Concordance of professional ethical practice standards for the domain of Data Science: A white paper

Rochelle E. Tractenberg

Rochelle Tractenberg
Building D, Suite 207
Georgetown University Medical Center
4000 Reservoir Rd. NW
Washington, DC 20057 USA
rochelle.tractenberg@gmail.com

Tractenberg, RE. (2020, 20 February). Concordance of professional ethical practice standards for the domain of Data Science: A white paper. Published in the *Open Archive of the Social Sciences (SocArXiv)*, DOI 10.31235/osf.io/p7rj2

Shareable under a CC-by attribution-noncommercial-no derivatives 4.0 license

Running Head: Professional ethical practice standards for data science

Acknowledgement: While this white paper is intended to be *factual* - and the facts reported are correct - any opinions expressed are my own, and do not represent the perspectives of the American Statistical Association, Association for Computing Machinery, or Georgetown University.

Abstract

Data science is a discipline that has emerged at the intersection of computing and statistics – two disciplines with long standing guidance for ethical practice that feature professional integrity and responsibility. The 2018 National Academies of Science Report on Envisioning the Data Science Discipline recommends that “*The data science community should adopt a code of ethics*”, but due to its recency and to the diversity of paths into data science as a discipline, there is no real “community” that can do or organize this adoption. To support this recommendation, this white paper is an effort to document concordance across professional association practice standards, intended to support *the ethical practice* of data science by appealing to the consensus of these professional organizations on what constitutes ethical practice. The American Statistical Association (ASA) and the Association of Computing Machinery (ACM) recently revised their professional ethical practice standards in 2018. Both sets of guidance represent the perspectives of experienced professionals in their respective domains, but both organizations explicitly state that the guidelines apply to – should be utilized by – *all* who employ the domain in their work, irrespective of job title or training/professional preparation. Given that both statistics and computing are essential foundations for data science, their ethical guidance should therefore be a starting point for the community as it contemplates what “ethical data science” looks like.

The work of analyzing concordance in ethical guidance begins with a qualitative examination of the overlap (similarly worded principles), alignment (thematically similar principles), and gaps (dissimilar principles) that exist between existing sets of standards. To that end, the ethical practice guidance has been thematically analyzed from the standards outlined by the ASA, ACM, the International Statistics Institute, Royal Statistical Society, and the *Ethics in Action* guidance drafted by the Institute of Electrical and Electronics Engineers (IEEE) Initiative on Ethics of Autonomous and Intelligent Systems. This synthesis is intended to capture similarities and differences in relevant practical guidelines, integrating professional organizational perspectives on what constitutes ethical practice in data science to support and strengthen the domain. Ultimately, guidelines for ethical data science that reflect the concordance of cognate disciplines can ensure coherent integration of the features of ethical practice into training of data scientists - for both the practitioner and those who use data science, or its outputs, in their work.

Introduction

According to the National Academies of Sciences, Engineering, and Medicine (NAS, 2018), “data science spans a broad(er) array of activities that involve applying principles for data collection, storage, integration, analysis, inference, communication, and ethics”. (p. 1) With respect to “ethics”, the Committee on Envisioning the Data Science Discipline “... underscores the centrality of studying the many ethical considerations that arise as workers engage in data science. These considerations include deciding what data to collect, obtaining permissions to use data, crediting the sources of data properly, validating the data’s accuracy, taking steps to minimize bias, safeguarding the privacy of individuals referenced in the data, and using the data correctly and without alteration. It is important that students learn to recognize ethical issues and to apply a high ethical standard.” (p. 2) This echoes the 2013 undergraduate guidelines for computer science curricula, “(c)urricula must prepare students for lifelong learning and must include professional practice (e.g., communication skills, teamwork, ethics) as components of the undergraduate experience.” (ACM/IEEE-CS Joint Task Force, 2013: p.21). Two of the 2018 NAS action items follow directly from the Committee’s focus on the centrality of ethical professional practice in data science:

Recommendation 2.4: Ethics is a topic that, given the nature of data science, students should learn and practice throughout their education. Academic institutions should ensure that ethics is woven into the data science curriculum from the beginning and throughout.

Recommendation 2.5: The data science community should adopt a code of ethics; such a code should be affirmed by members of professional societies, included in professional development programs and curricula, and conveyed through educational programs. The code should be reevaluated often in light of new developments.
(p. 3)

These recommendations are not unique to the discipline of data science: Golde & Walker (2006) stated that “**Upon entry into practice, all professionals assume at least a tacit responsibility for the quality and integrity of their own work and that of colleagues. They also take on a responsibility to the larger public for the standards of practice associated with the profession.**” (emphasis added; p. 10). This quote articulates an *assumption* about professional practice – namely, that everyone who enters a profession has been sufficiently prepared to take on the responsibilities for the standards that define that profession or discipline. The assertion was made (in the 2006 book on doctoral level preparation) in the context where individuals will have been formally trained as *stewards* of a particular discipline – those to whom the vigor, quality, and integrity of a particular discipline can be entrusted (Golde & Walker 2006: p. 5; see also Rios et al. 2019); however, many who enter into the domain of data science do so after training in a variety of *other* fields (Hogan et al. 2017). Moreover, while there are many new jobs emerging every day with the specific title/duties of “data science” that are consistent with the NAS definition, many more jobs exist, or will evolve, where data science tools, techniques, and methods are only a part of daily practice. Thus, specific features of the NAS recommendations are especially salient: that academic institutions should weave ethical problem solving throughout data science curricula as well as individual courses; and that professional societies

should affirm the importance and relevance of generally accepted, consensus-based and authentic, ethical codes for all who engage in data science.

