

# ONE HUNDRED YEARS OF PRIVACY LAW: A LATENT SEMANTIC ANALYSIS

By

Robert Sprague,<sup>\*</sup> Kevin Grauberger,<sup>\*\*</sup> and Nicole Barberis<sup>\*\*\*</sup>

*Privacy, like an elephant, is perhaps more readily recognized than described.*<sup>1</sup>

## I. INTRODUCTION

Privacy, being an evolutionary product of social development,<sup>2</sup> has been a human need and desire for millennia. Privacy law scholarship, in contrast, is a relatively recent phenomenon. Of all the privacy-related law review articles published in the history of the United States, for example, over ninety percent were published after 1990—and half that amount in the past decade. Within this recent profusion of scholarship lies a conundrum: there is no clear definition of privacy;<sup>3</sup> there is not even consensus of what would constitute an adequate description. Fundamental concepts associated with privacy have been identified and analyzed—for example, seclusion, intimacy, surveillance, anonymity, and control of information. But, as Nissenbaum has noted, most calls for privacy arise from context, as well as advancing technologies,<sup>4</sup> meaning the legal system often has difficulty identifying and protecting rights to privacy. Without a coherent construction of privacy principles shared by the community of scholars, the legal discipline will never explicitly articulate those principles.<sup>5</sup>

This paper reports preliminary results from a research project aimed at identifying fundamental privacy law principles derived from the writings of legal scholars and commentators using probabilistic topic modeling, which is comprised of a suite of algorithms that attempt to discover hidden thematic structures in large archives of documents.<sup>6</sup> Topic modeling algorithms are statistical methods that analyze the words of texts to discover topics (themes) contained within, how those topics are connected to each other, and how they change over time.<sup>7</sup> A latent Dirichlet allocation process, which identifies sets of terms that more tightly co-occur, is incorporated into the topic modeling analysis to identify words most closely associated with each identified topic.<sup>8</sup> The latent Dirichlet allocation therefore provides insight into the context in which each identified topic occurs.

---

<sup>\*</sup> J.D., M.B.A. Associate Professor, University of Wyoming College of Business Department of Management & Marketing.

<sup>\*\*</sup> J.D., University of Nebraska College of Law, 2014.

<sup>\*\*\*</sup> M.S. in Applied Statistics, University of Wyoming, 2013. Nicole Barberis is a data scientist and statistician.

The authors are extremely grateful for the support and assistance they have received from: Professor Ken Gerow, Chair, University of Wyoming Department of Statistics; Abigail Fournier, J.D., University of Wyoming College of Law, 2013, and Kellsie Jo Nienhuser, J.D., University of Wyoming College of Law, 2015, for their long-suffering work in converting law articles from pdf to plain text format; as well as the University of Wyoming College of Business.

Most published law review articles that cite to Samuel Warren's and Louis Brandeis's seminal article, *The Right to Privacy*<sup>9</sup> (some 3000 articles) are being converted to plain text to build a corpus of publications focused on the law of privacy. *The Right to Privacy* was selected as the focal point of the document collection because it is the original published scholarly call for a formal legal right to privacy in the United States; hence, the vast majority of privacy law publications cites to it. Probabilistic topic modeling using latent Dirichlet allocation is being applied to the document collection in time slices to reveal the evolution of fundamental privacy law concepts expressed in the legal literature published from 1890 to the present. Studies in different disciplines have demonstrated the ability of latent Dirichlet allocation to analyze the rich underlying structures of a particular domain—depicting emerging and sustained trends in a given discourse.<sup>10</sup> The ultimate goal of this project is to identify the fundamental conceptual structure of privacy law in the United States as reflected by over a century of published law review and journal articles.

This paper first presents an overview of the origins of privacy law to provide a backdrop for the motivations underlying this project. It then provides an overview of the topic modeling process using latent Dirichlet allocation to explain and validate the underlying analytical methodology. Preliminary results from applying the statistical modeling to a collection of published law review and journal articles are then presented and discussed—specifically, fundamental topics identified from published articles in the following time slices: 1891 – 1950, 1951 – 1960, 1961 – 1970, 1971 – 1980, 1981 – 1990, and the full corpus (to date), 1891 – 1990. We identify four principal themes in privacy law scholarship—Appropriation, Disclosure, Intrusion, and Observation—which we link with each identified topic.

## II. ORIGINS OF PRIVACY LAW

What is privacy? And, in particular, what legal rights to privacy does U.S. law protect? What legal rights to privacy should U.S. law protect? This conundrum has vexed legal scholars and commentators for over a century.

There have been attempts to describe, more than define, privacy. Warren and Brandeis, of course, most famously asserted the right to privacy as the right to be let alone.<sup>11</sup> In 1905, the Georgia Supreme Court described the right of privacy as “the right of a person to be secure from invasion by the public into matters of a private nature.”<sup>12</sup> A half-century later, Alan Westin described privacy as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.”<sup>13</sup> Westin identified four states of privacy: (1) solitude, where an individual is separated from the group and freed from the observation of others; (2) intimacy, where an individual exercises seclusion to achieve a closer relationship with others; (3) anonymity, the freedom from identification and surveillance in public places or while performing public acts; and (4) reserve, the creation of a psychological barrier against unwanted intrusions.<sup>14</sup> Contemporaneously, Margaret Mead described privacy as “the right to live part of one's life out of the public eye, according to one's own choice, and free from interference by others.”<sup>15</sup> More recently, Solove identified six types of privacy: (1) the right to be let alone; (2) limited access to the self; (3) secrecy; (4) control over personal information; (5) personhood; and (6) intimacy.<sup>16</sup>

These notions of seclusion, withdrawal from observation, and intimacy reflect human concerns since at least the second century A.D.<sup>17</sup> By the end of the first millennium, privacy had evolved to the concept “that there are acts, individuals, and objects by law not subject to public

authority; for that reason they are situated in a precisely delineated sphere designed to thwart any attempt at intrusion.”<sup>18</sup> By this time the home had become an “intimate haven.”<sup>19</sup> This trend continued throughout the second millennium as society became more urbanized and, hence, public life more anonymous; home and family became the center of seclusion from others.<sup>20</sup> Ultimately, social taboos against a desire for solitude began to fade by the end of the seventeenth century.<sup>21</sup>

In the formative years of the United States, eavesdropping was the principal form of privacy concern—preventing someone from standing under the eaves of a home and listening to the conversation inside.<sup>22</sup> However, by the late nineteenth century, two developments combined to threaten further individuals’ privacy—the first was “keyhole journalism,” reflecting the growth of newspapers printing salacious material;<sup>23</sup> the second was the advent of technologies such as audio recording devices and instantaneous photography.<sup>24</sup> These developments piqued one individual in particular:

On January 25, 1883, [Samuel] Warren had married Miss Mabel Bayard, daughter of Senator Thomas Francis Bayard, Sr. They set up housekeeping in Boston’s exclusive Back Bay section and began to entertain elaborately. The *Saturday Evening Gazette*, which specialized in “blue blood items,” naturally reported their activities in lurid detail. This annoyed Warren, who took the matter up with [Louis] Brandeis.<sup>25</sup>

The rest, as they say, is history.<sup>26</sup> One can see reflected in *The Right to Privacy* concerns of intrusive technology recording and publishing intimate private life.<sup>27</sup> Warren and Brandeis were searching for some form of legal recourse to prevent such intrusions. They recognized that traditional libel and slander torts would not address the type of harm they had in mind—“our law recognizes no principle upon which compensation can be granted for mere injury to the feelings.”<sup>28</sup> Recognizing that “[t]he common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others,”<sup>29</sup> Warren and Brandeis argued that traditional intellectual property rights, such as copyright, were insufficient.<sup>30</sup> Instead, they argued, “the right to privacy, [w]as a part of the more general right to the immunity of the person,—the right to one’s personality.”<sup>31</sup>

It is not as if the notion of privacy did not exist in the United States prior to the publication of *The Right to Privacy*. As discussed above, eavesdropping had long been recognized as an actionable harm.<sup>32</sup> In 1811, a Louisiana court prohibited a newspaper editor from publishing the contents of a letter without permission from the sender,<sup>33</sup> and in an 1881 case in which a doctor allowed an unmarried man with no medical training to be present while a woman gave birth, the court acknowledged the woman’s right to privacy during “a most sacred” occasion.<sup>34</sup>

But while Warren and Brandeis are recognized for putting into motion a formal judicial recognition of the right to privacy,<sup>35</sup> they did not end the discussion of exactly what constitutes a legally protected right to privacy. As Zimmerman argues, “Lawyers and philosophers have generated a vast literature on the subject without being able to agree upon some core of values or interests common to each of the cases in which the ‘right to privacy’ has been applied.”<sup>36</sup> Zimmerman concludes, “A part of the difficulty is that the phrase today is a catch-all, attached to a broad range of interests which often have little or nothing to do with the tort originally envisioned by Warren and Brandeis.”<sup>37</sup> For example, Warren and Brandeis were concerned with intimate details being revealed to the public. What of today’s near-ubiquitous electronic tracking and surveillance? “Increasingly, what we are worried about are practices that involve collecting,

using and disclosing information that is not sensitive or intimate and that is increasingly collected in public—concerns that do not easily fall within the domain of traditional privacy theory.”<sup>38</sup>

The frustration is clearly reflected in the literature. Robert Blakey wonders if privacy is “an idea of which there are two hundred definitions,”<sup>39</sup> “better suited to literary than legal analysis.”<sup>40</sup> One commentator opined nearly fifty years ago:

No other area in the law has caused so much confusion to the courts than that dealing with the right of privacy. The doctrine of privacy is still very much in its infancy seventy-five years after its conception. Only the barest outlines of its nature and extent have as yet been sketched by judicial decisions.<sup>41</sup>

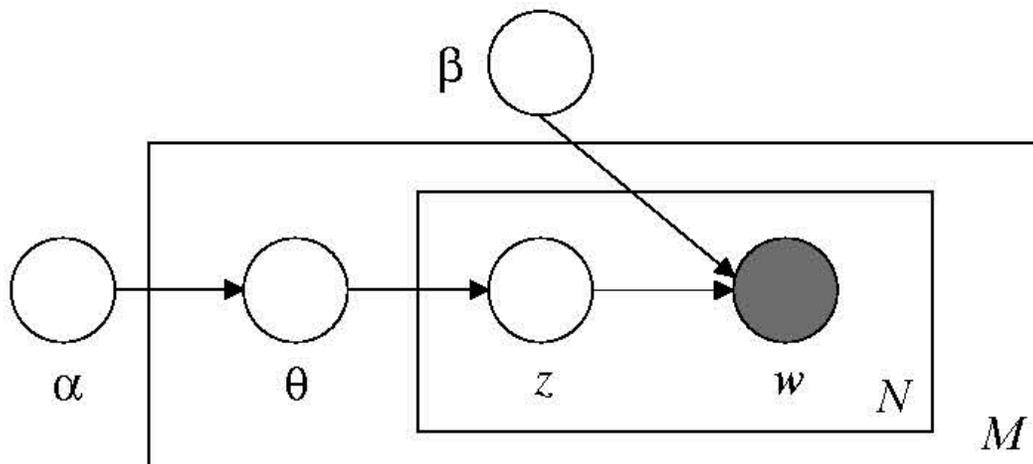
More recently, regarding being named with Daniel Solove to lead a project to compile a *Restatement (Third) Information Privacy Principles*, privacy scholar Paul M. Schwartz stated “that the project is designed to bring clarity to a body of common law that has become a bewildering array of conflicting and overwhelming rules differing from state-to-state and the federal government.”<sup>42</sup>

There have been approximately 3500 law review articles published in the United States that cite to *The Right to Privacy*,<sup>43</sup> and nearly 5000 law review articles contain the word “privacy” in their titles.<sup>44</sup> Owing to the quantity—not to ignore quality—of publications on the topic, it is clear there is much to be said about privacy. It is not a settled, nor static, issue.

This paper lays the foundation, and presents preliminary data, not for a definition or categorization of privacy, but for an ontology of privacy—an interpretation of the scope of privacy without assuming that anything exists beyond written expressions; namely, from published law review and journal articles.<sup>45</sup>

### III. BUILDING AN ONTOLOGY USING PROBABILISTIC TOPIC MODELING

Probabilistic topic modeling represents a computational method to understand digitized data.<sup>46</sup> The basic intuition behind latent Dirichlet allocation (LDA), the simplest form of topic modeling, is that documents exhibit multiple topics.<sup>47</sup> LDA is a statistical model of document collections (corpora) that tries to capture this intuition.<sup>48</sup> More formally, LDA is a generative probabilistic model of a corpus in which documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words.<sup>49</sup> See Appendix A for a conceptual and algorithmic description of the computational process.



