Comparing Ambiguity in Privacy Policies and the Impact of Regulation

Joel R. Reidenberg, Jaspreet Bhatia, Travis D. Breaux, and Thomas B. Norton*

Table of Contents

I. Introduction 2
II. Defining and Measuring Ambiguity 3
A. Taxonomy of Vague and Ambiguous Terms 3
B. Grounded Analysis and Emergent Theory 8
C. Comparative Levels of Ambiguity 10
D. Ambiguity through Incompleteness 13
III. Scoring Ambiguity 14
A. The Landscape of Ambiguity in Privacy Policies 15
B. The Scoring Model 16
IV. Comparative Scores and the Impact of Regulation 18
A. Benchmark Scores for Regulated Disclosures 18
B. Company Scores for Unregulated Disclosures 20
C. Normative Role of Privacy Notice Regulation 22
V. Public Policy Considerations: Technological Tools, Linguistic Guidelines and Reporting 23
A. Technical Tools 24
B. Linguistic Guidelines 25
C. Reporting Framework 26

* Respectively, Stanley D. and Nikki Waxberg Professor of Law, Fordham University; Ph.D. candidate, Carnegie Mellon University; Assistant Professor of Computer Science, Carnegie Mellon University; Privacy Fellow, Fordham CLIP. Work on this project was supported in part by NSF Awards #1330596 and # 1330214 and a Fordham Faculty Fellowship. The authors would like to thank Stephen Broomell for his advice on testing validity and Stephanie Tallering for her assistance with coding.

I. Introduction

While scholars have shown weaknesses in the readability of privacy policies[1] and weaknesses in the substantive protections,[2] they have not focused carefully on policy ambiguity. Privacy policies often contain ambiguous language describing website practices for data processing activities such as collection, use, sharing, and retention. Ambiguity regarding these practices undermines the purpose and value of a privacy policy for website users. Without clear affirmative statements, privacy policies are, in effect, meaningless. They would convey no true indication to users of the website's actual practices and they would provide declarations that would be unenforceable. On a practical level, ambiguity also challenges the usability of privacy technologies for user empowerment.[3]

This paper will explore the problem of ambiguity in policy language. In Part II, we develop a theory for the definition of vague and ambiguous terms and for the measurement of such terms. In Part III, we develop a scoring method to compare the relative vagueness of different privacy policies. In Part IV, we apply the theory and method using natural language processing (NLP) techniques to score a set of privacy policies for clarity and comparison, We then use these comparative rankings to test whether regulation improves the clarity of privacy policies. To test the impact of regulation, we selected two benchmarks to show whether government-mandated privacy disclosures result in notices less ambiguous than those emerging from the market. The two benchmarks are the model financial services policy approved by federal regulators and a set of privacy policies adopted under the US-EU safe harbor inter-governmental agreement. The results provide normative insight on the role of privacy notice regulation. In Part V, we address a number of practical public policy considerations resulting from our scoring. The techniques and corresponding technical tools can provide companies with a useful mechanism to improve the drafting of their policies. At the same time, automated tools embodying our theory and scoring method enables regulators to easily scan industries and companies for poor language in their privacy policies. Such

[1] *See e.g.,* A. M. McDonald and L. F. Cranor, *The cost of reading privacy policies*, 4(3) I/S – A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY 543 (2008); Irene Pollach, *What's wrong with online privacy policies?*, 50 Comm. of ACM,103, 103 (2007); Carlos Jensen and Colin Potts, *Privacy policies as decision-making tools: an evaluation of online privacy notices*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04), available at* http://doi.acm.org/10.1145/985692.985752.

[2] *See* Florencia Marotta-Wugler, *Does "Notice and Choice" Disclosure Regulation Work? An Empirical Study of Privacy Policies* (Working Paper, Jan. 15, 2015); Pollach, *supra* note 1.

[3] For example, the joint Carnegie Mellon University, Fordham University and Stanford University usable privacy project seeks to combine crowd sourcing, natural language processing and machine learning to develop browser plug-in technologies that will automatically interpret privacy policies for users. See http://usableprivacy.org. If policies are too ambiguous, automated processing will be frustrated.

inexpensive scans revealing problems with privacy policy language then empowers regulators to more effectively target defective privacy policies for remedial action.

II. Defining and Measuring Ambiguity

A. Taxonomy of Vague and Ambiguous Terms

Vagueness and ambiguity arise when a statement is incomplete and missing relevant information, or when a word or phrase has more than one possible interpretation and the reader is uncertain about which interpretation the author intended. In contract theory, vagueness connotes a distribution around a norm without a clear delineation while ambiguity refers to situations where a word may have at least two meanings.[4] In each case, multiple interpretations can arise when a statement is incomplete, or when a generic word or phrase is used in place of a more specific word or phrase. When a website privacy policy uses vague or ambiguous terms, the language choices dilute the ability of a policy to describe the website's actual practices.

Because privacy policies summarize an organization's data practices, it is not surprising that policies include vagueness. There are at least two motivations for introducing vagueness: (1) the practices include divergent or separate situations where actions do and do not occur, in which case the action "may" occur, depending on what situation the individual encounters; and (2) there are foreseeable, yet unrealized actions that "may" occur in the future, and the policy authors wish to be flexible to accommodate those future actions without changing the policy. In the case of the first motivation, we believe changes to some policy statements can clarify under what situations the action does or does not occur, resulting in a less vague policy. However, the second motivation to accommodate flexibility is at best a form of inaccuracy and at worst misleading and misrepresentative.

To demonstrate this effect with an illustrative case study, we show a few examples from the Barnes and Noble privacy policy concerning *personal information* and we discuss the vague terms used in these statements along with their effects.[5] The Barnes and Noble policy includes two statements that describe the possibility of collection:

[4] *See e.g.*, E. Alan Farnsworth, Contracts, § 7.8 (3rd ed. 1999).
[5] The Barnes & Noble Privacy Policy, Barnes & Noble Privacy Policy, http://www.barnesandnoble.com/u/nook-for-windows-privacy-policy/379003719, is used here as an example of common drafting problems found in many other privacy policies.

(1) *"**Depending** on how you choose to interact with the Barnes & Noble enterprise, we **may** collect personal information from you . . . ."*[6]
(2) *"We **may** collect personal information and other information about you from business partners, contractors and other third parties."*[7]

In statement (1), the collection is conditioned upon how the user interacts with the company. This is vague, because the statement summarizes multiple situations, some of which will include the collection of personal information and some of which will not. To achieve clarity, however, it seems reasonable to exclude those situations where personal information is not collected, and to focus on where personal information "will be collected." Subsequent parts of the policy would then address situations were non-personal information "will be collected." In contrast, statement (2) is vague because the conditional situations are not described, thus all third-party transactions are summarized into a single statement. By separating these statements and iterating over the different categories, the policy authors can exclude prospective collections (envisioned, but not actual collections) and those situations where personal information is not collected.