Data science is a new discipline, but it arises from two disciplines with long-standing commitments to ethical practice: computing and statistics. Ethical guidelines have been developed over several decades to support ethical professional practice with – as well as the application of – tools, techniques, and methods from both statistics and computing (Tractenberg et al. 2015; see also Tractenberg 2019-a). These substantial efforts can simplify the accomplishment of both NAS recommendations. Instead of creating a new code of ethics, existing codes should first be explored for their relevance and potential contributions to ethical data science. Wolpe (2006) noted, “From stem cells and cloning to genetic engineering to the sale of organs for transplant, there is no dearth of contentious bioethical debates. Sometimes the debates seem intractable, with all sides convinced of the validity of their ethical position. *It is easy to conclude that ethics is essentially arbitrary...* In fact, however, there is widespread consensus on a host of ethical issues in science policy. Consensus tends to be hidden because it is taken for granted; only the controversies make the headlines.” (emphasis added; p. 1024) If all of the professional organizations in the world that seek to represent the domain of data science create their own ethical guidelines, then this may contribute to the perception that Wolpe identified, namely, that “ethical data science” is essentially arbitrary. Another concern deriving from a wide variety of ‘ethical guidelines’ for data science is the argument that “Ethical principles are worse than useless if we don’t allow them to change our practice, if they don’t have any effect on what we do day-to-day.” (Loukides et al. 2018); proliferating guidelines are unlikely to be used in daily practice, but may also possibly be just insufficiently specific (e.g., Baumer et al. 2020, pp. 5-8).

The National Academies (National Academies of Sciences and Engineering and the Institute of Medicine (NASEM) called attention to “disruptive research practices” in their 2017 report, “Fostering Integrity”:

“Actions such as failing to retain or share data and code supporting published work *in accordance with disciplinary standards*, practices such as honorary or ghost authorship, and using inappropriate statistical or other methods of measurement and data presentation to enhance the significance of research findings are clearly detrimental to the research process and may impose comparable or even greater costs on the research enterprise than those arising from research misconduct.” NASEM 2017, p. 206 (emphasis added).

The NASEM report urges greater attention to be placed on *training new scientists* to avoid these practices; however, researchers are not the only users of data science, statistics, and computing who need to work in accordance with disciplinary standards. Moreover, the ASA and ACM could argue that their disciplinary standards are the ones to be followed, since some non-statistics/non-data science disciplinary norms (e.g., Wang et al. 2018) might actually reinforce the disruptive practices the NASEM report discusses. Thus, whatever ethical data science training paradigm(s) are created need to be effective both across the curriculum (NAS Recommendation 2.4) and also across disciplines. The domain of data science is dependent on statistics and computing, and as both the American Statistical Association (ASA), representing roughly 18,000 practitioners worldwide. The Association for Computing Machinery (ACM),

representing roughly 100,000 computing professionals worldwide, assert that their ethical practice guidance should pertain to members and non-members alike who utilize their methods and techniques – promoting the ethical use of statistical and computing practices; thus, these are a natural first step for understanding and describing ethical data science.

ASA Ethical Guidelines

The ASA Ethical Guidelines for Statistical Practice (ASA, 2018) includes 52 items under eight general areas:

- A.** Professional Integrity & Accountability (7)
- B.** Integrity of data and methods **(11)**
- C.** Responsibilities to Science/Public/Funder/Client (5)
- D.** Responsibilities to Research Subjects (7)
- E.** Responsibilities to Research Team Colleagues (4)
- F.** Responsibilities to Other Statisticians or Statistics Practitioners (4)
- G.** Responsibilities Regarding Allegations of Misconduct (6)
- H.** Responsibilities of Employers/Clients Employing Statistical Practitioners (8)

Two things are highlighted in the list above: firstly, that there are 11 different elements in the principle “integrity of data and methods” – more than any other principle - and secondly, that there is a principle that is entirely dedicated to what employers of anyone who does/uses statistics should know, understand, and expect. These principles highlight the relevance of the ASA Guidelines in any activities that involve **data** - whether it is big data or small data, and also, whether the user is the “statistician”/data analyst or the employer/client of that statistician or data analyst. The ASA Guidelines are included in full in the Appendix. Also included in the Appendix are alignment tables that document the near-complete alignment between the ASA Ethical Guidelines and those for the Royal Statistical Society (RSS, 2014) and International Statistics Institute (ISI, 2010).

It is important to note that, while both the RSS and ISI have invested significant effort in their respective professional ethical statements, neither has articulated that their ethical guidelines pertain to non-members or to all those who utilize statistical methods and/or techniques. By contrast, the ASA Guidelines explicitly articulate that, “(t)he principles expressed here should guide both those whose primary occupation is statistics and those in all other disciplines who use statistical methods in their professional work... comprising statisticians at all levels of the profession and members of other professions who utilize and report statistical analyses and their implications.” The Guidelines promote ethical science and ethical evidence-driven decisions, and are intended “to help statistics practitioners <including those who use and report statistical analyses and their implications> make decisions ethically and communicate their rationale (s) honestly”. The ASA Ethical Guidelines for Statistical Practice represent a “virtue ethics” perspective: ethical statistical practice is described as “what the ethical practitioner would do” or how the ethical practitioner does his/her work. The ASA Guidelines represent a formal structure to promote the conduct, aims, and qualities that characterize or mark ethical practitioners who use statistics (but are not statisticians themselves), as well as ethical professional statisticians. With more than 50 specific elements, the ASA Guidelines are also more extensive than those of

the other statistical organizations. Because of the near-total alignment between the ASA Guidelines and those of the RSS and ISI, the ASA Guidelines are the ones featured in the alignment explorations with the ACM Code that follow.

The ASA Guidelines were revised (2016) to support professionalism, which is defined by Merriam Webster as, “...the skill, good judgment, and polite behavior that is expected from a person who is trained to do a job well.” ASA members and nonmembers alike are obliged to promote understanding, through trust in their professionalism and judgment, when they engage in data analysis. Specifically, the ASA Guidelines articulate, in the Preamble, that they exist “(t)o inform those relying on statistical analysis, including employers, colleagues and the public, of the standards that they should expect.”