**Figure 1**

Figure 1 is a graphical model representation of LDA.<sup>50</sup> The outer box represents documents (with  $M$  representing the corpus of documents), while the inner box represents the repeated choice of topics and words within a document (with  $N$  representing the total number of words);  $\alpha$  is the parameter needed for the Dirichlet distribution;  $\theta$  is a vector of topic proportions (one vector for each document);  $Z$  represents the topic that a word was assigned to;  $w$  is the observed word in the text; and  $\beta$  represents the discovered topics (themes).

For our project, this process is applied to document collections comprised of published law review and journal articles in order to discern fundamental privacy law concepts.

### **A. Methodology**

The initial document collection is comprised of relevant published law review and journal articles that cite to Warren’s and Brandeis’s *The Right to Privacy*,<sup>51</sup> as identified within the Westlaw and HeinOnline collections.<sup>52</sup> *The Right to Privacy* was selected as the initial means to identify relevant articles because, as noted previously,<sup>53</sup> it is recognized as the seminal scholarly article advocating the recognition of a formal legal right to privacy within U.S. law; consequently the majority of published articles which address privacy cite to *The Right to Privacy*.<sup>54</sup> All documents selected for the collection are converted to plain text. In addition, all titles, author names, section headings, footnotes and endnotes, and supplemental materials are removed in an effort to create a collection limited to addressing substantive privacy law issues.<sup>55</sup> As of the writing of this paper, only approximately one-third of the anticipated initial collection, representing privacy law articles published up to and through 1990, has been converted. The current document collection has been divided into the following time-slice corpora<sup>56</sup>: 1891 – 1950 (which includes *The Right to Privacy*),<sup>57</sup> 1951 – 1960,<sup>58</sup> 1961 – 1970,<sup>59</sup> 1971 – 1980,<sup>60</sup> and 1981 – 1990.<sup>61</sup> A cumulative corpus has also been created for the time period 1891-1990.<sup>62</sup> Work is continuing to build complete corpora for: 1991 – 2000, 2001 – 2010, and 2011 – present.

The LDA topic model algorithms are then applied to the corpora using the popular—at least within the topic model community—software implementation, MALLET.<sup>63</sup> Data, in the

form of plain text, are first imported into MALLET. Once the data are imported, MALLET provides a number of output options. The critical output files used in this project include:

- A “keys” file that provides an overview of the topics identified within the data; it contains a “key” consisting of the top k words for each topic (where k is the number of most probable words for each topic after model estimation). This file also reports the Dirichlet parameter of each topic that will be roughly proportional to the overall portion of the collection assigned to a given topic—in other words, the weight of the latent topic and the words that compose that topic (the cluster of words that define the latent topic).
- A “composition” file that lists the source documents and reports the rank of topics identified with each document.

The reported weight of each topic represents the strength of likelihood that the collection of words represented within the topic conveys a meaningful concept. A lower weight generally means that the words associated with the topic do not appear very frequently within the corpus or they co-occur less frequently. This notion is also reflected in the “composition” file where one generally sees more documents linked to higher weighted topics; in other words, a topic is more significant because it is discussed more frequently (using the words that appear in the topic), which supports a key concept behind this project—the essence of privacy is identified by the words of the authors themselves.

There are three variables that were used to define the parameters of each MALLET topic model “run.” The number of iterations was increased from the default of 1000 to 2000; this variable controls how many samples will be taken of the corpus. According to MALLET, increasing this variable can increase the quality of the topic model. We also chose to create topics that listed just the first twenty identified words. The MALLET “keys” file lists the words in each topic in descending order of frequency of occurrence; we found that for the most part the fifteenth to twentieth words reflected very low frequencies.<sup>64</sup>

Perhaps the most critical variable to select in any topic modeling exercise is the number of topics to identify. There is no consensus as to the optimum number of topics.<sup>65</sup> Since our entire corpus is relatively small at this point (1182 documents), we were able to run multiple topic models with 20, 25, 50, 100, 200, and 400 topics. We determined that twenty-five topics appeared to be optimal; less than twenty-five underfit the data (i.e., we risked missing some topics), while models with fifty or more topics overfit the data—we found the topics were diffused, as if MALLET were trying to identify too many topics.<sup>66</sup>

We ran three topic models on each corpus. Since topic models are unsupervised, meaning no prior assumptions are made about the data,<sup>67</sup> each model begins with a random assignment of topics that is then honed through each iteration. As a result, while they may be quite similar, no two topic models run against the same data will be identical, both in terms of the topic weights as well as the exact words comprising each topic. In order to identify what appeared to be the most relevant topics among the runs, we developed a scoring system that first identified words in each topic that appeared in all three runs. We then scored each topic, among the three runs, based on the number of three-word-matches, the weight of each topic, and how close the individual weights were to each other.<sup>68</sup> The scoring was not the definitive determination of whether a topic was identified for our results; it did, however, significantly aid in identifying the results among the topics within the three runs.

Finally, we employed a stop words list to filter out common prepositions and conjunctions. We customized the stop words list with a few additional generic legal terms such as plaintiff, defendant, law, case, court, held, etc.<sup>69</sup>

## **B. Results**

As noted previously, LDA topic modeling does not identify single-word topics; rather, it identifies probable topics consisting of multiple, clustered terms. As such, identifying and reporting in this paper substantive privacy law topics from the corpus involve filtering with subject-matter expertise. Many highly weighted topics were often generic from a legal sense—for example, comprised of terms such as decision, supreme, rights, constitutional, etc., but not terms that could be directly linked to a privacy-related concept. In addition, when a privacy-related topic was identified it was filtered itself by removing some extraneous terms, such as action, relief, act, etc. We also removed the word “privacy” from any topics in which it occurred since every document in the corpus theoretically addresses privacy. We also rearranged a few words within some topics to make them more comprehensible—e.g., the words “amendment” and “fourth” may have occurred in that order and perhaps been separated by a word but are represented in our reported topics as “Fourth Amendment”; the same with “false light.”

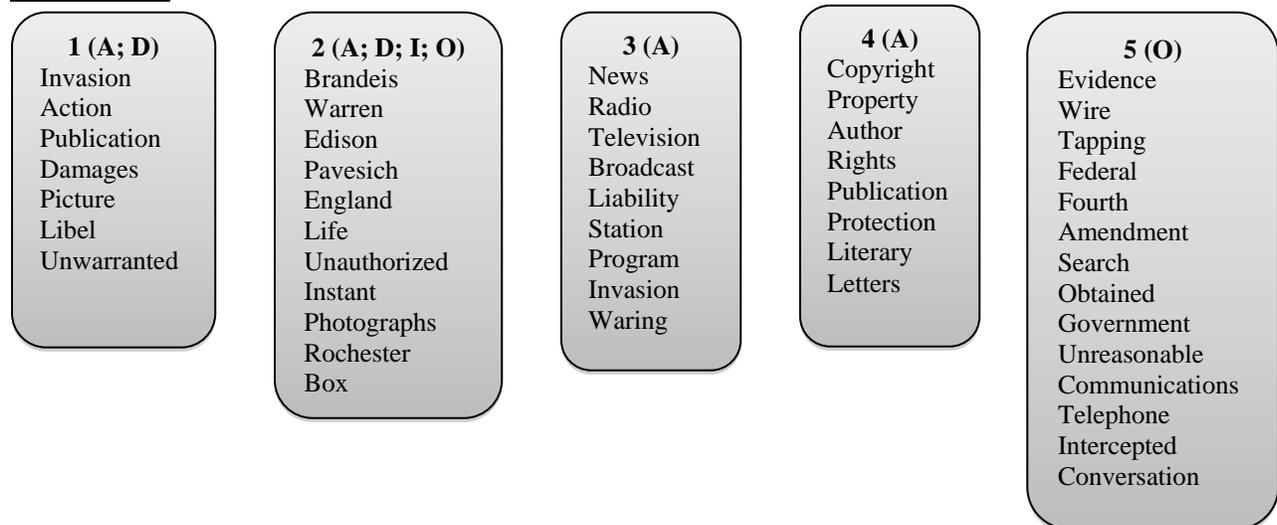
Our identified topics are presented below in within their respective time slices. The topics are numbered by their relative weight within each slice and the words are presented top to bottom essentially in order of frequency of occurrence, from highest to lowest. As such, one can view the topics as ranked in order of prevalence within each slice, and the terms comprising each topic as ranked in order of frequency of occurrence with their respective topic.

## **C. Analysis of Results**

Presented below is a discussion of the topics identified by the model; first categorized by the time slices and then concluding with respect to the entire corpus (to date). We elaborate on the identified topics within the time slices by noting privacy law developments and referencing representative publications.

We identify four principal themes revealed by the topic models: Appropriation—unauthorized use of a person’s likeness or intellectual works; Disclosure—often of personal, intimate facts, either generally or with resulting defamation (false light); Intrusion—not just upon seclusion per se but also into personal, intimate decisions, as well as compelling a person to testify; and Observation—including surveillance as well as collecting personal information. To some extent, some or all of these themes can be interrelated and an identified topic can certainly be linked to more than just one theme. The topics presented and discussed below include the linked theme(s)<sup>70</sup> as part of the topic number.<sup>71</sup>

## 1891 – 1950



Topics 1 (A) and 2 (A; D; I; & O) reveal that during the first sixty years, privacy law articles centered on the publication of pictures and photographs and particularly the unauthorized use of a person's likeness; also included are concepts of general discussion vis-à-vis the developing right to privacy in America. While Warren's and Brandeis's *The Right to Privacy* focused on threats to privacy posed by a gossip-hungry press aided by recording devices and instantaneous photography,<sup>72</sup> the two leading privacy cases early in the twentieth century both addressed the commercial use of a person's likeness without authorization—and reached opposite conclusions. In *Roberson v Rochester Folding Box Company* the New York Court of Appeals refused to recognize a common law right of privacy arising from the unauthorized publication of the plaintiff's likeness on flour boxes.<sup>73</sup> In contrast, the Georgia Supreme Court held that the unauthorized publication of person's picture for commercial purposes constituted an invasion of privacy in *Pavesich v. New England Life Insurance Company*.<sup>74</sup> As a result of *Roberson*, the New York legislature passed the country's first privacy legislation, but it was limited to the commercial appropriation of a person's likeness.<sup>75</sup> As such, much of the early published commentary focused on the development of privacy case law vis-à-vis *The Right to Privacy*. For example, a Comment in a 1920 *Cornell Law Quarterly* contrasted *Roberson* with two Louisiana cases that recognized a right of privacy involving the publications of individuals' pictures in a "rogues" gallery after they had been arrested but before conviction.<sup>76</sup> The Comment summarized the state of the right to privacy in America as of 1920: "Without exception. . . all jurisdictions have refused to protect this so-called right of privacy, unless there has been also an infringement of a property right or a breach of trust and confidence upon which the court could base the right to injunctive relief."<sup>77</sup>

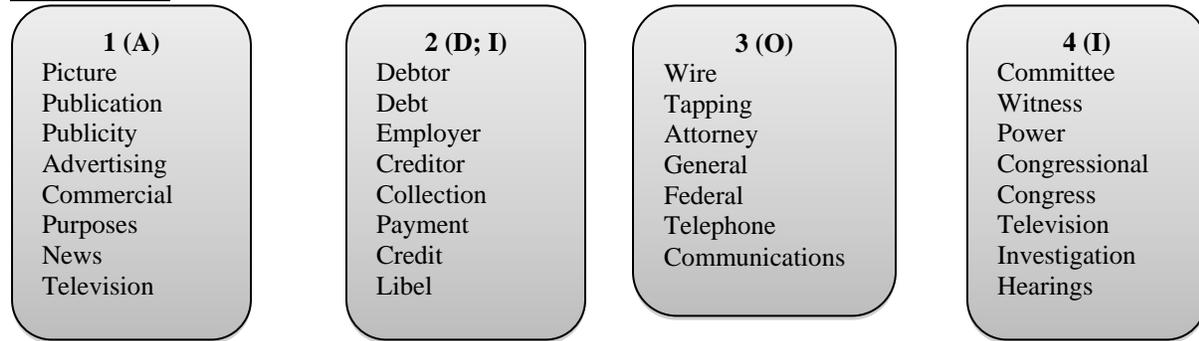
Closely related to this topic is an individual's right to privacy in broadcasts (Topic 3 (A)), which began to be addressed in the 1930s.<sup>78</sup> One comment succinctly rephrased the state of the right to privacy with respect to publications and broadcasts:

A reading of the decisions dealing with the subject leaves an impression of substantial lack of uniformity in the reasoning employed to support the right of privacy. Some recognize its existence only when property is involved, a contractual relationship can be implied, or when a relationship of trust can be established. A second group place recognition on a constitutional basis. Others

say an independent legal right of privacy exists and seek no further for its support.”<sup>79</sup>

Topic 4 (A) is associated with property rights in letters and other literary property—specifically, whether the right to privacy could prevent the unauthorized publication of private letters or other writings.<sup>80</sup> Finally, Topic 5 (I; O) reveals concepts well associated with privacy: wire tapping, the Fourth Amendment, unreasonable searches, and interception of communications.<sup>81</sup>