In statement (3), below, it is difficult to envision situations where personal information would not be collected to make purchases or rentals. Presumably, payment information is necessarily collected at some point, but other kinds of personal information, such as one's height and weight, are not likely generally collected for completing a purchase. In contrast, sentence (4) shows clearly that the enrollment into Member Loyalty programs would require submissions of personal information.

(3) *"When you make a purchase, rent . . . you **may** need to submit personal information to us."*[8]
(4) *"Similarly, when you enroll in our Member Loyalty Program, we **will** ask you to submit personal information . . . ."*[9]

Another attribute of vagueness concerns the vague conditions and purposes under which information is used. In statement (5), below, the policy links the collection to a broad purpose (improving customer experience) under a general assumption about the "necessary" situations that define this broad purpose. An alternative statement would replace the phrase "as necessary" with specific purposes intended to improve customer service.

---

[6] *See* Barnes & Noble Privacy Policy, http://www.barnesandnoble.com/u/nook-for-windows-privacy-policy/379003719.
[7] *Id.*
[8] *Id.*
[9] *Id.*

(5) "*We collect your personal information in an effort to provide you with a superior customer experience and, **as necessary**, to administer our business."*[10]

Vagueness pertaining to standard third-party transactions is also evident in the following two statements from the Barnes and Noble privacy policy:

(6) *"In addition, we disclose **certain** personal information to the issuer of the MasterCard . . . ."*[11]
(7) *"If you are accessing our goods and services using a Microsoft account, Microsoft **may** share your personal information with us . . . ."*[12]

In statements (6) and (7), the mechanisms for exchanging personal information are coded in software. In the case of MasterCard, while their transaction processing rules are updated from time to time, the technical details of electronic payment processing are less likely to change and presumably the policy could restrict the kind of personal information to payment information or information for the purpose of completing a purchase. In statement (7), Microsoft's Live Connect API for OAuth 2.0 access to the Microsoft account is very explicit about what information "may" be shared (first and last name, email address, gender, age) and Barnes and Noble can further commit to which of these information types they "will" collect as encoded by their software.

As the case study reveals, the contours of ambiguity are very complex. The measurement of ambiguity, thus, becomes a valuable marker to signal whether a privacy policy is a meaningful notice of a website's actual policies and practices and a notice that might give rise to a contractual commitment. The first step in the measurement of ambiguity in privacy policies is the development of a rigorous and validated taxonomy of terms that can be used to examine a diverse set of online sectors such as shopping, news and financial services.

Linguistic scholars have identified various forms of ambiguity in the use of language.[13] Some terms may have inherent ambiguity. For example, many privacy policies use the modal verb "might" to describe data processing activities ("we might collect . . . ") that may or may not occur in the future. In addition, policies use conditional phrases, such as "when", "upon", and "during", that indicate an event upon which a particular statement becomes true ("upon consent, we will share . . . ") When multiple modal verbs and conditional terms are used together, readers cannot

[10] *Id.*
[11] *Id.*
[12] *Id.*
[13] *See e.g.*, Paul Hoffman, Matthew A. Lambon Ralph, and Timothy T. Rogers, *Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words*, 45 BEHAV. RES. 718 (2013).

actually determine if the described practices occur, or in what combination, or under which specific conditions or how to satisfy those conditions. Similarly, many privacy policies use elastic terms like "partners" or "affiliates" to describe potential recipients of user data. These terms are "elastic" because they can encompass different meanings in different contexts.

Scholars, though, have classified textual ambiguity in various ways. For example, one study of the language used in a regulation in the health care field looked at six types of ambiguity:

- *Lexical* – "a word or phrase with multiple, valid meanings"
- *Syntactic* – "a sequence of words with multiple valid grammatical interpretations regardless of context
- *Semantic* – "a sentence with more than one interpretation in its provided context"
- *Vagueness* – "a statement that admits borderline cases or relative interpretation"
- *Incompleteness* – "a grammatically correct sentence that produces too little detail to convey a specific or needed meaning"
- *Referential* – "a grammatically correct sentence with a reference that confuses the reader based on the conduct"[14]

These types were based on categories drawn from the field of literary criticism.[15] These categories were used to evaluate a legal text for the purposes of requirements engineering.

Another study that focused on privacy policies used four categories:

- *Lexical Choice*. Looking at the systematic use or avoidance of words;
- *Syntactical Transformation*. Exploring the use of passive voice and nominalizations;
- *Negation*. Examining which issues are denied; and
- *Modality*. Assessing the certainty of the speaker about the content of an utterance.[16]

These categories were used with respect to the substantive content of the online privacy policies.

For a rigorous analysis of textual ambiguity, the starting point is, thus, the establishment of a typology of ambiguous terms. Since our objective is to provide a

---

[14] Aaron Massey et al., *Identifying and Classifying Ambiguity for Regulatory Requirements*, Requirements Engineering Conference (RE), 2014 IEEE 22nd International, 83, at 86.
[15] *Id.* at 85 (citing W. Empson, Seven Types of Ambiguity. New Directions (1966)).
[16] Pollach, *supra* note 1.

qualitative rating of ambiguity, we have chosen to focus on the subset of vagueness. This means that our scoring will provide relational comparability, but underrate overall ambiguity. We define our typology with four categories in Table 1. From a legal perspective, *conditional* terms are inherently vague because the performance of a stated action or activity will be dependent on a variable trigger.[17] Similarly, *generalizations* are terms that describe multiple, more specific variants ("personal information" includes "full name," "birthdate," and so on). From the linguistic perspective, *modality* (modal verbs, adverbs and non-specific adjectives) creates uncertainty with respect to actual action;[18] this includes whether an action is possible, likely, permitted or obligatory, among others. If the action is only permitted, it may never occur, whereas obligatory actions are expected to occur in the future (the difference between "we may" and "we will"). Similarly, numeric quantifiers that are non-specific create ambiguity as to the actual measure. To assure the completeness of the typology, three researchers reviewed 15 policies first in their entirety and then statement by statement to identify vague phrases and determine if they fit these categories or if new categories were required.[19]

These categories are represented below:

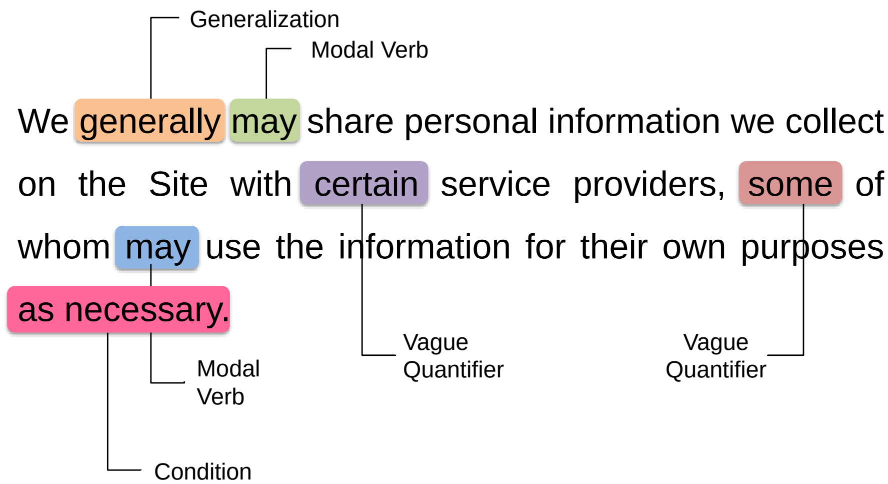
Table 1

Categories of Ambiguous Terms	
Category	**Description**
Condition	Action(s) to be performed are dependent on a variable or unclear trigger
Generalization	Action(s)/Information Types are vaguely abstracted with unclear conditions
Modalilty (including modal verbs)	Vague likelihood of action(s) or ambiguous possibility of action or event
Numeric quantifier	Vague quantifier of action/information type

To see how a sentence may reflect these categories, the phrase "we generally may share personal information we collect on the Site with certain service providers, some of whom may use the information for their own purposes as necessary"

[17] *See* Farnsworth, *supra* note 4 at § 8.2.
[18] *See* Kai von Fintel, *Modality and Language* in Encyclopedia of Philosophy (Donald M. Borchert, ed.) (2006), *available at* http://mit.edu/fintel/www/modality.pdf.
[19] This also included an evaluation of cases where the language might appear as a borderline classification.

contains a condition, generalization, modal verbs and numeric quantifiers.[20] These ambiguous terms are annotated in the sentence as follows:



In combination, these six forms of vagueness and ambiguity combine to allow any organization sharing personal information under this statement to share it with anyone for any purpose, as long as the recipient is a service provider. The combination of these six terms further leaves unclear the conditions under which information is shared, and the number or proportion of service providers that engage in this practice.

B. Grounded Analysis and Emergent Theory

Based on the taxonomy of vagueness, we analyzed a set of policies across a variety of sectors to identify and classify terms for each category. To assure saturation within each category, we chose to examine three diverse sectors (shopping, telecommunications and employment) and five policies within each sector.[21] We chose privacy policies of major sites that are visited by large numbers of users. Our shopping category consists of a major book seller (Barnes & Noble), warehouse merchandiser (Costco), clothing seller (JC Penny), hardware store (Lowe's), and online surplus and liquidator (Overstock). This reflects a diversity of types of shopping sites. Our telecommunications category consists of mobile (AT&T, Verizon), cable (Time Warner) and land websites (AT&T, Verizon, Charter). This reflects a diversity of types of telecommunications service providers. And, lastly,

---

[20] The original sentence was extended to include "generally" and "as necessary" for illustration purposes; the sentence without these additions is found in Lowe's website privacy policy on April 27, 2015 at: http://www.lowes.com/en_us/l/privacy-and-security-statement.html.

[21] We considered examining policies from top site rankings, but the various rankings did not assure diversity of sectors.

our job category consists of major sites CareerBuilder, GlassDoor, SimplyHire, Monster and Indeed. These reflect the major online job search sites.[22]

Table 2 shows these sectors and the specific policies reviewed to identify terms for each of the categories. The last policy updates were taken directly from the policy, where available.

Table 2

Type of Site	Policy	Last policy update
Shopping	Barnes and Noble	05/07/2013
Shopping	Costco	12/31/2013
Shopping	JC Penny	05/22/2015
Shopping	Lowes	04/25/2015
Shopping	Over Stock	01/09/2013
Telecommunications	AT&T	09/16/2013
Telecommunications	Charter Communication	05/04/2009
Telecommunications	Comcast	03/01/2011
Telecommunications	Time Warner	09/2012
Telecommunications	Verizon	10/2014
Employment	Career Builder	05/18/2014
Employment	Glassdoor	09/09/2014
Employment	Indeed	2015
Employment	Monster	03/31/2014
Employment	SimplyHired	4/21/2010

The analysis was performed manually on these policies by three researchers using coding.[23] This technique is a qualitative research method from the social sciences that is aimed at identifying instances of data in text, images or video.[24] This form of analysis is also called *grounded theory*, because the theory (or taxonomy) is grounded in the dataset. In our application, we began with an established coding frame based on our taxonomy in Table 1. This approach is an accepted practice for grounded analysis,[25] though it contrasts with another approach that takes the coding frame as solely emerging from the data.[26] In using the grounded theory form

[22] The small number of policies in each category preclude broad generalizations within and across categories, but do enable us to show the value of a score for comparison purposes, including comparison against the financial services benchmark

[23] *See* J. SALDAÑA. THE CODING MANUAL FOR QUALITATIVE RESEARCHERS (Sage Pubs., 2012).

[24] *Id.*

[25] *See* J. CORBIN AND A. STRAUSS, BASICS OF QUALITATIVE RESEARCH: TECHNIQUES AND PROCEDURES FOR DEVELOPING GROUNDED THEORY (Sage Pubs., 2007).

[26] *See* GLASER AND STRAUSS, THE DISCOVERY OF GROUNDED THEORY: STRATEGIES FOR QUALITATIVE RESEARCH (Sage Pubs., 1999).

of natural language analysis, the number of new terms or phrases that fall into a category will begin to diminish until no new instances are found after multiple observations are coded by the researcher. This condition is called saturation, and strengthens the claim that the taxonomy is reaching completeness for a specific domain (in our case, privacy policies). In this study, we reached saturation after analyzing 5 policies (Barnes and Noble, Lowes, Costco, AT&T, and Comcast).

The review of these 15 policies resulted in a set of terms for the taxonomy as shown in Table 3.