The ASA Ethical Guidelines for Statistical Practice are intended to support every practitioner in the quantitative sciences –whether trained as a statistician or as a scientist or other professional with quantitative expertise. The 2016 revision of the Guidelines (maintained in the 2018 revision) represent the habits of mind that characterize the experienced – ethical - quantitative practitioner and as such, they were specifically intended to be disseminated and used in teaching and training, and can promote professionalism, and/or the development of professional identity, in all who use statistics.

ACM Code of Ethics

The ACM Code of Ethics was first published in 1992 and most recently revised in 2018 (ACM, 2018). The ACM Code represents a “utilitarian ethics” perspective: it directs “computing professionals” to focus attention on the positive and negative effects of their decisions – and emphasizes decisions that avoid or minimize harms. In the preamble, it is stated that computing professionals act responsibly when “consistently supporting the public good”. However, because the ACM Code also specifies what “a computing professional *should*” do (in order to comply with the Code), it can also be interpreted from a virtue perspective, i.e., by describing the ideal professional (an ethical one). There are four main areas with specific points elaborated in terms of what “a computing professional should” do:

1. General ethical principles (7)
2. Professional responsibilities (9)
3. Professional leadership principles (7)
4. Compliance with the code (2)

Like the ASA Guidelines, the ACM Code seeks to support ethical decision making rather than to establish punishment, or describe explicitly what is right or wrong about aspects of practice. In the preamble, it is noted that,

“The Code as a whole is concerned with how fundamental ethical principles apply to a computing professional’s conduct. The Code is not an algorithm for solving ethical problems; rather it serves as a basis for ethical decision-making. When thinking through a particular issue, a computing professional may find that multiple principles should be taken into account, and that different principles will have different relevance to the issue.”

Additionally, the Code states that (it is) “designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way.” Following the 2018 revision of the ACM Code, a book was compiled to support teaching computing professionals about the Code; this is freely available on the ACM website (<https://www.acm.org/binaries/content/assets/about/acm-code-of-ethics-booklet.pdf>; individual case studies are also available, <https://www.acm.org/code-of-ethics/case-studies>). The ACM Code of Ethics is included (in full) in the Appendix. Also, in Table 1, each of the 25 ACM elements is examined for thematic alignment with the eight principles of the ASA Guidelines.

As they are representative of two essential constituent disciplines for data science, Table 1 explores the thematic alignment of the ACM Code (rows) and the ASA Guideline Principles (columns). ACM Principle elements are given, and alignment with the ASA Principle is indicated in the cell, with the specific ASA element (“A” means concordance is with the A Principle statement; “A1” means concordance is exact or directly comparable with the first element under Principle A, etc.). The specific ASA element or Principle is indicated if there is a thematic or exact match to the specific element of the ACM Code. If the alignment is abstract (e.g., if the core idea is “professional integrity” but this is demonstrated substantially differently in statistics -e.g., with alternative sampling and analysis plans; and in computing machinery -e.g., with explicit evaluation of risks), then the alignment with the particular ASA element is indicated by parenthesis “(*element*)”. That is, *abstract* concordance means the ACM Principle is described in terms of computing specifically, and would be demonstrated by/is described fundamentally differently by the ASA Guidelines. If there is neither exact or thematic alignment, then the cell is left blank.

Table 1. Association of Computing Machinery Ethics (2018) Code v3 vs American Statistical Association Ethical Guidelines (2018)

ASA:	A Professional Integrity and Accountability	B Integrity of data and methods	RESPONSIBILITIES TO/REGARDING:				G Allegations of Misconduct	H Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners
			C Science/ Public/ Funder/ Client	D Research Subjects	E Research Team Colleagues	F Other Statisticians or Statistics Practitioners		
ACM:								
1. GENERAL MORAL PRINCIPLES. <i>A computing professional should...</i>								
1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.	(A; A1, A2; and preamble)	(ALL B)	(C; C4)	(D)	(E4)	(F1, F3)		
1.2 Avoid harm. <i>In this document, "harm" means negative consequences to any stakeholder, especially when those consequences are significant and unjust.</i>	(A; A1, A2; preamble)	(ALL B)	(C; C5)	(ALL D)	(E2, E3)	(F4)	G1, G5, G6	(H2, H3, H4, H5, H6, H7, H8)
1.3 Be honest and trustworthy.	ALL A	ALL B	(C; C1, C2, C3)	(ALL D)	(E2, E3, E4)	(F4)	G1	(H2, H3, H4, H5, H6, H7, H8)
1.4 Be fair and take action not to discriminate.	Preamble; A7	(B4, B9)	(C; C1, C2, C3)	(D2, D3, D4, D5, D6, D7)	(E)	(F2, F4)	G1, G6	ALL H
1.5 Respect the work required to produce	A3	B3	(C5)		(E1)	(F2)	(G2)	ALL H

new ideas, inventions, creative works, and computing artifacts.								
1.6 Respect privacy.		(B10)	C5	(D3, D4, D5, D6)			(G4)	
1.7 Honor confidentiality.		(B10)	C5	(D3, D4, D5, D6)			(G3, G4)	
2. PROFESSIONAL RESPONSIBILITIES. <i>A computing professional should...</i>								
2.1 Strive to achieve high quality in both the process and products of professional work.	ALL A	ALL B	ALL C		E2, E3, E4	All F	G1	ALL H
2.2 Maintain high standards of professional competence, conduct, and ethical practice.	ALL A	ALL B	ALL C	ALL D	E2, E3, E4	ALL F	G1	ALL H
2.3 Know, respect, and apply existing rules pertaining to professional work.	A; A7	ALL B	ALL C	(D1, D2, D4, D5, D6)	ALL E	ALL F	(ALL G)	ALL H
2.4 Accept and provide appropriate professional review.	A; A6	(B10; B11)	(C)	(D5, D6)	(E2)	F2		(H7)
2.5 Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks.	(A1; A3)	(B; B1)	(C; C1, C2)	(D3, D5, D6, D7)	(E2; E4)	(F1; F2)		(H7)
2.6 Have the necessary expertise, or the ability to obtain that expertise, for completing a work assignment before	A; A1, A2; A6	ALL B	(C; C1, C2, C3)	(D2)	(E3)		(G1)	