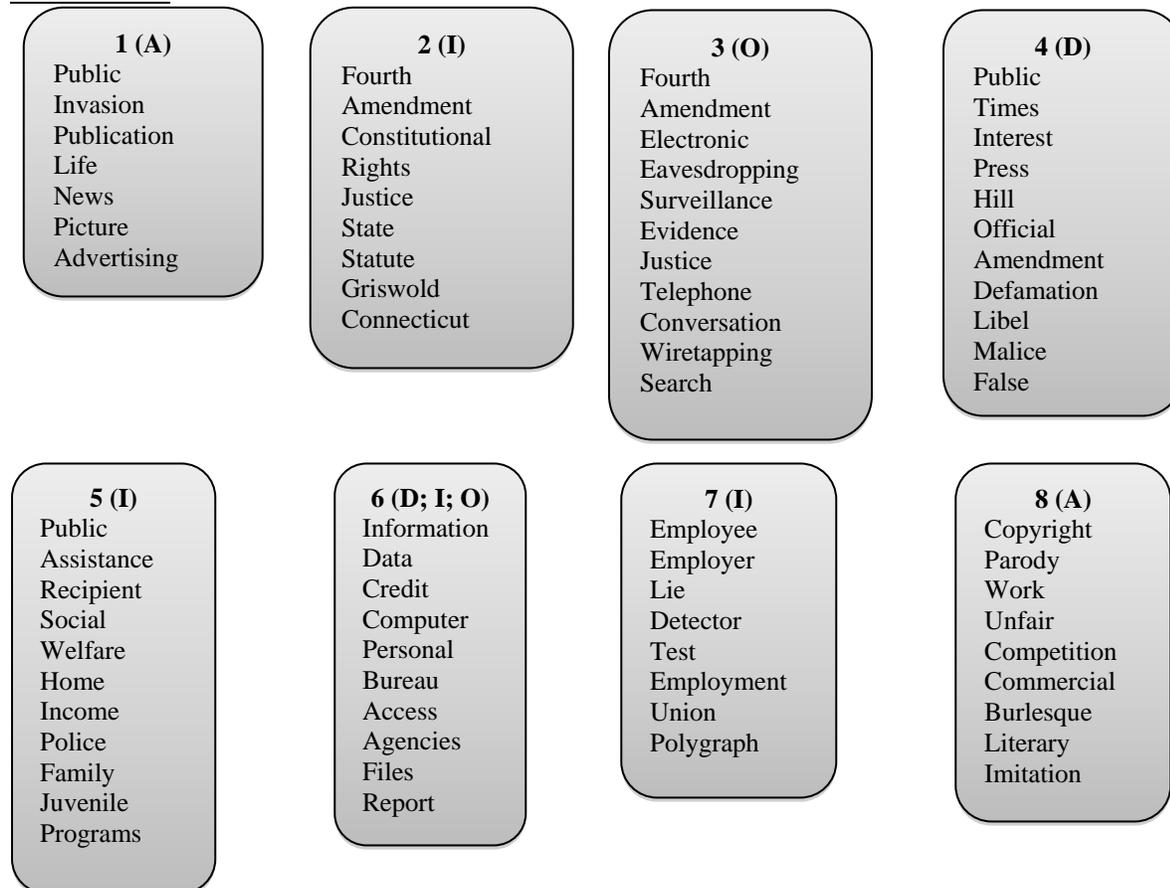
### 1951 – 1960



Continuing a major theme in the previous sixty years, articles published from 1951 – 1960 primarily focused on the unauthorized commercial use of a person’s name or likeness (Topic 1 (A)).<sup>82</sup> However, it was also in 1960 that the *California Law Review* published Prosser’s seminal article, *Privacy*, in which he articulated the four elements of an invasion of privacy that had developed in the states since the publication of Warren’s and Brandeis’s *Right to Privacy*: “1. Intrusion upon the plaintiff’s seclusion or solitude, or into his private affairs. 2. Public disclosure of embarrassing private facts about the plaintiff. 3. Publicity which places the plaintiff in a false light in the public eye. 4. Appropriation, for the defendant’s advantage, of the plaintiff’s name or likeness.”<sup>83</sup> The topics most frequently identified by the model through 1960 relate, however, primarily to the fourth element above: Appropriation, for the defendant’s advantage, of the plaintiff’s name or likeness.

The model also reveals discussion, once again, of wire tapping (Topic 3 (I; O));<sup>84</sup> though it is less prevalent in the literature.<sup>85</sup> Two “new” topics emerge in this decade. The first (Topic 2 (D; I)) relates to debt collections; specifically, whether multiple phone calls during the day and evening to the debtor’s residence invades the debtor’s privacy.<sup>86</sup> The second (Topic 4 (I)) relates to whether compelling a witness to testify before televised Congressional hearings violates the witness’s right of privacy.<sup>87</sup>

1961 – 1970

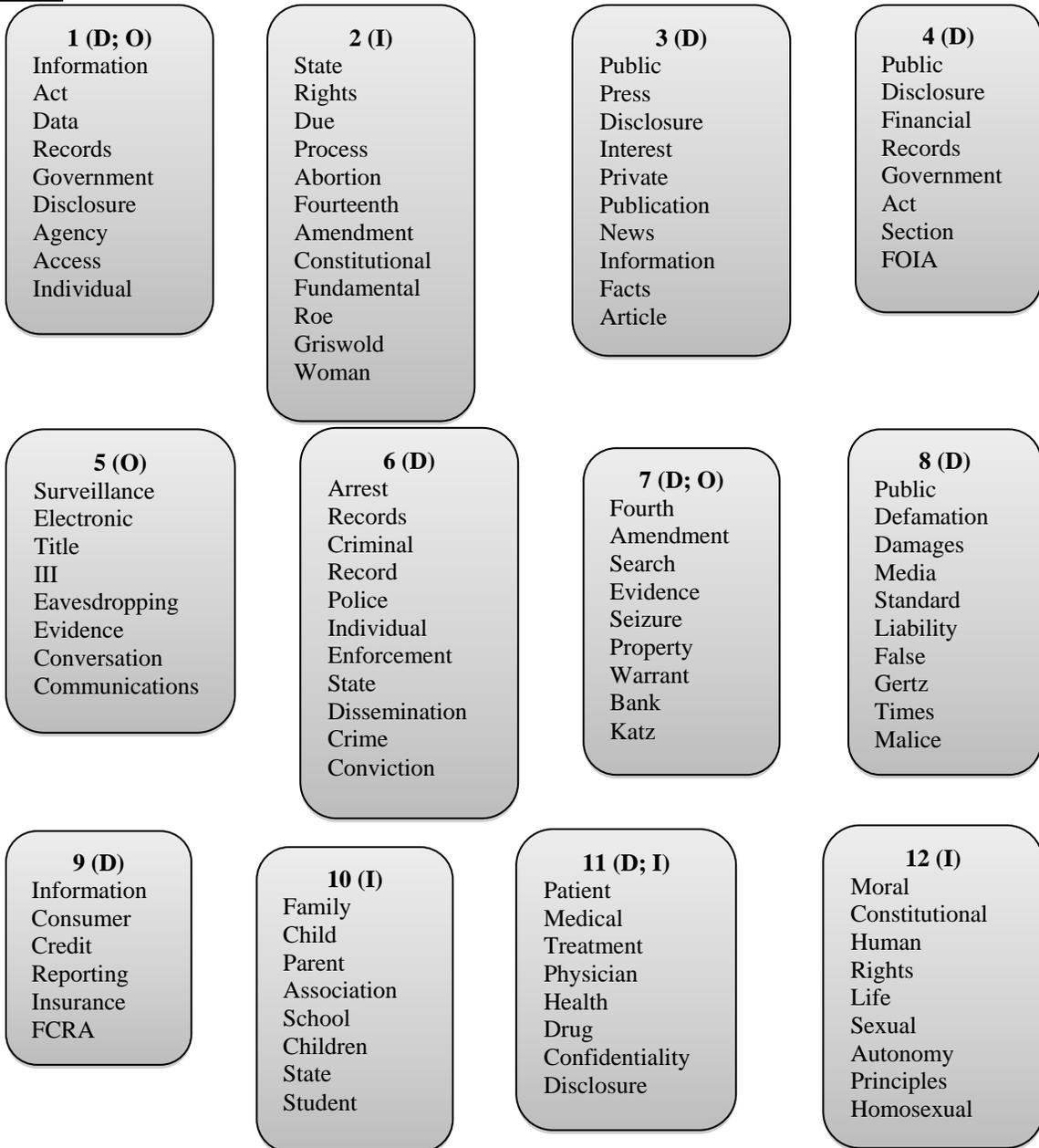


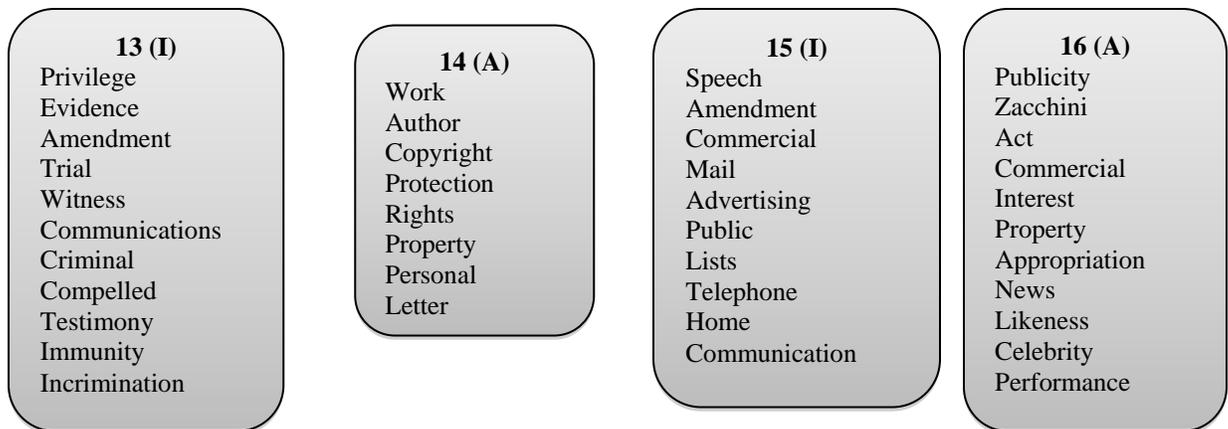
Once again, the most heavily weighted topic for this decade (Topic 1 (A)) relates to publication of a person’s name or likeness without his or her consent.<sup>88</sup> Topic 2 (I) arises directly from *Griswold v. Connecticut*,<sup>89</sup> in which the U.S. Supreme Court recognized a “zone of privacy created by several fundamental constitutional guarantees.”<sup>90</sup> Fourth Amendment searches (Topic 3 (I; O)), specifically through wiretapping, is heavily weighted due to the Supreme Court’s holding in *Katz v. United States*.<sup>91</sup>

Topic 4 (D) arises from discussion of *Time, Inc. v. Hill*,<sup>92</sup> which held that constitutional protections of free speech and the press preclude application of New York’s privacy statute to redress false reports of matters of public interest in the absence of proof that the defendant published the report with knowledge of its falsity or in reckless disregard of the truth, as compared with *New York Times v. Sullivan*,<sup>93</sup> which limited libel actions under the First Amendment.<sup>94</sup> Topic 5 (I) introduces the concept of privacy with respect to home visits by representatives of government agencies, such as for public assistance.<sup>95</sup>

Topic 6 (D; I; O) is the first identified topic directly related to computers and data collection, though it also continues Topic 2 (D; I) from 1951 – 1960 regarding credit.<sup>96</sup> Workplace privacy makes its first appearance in Topic 7 (I), reflecting concerns over employee polygraph tests.<sup>97</sup> The final topic in this decade (Topic 8 (A)) appears to be similar to the ongoing topic related to unauthorized use of name or likeness (Topic 1 (A) in both the 1891 – 1950 and 1951 – 1960 time slices), this time more directly related to imitations and parody.<sup>98</sup>

1971 – 1980





The heaviest-weighted topic (Topic 1 (D; O)) now relates to information. In particular, this decade was a time when scholars and commentators wrote about “data bank privacy”<sup>99</sup> and first raised concerns about computer matching programs<sup>100</sup> and digital dossiers.<sup>101</sup> Discussion also included the Privacy Act of 1974.<sup>102</sup> Topic 2 (I) reflects constitutional protection of intimate, personal decisions, first articulated in *Griswold v. Connecticut*,<sup>103</sup> as extended by not only *Roe v. Wade* in 1973<sup>104</sup> but also by *Carey v. Population Services International*.<sup>105</sup> Topics 3 (D) and 4 (D) both relate to public disclosure of personal facts; Topic 3 reflecting First Amendment and tort issues associated with disclosures by the press,<sup>106</sup> and Topic 4 reflecting disclosures by the government.<sup>107</sup>