Table 3

Results from Applying Taxonomy to Privacy Policies	
Category	**Key Words and Phrases**
Condition	depending, necessary, appropriate, inappropriate, as needed
Generalization	generally, mostly, widely, general, commonly, usually, normally, typically, largely, often
Modality (including modal verbs)	may, might, can, could, would, likely, possible, possibly, unsure,
Numeric quantifier	anyone, certain, everyone, numerous, some, most, few, much, many, various

C. Comparative Levels of Ambiguity

While the taxonomy results in Table 3 present the terms that obscure the clarity of the policy descriptions, the taxonomy does not address the relative levels of ambiguity among the terms. For example, the following two statements appear to have different levels of ambiguity:

1) "*we may generally collect …*"
2) "*we may collect as necessary …*"

Each uses a modal verb ("may"), but the first statement containing the generalization "*generally*" seems less clear than the second statement containing the condition "*as necessary.*" The practice described "as necessary" suggests that collection will only occur in exceptional case while "generally" suggests that collection is likely to occur under broader circumstances. This qualitative difference in clarity may also be linked to the degree of flexibility that the textual language provides to the website. Language designed to give websites greater flexibility is likely to be perceived as more ambiguous. The statement "*may collect generally*" provides greater flexibility to the website than "*as necessary*." Consequently, the generalization term "*generally*" obscures for the user the website's activities more than the conditional term "*as necessary*."

In addition to variations in clarity among the categories, the combination of terms from different categories in the same sentence may also affect the level of ambiguity perceived in descriptions of privacy practices. Table 4 shows our initial hypothesis regarding the possible cumulative effect of vague terms represented as a lattice.

Table 4

The lattice begins with a modal statement "we *may* collect" and then in the next row adds a term from each of the remaining three different categories: a generalization term "generally," a conditional term "as needed" and a numeric quantifying term "some." Our initial assumption was that additional terms would increase the vagueness of the statement, i.e. reduce the clarity of the description of the data collection practice. With each successive combination of vague terms, from the first to the second, third and fourth rows in Table 4, vagueness would increase until some degree of saturation would occur (i.e., adding additional vague terms would have no significant impact on increasing vagueness).

The relative impact of each possible combination is critical to the development of an accurate score for a privacy policy's ambiguity. To determine this impact we conducted a paired comparison survey.[27] The survey results show the relationship

[27] A paired comparison survey is a standard statistical technique that collects multiple preferences between two statements from multiple judges and, through the aggregation of the results, establish a matrix of rating comparisons for all possible combinations of the terms being studied. An example of the paired comparison is:

> For each numbered question, please read each pair of statements, and identify which of the two statements best represents ***a more clear description*** of the company's treatment of personal information.

of combinations of terms on the level of ambiguity, enables the assignment of relative weights to different combinations of terms from one or more categories. We used the Bradley-Terry model to calculate the weights from the paired comparison data.[28]

These results are presented in Table 5. The table shows the Bradley-Terry coefficients for the combinations of conditions (C), numeric quantifiers (N), modal terms (M) and generalizations (G).

Table 5[29]

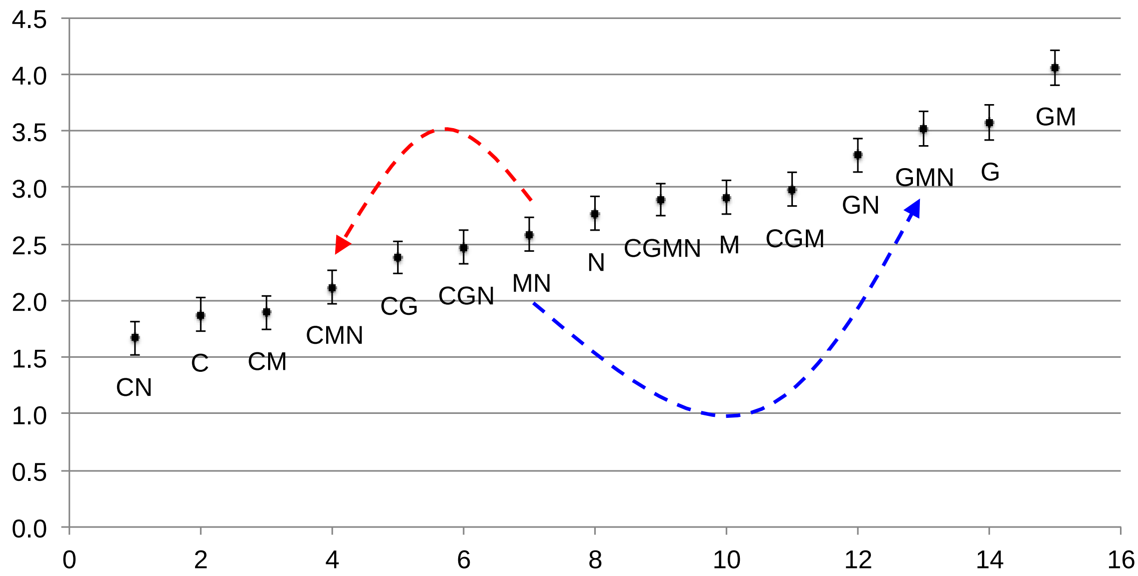
- We share your personal information *as needed*.
- We *generally may* share *some* of your personal information (Italics added to show terms.)

For more details on the survey, see Jaspreet Bhatia & Travis Breaux, *Technical Report: Automated Measurement of Privacy Policy Ambiguity* (work-in-progress)[hereinafter "Technical Report"].

[28] *See id.* The Bradley-Terry model scales preferences among different pair comparisons. *See generally* Heather Turner, David Firth, *Terry Models in R: The BradleyTerry2 Package*, 48(9). J. of Statistical Software 1 (2012). *available at* http://www.jstatsoft.org/v48/i09/.

[29] See Technical Report, *supra* note 27. The actual values represented in the graph are as follows:

Category of Vagueness	Coefficient	Standard Error
CN	1.667	0.1477
C	1.877	0.1472
CM	1.893	0.1473
CMN	2.119	0.1470
CG	2.382	0.1470
CGN	2.470	0.1471
MN	2.585	0.1474
N	2.765	0.1480
CGMN	2.892	0.1481
M	2.910	0.1483
CGM	2.978	0.1484
GN	3.286	0.1499
GMN	3.517	0.1513
G	3.571	0.1521
GM	4.059	0.1570



These results show the quantity that each combination of vague terms contributes to the overall concept of vagueness in the survey: that data practices described with combinations at the left of the chart have greater clarity than those practices described with combinations at the right of the chart. While phrases with both a conditional term and a vague numeric quantifier (CN) are indistinguishably clear from phrases with just a conditional term alone (C), we can observe how the vagueness taxonomy influences overall vagueness. The red arrow in the chart shows that condition terms increase clarity and reduce vagueness: e.g., statements with both a modal term and numerical quantifier (MN) are significantly more vague than similar statements with the addition of a conditional term (CMN). The blue arrow in the chart illustrates how generalizations significantly increase vagueness: e.g., the MN statements with the addition of a generalization (GMN) are more vague. By comparison, statements with a modal term and generalization are twice as vague as statements with a modal term and condition.