accepting it. Once accepted, that commitment should be honored.								
2.7 Improve public awareness and understanding of computing, related technologies, and their consequences.	(A2)	(ALL B)	(C4)		(E4)	(F1)		
2.8 Access computing and communication resources only when authorized to do so.		(B10)	(C5)	(D3, D5, D6)				(H3, H7)
2.9 Design and implement systems that are robustly and usably secure.	(A; A1, A2)	(ALL B)	(ALL C)	(D3, D4, D5, D6)	(E3; E4)	(F1)		
<p align="center">3. PROFESSIONAL LEADERSHIP PRINCIPLES.</p> <p align="center"><i>In this section, “leader” means any member of an organization or group who has influence, educational responsibilities, or managerial responsibilities. These principles generally apply to organizations and groups, as well as their leaders.</i></p> <p align="center"><i>A computing professional should...</i></p>								
3.1 Ensure that the public good is the central concern during all professional computing work.	(A2)	(B; B10, B11)	(C4)	(D2; D4, D5, D6, D7)	(E2)	(F1)		
3.2 Articulate, encourage acceptance of, and evaluate fulfillment of the social responsibilities of members of an organization or group.								
3.3 Manage personnel and resources to enhance the quality of working life.								

3.4 Articulate, apply, and support policies and processes that reflect the principles in the Code.	(ALL A)	(ALL B)	(C1, C2, C3)	(D3, D5, D6)	(E1, E2)	(F4)	(ALL G)	(ALL H)
3.5 Create opportunities for members of the organization or group to learn and be accountable for the scope, functions, limitations, and impacts of systems.	(A3, A4, A6)	(ALL B)	(C4)	(D5, D6)	(E4)	(F2, F3)		
3.6 Retire legacy systems with care.								
3.7 Recognize when a computer system is becoming integrated into the infrastructure of society, and adopt an appropriate standard of care for that system and its users.								
4. COMPLIANCE WITH THE CODE. <i>A computing professional should...</i>	(* suggested in the ASA preamble; included <i>by implication</i> only (that “the ethical statistician...” follows the GLs))							
4.1 Uphold, promote, and respect the principles of the Code.								
4.2 Treat violations of the Code as inconsistent with membership in the ACM.								

Table 1 shows a high level of *thematic concordance* between the two practice standards. That is, while the wording (and some of the manifestations of the specific feature of ethical practice) is different, both organizations specify similar features as foundational in ethical practice with the methods and materials of their respective disciplines. Several ACM Code elements are aligned with every ASA principle; these are highlighted in yellow.

Of the 25 ACM Code elements, six of them are not aligned with any aspect of the ASA Guidelines (highlighted in pink in Table 1), but every ASA Guideline has at least some thematic alignment with all of the other ACM elements. While ACM Principle 4, “compliance with the code” has some overall suggestion of a match to the ASA preamble (“...in all cases, stakeholders have an obligation to act in good faith, to act in a manner that is consistent with these guidelines, and to encourage others to do the same.”), as well as some suggestions about not letting others or their guidelines conflict with the ASA GLs (Principle F/F1), the ACM Code goes further, specifically stating that violations of their Code should be considered “inconsistent with membership in the ACM”. The ASA Committee on Professional Ethics considered including one or both of these elements in the revisions that were approved by the ASA Board in 2016, but they were not approved by the Committee; since the Guidelines are intended for all “who use statistical methods in their professional work” and not just statisticians, it may not matter to some practitioners whether failing to follow the ASA Guidelines is inconsistent with membership in the ASA. Another suggestion has been made for the 2021 revisions of the ASA Ethical Guidelines to include a new item under Principle A, (the ethical statistician) “Upholds, promotes, and respects these Guidelines.” Both organizations articulate the need to be professional and respectful in practice, whether users of statistics or computing are members or not, or hold specific job titles (“statistician”, “computing professional”), or instead use these domains in their work.

There are several features of ACM principle 3, Professional Leadership, which also have no alignment with ASA Guideline Principles in terms of four specific elements:

- 3.2 Articulate, encourage acceptance of, and evaluate fulfillment of the social responsibilities of members of an organization or group.
- 3.3 Manage personnel and resources to enhance the quality of working life.
- 3.6 Retire legacy systems with care.
- 3.7 Recognize when a computer system is becoming integrated into the infrastructure of society, and adopt an appropriate standard of care for that system and its users.

With respect to the domain of data science, 3.6 and/or 3.7 may be suggested, with some revisions, for consideration by the ASA Committee on Professional Ethics for inclusion in the 2021 revision of the Guidelines for Ethical Statistical Practice (because as articulated by the ACM, there is no alignment). While the ACM naturally focuses on legacy *computing systems*, the ASA membership may consider legacy *models and algorithms* more relevant for both data science and general statistical practice; as organizations utilize models and algorithms, sensitivity checks are needed over time to ensure that these models and algorithms support decision making effectively, and that assumptions are tested and met. When those sensitivity checks are failed, then those models and algorithms certainly need to be reconsidered by the ethical statistician and data scientist, especially if they are to be “retired”. This suggestion has been made for the

planned 2021 revision of the ASA Guidelines through the new online system created specifically to elicit such recommendations from the ASA member community (i.e., available only to ASA members in good standing at <https://community.amstat.org/communities/community-home?CommunityKey=b482848a-43f9-441c-84cb-96bfe4f732c7>).