Topics 5 (I; O) and 7 (D; I; O) reveal issues of electronic surveillance, eavesdropping, and Fourth Amendment searches,<sup>108</sup> with Topic 5 more directly reflecting electronic surveillance.<sup>109</sup> Topic 7 also introduces the third-party doctrine expressed in *United States v. Miller*,<sup>110</sup> which held that the Fourth Amendment does not require the government to obtain a warrant to seize bank records.<sup>111</sup> Topic 6 (D) reveals concepts associated with the disclosure and retention of arrest records.<sup>112</sup> Topic 8 (D) reflects defamation and privacy under the First Amendment, particularly by extending the discussion of *New York Times v. Sullivan*<sup>113</sup> and *Gertz v. Robert Welch, Inc.*<sup>114</sup>

Topic 9 (D) focuses on the interplay between privacy and credit reporting, with particular emphasis on the Fair Credit Reporting Act.<sup>115</sup> Topic 10 (I) contains a mixture of themes; including privacy of student records, desegregation, and rights of association (the model apparently linked students and zoning with respect to desegregation).<sup>116</sup> While Topic 11 (D; I) is similar to Topic 2 (I), it extends medical decisional privacy and autonomy to categories such as minors’ rights to abortion,<sup>117</sup> the right to refuse medical treatment,<sup>118</sup> and disclosure of health data.<sup>119</sup> Topic 12 (I) reveals the beginnings of discussions regarding sexual autonomy.<sup>120</sup>

Topic 13 (I) reveals concepts related to witness testimony and immunity.<sup>121</sup> Topics 14 (A) and 16 (A) are closely related; Topic 14 reveals the interplay between copyright and privacy,<sup>122</sup> while Topic 16 deals more directly with the right of publicity.<sup>123</sup> Finally, Topic 15 (I) reveals issues associated with commercial speech and privacy.<sup>124</sup>

1981 – 1990

<p><b>1 (D; O)</b> Information Data Act Individual Personal Government Records Disclosure Access</p>	<p><b>2 (O)</b> Fourth Amendment Search Government Expectation Surveillance Warrant Police Katz Reasonable</p>	<p><b>3 (I)</b> Public Information Media Press Victim Rape Trial</p>	<p><b>4 (D)</b> Public False Light Defamation Speech Private Amendment Standard Malice Publication Libel</p>
<p><b>5 (D; I)</b> Privilege Patient Information Disclosure Confidentiality Child Physician Communications Attorney Client</p>	<p><b>6 (D; I)</b> Information Patient Disclosure Health Medical Confidentiality Physician Records Care</p>	<p><b>7 (D)</b> Bank Records Financial Customer Information Business Disclosure Miller</p>	<p><b>8 (I)</b> Sexual Homosexual State Sodomy Constitutional Conduct Marriage Consensual</p>
<p><b>9 (I)</b> Abortion State Woman Decision Child Parents Minor Parental Consent Roe Pregnancy</p>	<p><b>10 (A)</b> Publicity Celebrity Rights Commercial Property Public Likeness Appropriation Exploitation</p>	<p><b>11 (I)</b> Employee Testing Drug Employer Polygraph</p>	<p><b>12 (I)</b> Treatment Patient Life State Medical Death Suicide Consent Refuse Hunger Prison</p>

**13 (A)**

Property  
Copyright  
Work  
Author  
Protection  
Rights  
Fair  
Personality  
Picketing  
Ordinance  
Salinger  
Letters

**14 (I; O)**

Cable  
Television  
Telephone  
Service  
Caller  
Electronic  
Subscriber  
Regulation  
Identification  
FCC

As with the previous decade, information-related privacy is the most heavily weighted topic (Topic 1 (D; O)), with issues associated with the computerized collection and storage of information, both privately and by the government.<sup>125</sup> The next most heavily weighted topic (Topic 2 (I; O)) reflects Fourth Amendment searches.<sup>126</sup> Topic 3 (I) reflects almost exclusively discussion of *The Florida Star v. B.J.F.*,<sup>127</sup> which held that a Florida newspaper reporting the name of a rape victim, which had been obtained from public records, in violation of a Florida statute and the newspaper's own internal policy was protected by the First Amendment.<sup>128</sup> Topic 4 (D) identifies concepts associated with false light, defamation, and privacy.<sup>129</sup>

Topic 5 (D; I) reveals concepts associated with privilege,<sup>130</sup> specifically related to family testimonial<sup>131</sup> and rape victim privileges.<sup>132</sup> Topic 6 (D; I) relates to medical confidentiality,<sup>133</sup> including AIDS-related medical records.<sup>134</sup> Bank and business records privacy is reflected in Topic 7 (D).<sup>135</sup> Topic 8 (I) reveals concepts associated with sexual privacy,<sup>136</sup> while Topic 9 (I) is associated with reproductive autonomy.<sup>137</sup>

Topic 10 (A) relates to an extension of the original unauthorized commercial appropriation action, the right of publicity.<sup>138</sup> Employee drug and polygraph testing is revealed in Topic 11 (I).<sup>139</sup> Topic 12 (I) reveals various concepts associated with medical decisions, including sustaining life and euthanasia,<sup>140</sup> as well as prisoner hunger strikes.<sup>141</sup> Topic 13 (A) reveals the interplay between privacy and copyright,<sup>142</sup> while Topic 14 (I; O) is associated with emerging technological innovations such as two-way cable,<sup>143</sup> and Caller ID.<sup>144</sup>

1891 – 1990

**1 (A; D; I; O)**

Information  
Private  
Individual  
Public  
Disclosure  
Interest  
Personal  
Protection  
Invasion  
Intrusion  
Government  
Protected  
Rights  
Constitutional

**2 (A)**

Picture  
Person  
Invasion  
Statute  
Publication  
Consent  
Photograph  
Public  
Property  
Advertising  
Article  
Brandeis  
Warren

**3 (D)**

Tort  
Action  
Privacy  
Damages  
Invasion  
False  
Light  
Liability  
Distress  
Defamation  
Emotional  
Mental  
Injury  
Intrusion  
Prosser

**4 (I)**

State  
Rights  
Constitutional  
Fourteenth  
Amendment  
Justice  
Fundamental  
Due  
Process  
Protection  
Liberty  
Government  
Griswold

**5 (I; O)**

Information  
Data  
Records  
Individual  
Computer  
Record  
Agencies  
Personal  
System  
Agency  
Government  
Access  
Systems  
Files

**6 (D)**

False  
Light  
Public  
Defamation  
Standard  
Speech  
Times  
Malice  
Media  
Hill  
Gertz  
Libel  
Press

**7 (D)**

Criminal  
Arrest  
Police  
Evidence  
Rule  
Crime  
Records  
Enforcement  
Witness

**8 (O)**

Surveillance  
Electronic  
Eavesdropping  
Conversations  
Evidence  
Telephone  
Communications  
Fourth  
Amendment  
Devices  
Wiretapping

**9 (O)**  
 Amendment  
 Fourth  
 Search  
 Privacy  
 Government  
 Warrant  
 Searches  
 Seizure  
 Reasonable  
 Expectation  
 Katz

**10 (A)**  
 Property  
 Rights  
 Copyright  
 Protection  
 Author  
 Personal  
 Works  
 Contract  
 Equity  
 Publication  
 Letters

**11 (D)**  
 Public  
 Press  
 Information  
 Victim  
 Media  
 Trial  
 Rape  
 Florida  
 Interest  
 News  
 Access

**12 (D)**  
 Privilege  
 Patient  
 Information  
 Disclosure  
 Confidentiality  
 Physician  
 Medical  
 Communications  
 Confidential  
 Relationship  
 Health

**13 (I)**  
 Abortion  
 State  
 Roe  
 Decision  
 Pregnancy  
 Women  
 Minor  
 Consent  
 Interest  
 Medical  
 Rights  
 Constitutional  
 Health

**14 (D)**  
 Bank  
 Records  
 Financial  
 Customer  
 Information  
 Business  
 Subpoena  
 Tax  
 Disclosure  
 Account  
 Documents  
 Miller

**15 (A)**  
 Publicity  
 Rights  
 Celebrity  
 Commercial  
 Property  
 Likeness  
 Appropriation  
 Interest  
 Public  
 Copyright  
 Zacchini  
 Exploitation

**16 (I)**  
 Employee  
 Testing  
 Drug  
 Employment  
 Test  
 Employers  
 Polygraph  
 AIDS

**17 (I)**  
 Sexual  
 Homosexual  
 Sodomy  
 State  
 Conduct  
 Private  
 Marriage  
 Constitutional  
 Statute  
 Fundamental  
 Individual  
 Rights

**18 (I; D)**  
 Speech  
 Public  
 Ordinance  
 Interest  
 Commercial  
 Picketing  
 Home  
 Residential  
 Communication  
 Expression  
 Advertising  
 Content  
 Subliminal

**19 (I)**  
 Patient  
 Treatment  
 Medical  
 Life  
 State  
 Interest  
 Consent  
 Suicide  
 Individual  
 Death  
 Research  
 Prison

The topics identified in this time span represent fundamental privacy concepts from a century of publications. We have linked the four principal privacy themes to the identified topics as follows:

Appropriation—Topic 1 (fundamental concepts associated with privacy); Topic 2 (commercial use of a person’s likeness); Topic 10 (privacy and copyright); Topic 15 (right of publicity).

Disclosure—Topic 1 (fundamental concepts associated with privacy); Topic 3 (false light and defamation); Topic 6 (false light and the First Amendment); Topic 7 (arrest records); Topic 11 (victim and witness privacy); Topic 12 (privileges); Topic 14 (bank and business records privacy); Topic 18 (free and commercial speech and privacy).

Intrusion—Topic 1 (fundamental concepts associated with privacy); Topic 4 (decisional autonomy); Topic 5 (collection of computerized information); Topic 8 (electronic surveillance); Topic 9 (Fourth Amendment searches); Topic 13 (reproductive decisional autonomy); Topic 16 (employee drug and polygraph testing); Topic 17 (decisional autonomy); Topic 18 (free and commercial speech and privacy); Topic 19 (medical decisional autonomy).

Observation—Topic 1 (fundamental concepts associated with privacy); Topic 5 (collection of computerized information); Topic 8 (electronic surveillance); Topic 9 (Fourth Amendment searches).

#### IV. CONCLUSION

This paper has presented preliminary results from applying latent Dirichlet allocation probabilistic topic modeling algorithms to a document collection comprised of published law review and journal articles from 1891 through 1990—all citing to Warren’s and Brandeis’s *The Right to Privacy* and substantively discussing privacy law issues. Our initial interpretation of the results reveals four principal privacy themes: Appropriation, Disclosure, Intrusion, and Observation. Once our initial data collection has been completed, by filling out the years 1991 through 2013, we will be able to discern whether these concepts survive into the twenty-first century as well as how they may have evolved over time, allowing us to develop a full ontology of privacy law based upon the topics revealed in published law review and journal articles.

---

<sup>1</sup> John B. Young, *Introduction: A Look at Privacy*, in PRIVACY 1, 2 (John B. Young ed., 1978).

<sup>2</sup> BARRINGTON MOORE, JR., PRIVACY: STUDIES IN SOCIAL AND CULTURAL HISTORY 268 (1984).

<sup>3</sup> Leilani A. Roberts, *The Ontology of Privacy* 7 (Oct. 1993) (unpublished Ph.D. dissertation, University of Oregon) (on file with author).

<sup>4</sup> See generally HELEN NISSENBAUM, PRIVACY IN CONTEXT (2010) (arguing that a framework of contextual integrity can determine when people will perceive new information technologies and systems as threats to privacy).

<sup>5</sup> See Roberts, *supra* note 3, at 9.

<sup>6</sup> See generally David M. Blei, *Probabilistic Topic Models*, 55 COMM. ACM 77 (2012) (providing an overview of probabilistic topic modeling); Mark Steyvers & Tom Griffiths, *Probabilistic Topic Models*, in HANDBOOK OF LATENT SEMANTIC ANALYSIS 427 (Thomas K. Landauer et al. eds., 2007) (providing same).

<sup>7</sup> See Blei, *supra* note 6, at 77-78.

<sup>8</sup> See *id.*

<sup>9</sup> Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890).

<sup>10</sup> See Cassidy R. Sugimoto, *The Shifting Sands of Disciplinary Development: Analyzing North American Library and Information Science Dissertations Using Latent Dirichlet Allocation*, 62 J. AM. SOC’Y INFO. SCI. & TECH. 185, 187 (2011); see also generally Olessia Koltsova & Sergei Koltcov, *Mapping the Public Agenda with Topic Modeling: The Case of the Russian LiveJournal*, 5 POL’Y & INTERNET 207 (2013); Aaron K. Massey, et al., *Automated Text Mining for Requirements Analysis of Policy Documents*, (Sixth Int’l Workshop on Requirements Engineering and Law, 2013), available at <http://www.cc.gatech.edu/~jeisenst/papers/re13rt-p085-p-18125-preprint.pdf>.

