason is two-fold. In the first instance, providing the "free" storage ensures that the customers will continue to subscribe to the database to maintain access to their research results. In the second place, these family trees provide additional information to the database companies that can then be exploited to provide another record set to serve the subscribers and also information to increase the availability and accuracy of research hints, or suggestions of related records for further exploration. Even although family trees may be designated as private trees, so the results are not shared directly with other searchers, some databases do provide hints that reveal the existence of matching data within private trees – and provide a means for contacting the owner of that family tree.

For trees that are not designated as private, McClure (2003) notes that once a tree is released to one person, it is released to all. She also discusses the problems of the lack of appropriate attribution and citation, where public trees are incorporated complete into other family trees, with no acknowledgement or permission. Publication is interpreted as consent to replicate

a republish and the original researcher has lost control of their work.²

7. Individual vs. Family Consent

Wallace et al (2015) conducted an extensive study of online entities offering genealogical services including DNA testing, record discovery and family tree storage and publishing. The authors were concerned with the idea of consent to publish and the differences between the consent to publish information concerning only the one person granting the permission, and the consent to publish that included the personal information of a large number of third parties who were not party to the agreement. They noted that while some entities, primarily those offering record discovery and family tree storage and publishing without DNA testing, offered more specific warnings on the need to ensure agreement of all affected individuals in the publication of any family genealogical information, those companies that only offered or included DNA testing services focused much less of the consent of others potentially affected and concentrated on ownership and usage rights.

The authors raise the important issue of privacy concerns going beyond a single individual who may be in possession of genealogical information to including all those living (and perhaps future?) family members whose privacy may be compromised by the creation and publication of family trees, including those containing DNA data. To what extent should consent to publish be required of all potentially involved parties? Is it sufficient to leave this up to the individual in possession of the family tree with the assumption that submission for publication through a particular genealogical website implies agreement of all involved parties, or should a more robust consent process, what they term “generational consent” be created? (2015:9).

2 This wholesale copying without attribution can also lead to the propagation and multiplying of error. The more an inaccuracy is repeated in print, and the more trees in which it appears, the greater its credibility and the harder it is to correct.

8. Conclusion

All genealogical endeavors are concerned with the discovery of personal data, some of it of people long since dead, but much of it of people still living. Lack of privacy legislation that addresses online data or DNA information, as well as inconsistencies in the handling of privacy of public records lead to a confusing situation where recreational genealogists in particular have little real understanding of the long-term implications of their sharing and publishing actions, especially where this affects individuals other than themselves. More education in privacy rights and the consequences of renouncing what rights an individual possesses could form the basis for more informed decision making, including the consideration of the potential harms to other individuals.

Sources and Literature

- American Library Association. "Professional Ethics: Code of Ethics". 2020. <http://www.ala.org/tools/ethics>
- Carpenter, Todd A. "NISO: Striving for a Consensus Framework for Patron Privacy." *Computers in Libraries*, vol. 35, no. 5, June 2015, pp. 4–7.
- CILIP: The library and information association. "Code of Professional Practice for Library and Information Professionals." 2012. https://archive.cilip.org.uk/sites/default/files/media/document/2017-10/code_of_professional_practice_for_library_and_information_professionals_oct_2012.pdf
- Davidowitz, Shanni. "23andEveryone: Privacy Concerns With Law Enforcement's Use Of Genealogy Databases To Implicate Relatives In Criminal Investigations." *Brooklyn Law Review*, vol. 85, no. 1, Fall 2019, pp. 185–215.
- European Union. "Charter Of Fundamental Rights Of The European Union." 2012. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>
- General Register Office. "Research Your Family History Using the General Register Office." <https://www.gov.uk/research-family-history> Accessed 28 Apr. 2020.

- Greytak, Ellen M., et al. "Privacy and Genetic Genealogy Data." *Science* (New York, N.Y.), vol. 361, no. 6405, Aug. 2018, p. 857. doi:10.1126/science.aav0330.
- Libraries and Archives Canada. "Censuses." 2019. <https://www.bac-lac.gc.ca/eng/census/Pages/census.aspx>
- McClure, Rhonda R. "Ethics in Genealogy." 27 Mar. 2003. <https://www.genealogy.com/articles/twigs/rhonda032703.html>
- The National Archives. "How to look for records of... The 1939 Register." <https://www.nationalarchives.gov.uk/help-with-your-research/research-guides/1939-register/> Accessed 28 Apr. 2020.
- National Genealogical Society. "Guidelines for Sharing Information With Others." 2016. https://www.ngsgenealogy.org/wp-content/uploads/NGS-Guidelines/Guidelines_SharingInfo2016-FINAL-30Sep2018.pdf
- National Genealogical Society. "Guidelines for Sound Genealogical Research." 2016. https://www.ngsgenealogy.org/wp-content/uploads/NGS-Guidelines/Guidelines_Sound-Research2016-FINAL-30Sep2018.pdf
- National Records of Scotland. "National Health Service Central Register (NHSCR): About the Register." 2020. <https://www.nrscotland.gov.uk/statistics-and-data/nhs-central-register/about-the-register/the-history-of-the-register>
- ScotlandsPeople. "Guides: Census returns." <https://www.scotlandspeople.gov.uk/guides/census-returns> Accessed 28 Apr. 2020.
- Scotlands People. "Guides: Statutory Registers." <https://www.scotlandspeople.gov.uk/guides/statutory-registers> Accessed 28 Apr. 2020.
- State Library Victoria. "Early Australian Census Records: Records After 1901." 21 Apr. 2020. <https://guides.slv.vic.gov.au/earlycensus/after1901>
- StatsNZ. "A History of Census-taking in New Zealand." <http://archive.stats.govt.nz/Census/2013-census/info-about-the-census/intro-to-nz-census/history/history-summary.aspx> Accessed 28 Apr. 2020.
- Tashea, Jason. "Genealogy Sites Give Law Enforcement a New DNA Sleuthing Tool, but the Battle over Privacy Looms." *ABA Journal*, vol. 105, no. 8, Winter 2019/2020 2019, p. N.PAG.
- United States Census Bureau. "When will census records be available?" https://www.census.gov/history/www/faqs/genealogy_faqs/when_will_census_records_be_available.html Accessed 28 Apr. 2020.

Wallace, Susan E., et al. "Family Tree and Ancestry Inference: Is There a Need for a 'Generational' Consent?" *BMC Medical Ethics*, vol. 16, Dec. 2015. DOI 10.1186/s12910-015-0080-2

Zeevi, Jamie M. "Dna Is Different: An Exploration of the Current Inadequacies of Genetic Privacy Protection in Recreational Dna Databases." *St. John's Law Review*, vol. 93, no. 3, July 2019, pp. 787–808.

Andrew J. M. Smith

ПРИВАТНОСТ ИНФОРМАЦИЈА У ОБЛАСТИ ГЕНЕАЛОШКИХ ИСТРАЖИВАЊА

Овај рад истражује питања приватности појединаца на која се наилазило током генеалošких истраживања. Расправља се о објављивању јавних записа, укључујући виталне записе и пописне податке, као и о све већој улози генеалošких ДНК профила у истраживању породичне историје. Такође се разматра улога библиотека у одржавању приватности уз потицање генеалošких истраживања.

Кључне речи: приватност информација, генеалогија, витални записи, генетичка генеалогија, информациона политика, сагласност.