Given the concordance of these two sets of practice standards, they represent an excellent – coherent – resource for beginning to describe what constitutes ethical data science. The features in the ACM Code that could be relevant for guiding ethical practice by statisticians and data scientists, but are not (yet) included in the ASA Ethical Guidelines, have been suggested for consideration in the upcoming ASA revisions planned for 2021 (using the link above). However, *this is not to suggest that the organizations should modify their ethical practice standards to accommodate the other*; instead, the dissemination and teaching/training endeavors by both organizations, and those who seek to incorporate training in ethical practice standards into data science (and other) curricula, can benefit from consideration and discussion of the similarities and differences in the standards.

The practice standards of the ASA, ACM, ISI and RSS (among other organizations) are not the only resources that ethical data scientist can utilize. Two tools that were intended to promote ethical work with data/data science were published in 2018: the Data Science Ethics Checklist (DSEC; <http://deon.drivendata.org/>), developed to support decision-making in data science projects; and the UK government’s Data Ethics Framework and Workbook (DEFW; <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>), which has a more specific target readership than DSEC: “Public sector organisations should use the Data Ethics Framework to guide the appropriate use of data to inform policy and service design.” These tools ask those engaged in projects that involve data science (or require “data ethics”) to consider features of those projects. Neither of these tools refers to the ethical practice guidelines available for the domains of statistics or computing; and they also do not provide guidance on what might be an ethical answer to any of the questions in their respective lists. Given the origins of these tools from what might be considered “the data science community”, their alignment with the ASA and ACM guidance should be explored. The following tables summarize Data Ethics Framework (DEFW, Table 2) and Data Science Ethics Checklist (DSEC, Table 3) coverage of ethical practice guidelines contained in the 52 ASA and 25 ACM guideline elements.

Table 2. Alignment of DEFW with ASA and ACM

	Principle 1. Start with clear user need and public benefit	Principle 2. Be aware of relevant legislation and codes of practice	Principle 3. Use data that is proportionate to the user need	Principle 4. Understand the limitations of the data	Principle 5. Ensure robust practices and work within your skillset	Principle 6. Make your work transparent and be accountable	Principle 7. Embed data use responsibly
ASA A	A1, A2		A2	A1, A2	A2	A2, A6	A2
ASA B	B8		B4	B1, B2, B4-B9	B1, B2, B4-B6, B9	B1, B2, B4-B6, B9	B1, B2, B4-B6, B9
ASA D		D1	D2				
ASA E		E1				E4	
ACM 1			1.6	1.3	1.3	1.2	
ACM 2		2.3		2.5	2.5, 2.6	2.1, 2.4	
ACM 3							3.6, 3.7

In this table, the DEFW principles (columns) are shown to align with four of the eight ASA Principles (**A.** Professional Integrity & Accountability; **B.** Integrity of data and methods; **D.** Responsibilities to Research Subjects; and **E.** Responsibilities to Research Team Colleagues) and three of the four ACM Principles (all but compliance with the code, ACM 4). Every DEFW principle has at least one ASA or ACM principle that also addresses it.

In each cell of Table 2 is given the specific element of the Principle (ASA or ACM) that addresses - or, helps practitioners to answer - the questions that are asked in the DEFW worksheets. There are 52 different elements in the ASA GLs; 15 of these are common with DEFW, meaning that the ASA Guidelines address $52-15=37$ *other* aspects of ethical practice that the DEFW does not address. Specifically, the DEFW does not address **C.** Responsibilities to Science/Public/Funder/Client; **F.** Responsibilities to Other Statisticians or Statistics Practitioners; **G.** Responsibilities Regarding Allegations of Misconduct; or **H.** Responsibilities of Employers/Clients Employing Statistical Practitioners.

Similarly, there are 25 ACM elements and the DEFW items are addressed by 10 of these, leaving 15 aspects of ethical computing that are not considered in the DEFW worksheets. These gaps are, perhaps, not surprising because the DEFW is a checklist specifically intended to support project-specific ethical collection, use, and storage of data – and to support conversations about what is ethical about a given project, rather than provide guidance about ethical practice overall. Another important thing about Table 2 is the level of overlap between the ASA and ACM Principles; they both include relevant material to support all of the DEFW principles except DEFW Principle 1, which is only addressed by the ASA Guidelines.

The DSEC areas are more concrete than DEFW Principles (i.e., all leading to yes/no answers), and are also intended for “data science ethics” rather than “data ethics”, as the DEFW is. The DSEC checklist includes 20 questions about five different aspects of data science practice: data collection (A), data storage (B), analysis (C), modeling (D), and deployment (E). In these tables, the most directly-aligned element from the ASA and ACM materials is listed (unless all of the

items within a principle are aligned, as they are for ASA Principle B with DSEC A.2). If there are thematic but not exact alignments, these are included in the table in parenthesis, as with ACM 2.5 with DSEC C5.

Table 3. Alignment of DSEC areas A-E with ASA and ACM.

DSEC A: Data collection

	A.1 Informed consent: If there are human subjects, have they given informed consent, where subjects affirmatively opt-in and have a clear understanding of the data uses to which they consent?	A.2 Collection bias: Have we considered sources of bias that could be introduced during data collection and survey design and taken steps to mitigate those?	A.3 Limit PII exposure: Have we considered ways to minimize exposure of personally identifiable information (PII) for example through anonymization or not collecting information that isn't relevant for analysis?
ASA A		A1	
ASA B		ALL B	
ASA C			C5
ASA D	D1, D5, D6	D7	D3
ACM 1	1.6	1.4	1.6
ACM 2		2.5	

DSEC B: Data storage

	B.1 Data security: Do we have a plan to protect and secure data (e.g., encryption at rest and in transit, access controls on internal users and third parties, access logs, and up-to-date software)?	B.2 Right to be forgotten: Do we have a mechanism through which an individual can request their personal information be removed?	B.3 Data retention plan: Is there a schedule or plan to delete the data after it is no longer needed?
ASA B	B10	B10	
ASA C	C5	C5	C5
ASA D	D3	D3	D3
ASA F	F1	F1	
ACM 1		1.6	1.6
ACM 2	2.9		
ACM 3	3.7		