<sup>11</sup> Warren & Brandeis, *supra* note 9, at 195. Warren and Brandeis were quoting THOMAS M. COOLEY, TREATISE ON THE LAW OF TORTS OR THE WRONGS WHICH ARISE INDEPENDENT OF CONTRACT 29 (1880) (stating that “[t]he right to one’s person may be said to be a right of complete immunity: to be let alone”); see also *Union Pac. Ry. Co. v. Botsford*, 141 U.S. 250 (1891) (affirming the lower court’s denial of Union Pacific’s motion to surgically examine Botsford, who had brought a personal injury action against Union Pacific). “No right is held more sacred, or is more

carefully guarded, by the common law, than the right of every individual to the possession and control of his own person, free from all restraint or interference of others, unless by clear and unquestionable authority of law.” *Id.* at 251.

<sup>12</sup> *Pavesich v. New Eng. Life Ins. Co.*, 50 S.E. 68, 71-72 (Ga. 1905).

<sup>13</sup> ALAN F. WESTIN, *PRIVACY AND FREEDOM* 7 (1967).

<sup>14</sup> *Id.* at 31-32.

<sup>15</sup> Margaret Mead, *Margaret Mead Re-examines Our Right to Privacy*, REDBOOK, Apr. 1965, at 15.

<sup>16</sup> DANIEL J. SOLOVE, *UNDERSTANDING PRIVACY* 13 (2008).

<sup>17</sup> *See, e.g.*, Peter Brown, *Late Antiquity*, in *I A HISTORY OF PRIVATE LIFE* 254 (Paul Veyne ed., Arthur Goldhammer trans., 1987) (noting that by the second century A.D. “[w]hat was most private in the individual, his or her most hidden feelings and motivations” were beginning to be acknowledged). Notions of privacy did exist to some extent in earlier times. Ancient Greece and Rome recognized legally protected rights of personality, more for protecting honor than the physical person. *See* Roscoe Pound, *Interests in Personality*, 28 HARV. L. REV. 343, 357 (1915). Other approaches addressed when something akin to privacy could or should be lost. *See, e.g.*, WARD JUST, *AN UNFINISHED SEASON* 153 (2004) (“Odysseus wept when he heard the poet sing of his great deeds abroad because, once sung, they were no longer his alone. They belonged to anyone who heard the song.”); STANLEY ROSEN, *PLATO’S REPUBLIC* 129 (2005) (noting that Plato considered the guardians of the city to be subject to complete publicity).

<sup>18</sup> Georges Duby, *Introduction*, in *II A HISTORY OF PRIVATE LIFE* 6 (Georges Duby ed., Arthur Goldhammer trans., 1987).

<sup>19</sup> Evelyn Patlagean, *Byzantium in the Tenth and Eleventh Centuries*, in *I A HISTORY OF PRIVATE LIFE*, *supra* note 17, at 571.

<sup>20</sup> *See* Michel Rouche, *The Early Middle Ages in the West*, in *I A HISTORY OF PRIVATE LIFE*, *supra* note 17, at 415; Georges Duby et al., *Portraits*, in *II A HISTORY OF PRIVATE LIFE*, *supra* note 18, at 157; Philippe Ariès, *Introduction*, in *III A HISTORY OF PRIVATE LIFE* 9 (Roger Chartier ed., Arthur Goldhammer trans., 1987). Respect for the home has a long tradition. Historians and anthropologists consider the development of permanent residences as the precursor to initial human desires for privacy. Hofstadter and Horowitz claim that the first legally protected right of privacy is found in third century Jewish law in a rule regulating the height of house walls so as to minimize the risk of neighbors peering into each others’ houses. SAMUEL H. HOFSTADTER & GEORGE HOROWITZ, *THE RIGHT OF PRIVACY* 9 (1964). Hofstadter and Horowitz also quote an 1180 restatement of Jewish law explicitly stating that “the harm of being seen in privacy is a legal wrong.” *Id.* Roman law also recognized a home-related right to privacy: “To enter a man’s house against his will, even to serve a summons, was regarded as an invasion of his privacy.” W.A. HUNTER, *A SYSTEMATIC AND HISTORICAL EXPOSITION OF ROMAN LAW* 149 (J. Ashton Cross trans., 4th ed. 1803). And the notion that “a man’s home is his castle” was quickly adopted in America and remains a protected zone, at least against government searches. *See* 1 *LEGAL PAPERS OF JOHN ADAMS* 137 (L. Kinvin Wroth & Hiller B. Zobel eds., 1965) (“[E]very member of Society has entered into a solemn Covenant with every other that he shall enjoy in his own dwelling House as compleat a security, safety and Peace and Tranquility as if it was surrounded with Walls of Brass, with Ramparts and Palisadoes and defended with a Garrison and Artillery.”); *Kyllo v. United States*, 533 U.S. 27, 38-39 (2001) (concluding that detecting heat emanating from a home using thermal imaging constituted a search within the meaning of the Fourth Amendment because, in part, “intimate activities” within the home may be disclosed).

<sup>21</sup> *See* Ariès, *supra* note 20, at 5.

<sup>22</sup> *See* DAVID J. SEIPP, *THE RIGHT TO PRIVACY IN AMERICAN HISTORY* 2 (1978) (discussing how home construction made listening to private conversations possible). Eavesdropping was considered an actionable public nuisance in England by the beginning of the nineteenth century. *See* SIR WILLIAM BLACKSTONE, *COMMENTARIES ON THE LAWS OF ENGLAND* 168-69 (Thomas B. Wait, & Co. 1807) (classifying eavesdropping as one of seven wrongs whose punishment is short of death). “Eaves-droppers, or such as listen under walls or windows or the eaves of a house, to hearken after discourse, and thereupon to frame slanderous and mischievous tales, are a common nuisance and presentable at . . . [court].” *Id.* Eavesdropping was recognized as an actionable wrong by American courts by the early nineteenth century. *See, e.g.*, *Commonwealth v. Lovett*, 4 Pa. L.J. Rpts. (Clark) 226, 226 (Pa. 1831) (“No man has a right to pry into your secrecy in your own house.”); *see also* Lance E. Rothenberg, *Re-Thinking Privacy: Peeping Toms, Video Voyeurs, and Failure of the Criminal Law to Recognize a Reasonable Expectation of Privacy in the Public Space*, 49 AM. U. L. REV. 1127, 1140-42 (2000) (describing the origin of the term “peeping tom” as well as its early prosecutions as a breach of the peace).

<sup>23</sup> FRANK LUTHER MOTT, *AMERICAN JOURNALISM: A HISTORY OF NEWSPAPERS IN THE UNITED STATES THROUGH 250 YEARS 1690 TO 1940*, at 444 (1941) (describing keyhole journalism as an “invasion of privacy by prying reporters”).

<sup>24</sup> See ROBERT ELLIS SMITH, *BEN FRANKLIN’S WEB SITE: PRIVACY AND CURIOSITY FROM PLYMOUTH ROCK TO THE INTERNET* 123 (2000) (listing technological developments in the latter half of the nineteenth century); see also Robert E. Mensel, “*Kodakers Lying in Wait*”: *Amateur Photography and the Right to Privacy in New York, 1885-1915*, 43 AM. Q. 24, 25 (1991) (discussing the impact of unprecedented technological change upon the public’s psyche).

<sup>25</sup> ALPHEUS THOMAS MASON, *BRANDEIS: A FREE MAN’S LIFE* 70 (1946). Warren’s father-in-law, Thomas Francis Bayard, was a U.S. Senator, one-time presidential candidate, and Secretary of State under Grover Cleveland. LEWIS J. PAPER, *BRANDEIS* 35 (1983). There is some dispute as to what specific event triggered Warren’s call to action. Prosser reported it was “when the newspapers had a field day on the occasion of the wedding of a daughter, and Mr. Warren became annoyed.” William L. Prosser, *Privacy*, 48 CAL. L. REV. 383, 383 (1960) (citing MASON, *supra* at 70 (stating that the *Saturday Evening Gazette* reported the Warrens’ activities in “lurid detail”). *But see* James H. Barron, *Warren and Brandeis*, *The Right to Privacy*, 4 *Harv. L. Rev.* 193 (1890): *Demystifying a Landmark Citation*, 13 SUFFOLK U. L. REV. 875, 894 (1979) (suggesting the triggering event was the reporting of the wedding of a cousin since the Warrens’ oldest daughter could not have been more than seven years old in 1890; asserting the public disclosure that there were “no bridesmaids” was the embarrassing incident that “launched a thousand lawsuits”) (paraphrasing Prosser, *supra* at 423 (suggesting that Warren’s daughter “was the face that launched a thousand lawsuits”). Finally, a Brandeis biographer suggests Warren called upon Brandeis to write the article with him after becoming “outraged when photographers invaded his babies’ privacy and snapped perambulator pictures.” ALFRED LIEF, *BRANDEIS: THE PERSONAL HISTORY OF AN AMERICAN IDEAL* 51 (1936).

<sup>26</sup> *See* Barron, *supra* note 25, at 876-77 (noting the article has “assumed a hallowed place in both legal literature and history” and the “near unanimity among courts and commentators that the Warren-Brandeis conceptualization created the structural and jurisprudential foundation of the tort of invasion of privacy”).

<sup>27</sup> “Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that ‘what is whispered in the closet shall be proclaimed from the house-tops.’” Warren & Brandeis, *supra* note 9, at 195.

<sup>28</sup> *Id.* at 197.

<sup>29</sup> *Id.* at 198.

<sup>30</sup> *Id.* at 200 (“But where the value of the production is found not in the right to take the profits arising from publication, but in the peace of mind or the relief afforded by the ability to prevent any publication at all, it is difficult to regard the right as one of property, in the common acceptance of that term.”)

<sup>31</sup> *Id.* at 207; *see* Neil M. Richards & Daniel J. Solove, *Privacy’s Other Path: Recovering the Law of Confidentiality*, 96 GEO. L.J. 123, 128-31 (2007) (succinctly analyzing Warren’s and Brandeis’s arguments in *The Right to Privacy*); Diane L. Zimmerman, *Requiem for a Heavyweight: A Farewell to Warren and Brandeis’s Privacy Tort*, 68 CORNELL L. REV. 291, 295 (1983) (asserting that Warren and Brandeis appeared “to believe that the details of one’s personal life ‘belonged’ in some sense to the individual and could not be ‘used’ by others without permission”).

<sup>32</sup> *See supra* note 22 and accompanying text.

<sup>33</sup> *Dennis v. LeClerc*, 1 Mart. (o.s.) 297 (Orleans 1811); *see also* *Woolsey v. Judd*, 11 How. Pr. 49 (N.Y. Super. Ct. 1855 (holding same)); Richards & Solove, *supra* note 31, at 144 n.43 (describing nineteenth-century cases preventing the disclosure of confidential letters).

<sup>34</sup> *De May v. Roberts*, 9 N.W. 146, 149 (Mich. 1881).

<sup>35</sup> *See* Prosser, *supra* note 25, at 386 (impliedly crediting Warren and Brandeis for the fact that by 1960 the right to privacy was recognized in one form or another by an “overwhelming majority of the American courts”). However, Warren and Brandeis were not the first to publicly call for the recognition of a right to privacy. *See, e.g.*, E. L. Godkin, *The Rights of the Citizen. I.—To His Own Reputation*, SCRIBNER’S MAG., July-Dec. 1890, at 58, 65 (claiming it was a “natural right” to decide how much the public may know of an individual’s personal thought and feeling, tastes, habits, and private doings and affairs, including those of the family living under one’s roof).

<sup>36</sup> Zimmerman, *supra* note 31, at 294 & n.9 (citing various scholars struggling with the ideal of a definition of privacy).

<sup>37</sup> *Id.* at 294-95.

<sup>38</sup> Lisa Austin, *Privacy and the Question of Technology*, 22 LAW & PHIL. 119, 120 (2003).

<sup>39</sup> G. Robert Blakey, *The Rule of Announcement and Unlawful Entry: Miller v. United States and Ker v. California*, 112 U. PA. L. REV. 499, 554 (1964) (paraphrasing BARON JOHN EMERICH EDWARD DALBERG ACTON ACTON, ESSAYS ON FREEDOM AND POWER 36 (Meridian ed. 1960)).

<sup>40</sup> *Id.* (quoting William M. Beany, *The Constitutional Right to Privacy in the Supreme Court*, 1962 SUP. CT. REV. 212, 214 (1962)); see also Young, *supra* note 1, at 4 (“Privacy means different things to different men. It is an elusive concept, difficult to define.”).