D. Ambiguity through Incompleteness

Lastly, silence in a privacy policy can often introduce ambiguity.[30] For example, if the policy is silent on sharing data with third parties, then the policy fails to convey whether and under what conditions data may be transferred to others. As a result, completeness of the privacy policy will have an impact on the scoring of ambiguity. While there are no legal requirements spelling out all the terms that must be contained in a privacy policy, various templates might be used to determine

[30] See Marotta-Wurgler, supra note 2.

completeness.[31] For purposes of this analysis, four elements will be addressed: collection, retention, sharing and use. These elements reflect the most significant privacy harms demonstrated litigation and that can be resolved by unambiguous privacy policy statements.[32]

III. Scoring Ambiguity

With the vagueness taxonomy populated using key words and phrases corresponding to each category, a comparative classification and a completeness indicator can be constructed to score the degree of affirmation or certainty associated with data practices for specific types of personal information. Privacy policy statements about companies that "might collect" are less certain than statements that they "will" or "will not collect" a particular information type. Highly uncertain statements can more easily accommodate a company's future practices, thus providing these companies more flexibility in the interim to alter those practices. However, highly uncertain statements allow for interpretations that may be untrue, thus giving users a false sense of privacy. By contrast, if an organization has a policy that is more certain, particularly with more restrictive practices, any new unstated practices would require a change in the policy. Such changes would trigger opportunities for users to re-evaluate their relationship with those

[31] For example, the Gramm-Leach-Bliley Act only stipulates that financial service companies provide notice to customers of their privacy policies and that the notice at a minimum contain certain types of disclosures. *See* Gramm-Leach-Bliley Act, 15 U.S.C. § 6803 (1999). The Federal Trade Commission and the Department of Commerce have each articulated several sets of fair information practices. *See e.g.*, Federal Trade Commission, *Privacy Online: Fair Information Practices in the Electronic Marketplace* (May 2000), *available at* https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission-report/privacy2000.pdf; Federal Trade Commission, *Privacy Online: A Report to Congress* (June 1998), *available at* https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-report-congress/priv-23a.pdf; U.S. Dept. of Commerce, Int'l Trade Adm, Notice: Issuance of Safe Harbor Principles and Transmission to European Commission, 65 Fed. Reg. 45665-45686 (July 24, 2000) also available at http://www.export.gov/safeharbor (last visited Oct. 5, 2015). See also Marotta-=Wurgler, supra note 2.

[32] *See* Joel R. Reidenberg, N. Cameron Russell, Alexander Callen, and Sophia Qasir, *Privacy Harms and the Notice and Choice Framework*, 11 I/S J. OF LAW & POLICY FOR THE INFO. SOC'Y (forthcoming 2015). Class action litigation and FTC enforcement actions show four categories of harm: surreptitious collection, unauthorized disclosure, wrongful retention and inappropriate security. *See id.* Statements on collection address the harm of surreptitious collection; statements on sharing and use address the harm of unauthorized disclosure; statements on retention address wrongful retention. We do not look for statements on data security because the failure to deploy adequate security cannot be resolved by notice. *See id.*

companies under the new practices. This opportunity to evaluate policy changes is necessary if the privacy principle of user consent is to have any meaning. Policies containing more certain statements are more likely to increase the opportunity for choice, since those policies will need to be revised each time a new practice is to be covered.

To score privacy policies, the first step is to determine if ambiguous terms are common in privacy policies through an analysis of the landscape of ambiguity found in privacy policies. The frequent existence of ambiguous terms leads to the definition of a scoring model that can then be applied to benchmarks and other privacy policies to rank their ambiguity against each other.

A. The Landscape of Ambiguity in Privacy Policies

When the taxonomy is applied to the set of 15 privacy policies shown in Table 2, every policy in the data set contains ambiguous terms. Table 6 shows the frequency of terms in each category that appear in the respective privacy policies.[33] The table also shows that the number and type of ambiguous terms vary across the policies.

Table 6
Frequency of Relevant Vague Terms by Category and Policy[34]

	Policy	Condition	General-ization	Modality	Numeric Quantifier	Complete-ness
Shopping	Barnes & Noble	6	0	33	5	0
	Costco	1	2	17	1	0
	JC Penny	1	0	15	2	0
	Lowes	2	0	18	4	0
	OverStock	0	1	13	1	0
Telecom	AT&T	0	0	27	4	0
	Charter Comm.	3	0	26	11	0
	Comcast	9	2	31	23	0
	Time Warner	0	0	21	4	0
	Verizon	2	0	36	6	0
plo ym	Career Builder	1	1	16	3	0
	GlassDoor	0	2	13	3	0

[33] *See* Technical Report, *supra* note 27.
[34] These frequency counts reflect each instance in which a word in the taxonomy represented in Table 3 is used in association with a type of personal information and a data processing action applied to the information. *See* Travis D. Breaux and Florian Schaub, *Scaling Requirements Extraction to the Crowd: Experiments on Privacy Policies*, 22nd IEEE International Requirements Engineering Conference (RE'14), Karlskrona, Sweden, pp. 163-172, Aug. 2014.

	Indeed	0	1	23	3	0
	Monster	1	0	12	2	0
	SimplyHired	2	1	22	5	0

In terms of distribution, the most frequently observed ambiguous terms are modal verbs followed by numeric quantifiers (MN). Conditions and generalizations lag far behind. Table 7 shows the distribution of ambiguous terms in these policies.[35]

Table 7

Distribution of Ambiguous Terms (% = vague terms in policy belonging to category / total number of vague terms in 15 policies)	
Category	**Terms**
Condition	6.4%
Generalization	2.3%
Modality	73.7%
Numeric quantifier	17.6%

This suggests that the use of modal terms will dominate all other terms in the calculation of overall vagueness in a privacy policy. However, the Bradley Terry coefficients in Table 5 show the significant impact that conditions, generalizations and numerical quantifiers have on modality: while modality along (M) scores at 2.910±0.148, the addition of generality to modality (GM) scores more vague at 4.059±0.157, and the addition of conditions to modality (CM) scores less vague at 1.667±0.148. In the extreme case that these additional categories of vagueness always appear with modal terms, then over one-third of the total 73.7% of modal terms will score well above or well below the coefficient for modality, alone. This can lead to significant differences between policies in terms of overall vagueness and especially pronounced differences within a single category of data practice (e.g., collection, retention, sharing, etc.) in the event that vagueness is concentrated in one area of the policy.