DSEC C: Analysis

	C.1 Missing perspectives: Have we sought to address blindspots in the analysis through engagement with relevant stakeholders (e.g., checking assumptions and discussing implications with affected communities and subject matter experts)?	C.2 Dataset bias: Have we examined the data for possible sources of bias and taken steps to mitigate or address these biases (e.g., stereotype perpetuation, confirmation bias, imbalanced classes, or omitted confounding variables)?	C.3 Honest representation: Are our visualizations, summary statistics, and reports designed to honestly represent the underlying data?	C.4 Privacy in analysis: Have we ensured that data with PII are not used or displayed unless necessary for the analysis?	C.5 Auditability: Is the process of generating the analysis well documented and reproducible if we discover issues in the future?
ASA A	A1	A1			A2
ASA B	ALL B	ALL B	B1, B2, B4-B6	B10	ALL B
ASA C				C5	
ASA D	D1	D1			
ASA E			E4		
ASA F				F1	
ACM 1	1.4	1.4		1.6	
ACM 2			(2.1)		(2.4)

DSEC D: Modeling

	D.1 Proxy discrimination: Have we ensured that the model does not rely on variables or proxies for variables that are unfairly discriminatory?	D.2 Fairness across groups: Have we tested model results for fairness with respect to different affected groups (e.g., tested for disparate error rates)?	D.3 Metric selection: Have we considered the effects of optimizing for our defined metrics and considered additional metrics?	D.4 Explainability: Can we explain in understandable terms a decision the model made in cases where a justification is needed?	D.5 Communicate bias: Have we communicated the shortcomings, limitations, and biases of the model to relevant stakeholders in ways that can be generally understood?
ASA A	A2	A1, A2	A2	A2	
ASA B	B1, B2, B5, B6	B1, B2, B4-B6, B9, B11	B1, B2, B5, B6, B9, B11	ALL B	ALL B
ACM 1		1.2, 1.4	1.2, 1.4		
ACM 2	2.4, 2.5			2.5	2.5

DSEC E: Deployment

	E.1 Redress: Have we discussed with our organization a plan for response if users are harmed by the results (e.g., how does the data science team evaluate these cases and update analysis and models to prevent future harm)?	E.2 Roll back: Is there a way to turn off or roll back the model in production if necessary?	E.3 Concept drift: Do we test and monitor for concept drift to ensure the model remains fair over time?	E.4 Unintended use: Have we taken steps to identify and prevent unintended uses and abuse of the model and do we have a plan to monitor these once the model is deployed?
ASA B	B11	B1, B2, B4-B6, B11	B1, B2, B4-B6, B9	ALL B
ASA D				D5, D6
ACM 2	2.5	2.5, 2.9	2.5	2.5, 2.9

Like with the DEFW, every DSEC area and item has at least one ASA or ACM principle that also addresses it. The alignment of DSEC areas with the ASA guidance standards is strongest with ASA Principles A, B, and D (**A.** Professional Integrity & Accountability; **B.** Integrity of data and methods; **D.** Responsibilities to Research Subjects); there is also alignment with some features of C, E, and F (**C.** Responsibilities to Science/Public/Funder/Client; **E.** Responsibilities to Research Team Colleagues; **F.** Responsibilities to Other Statisticians or Statistics Practitioners) but no alignment with G or H (**G.** Responsibilities Regarding Allegations of Misconduct; or **H.** Responsibilities of Employers/Clients Employing Statistical Practitioners). Like the DEFW, the DSEC areas align with three of the four ACM Principles (all but compliance with the code). Importantly, the items on the DSEC Checklist are all yes/no questions. Both the DSEC and to a lesser extent the DEFW imply that “any “NO” answer creates an ethical problem”, but what to do if there is a NO answer is unspecified. In fact, the ACM CE and ASA GLs also do not offer exact specifications on what to do— that is left up to the practitioner’s judgment, ability level, and the specific context of practice.

Tables 1-3 are all summarized in Figure 1.

Figure 1. The universe of ethical statistical and data science practice

As noted by Loukides et al. (2018), “After the ethical principle, we have to think about the implementation of the ethical principle. That isn’t easy; it encompasses everything from user experience design to data management.” However, these authors also criticize efforts (the DEFW, specifically) to achieve this implementation: “Our criticism is that their process imposes a lot of overhead. While anyone going through their entire process will certainly have thought carefully about ethical issues, in practice, asking developers to fill out a workbook with substantive answers to 46 questions is *an effective way to ensure that ethical thought doesn’t happen.*” (emphasis added) Loukides et al. offer an alternative: a 13-item checklist, about which they state, “Granted, it’s still possible to take shortcuts: your analysis might be inadequate and your consent mechanism might be flawed, but *you’ve at least gone on record for saying that you’ve done it.*” (emphasis added) By contrast, data science practitioners and teams who work through the DEFW workbook for each project might be less likely to take shortcuts, and rather than only going on record for saying that they have done the analysis, they will very probably “have thought carefully about ethical issues” and would also be able to document this.

Table 2 shows that even this careful thinking using the DEFW will be specific to a given project, and may not promote ethical data science practice more generally. DSEC items may not require as much careful thought as DEFW questions do; but the ASA and ACM practice standards can support either of these checklist approaches within ethical data science practice -and go beyond the single-project consideration of ethical issues to promote ethical data science practice generally (as shown in Figure 1). Figure 1 – which is not to scale - shows that the universe of ethical practice in data science (and “statistics and data science”, for programs with that title/theme) is supported and described by the ASA Guidelines and the ACM Code. The figure and Tables 2 and 3 show that there is no aspect of ethical practice identified by DSEC or DEFW that is not covered by the ASA, ACM, or both standards. The figure shows that DSEC and DEFW coverage tends to represent what BOTH of these practice standards also jointly cover, but since the ACM code does not focus on analysis, modeling, interpretation, and collaborating with

others in the way that statisticians and data scientists must do, there are clear areas of DSEC and DEFW that are more specifically addressed by the ASA Guidelines than are addressed jointly by ASA and ACM; and no DSEC or DEFW element is addressed by the ACM Code alone.