<sup>41</sup> Note, *Right of Privacy—Property Rights—Unfair Competition—Constitutional Guarantees—Personal Rights*, 11 N.Y. L.F. 120, 120 (1965) (footnote omitted). Indeed, D.N. MacCormick even argues there is a significant distinction between defining “privacy” and the “right to privacy.” D.N. MacCormick, Comment, *Privacy: A Problem of Definition?*, 1 BRIT. J.L. & SOC’Y 75, 75 (1974).

<sup>42</sup> David McAuley, *ALI to Commence Restatement Project Compiling Information Privacy Principles*, BLOOMBERG BNA ELECTRONIC COM. & L. REP., 18 ECLR 684 (Apr. 10, 2013).

<sup>43</sup> According to a study by Shapiro and Pearse published in 2012, *The Right to Privacy* is the second most-cited law review article of all time, behind Ronald Coase’s *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960). See Fred R. Shapiro & Michelle Pearse, *The Most-Cited Law Review Articles of All Time*, 110 MICH. L. REV. 1483 (2012). Admittedly, a portion of articles that cite to *The Right to Privacy* do so for reasons other than a substantive discussion of privacy. For example, HeinOnline records 1753 articles that cite to *The Right to Privacy* between 1890 and 1990. Upon review, 469 (approximately thirty-two percent) of those articles were excluded from this project’s analysis because they were bar journals, not published in the United States, or only tangentially cited to *The Right to Privacy*.

<sup>44</sup> This is an admittedly crude method to calculate the number of published law review articles substantively discussing privacy, particularly, as indicated *supra* note 43, some articles may address privacy only tangentially. However, until a more complete review of the articles is completed, it at least gives some indication of the volume of published work from the legal discipline on the subject.

<sup>45</sup> Here we are borrowing liberally from Leśniewski’s system of ontology as a general principal of being. See John T. Kearns, *The Contribution of Leśniewski*, 8 NOTRE DAME J. FORMAL LOGIC 61, 62 (1967). In this context, “extensionality requires that any expression must be considered only in terms of those existing objects to which it can be related.” *Id.* at 72. Fundamentally, Leśniewski’s ontology addressed syntactic paradoxes based on a theory of distributive predication. See John Thomas Canty, *Ontology: Leśniewski’s Logical Language*, 5 FOUND. LANGUAGE 455, 455 (1969).

<sup>46</sup> See Blei, *supra* note 6, at 77.

<sup>47</sup> *Id.* at 78.

<sup>48</sup> *Id.*

<sup>49</sup> David M. Blei et al., *Latent Dirichlet Allocation*, 3 J. MACHINE LEARNING RES. 993, 996 (2003).

<sup>50</sup> *Id.* at 997.

<sup>51</sup> See *supra* note 9.

<sup>52</sup> Not every published article that cites to *The Right to Privacy* substantively addresses privacy; for example, some articles cite to *The Right to Privacy* as an example of a law review article’s influence on the development of law, while others may address privacy only as a minor, secondary topic. These articles have been excluded from the document collection. Of approximately 3500 articles that cite to *The Right to Privacy*, it is anticipated that the initial document collection will comprise approximately 3000 articles. The selected articles include scholarly articles, student notes and comments, and analyses of recent decisions.

<sup>53</sup> See *supra* notes 35 & 43.

<sup>54</sup> Citing to *The Right to Privacy* does not automatically qualify a published article for inclusion in the document collection. As noted *supra*, note 52, articles were reviewed for relevancy in an attempt to collect only documents that substantively address privacy law. It is also acknowledged that many privacy law articles will not necessarily cite to *The Right to Privacy*. Limiting document collection to articles citing to *The Right to Privacy* is an initial methodology to quickly identify a large collection of documents. It is hoped that a later phase of this project will identify and include those additional articles not identified through the initial methodology.

<sup>55</sup> This is a fairly daunting task, as the majority of law review articles published prior to 1990 are available only through HeinOnline, and stored as pdf replicas of the original articles. Conversion to plain text often leaves garbled words, as well as footnotes which must be removed on a page-by-page basis. In addition, all formatting codes—particularly paragraph codes—must be stripped from each document.

<sup>56</sup> Fundamentally, the approach is to track the evolution of privacy law decade by decade; however, due to the relatively few articles published prior to 1950, the first time slice covers the sixty-year span from 1891 – 1950.

<sup>57</sup> Comprising 159 documents.

<sup>58</sup> Comprising 109 documents.

<sup>59</sup> Comprising 160 documents.

<sup>60</sup> Comprising 326 documents.

<sup>61</sup> Comprising 428 documents.

<sup>62</sup> Comprising 1182 documents.

<sup>63</sup> Andrew Kachites McCallum, *MALLET: A Machine Learning for Language Toolkit* (2002),

<http://mallet.cs.umass.edu>. There are a variety of additional open-source software packages for topic modeling. See <http://www.cs.princeton.edu/~blei/topicmodeling.html> for a listing of topic modeling software packages in various programming languages.

<sup>64</sup> This is validated using another MALLET output file that records the number of times a specific word is associated (if at all) with each topic.

<sup>65</sup> Techniques have been applied to confirm an optimum number of topics. For example, Massey et al. performed multiple topic models on a sample of their corpus and then calculated the perplexity—a measure of the predictive likelihood of a model of text—of the models. See Massey et al., *supra* note 10; see also Steyvers & Griffiths, *supra* note 6, at 441 (discussing various methods to determine the number of topics). See generally, Airoidi et al., *Reconceptualizing the Classification of PNAS Articles*, 107 PROC. NAT'L ACAD. SCI. 20899 (2010) (explaining their methodology to determine twenty to forty topics were optimal for their data set, which comprised documents from multiple disciplines).

<sup>66</sup> See Steyvers & Griffiths, *supra* note 6, at 441 (noting that “a solution with too many topics will result in uninterpretable topics that pick out idiosyncratic word combinations”).

<sup>67</sup> See Zoubin Ghahramani, *Unsupervised Learning* (Sept. 16, 2004), available at

<http://mlg.eng.cam.ac.uk/zoubin/papers/ul.pdf>.

<sup>68</sup> The final score “number” is arbitrary as it is meant merely to rank the topics according to how strongly they fit the parameters.

<sup>69</sup> We did not, however, stem any words. As such, we see in our models singular and plural versions of some words, though one or the other has been removed from the results reported here. We are not yet at a point where we are confident we can stem certain words without possibly limiting the relevance of topics.

<sup>70</sup> The themes will be abbreviated as follows: Appropriation (A); Disclosure (D); Intrusion (I); and Observation (O).

<sup>71</sup> As mentioned previously, the topic numbering denotes the relative weight of each topic in descending order. As such, they are presented and discussed in order of prevalence within each slice (and the terms comprising each topic are ranked in order of frequency of occurrence with their respective topic).

<sup>72</sup> Warren & Brandeis, *supra* note 9, at 195 (“Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that ‘what is whispered in the closet shall be proclaimed from the house-tops.’”).

<sup>73</sup> 171 N.Y. 538 (1902).

<sup>74</sup> 50 S.E. 68, 77-81 (Ga. 1905); see also *Edison v. Edison Polyform Mfg. Co.*, 67 A. 392 (N.J. Ch. 1907) (granting Thomas Edison an injunction against a company from using his name in association with its business and products).

<sup>75</sup> See N.Y. CIV. RIGHTS LAW § 50 (McKinney 2009) (“A person, firm or corporation that uses for advertising purposes, or for the purposes of trade, the name, portrait or picture of any living person without having first obtained the written consent of such person . . . is guilty of a misdemeanor.”) (held constitutional, *Rhodes v. Sperry & Hutchinson Co.*, 85 N.E. 1097 (N.Y. 1908), *aff'd*, *Sperry & Hutchinson Co. v. Rhodes*, 220 U.S. 502 (1911)).

<sup>76</sup> Warren G. Willsey, Comment, *Equity: Right of Privacy: Injunction to Protect Personal Rights*, 5 CORNELL L.Q. 177, 178 (1920).

<sup>77</sup> *Id.*; see also Comment, *Property Rights in Letters*, 46 YALE L.J. 493 (1937).

<sup>78</sup> See, e.g., Louis H. Beard, Recent Cases, *Torts—Right of Privacy—Unauthorized Use of Photograph for Advertising*, 21 TEX. L. REV. 103 (1942); W.H. Brewster, Recent Decision, *Torts—Invasion of Personal Safety, Comfort or Privacy—Whether Publication of Photograph Without Permission Constitutes Violation of Right of Privacy*, 27 CHI.-KENT L. REV. 169 (1948); Note, *Broadcasting—Invasion of Privacy—Broadcast of Current News Events—Injunctive Relief*, 11 AIR L. REV. 83 (1940); Recent Cases, *Right of Privacy—Nature and Extent—Biographical Motion Pictures*, 44 HARV. L. REV. 1146 (1931) (discussing *Melvin v. Reid*, 297 P. 91 (Cal. Dist. Ct. App. 1931) (holding that release of an unauthorized biographical motion picture of plaintiff’s past life of shame violated her right to privacy)); Recent Cases, *Torts—Invasion of the Right to Privacy by Radio Broadcast*, 88 U. PA. L. REV. 374 (1940) (discussing *Mau v. Rio Grande, Inc.*, 28 F. Supp. 845 (N.D. Cal. 1939) (ruling plaintiff’s right to

privacy had been violated by broadcast of plaintiff's name without consent in radio dramatization of an actual robbery in which plaintiff was the victim) (applying California law)).

<sup>79</sup> Brewster, *supra* note 78, at 170-71; *see also* Howard M. Finkelstein, Recent Decisions, *Television and the Right of Privacy*, 3 SYRACUSE L. REV. 208 (1951).

<sup>80</sup> *See, e.g.*, Marilyn C. Lazar, Case Note, *Property Rights—Ideas*, 18 S. CAL. L. REV. 256 (1945); *Property Rights in Letters*, *supra* note 77; *see also* Decisions, *Right of Privacy—Comic Books—Liability for Reproduction of Actual Episode in Cartoon Form*, 26 N.Y.U. L. REV. 217 (1951) (discussing *Moloney v. Boy Comics Publishers, Inc.*, 98 N.Y.S. 2d 119 (N.Y. App. Div. 1950) (holding New York's privacy statute did not protect against reproduction of cartoon)).

<sup>81</sup> *See* Note, *Wire Tapping and Law Enforcement*, 53 HARV. L. REV. 863 (1940) (discussing *Nardone v. United States*, 308 U.S. 338 (1939), and *Weiss v. United States*, 308 U.S. 321 (1939); comparing with *Olmstead v. United States*, 277 U.S. 438 (1928); noting all cases found intercepting conversations not to be an unreasonable search under the Fourth Amendment). *But cf.* *Rhodes v. Graham*, 37 S.W.2d 46 (Ky. 1931) (ruling wire tapping by private individuals violated plaintiff's right to privacy).

<sup>82</sup> *See, e.g.*, Charles F. Graney, Recent Cases, *Manner of Publication Determinative of Action for Invasion of Privacy*, 9 BUFF. L. REV. 362, 363 (1960) ("The right of privacy. . . is a strictly personal right based on the interest one has in living his life out of the public view. . . . The right, however, is not absolute. The principal limitation on the right to be free from undesired publicity is the defendant's privilege to publish matter which is of sufficient public interest.").

<sup>83</sup> Prosser, *supra* note 25, at 389.

<sup>84</sup> *See, e.g.*, Herbert Brownell, Jr., *The Public Security and Wire Tapping*, 39 CORNELL L.Q. 195 (1954) (discussing various wire tapping bills pending in Congress).

<sup>85</sup> In other words, wire tapping-related topics received relatively low weights.

<sup>86</sup> *See, e.g.*, H. Edward Keefe, Jr., Recent Decisions, *Torts—Harassment Held Actionable as Invasion of Right of Privacy*, 1956 U. ILL. L.F. 522 (1955) (discussing *Housh v. Peth*, 133 N.E.2d 340 (Ohio 1955)); Russell A. McNair, Jr., Recent Decisions, *Torts—Invasion of Privacy—Conduct of a Debt Collector*, 58 MICH. L. REV. 484 (1960) (discussing *Biederman's of Springfield, Inc. v. Wright*, 322 S.W.2d 892 (Mo. 1959) (holding that degrading and humiliating the defendant in public constituted an invasion of privacy)). *But see* James Magavern, Recent Decisions, *Invasion of Privacy in Debt Collection*, 7 BUFF. L. REV. 327 (1958) (discussing *Gouldman-Taber Pontiac, Inc. v. Zerbst*, 100 S.E.2d 881 (Ga. 1957) (holding that creditor who wrote a letter to an employer notifying him that his employee was indebted to creditor and seeking employer's aid in collection of debt did not violate the employee's right of privacy)).

<sup>87</sup> *See, e.g.* Joseph M. Snee, *One for the Money, Two for the Show: The Case Against Televising Congressional Hearings*, 42 GEO. L.J. 1 (1953); Robert M. Zweiman, *Trial by Wager of Television*, 7 INTRAMURAL L. REV. N.Y.U. 75 (1952); Comment, *Television and Congressional Investigations*, 1 DEPAUL L. REV. 112 (1951).