B. The Scoring Model

Simply counting the number of ambiguous terms in a privacy policy will not provide an adequate measure of ambiguity. For example, the AT&T policy contains 31 vague

---

[35] *See* Technical Report, *supra* note 27.

phrases, which places it well above the median of 25 vague phrases and slightly above Simply Hired, which has 30 vague phrases. But this frequency count does not indicate the relative context. Context matters, and a granular scoring model needs to take into account three key variables: 1) the existence of vague terms and their relation to specific categories of data practice (e.g., collection, retention, sharing, and usage); 2) the relative impact that a combination of vague terms may have on overall ambiguity; and, 3) the completeness of the policy.

To accomplish this goal, we propose a scoring model based on a relative comparison of vagueness in phrases for each policy. This score is based on a statistical measure that scales the overall vagueness of individual statements in each policy based on the Bradley-Terry model for paired comparisons.[36]

The coefficients in Table 5 that were computed by this method serve for these calculations to rank the vagueness of every phrase in each policy containing a vague term or combinations of vague terms associated with an action-information pairing where a data practice (action) is applied to a type of information (information).[37] The vagueness scores appropriately ignore phrases that do not specifically describe a data processing activity or that do not contain any vague terms. This means that non-relevant language, such as a corporation's philosophy relating to privacy, or unambiguously described data practices will not factor into the vagueness score.

For each policy, we can then calculate an aggregate vagueness score by taking the sum of the coefficients for each action-information pair containing vague terms. This policy-specific aggregate score is not, however, sufficient to compare two policies. For example, if a policy is long, it may contain more action-information pairs containing vague terms than a shorter policy, but proportionately be much clearer. To account for this situation, we normalize the aggregate vagueness score by dividing the aggregate score by the total number of action-information pairs in the policy; we call this normalized score the *vagueness score*. The vagueness score reflects positively if a policy has more action-information pairs that clearly describe data practices and negatively if the policy has more pairs that include vague terms. Moreover, it reflects the total unit vagueness independent of policy length, but relative to the level of contribution to vagueness by each category of vague terms in Table 3.

Lastly, in case a policy has a high level of ambiguity in paragraphs pertaining to key elements that may be masked by clear language elsewhere in the policy, we calculate the vagueness scores for the collection of policy statements addressing the

[36] *See* Table 5 and corresponding discussion, supra.
[37] This calculation only looked at one set of vague terms per statement in the policy so it may be biased downward where policies contain two separate pairings within the same statement and each pairing has associated vague terms.

four key data practices: collection, retention, sharing and usage. These scores are calculated in the same manner as those for the overall policy.

Separately, we report on the completeness of the privacy policies using a scale of 0 to 4. For each element missing from the four data practices (collection, retention, sharing and use), the policy is assigned one point. Thus, a policy containing descriptions for all four elements will score a 0 and a policy missing all four elements will score a 4.

IV. Comparative Scores and the Impact of Regulation

A. Benchmark Scores for Regulated Disclosures

Because the score ratios are designed to compare the clarity of policies against each other and do not provide a minimum level of acceptability for ambiguity, the *Model Privacy Form under the Gramm Leach Bliley Act*[38] can serve as an informative target benchmark. This model form was adopted by regulatory agencies after careful analysis and testing of language options.[39] In fact, eight federal financial service regulatory agencies approved the language used in this standardized privacy disclosure statement.[40] Financial service providers may use the model form to satisfy their obligations under the Gramm-Leach-Bliley Act, though they are not required to adopt its language.

Table 8 presents the vagueness score calculations for a set of financial institutions' privacy policies based on the *Model Privacy Form.*[41] Where the ratios are in proximity to each other, they indicate that those policies have similar levels of ambiguity. Where a ratio is double another, the ratios indicate that the policy with the higher ratio is twice as vague as the policy with the lower ratio.

Table 8

[38] *See* Office of the Comptroller of the Currency, Treasury (OCC), Board of Governors of the Federal Reserve System (Board), Federal Deposit Insurance Corporation (FDIC), Office of Thrift Supervision, Treasury (OTS), National Credit Union Administration (NCUA), Federal Trade Commission (FTC), Commodity Futures Trading Commission (CFTC), and Securities and Exchange Commission (SEC), *Final* Model Privacy Form *under the Gramm-Leach-Bliley Act*, 74 Fed. Reg. 62890-62994 (Dec. 1, 2009).
[39] *Se*e Alan Levy and Manoj Hastak, *Consumer Comprehension of Financial Privacy Notices: A Report on the Results of Quantitative Testing*, Interagency Notice Project (Dec. 15, 2008), *available at* http://www.sec.gov/comments/s7-09-07/s70907-21-levy.pdf.
[40] *See supra* note 38.
[41] *See* Technical Report, *supra* note 27. Because the *Model Privacy Form* must be applied in context, we chose a set of large, national financial institutions that used the model form as the basis for their privacy policies.

Financial Services' Vagueness Scores

Privacy Policy	Vague-ness Score	Collect	Retain	Share	Use	Complete-ness
Bank of America	0.80	0.42	0.00	0.96	0.00	0
Capital One	0.51	0.00	2.90	0.52	0.00	0
Citi Group	0.41	0.00	*	0.57	0.00	1
JP Morgan	0.12	0.00	0.00	0.31	0.00	0
Mean	0.46	0.42	0.97	0.59	0.00	.25

* The policy does not talk about retention practices. The mean score for retention is computed excluding this policy.

While the model policy language is designed for financial services, the vagueness score of these policies represents a standard of clarity that is acceptable to federal government agencies. A score lower than the ratio for the *Model Privacy Form* means that the evaluated policy uses clearer language than the government's standard. A score higher than the *Model Privacy Form* means that the evaluated policy does not meet the acceptable government standard.[42]

Another benchmark can be derived from the US-EU Safe Harbor Agreement [EU Safe Harbor].[43] The EU Safe Harbor identifies data practices that must be contained and described in a privacy policy to satisfy European data export requirements, but stops short of providing model language, as does the *Model Financial Privacy Policy*.[44] The EU Safe Harbor terms were, however, negotiated between the US Department of Commerce and the European Commission and approved by the US Department of Commerce.[45] Companies may benefit from the EU Safe Harbor if they

---

[42] Some parts of the *Model Privacy Form* are boilerplate descriptions of the importance of privacy rather than the website's practices. These sentences would increase the denominator and thereby lower the vagueness rating. Since other privacy policies have similar boilerplate language, we believe that this measure is still valid for purposes of comparing one policy's vagueness against another's. In any case, if there is bias in the overall policy vagueness score, this will be revealed by the analysis of specific data practice areas.