Figure 1 was created to highlight the fact that there is a great deal of the *universe of ethical practice in data science* that is neither covered nor addressed by either of these checklists. Thus, while the DSEC and/or DEFW may be useful characteristics for any given project, data scientists and those who take courses in the domain need to be aware of the complexities – and responsibilities – inherent in the practice of data science. The use of these checklists cannot be considered acceptable alternatives to training with the ethical practice standards outlined by the ASA and ACM. However, instruction in “ethical data science” may benefit from the inclusion of these checklists by engaging students in conversations about how exactly to use the Guidelines and Code to answer each checklist question. Utilizing the concordance table (Table 1) can also support those conversations. Together, the two ethical guideline documents and either or both checklists can be leveraged to create curriculum-spanning – evaluable – assignments and in-class discussions. As students progress through a curriculum that responds to NAS Recommendation 2.4 (whether data science, statistics, computing, or another curriculum where any of these quantitative disciplines will be taught or introduced), their level of sophistication should be documented as increasing. The same assignments can be re-used – more advanced students can be asked to reflect on, justify, or discuss alternative perspectives on, their earlier work (because reflection is a more cognitively complex activity than fill in the blank or choosing which guideline principle applies).

A final consideration in addressing the NAS recommendations is to think of those who use, or rely on, the *results* of data science. In their 2019 Ethics in Action guidance, drafted by the Institute of Electrical and Electronics Engineers (IEEE) Initiative on Ethics of Autonomous and Intelligent Systems, the following argument concerning the eight key features of ethical development of Autonomous and Intelligent Systems. To the extent that data science feeds into autonomous and intelligent systems, ethical practice in the domain of data science must conform to the IEEE ethically aligned design (EAD) features; while the IEEE EAD report (IEEE, 2018) does not specify the importance of following the ethical guidelines whether or not the individual is a member of the organization (like ASA and ACM materials do), it is helpful to triangulate the concordance of the ASA and ACM principles with those of the IEEE EAD report. This 2018 report was the result of a worldwide, crowd-sourced, effort and its alignment with the ASA and ACM practice standards – which were formulated by Ethical Practice Committees (rather than through a global response) – can be helpful in establishing some external validity to the argument that concordance between ASA and ACM practice standards can be leveraged to support ethical data science. The diversity in origins of the guidance documents underscores the importance that the key EAD principles can bring to a concordance effort seeking to represent agreement among the “data science community” that the NAS report refers to in Recommendation 2.5. The Appendix includes a table showing how the eight ethical features of the EAD report are aligned thematically with the ASA Ethical Guidelines and the ACM Code of Ethics. All but one of the eight EAD features are aligned with at least one Principle in the ASA practice standards; the exception being the one EAD feature regarding “well-being”: “A/IS creators shall adopt increased human well-being as a primary success criterion for development.” This is thematically aligned with the ACM Code (Preamble and Principle 1), but there is no

direct thematic link with the ASA Guidelines. Every one of the EAD features are aligned with at least one Principle in the ACM practice standards except the ACM standards relating to compliance with the ACM Code (Principle 4). It may be argued that the well-being feature of the EAD guidance is *indirectly* aligned with the ASA Guideline Preamble assertion that, “Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical.”

Implementation of ethical practice standards to support ethical data science

“In addition to teaching the special obligations of scientists and engineers, ethics instruction can prepare students for challenges that might arise in their professional lives. Ethics education programs typically help students develop the skills to recognize ethical problems, to reason about conflicts between important values, and to evaluate possible actions to address those problems.” (National Academy of Engineering, NAE, 2013: p2)

The proliferation of checklists (e.g., Loukides et al., the DSEC and DEFW, the “data ethics checklist, <https://eticasfoundation.org/data-ethics-checklist>; the “ethics checklist” from the University of Leicester, <https://www2.le.ac.uk/institution/ethics/resources/checklist>) may bring the “data science community” dangerously close to Wolpe’s assertion, “*It is easy to conclude that ethics is essentially arbitrary*” since there is no rational argument for one of these checklists over another, or against creating a new one for your own purposes. Baumer et al. (2020) created instructional modules (“patches”) for use at different levels, in response to NAS Recommendation 2.4, but without appealing to any guidelines (nor to any existing checklists). This may tend to contribute to the diversity of content, topical coverage, and learning outcomes identified by Feisler et al. (2020) in university courses featuring “technology ethics”. Accomplishing NAS Recommendation 2.5, documenting the existing consensus on what professionals in the communities of statistics and computing consider ethical practice, *would* support the selection and utilization of one or more checklists: the one that maps best to relevant professional ethical practice guidelines should be utilized (and those that do not comport with practice standards should not be used) in realistic, authentic, ways.

Loukides et al. (2018) argue that “hard work”, i.e., thinking carefully about ethical issues, as is required to utilize the DEFW, ensures that “ethical thought doesn’t happen”. However, to accomplish NAS Recommendation 2.4, would require a working understanding of the professional ethical practice standards in order to *select and effectively* implement the best checklist; otherwise, it might be more likely to take shortcuts and simply go on record as saying that the work has been done (per Loukides et al. 2018). To truly fulfill the Recommendation 2.4, students will need to know a) how to utilize the practice standards to figure out what both the “yes” and the “no” answers to any checklist item imply for the practitioner’s next steps; and b) what *else* constitutes “ethical data science”; namely, the standards themselves.