<sup>88</sup> *See, e.g.*, Eugene N. Aleinikoff, *Privacy in Broadcasting*, 42 IND. L.J. 373 (1967); Case Note, *Right of Privacy—Use of Newsworthy Photograph in Incidental Advertising of Magazine Held No Invasion of Privacy*, 31 FORDHAM L. REV. 394 (1962). *But cf.* Charles Simon, Case Note, *Torts—Invasion of Privacy—Unauthorized Use of Photograph*, 16 DEPAUL L. REV. 255 (1966) (noting that as of yet, no Massachusetts case had established a legally protected right of privacy).

<sup>89</sup> 381 U.S. 497 (1965) (ruling unconstitutional state statute prohibiting the use of contraceptives).

<sup>90</sup> *Id.* at 485. "[T]he right of marital privacy is protected. . . within the protected penumbra of specific guarantees of the Bill of Rights. . . ." *Id.* at 487 (Goldberg, J. concurring).

<sup>91</sup> 389 U.S. 347 (1967) (overruling *Olmstead v. United States*, 277 U.S. 438 (1928); limiting government's ability to wiretap and eavesdrop under the Fourth Amendment); *see also* Note, *From Private Places to Personal Privacy: A Post-Katz Study of Fourth Amendment Protection*, 43 N.Y.U. L. REV. 968, 968 (1968) (noting that since *Katz*, "judicial rationale has emphasized the privacy of the individual, as opposed to the privacy of places, in determining the scope of fourth amendment protection").

<sup>92</sup> 385 U.S. 374 (1967).

<sup>93</sup> 376 U.S. 254 (1964).

<sup>94</sup> *See generally* *The Supreme Court 1966 Term: Libel and Privacy Actions*, 81 HARV. L. REV. 110, 160 (1967).

<sup>95</sup> *See, e.g.*, Joel F. Handler & Margaret K. Rosenheim, *Privacy in Welfare: Public Assistance and Juvenile Justice*, 31 LAW & CONTEMP. PROBS. 377 (1966); Douglas Q. Wickham, *Restricting Home Visits: Toward Making the Life of the Public Assistance Recipient Less Public*, 118 U. PA. L. REV. 1188 (1970).

<sup>96</sup> See, e.g., Kenneth L. Karst, “*The Files*”: *Legal Controls over the Accuracy and Accessibility of Stored Personal Data*, 31 LAW & CONTEMP. PROBS. 342 (1966); Arthur R. Miller, *Personal Privacy in the Computer Age: The Challenge of a New Technology in an Information-Oriented Society*, 67 MICH. L. REV. 1089 (1969); Note, *Credit Investigations and the Right to Privacy: Quest for a Remedy*, 57 GEO. L. J. 509 (1969).

<sup>97</sup> See, e.g., Armando M. Menocal III & James McDonald Williams, Editorial Notes, *Lie Detectors in Private Employment: A Proposal for Balancing Interests*, 33 GEO. WASH. L. REV. 932 (1965).

<sup>98</sup> See, e.g., Victor S. Netterville, *Copyright and Tort Aspects of Parody, Mimicry and Humorous Commentary*, 35 S. CAL. L. REV. 225, 267 (1962) (“Practically every humorous take-off or commentary involving the person or profession of another presents privacy problems that are at best only tentatively answered by the law as it stands today.”). *But see* Recent Cases, *Torts—Defamation—Right of Privacy—Unfair Competition—Imitation of Comedian’s Voice in Television Commercial May Constitute Defamation or Unfair Competition But Not Invasion of Privacy*, 76 HARV. L. REV. 1685 (1963).

<sup>99</sup> Arthur Miller published his book, *The Assault on Privacy: Computers, Data Banks, and Dossiers*, in 1971, and in 1972 Alan Westin and Michael Baker published their book, *Data Banks in a Free Society*, resulting in law journal book reviews. See, e.g., Mary Kay Kane, Book Review, 24 BUFF. L. REV. 331 (1975) (reviewing ALAN F. WESTIN & MICHAEL A. BAKER, *DATA BANKS IN A FREE SOCIETY* (1972) and *RECORDS, COMPUTERS AND THE RIGHTS OF CITIZENS: REPORT OF THE SECRETARY’S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS* (1973)); Bernard A. Berkman, Book Note, 22 CASE W. RES. L. REV. 808 (1971) (reviewing ARTHUR R. MILLER, *THE ASSAULT ON PRIVACY: COMPUTERS, DATA BANKS, AND DOSSIERS* (1971)).

<sup>100</sup> See, e.g., Kenneth James Langan, *Computer Matching Programs: A Threat to Privacy?*, 15 COLUM. J.L. & SOC. PROBS. 143 (1979).

<sup>101</sup> See, e.g., Vern Countryman, *The Diminishing Right of Privacy: The Personal Dossier and the Computer*, 49 TEX. L. REV. 837 (1971).

<sup>102</sup> See, e.g., Jerome J. Hanus & Harold C. Relyea, *A Policy Assessment of the Privacy Act of 1974*, 25 AM. U. L. REV. 555 (1975-1976); Note, *The Privacy Act of 1974: An Overview and Critique*, 1976 WASH. U. L.Q. 667 (1976).  
<sup>103</sup> 381 U.S. 497 (1965).

<sup>104</sup> 410 U.S. 113, 153 (1973) (finding a Fourteenth Amendment right to privacy in a women’s decision to terminate her pregnancy).

<sup>105</sup> 431 U.S. 678 (1978) (finding restrictions on the sale of contraceptives unconstitutional) (plurality). “[O]ne aspect of the liberty protected by the Due Process Clause of the Fourteenth Amendment is a guarantee of certain areas or zones of privacy . . . [that] include personal decisions relating to marriage, procreation, contraception, family relationships, and child rearing and education.” *Id.* at 684-85 (citations omitted) (internal quotation marks omitted).

<sup>106</sup> See, e.g., Comment, *An Accommodation of Privacy Interests and First Amendment Rights in Public Disclosure Cases*, 124 U. PA. L. REV. 1385 (1976); Greg Trout, Note, *Invasion of Privacy: New Guidelines for the Public Disclosure Tort*, 6 CAP. U. L. REV. 95 (1976).

<sup>107</sup> See, e.g., R. Perry Sentell, Jr., *The Omen of “Openness” in Local Government Law*, 13 GA. L. REV. 97 (1978); Martin Conroy Foster, Comment, *Plante v. Gonzalez: The Elected Official’s Limited Right to Financial Privacy*, 15 NEW ENG. L. REV. 192 (1979); Thomas C. Taylor, Jr., Comment, *Financial Disclosure by Public Officials and Public Employees in Light of Carmel-by-the-Sea v. Young*, 18 UCLA L. REV. 534 (1971).

<sup>108</sup> See, e.g., Shirley M. Hufstедler, *Invisible Searches for Intangible Things: Regulation of Governmental Information Gathering*, 127 U. PA. L. REV. 1483 (1979); David M. O’Brien, *Reasonable Expectations of Privacy: Principles and Policies of Fourth Amendment-Protected Privacy*, 13 NEW ENG. L. REV. 662 (1978).

<sup>109</sup> See, e.g., John F. Decker & Joe Handler, *Electronic Surveillance: Standards, Restrictions and Remedies*, 12 CAL. W. L. REV. 60 (1975); David P. Hodges, Note, *Electronic Visual Surveillance and the Fourth Amendment: The Arrival of Big Brother?*, 3 HASTINGS CONST. L.Q. 261 (1976).

<sup>110</sup> 425 U.S. 435 (1976)

<sup>111</sup> *Id.* at 443 (“[I]nformation revealed to a third-party and conveyed by him to Government authorities, even if the information is revealed on the assumption that it will be used only for a limited purpose and the confidence placed in the third party will not be betrayed.”). *Miller’s* holding was limited by the Right to Financial Privacy Act of 1978, Pub. L. No. 95-630, 92 Stat. 3697 (1978) (codified as amended at 12 U.S.C. §§ 3401 – 3422 (2012)), by, for example, requiring the Government authority to notify the bank customer of the subpoena or summons served on the financial institution as well as the nature of the law enforcement inquiry to which the subpoena or summons relates. See also Note, *Government Access to Bank Records*, 83 YALE L.J. 1439 (1974); Susan Orzack Posen, Comment, *A Paper Chase: The Search and Seizure of Personal Business Records*, 43 BROOK. L. REV. 489 (1977).

<sup>112</sup> See, e.g., J. Taylor DeWeese, *Reforming Our "Record Prisons": A Proposal for the Federal Regulation of Crime Data Banks*, 6 RUTGERS-CAMDEN L.J. 26 (1974); Dayton Michael Cramer & Michael O. Sturm, Recent Developments, *FBI Rap Sheets—An Invasion of Constitutional Rights?*, 20 CATH. U. L. REV. 511 (1971); Gregory J. Lake, Comment, *The Expungement or Restriction of Arrest Records*, 23 CLEV. ST. L. REV. 123 (1974); Diane Becker Potash; Comment, *Maintenance and Dissemination of Criminal Records: A Legislative Proposal*, 19 UCLA L. REV. 654 (1972).

<sup>113</sup> 376 U.S. 254 (1964).

<sup>114</sup> 418 U.S. 323 (1974); see, e.g., Arthur N. Frakt, *The Evolving Law of Defamation: New York Times Co. v. Sullivan to Gertz v. Robert Welch, Inc. and Beyond*, 6 RUTGERS-CAMDEN L.J. 471 (1975); Alfred Hill, *Defamation and Privacy Under the First Amendment*, 76 COLUM. L. REV. 1205 (1976); Note, *The Invasion of Defamation by Privacy*, 23 STAN. L. REV. 547 (1971).

<sup>115</sup> Pub. L. No. 91-508, tit. VI, 84 Stat. 1114, 1127 (1970) (codified as amended at 15 U.S.C. §§ 1601 – 1616 (2012)); see, e.g., Comment, *Analysis of the Fair Credit Reporting Act*, 1 FORDHAM URB. L.J. 48 (1972); Lenore Cooper Garon, Note, *Protecting Privacy in Credit Reporting*, 24 STAN. L. REV. 550 (1972); Marsha Morrow McLaughlin & Suzanne Vaupel, Note, *Constitutional Right of Privacy and Investigative Consumer Reports: Little Brother Is Watching You*, 2 HASTINGS CONST. L.Q. 773 (1975).

<sup>116</sup> See, e.g., Lawrence R. Caruso, *Privacy of Students and Confidentiality of Student Records*, 22 CASE W. RES. L. REV. 379 (1971); Comment, *The Confidentiality of University Student Records: A Common Law Analysis*, 1976 BYU L. REV. 477 (1976); see also Tamila C. Jensen, *From Belle Terre to East Cleveland: Zoning, the Family and the Right to Privacy*, 13 FAM. L.Q. 1 (1979); Patricia Fitts Jacobson, Note, *The Collision of Zoning Ordinances and the Constitutional Rights of Privacy and Association: Critique and Prognosis*, 30 CASE W. RES. L. REV. 155 (1979); Comment, *Private School Desegregation under Section 1981*, 124 U. PA. L. REV. 714 (1976).

<sup>117</sup> See, e.g., Note, *The Minor's Right to Abortion and the Requirement of Parental Consent*, 60 VA. L. REV. 305 (1974).

<sup>118</sup> See, e.g., Patricia Kunze Davis, Recent Decision, *Constitutional Law—Right of Privacy—Qualified Right to Refuse Medical Treatment May Be Asserted for Incompetent Under Doctrine of Substituted Judgment*, 27 EMORY L.J. 425 (1978).

<sup>119</sup> See, e.g., John J. Fargo, Comment, *Medical Data Privacy: Automated Interference with Contractual Relations*, 25 BUFF. L. REV. 491 (1976).

<sup>120</sup> See, e.g., June Aline Eichbaum, *Towards an Autonomy-Based Theory of Constitutional Privacy: Beyond the Ideology of Familial Privacy*, 14 HARV. C.R.-C.L. L. REV. 361 (1979).

<sup>121</sup> See, e.g., Jennifer A. Clark, Note, *Questioning the Recognition of a Parent-Child Testimonial Privilege*, 45 ALB. L. REV. 142 (1980); Anne N. DePrez, Note, *Pillow Talk, Grimgribbers and Connubial Bliss: The Marital Communication Privilege*, 56 IND. L.J. 121 (1980).

<sup>122</sup> See, e.g., Jane Tucker Dana, *Copyright and Privacy Protection of Unpublished Works—The Author's Dilemma*, 13 COLUM. J.L. & SOC. PROBS. 351 (1977); James R. Kirby II, Comment, *An Artist's Personal Rights in His Creative Works: Beyond the Human Cannonball and the Flying Circus*, 9 PAC. L.J. 855 (1978); Alan Lee Zegas, Note, *Personal Letters: A Dilemma for Copyright and Privacy Law*, 33 RUTGERS L. REV. 134 (1980).