[43] *See* Safe Harbor Agreement, *supra* note 31 . On October 6, 2015, the Court of Justice of the European Union, invalidated the Safe Harbor Agreement for defects in the privacy rights afforded to EU data. Schrems v. Ireland, CJEU C-362/14 (Judgment of Oct. 6, 2015), http://tinyurl.com/occ4mvx. Although the Safe Harbor Agreement was struck down, the policies written under the agreement are nevertheless a valuable benchmark for a regulatory rule that does not provide explicit language for privacy notices.

[44] *Id.* Safe Harbor requires privacy policies to provide consumers with : (1) notice; (2) choice; (3) disclosures of onward transfer; (4) security ; (5) data integrity; (6) access; and (7) enforcement. Id.

[45] *Id.*

include these terms in their privacy notices and register with the US Commerce Department.

Of the 15 companies in our data set (Table 2), five are members of the EU Safe Harbor.[46] Table 9 applies the scoring model to these five privacy policies.[47]

Table 9
Safe Harbor Companies' Vagueness Scores

Privacy Policy	Vagueness Score	Collect	Retain	Share	Use	Completeness
Barnes & Noble	2.11	2.22	1.52	2.26	1.93	0
Career Builder	0.79	0.58	0.72	1.17	0.85	0
GlassDoor	1.55	1.46	1.45	1.57	1.65	0
Indeed	1.44	1.04	1.10	2.26	1.33	0
Monster	0.77	0.78	0.73	1.26	0.49	0
Mean	1.33	1.22	1.10	1.70	1.25	0

The mean vagueness score for the financial services policies is considerably lower than the Safe Harbor policies: 0.46 to 1.33. This striking three-fold difference means that financial services policies are three times clearer than Safe Harbor policies. Similarly, the vagueness scores show that the descriptions of the four data practices found in the financial services policies are much less vague than those found in the Safe Harbor policies. As a benchmark, the *Model Privacy Form* for the financial services industry holds privacy policies to a higher standard of clarity and allows less ambiguity than the US-EU Safe Harbor.

All the benchmark policies were complete with sole exception of Citi Group that was silent on data retention.

B. Company Scores for Unregulated Disclosures

Applying the scoring model to the unregulated companies results in the vagueness scores reported in Table 10.[48]

[46] *See* US. Dep't of Comm. US-EU Safe Harbor List, https://safeharbor.export.gov/companyinfo.aspx. One additional company, Time Warner , does subscribe to Safe Harbor, but not for customer information. *See* https://safeharbor.export.gov/companyinfo.aspx?id=28829
[47] *See* Technical Report, *supra* note 27.
[48] *See id.*

Table 10
Unregulated Companies' Vagueness Scores

Policy	Vagueness Score	Collect	Retain	Share	Use	Completeness
Costco	0.89	0.71	0.68	1.26	0.60	0
JC Penny	1.12	1.30	1.95	1.02	0.99	0
Lowes	1.29	1.19	1.66	1.16	1.33	0
OverStock	1.86	1.59	1.46	2.06	1.93	0
AT&T	1.00	0.86	0.47	1.23	0.95	0
Charter Comm.	1.55	1.53	0.98	1.57	1.73	0
Comcast	1.84	2.07	2.20	1.82	1.60	0
Time Warner	2.10	2.14	2.85	1.66	2.20	0
Verizon	1.36	1.53	1.09	1.44	1.20	0
Simply Hired	1.78	1.58	1.30	1.36	2.18	0
Mean	1.48	1.45	1.46	1.46	1.47	0

As shown in Table 10, the most ambiguous policies among the unregulated entities belong to Time Warner, with Overstock.com, Comcast, and Simply Hired closely following. The notably higher ratios for these policies indicate that they use significantly more ambiguous language than the other privacy policies. As observed in Table 6, each of these policies uses large numbers of ambiguous modal verbs and quantifiers. For example, Comcast describes sharing with third-parties in these terms:

> *"In **certain** situations, third party service providers **may** transmit, collect, and store this information on our behalf to provide features of our services."*

By contrast, Costco's language describing sharing with third parties is more direct:

> *"We do not otherwise sell, share, rent or disclose personal information collected from our pharmacy pages or maintained in pharmacist records unless you have authorized such disclosure, or such disclosure is permitted or required by law."*

By comparison to these most vague policies, the policies belonging to Costco, AT&T and JC Penny are almost twice as clear.

Table 10 also shows the vagueness scores for actions to collect, retain, share and use information. The overall mean vagueness across these four data actions varies little from 1.45-1.47; however, the variance is not homogenous across practices (collect variance =0.21, retain variance=0.52, share variance =0.10, and use variance=0.30). This variance across practices shows divergent uses of vague terms across companies, with the least consistency across policy descriptions of retention practices, and the most consistency around descriptions of sharing practices. Notably, companies such as Overstock, Comcast, and Time Warner score more vague than average in all four data practice categories. For the website user, however, Overstock's high vagueness score for sharing presents a more significant, or fundamentally different, privacy risk than Comcast's vagueness regarding collection and retention. Vagueness with respect to sharing is significant because third parties are rarely identified in privacy policies and most privacy policies disclaim responsibility for the data practices of the unnamed third parties. Vagueness with respect to collection and retention affords companies greater flexibility in broadening what kinds of information they are potentially collecting. This may or may not present heightened privacy risks. However, when combined with ambiguous sharing terms, website users will not be able to ascertain exactly what information may be at risk of sharing with third parties.

All the policies not subject to regulation were complete.

C. Normative Role of Privacy Notice Regulation

Comparing the vagueness scores for the benchmarks against the unregulated policies shows that more specific regulation of policy language has a positive impact on the clarity with which privacy policies describe data practices. Table 11 compares the mean vagueness scores of the two benchmarks to the privacy policies of the unregulated companies.

Table 11
Mean Vagueness Scores across Sector, Regulated and Unregulated Policies[49]

Privacy Policy	Vagueness Score	Collect	Retain	Share	Use
Financial Services	0.46	0.42	0.97	0.59	0.00
Employment	1.26	1.08	1.06	1.52	1.29
Shopping	1.45	1.40	1.45	1.55	1.35
Telecommunications	1.57	1.62	1.52	1.54	1.53
Safe Harbor	1.33	1.22	1.10	1.70	1.25
Unregulated	1.48	1.45	1.46	1.46	1.47
Mean	1.26	1.20	1.26	1.39	1.15

[49] *See supra* Table 8 (vagueness scores for financial services policies); Table 9 (vagueness scores for Safe Harbor policies). The unregulated policy mean is calculated from the scores shown in Table 10 excluding the Safe Harbor companies.