<sup>123</sup> See, e.g., Peter L. Felcher & Edward L. Rubin, *Privacy, Publicity, and the Portrayal of Real People by the Media*, 88 YALE L.J. 1577 (1979); Bennet D. Zurofsky, Case Comments, *Constitutional Law—Privacy Torts—First Amendment Does Not Privilege Violation of Right of Publicity*, 31 RUTGERS L. REV. 269 (1978) (discussing *Zacchini v. Scripps-Howard Broadcasting Co.*, 433 U.S. 562 (1977) (holding the First and Fourteenth Amendments did not immunize television broadcasting company when it broadcast "human cannonball" performer's entire act without his consent)).

<sup>124</sup> See, e.g., David P. Ballard, *Privacy and Direct Mail Advertising*, 47 FORDHAM L. REV. 495 (1979); Susan Burnett Lutten, Note, *Give Me a Home Where No Salesmen Phone: Telephone Solicitation and the First Amendment*, 7 HASTINGS CONST. L.Q. 129 (1979); Walker B. Lowman, Note, *Federal Pandering Advertisements Statute: The Right of Privacy Versus the First Amendment*, 32 OHIO ST. L.J. 149 (1971); Daniel M. Taubman, Comment, *Pickers at the Doorstep*, 9 HARV. C.R.-C.L. L. REV. 95 (1974).

<sup>125</sup> See, e.g., Warren Freedman, *The Right of Privacy in the Age of Computer Data and Processing*, 13 TEX. TECH L. REV. 1361 (1982); Spiros Simitis, *Reviewing Privacy in an Information Society*, 135 U. PA. L. REV. 707 (1987); C. Dennis Southard IV, *Individual Privacy and Governmental Efficiency: Technology's Effect on the Government's Ability to Gather, Store, and Distribute Information*, 9 COMPUTER/L.J. 359 (1989); John Barlow Spear, *Computers in the Private Sector: Right to Informational Privacy for the Consumer*, 22 WASHBURN L.J. 469 (1983); Jonathan P. Graham, Note, *Privacy, Computers, and the Commercial Dissemination of Personal Information*, 65 TEX. L. REV.

- 1395 (1987); Rhys A. Sterling, Comment, *Privacy, Computerized Information Systems, and the Common Law—A Comparative Study in the Private Sector*, 18 GONZ. L. REV. 567 (1982-1983).
- <sup>126</sup> Lewis R. Katz, *In Search of a Fourth Amendment for the Twenty-First Century*, 65 IND. L.J. 549 (1990); James J. Tomkovicz, *Beyond Secrecy for Secrecy's Sake: Toward an Expanded Vision of the Fourth Amendment Privacy Province*, 36 HASTINGS L.J. 645 (1985); Eric Dean Bender, Note, *The Fourth Amendment in the Age of Aerial Surveillance: Curtains for Curtilage?*, 60 N.Y.U. L. REV. 725 (1985); Note, *Protecting Privacy Under the Fourth Amendment*, 91 YALE L.J. 313 (1981).
- <sup>127</sup> 491 U.S. 524.
- <sup>128</sup> See, e.g., Ellen B. Fishbein, Comment, *Identifying the Rape Victim: A Constitutional Clash Between the First Amendment and the Right to Privacy*, 18 J. MARSHALL L. REV. 987 (1985).
- <sup>129</sup> See, e.g., Ruth Walden & Emile Netzhammer, *False Light Invasion of Privacy: Untangling the Web of Uncertainty*, 9 COMM/ENT L.J. 347 (1987) (contrasting *The Right to Privacy* with *Time, Inc. v. Hill*, 385 U.S. 374 (1967) (applying First Amendment rights to New York's privacy statute); Diane Leenheer Zimmerman, *False Light Invasion of Privacy: The Light that Failed*, 64 N.Y.U. L. REV. 364 (1989); Patricia A. Ohlsen, Comment, *Fictionalized Publications: When Should Defamation and Privacy Be a Bar?*, 1984 UTAH L. REV. 411 (1984).
- <sup>130</sup> See, e.g., Developments in the Law, *Privileged Communications*, 98 HARV. L. REV. 1450, 1480-86 (1985).
- <sup>131</sup> See, e.g., James H. Feldman & Carolyn Sievers Reed, *Silences in the Storm: Testimonial Privileges in Matrimonial Disputes*, 21 FAM. L.Q. 189 (1987); Ann M. Stanton, *Child-Parent Privilege for Confidential Communications: An Examination and Proposal*, 16 FAM. L.Q. 1 (1982); Deborah A. Ausburn, Note, *Circling the Wagons: Informational Privacy and Family Testimonial Privileges*, 20 GA. L. REV. 173 (1985).
- <sup>132</sup> See, e.g., William J. Swift, Comment, *Rape Crisis Counseling Dilemma: The Right to Privacy and the Right to Confrontation*, 4 PUB. L. FORUM 467 (1985).
- <sup>133</sup> See, e.g., Bernard R. Adams, *Medical Research and Personal Privacy*, 30 VILL. L. REV. 1077 (1985); Robert M. Gellman, *Prescribing Privacy: The Uncertain Role of the Physician in the Protection of Patient Privacy*, 62 N.C. L. REV. 255 (1984); Carole M. Cleaver, Note, *Privacy Rights in Medical Records*, 13 FORDHAM URB. L.J. 165 (1985).
- <sup>134</sup> See, e.g., Emily Campbell, *Mandatory AIDS Testing and Privacy: A Psycholegal Perspective*, 66 N.D. L. REV. 449 (1990).
- <sup>135</sup> See, e.g., John S. Applegate, *The Business Papers Rule: Personal Privacy and White Collar Crime*, 16 AKRON L. REV. 189 (1982); Joseph R. Mangan, Jr., *Reasonable Expectations of Privacy in Bank Records: A Reappraisal of United States v. Miller and Bank Depositor Privacy Rights*, 72 J. CRIM. L. & CRIMINOLOGY 243 (1981).
- <sup>136</sup> See, e.g., Shari A. Levitan, Case Note, *The Supreme Court as Super Legislature Emasculates the Right to Privacy*, 21 SUFFOLK U. L. REV. 853 (1987) (discussing *Bowers v. Hardwick*, 106 S. Ct. 2841 (1986) (upholding Georgia's sodomy statute)).
- <sup>137</sup> See, e.g., Susan Frelich Appleton, *Doctors, Patients and the Constitution: A Theoretical Analysis of the Physician's Role in "Private" Reproductive Decisions*, 63 WASH. U. L.Q. 183 (1985); Jonathan L. Rue, Comment, *The Distribution of Contraceptives to Unemancipated Minors: Does a Parent Have a Constitutional Right to Be Notified?*, 69 KY. L.J. 436 (1980-1981); Note, *The Evolution of the Right to Privacy After Roe v. Wade*, 13 AM. J.L. & MED. 365 (1987).
- <sup>138</sup> See, e.g., Anne Marie Fealey, *Privacy and Publicity, Then and Now*, 3 HOFSTRA PROP. L.J. 15 (1989); Larry L. Saret & Martin L. Stern, *Publicity and Privacy—Distinct Interests on the Misappropriation Continuum*, 12 LOY. U. CHI. L.J. 675 (1981); L. Lee Byrd, *Privacy Rights of Entertainers and Other Celebrities: A Need for Change*, 5 ENT. & SPORTS L.J. 95 (1988).
- <sup>139</sup> See, e.g., Edward M. Chen et al., *Common Law Privacy: A Limit on an Employer's Power to Test for Drugs*, 12 GEO. MASON U. L. REV. 651 (1990); Christine M. Wiseman, *Invasion by Polygraph: An Assessment of Constitutional and Common Law Parameters*, 32 ST. LOUIS U. L.J. 27 (1987); Steven E. Lake, Comment, *Unrestricted Private Employee Drug Testing Programs: An Invasion of the Worker's Right to Privacy*, 23 CAL. W. L. REV. 72 (1986).
- <sup>140</sup> See, e.g., Gregory Gelfand, *Euthanasia and the Terminally Ill Patient*, 63 NEB. L. REV. 741 (1984); Robert L. Lapointe, Note, *The Removal of Feeding Tubes: Has the Right to Die Reached Its Limit?*, 24 NEW ENG. L. REV. 185 (1989).
- <sup>141</sup> See, e.g., Steven C. Bennett, Note, *The Privacy and Procedural Due Process Rights of Hunger Striking Prisoners*, 58 N.Y.U. L. REV. 1157 (1983).
- <sup>142</sup> See, e.g., Judge Jon O. Newman, *Copyright Law and the Protection of Privacy*, 12 COLUM.-VLA J.L. & ARTS 459 (1988); Christopher A. Murphy, Note, *Salinger v. Random House: The Author's Interests in Unpublished Materials*, 12 COLUM.-VLA J.L. & ARTS 103 (1987).

<sup>143</sup> See, e.g., Kenneth M.H. Hoff, *Two-Way Cable Television and Informational Privacy*, 6 COMM/ENT L.J. 797 (1983-1984).

<sup>144</sup> See, e.g., Glenn Chatmas Smith, *We've Got Your Number! (Is it Constitutional To Give it Out?): Caller Identification Technology and the Right to Informational Privacy*, 37 UCLA L. REV. 145 (1989).

## Appendix A: Conceptual Description of the LDA Algorithm

The LDA algorithm uses Bayesian statistics to discover the latent topics in a corpus of documents.<sup>1</sup> Conceptually, the LDA topic model algorithm uses the following process:

- 1) The number of topics to identify must be pre-selected.<sup>2</sup> There is no set rule for how many topics are ideal; it depends on the context of the corpus.<sup>3</sup> The topics that are discovered are defined only by the cluster of words associated with them. The researcher may decide to choose a name for each topic that makes sense based on context and experience.
- 2) The algorithm reads each document in the corpus and repeatedly randomly assigns each word to a topic, so that eventually all the words in the document<sup>4</sup> are probabilistically assigned to the pre-selected number of topics.
- 3) Two values are calculated for each topic:
  - a)  $p(\text{topic } t_j | \text{document } d_i) =$  the proportion of words in  $d_i$  (document <sub>$i$</sub> ) that are currently assigned to topic  $t_j$  (topic <sub>$j$</sub> ). For example, if document  $d_i$  contains 500 words and 80 of those words are assigned to topic  $t_j$ , then  $p = 0.16$  (80/500).
  - b)  $p(\text{word}_1 | \text{topic } t_j) =$  the proportion of words from the entire corpus assigned to topic  $t_j$ . For example, if there are 10,000 words in the entire corpus, and 350 of those words are assigned to topic  $t_j$ ,  $p = 0.035$  (350/10,000).

The proportions in a) and b) are then multiplied:  $(p(\text{topic } t_j | \text{document } d_i) * p(\text{word}_1 | \text{topic } t_j)) = 0.16 * 0.035 = 0.0056$ . After repeating this procedure for each document in the corpus, the process converges on the values that represent topic assignments for each document and for the entire corpus.

Collapsed Gibbs sampling<sup>5</sup> is then used to infer the hidden (latent) thematic structure in the entire population of documents:

- a) The topics in the corpus;
- b) The per-document proportion of topics; and
- c) The per-document, per-word topic assignment.

The 3-level Bayesian model is represented mathematically as:

$$P(\theta_{1:D}, z_{1:D} | w_{1:D}, \beta, \alpha) = \frac{[\Gamma(\sum_{i=1}^k \alpha_i) / \prod_{d=1}^D \Gamma(\alpha_i)] * (\theta_1^{(\alpha-1)} \dots \theta_k^{(\alpha-1)}) * \prod_{j=1}^D \theta_j^N}{[\Gamma(\sum_{i=1}^k \alpha_i) / \prod_{d=1}^D \Gamma(\alpha_i)] * \sum_{i=1}^k [(\theta_1^{(\alpha-1)} \dots \theta_k^{(\alpha-1)}) * (\prod_{n=1}^N \sum_{i=1}^k \prod_{j=1}^V (\theta_i \beta_{ij})^w)]}$$

### Footnotes to Appendix A

<sup>1</sup> The Bayesian equation is a three-level nested probabilistic model.

<sup>2</sup> This is the  $\alpha$  parameter needed for the Dirichlet distribution, as represented in Figure 1, *supra*.

<sup>3</sup> See *supra* notes 65-66 and accompanying text.

<sup>4</sup> As noted *supra*, text accompanying note 69, some common words, particularly conjunctions and prepositions, are excluded through a stop words list.

<sup>5</sup> Because the corpora have large numbers of documents, each containing potentially tens of thousands of words, it would take an extremely long time to calculate all the possible ways of assigning words to topics. The LDA process therefore incorporates a statistical technique known as collapsed Gibbs sampling which samples distributions to obtain approximate identified topics in each document, substantially accelerating the entire process.