The overall mean score for vagueness across data collection, retention, sharing and usage is 1.26. While the number of policies per sector is presently small, we see a few mean differences when accounting for outliers or extreme differences. Among those surveyed, telecommunications companies with the exception of AT&T show higher vagueness about collection practices than employment or shopping.

The mean vagueness score for the financial services policies in Table 11 is 0.46, which is more than three times lower than the mean for the unregulated policies. Similarly, the mean vagueness score for the Safe Harbor polices is 1.33 which is slightly lower than the unregulated companies, but almost three times higher than the financial services benchmark. With respect to the descriptions of data practices (collection, retention, sharing, and use), the unregulated privacy policies have significantly less clarity than the financial services benchmark and with the exception of sharing, noticeably less clarity than the Safe Harbor policies.

The threefold difference between the means for financial services policies (0.46) and the Safe Harbor policies (1.33) indicates that a regulatory nudge providing specific model language (financial services) results in less ambiguous policies than a regulatory framework (Safe Harbor) only stating data practices that must be described in a policy.[50] The difference also shows that Safe Harbor does a poor job assuring clear descriptions of data practices when compared to the Model Privacy Form.

Overall, the comparisons to the benchmarks indicate that the market produces privacy policies that are more ambiguous than those subject to some form of regulation. The least ambiguous policies were those adopted by financial services companies after government regulators provided their imprimatur to specific language. While financial institutions were not required to adopt the *Model Privacy Form*, doing so assured compliance with their legal obligations under Gramm-Leach-Bliley. At the same time, the policies with the highest level of vagueness were those of the unregulated companies.

V. Public Policy Considerations: Technological Tools, Linguistic Guidelines and Reporting

Because the vagueness scores show that privacy policies for unregulated companies are notably ambiguous when compared to the federal benchmark privacy policy for

[50] Interestingly, the unregulated policies are less vague with respect to sharing than the Safe Harbor policies.

financial institutions, there is a critical need to improve the clarity of online privacy policies. The finding calls for three policy tools. First, technological mechanisms that can easily score the vagueness of large numbers of policies to identify those that are problematic must exist. Second, linguistic guidelines need to be developed so that drafters have a framework to reduce vagueness in policies and so that regulators can point to a set of norms that reduce ambiguity for users. And, lastly as well as a framework to assist privacy policy drafters with reducing vagueness and ambiguity in policies.

A. Technical Tools

The ability to identify and score large numbers of privacy policies for vagueness can help drafters recognize and improve the clarity of privacy policies and help regulators identify industries or companies for improvement or enforcement actions. Natural language processing (NLP) and machine learning (ML) tools can provide this functionality. The NLP tools can take entire privacy policies and, using the taxonomy in Table 3 and the vagueness scoring method based on the coefficients in Table 5, generate comparable vagueness scores for the policy.[51] ML tools can be trained and used to extract policy language relevant to specific data practices.[52] These extracted paragraphs can then be scored separately by NLP tools. These processes would enable drafters to easily identify ambiguity issues in their policies. The causes of vagueness may be due to a desire for flexibility, or it may be due to the policy author's incomplete knowledge about the actual data practices. In this last respect, a vagueness score could be used as motivation for conducting internal audits to rationalize the use of less vague language for specific practices. Such audits could improve internal traceability and transparency, resulting in a company's rise to a higher standard of care in regards to data protection and privacy.

Because these processes are automated, they can be easily executed on large numbers of privacy policies. With a database of thousands of privacy policies, scores could be generated automatically and outliers flagged for investigation by relevant regulatory authorities. For example, the Global Privacy Enforcement Network composed of data privacy regulators around the world conducted a manual sweep of internet sites to identify transparency issues for privacy.[53] This can be done on a much larger scale through these automated tools.

---

[51] These NLP tools were developed for this paper and will be released as part of the Usable Privacy Project. *See* http://usableprivacy.org.
[52] These ML tools are under development by the Usable Privacy Project and will be released once completed. *See id.*
[53] *See* Office of the Privacy Commissioner of Canada, Privacy enforcement authorities launch first ever international internet privacy sweep, May 6, 2013, https://www.priv.gc.ca/media/nr-c/2013/nr-c_130506_e.asp.

B. Linguistic Guidelines

We propose a set of guidelines designed to yield statements that will, through the use of these natural language processing tools, measurably show improvement in clarity. The goal of these guidelines is a reduction in the vagueness score.

For the improving the vagueness scores, four principles can be applied:

1) avoid terms that are shown to be problematic, specifically modal verbs and vague numeric quantifiers;

Vague terms, such as modal verbs, are easier to employ because they afford the greatest flexibility. However, better communication between policy authors and system developers can surface the various conditions under which data practices arise. These conditions can be categorized and the categories can be used to signal clearly in policies when specific types of personal information are used for which purposes. Most importantly, purposes for which use or sharing is not primary or expected can be highlighted to better inform users.

Specifically, policy drafters should avoid the terms in Table 3 to increase clarity in privacy policies. Evidence from our analysis of the financial service sector shows that companies can reduce vagueness and hedging and improve clarity.

2) use a glossary of key terms

In requirements engineering, glossaries are routinely part of disambiguation and reaching agreement among a broad class of stakeholders. By encoding a glossary in privacy policies, a company's legal counsel can standardize terminology with a broad range of software developers (website developers, mobile app developers, database administration, and backend office administration). Glossaries can also be used to identify sensitive information types, such as age or contact list, and ensure those types receive special protections across collection, retention, sharing and use.

3) signal when the meaning of terms change within a policy such that automated tools can readily detect the changed meaning

Based on our analysis, we observed that some policies use signals to indicate when the meaning of terms change. These signals include differentiation through sub-categories, such as distinguishing personal from technical information, or distinguishing identifiable from demographic information. In addition, distinguishing when interpretations of key terms should be excluded can help clarify changing meanings. Across different policies, for example, the IP address is and is not viewed as personal information. This distinction can be made explicit in the policy, as the classification can affect whether the information is shared with third parties when those practices refer only to personal information, generally.

C. Reporting Framework

Beyond these guiding principles, we propose a framework to enable public reporting of the vagueness scores to can enhance the public transparency of ambiguous privacy policies. The framework consists of two elements. First, we propose that the vagueness scores represent measurements that are comparable across privacy policies. These scores represent the vague quality of a policy rather than the actual practice of the website or the substantive content of the privacy protections being offered. Second, we propose that the scores for the *Model Privacy Form* be used as a benchmark standard for acceptable ambiguity against which other policies be measured. In the future, regulators may wish to present new scores and thresholds that companies can seek to achieve through better policy language that consumers can understand.

VI. CONCLUSION

TBD