While key *topics* of concern in ethical data science are covered in the six data science ethics patches described by Baumer et al. (privacy/confidentiality; responsibility in research; algorithmic bias; legal vs ethical considerations; social implications of data-mining by Internet companies; p. 7), the ethical practice standards can support much wider engagement with ethical dimensions of practice. They also support students developing the abilities to recognize ethical

challenges that are outside of these six patches – new ethical challenges that are currently unknown. Moreover, a focus on topics tends to ignore the responsibilities articulated by mature professions to strengthen the profession itself. These responsibilities are incorporated in the ASA and ACM guidelines, for example, responsibilities to accept and provide expert evaluation of work (e.g., ACM 2.4; ASA F2) or to encourage (ACM 2.1) professionally competent/ discourage incompetent (ASA G1) work. That is, the ASA and ACM guidance can lead towards the realization of Golde & Walker’s (2006) assumption, that **“Upon entry into practice, *all professionals* assume at least a tacit responsibility for the quality and integrity of their own work and that of colleagues. They also take on a responsibility to the larger public for the standards of practice associated with the profession.”**

Both ASA and ACM preambles state that their practice standards offer guidance, not rules; and that each practitioner – while responsible for following their disciplinary guidelines and following local rules/laws – must also utilize their judgement in order to practice ethically. While the DSEC and DEFW are compilations of important considerations, both are project-specific and do not offer any guidance on how to practice ethically or how to recognize an ethical practitioner. The practice standards that the ASA and ACM offer are useful in more situations than just specific projects; the rather limited alignment in DSEC and DEFW with ASA Principle A and ACM Principle 2 are telling in that A and 2 both discuss professional practice and integrity at all times, and not with respect to a specific project. The DSEC and DEFW are the most comprehensive of the checklists (that the author has reviewed); as can be inferred from Figure 1, less expansive examples (e.g., the checklist presented by Loukides et al.) will have even less relevance to the universe of ethical data science practice.

The foregoing alignment tables (and see Appendix for additional tables) highlight the fact that the ACM and ASA ethical practice standards can be used together to define, and subsequently promote, ethical data science. There is a high level of concordance in these independently-derived practice standard documents; they also clearly support the informed, and ethical, completion of checklists that are proliferating to help practitioners document their ethical practice considerations/decisions. To understand what is required in order for the NAS recommendations to be meaningfully implemented, it is helpful to consider how teaching “ethical practice” can be facilitated by inclusion of the constructs of *professionalism* and *professional identity*:

Professionalism is defined as:

- The skill, good judgment, and polite behavior that is expected from a person who is trained to do a job well.
- The conduct, aims, or qualities that characterize or mark a profession or a professional person. (Merriam-Webster).

Professional Identity is defined as:

“...the sense of being a professional...the use of professional judgment and reasoning ... critical self evaluation and self-directed learning...” (Paterson et al. 2002:7)

“Professional identity formation means becoming aware of ... what values and interests shape decision-making.” (Trede, 2012, p. 163).

If academic institutions seek to train their data science students “to do a job well”, particularly in preparation to become/take jobs as ethical data scientists, these are important constructs because they can motivate students to engage with a curriculum-spanning thread of “professional ethics”. As discussed by Paterson et al. (2002) and Trede (2012), “learning about the ethical guidelines” is clearly insufficient as a learning outcome unless it specifically involves the development of judgment. Moreover, to meet the NAE standard of “ethics education”, the engagement of data science students with the ASA and ACM professional practice guidelines must contribute to “the skills to recognize ethical problems, to reason about conflicts between important values, and to evaluate possible actions to address those problems.” (NAE 2013, p2). That is, the abilities to utilize the ethical practice standards in case analyses will require a degree of *ethical reasoning* (Briggle & Mitcham, 2012; see also Tractenberg & FitzGerald, 2012): “Ethics is the effort to guide one’s conduct with careful reasoning. One cannot simply claim “X is wrong.” Rather, one needs to claim “X is wrong because (fill in the blank)”.” (Briggle & Mitcham, 2012, p. 38) Dow et al. (2015) evaluated the theory and empirical work supporting case analyses for instruction in (professional) ethics. Integrating authentic cases (Tractenberg, 2019-b) together with professional practice guidelines to conduct the case analyses, unifies the pedagogic strengths of the case study method with the authenticity that practice guidelines can confer on a curricular focus on “ethical data science” or data science ethics. As students are ‘filling in the blank’ that Briggle & Mitcham identified, they will deepen their understanding, and gain practice with application, of ASA and ACM ethical practice principles. Thus, acknowledgement that ASA and ACM ethical practice standards support and define “ethical data science” accomplishes NAS Recommendation 2.5; teaching and engaging students in ethical reasoning with those practice standards across a curriculum can accomplish NAS Recommendation 2.4.

In order to encourage data scientists to utilize judgement, rather than memorization, as they learn about what constitutes ethical data science, instructors can leverage Bloom’s taxonomy of cognitive behaviors (Bloom et al. 1956). Specifically, there are six levels of complexity (from Bloom’s 1 “B1” to Bloom’s 6 “B6”): **B1** Remember/Reiterate (performance is based on recognition of previously seen example); **B2** Understand/Summarize (performance summarizes info already in question (and/or answers)); **B3** Apply/Illustrate (performance extrapolates from seen examples to (really) new examples); **B4** Analyze/Predict (performance requires application/following of a rule; organize/sort using concrete criteria; some use of judgement in conjunction with the rules is needed); **B5** Create/ Synthesize (find patterns, innovate – to fill gaps in knowledge); and **B6** Evaluate/Compare/Judge (Make reasoned judgments in the absence of “truth” or concrete criteria). Clearly the minimum level of performance that academic institutions need to strive for to meet the NAE definition of “ethics education”, and also to fully satisfy NAS Recommendation 2.4, is B4. This is shown in Figure 2, below.

Figure 2. Bloom’s cognitive complexity level 4 is the minimum required to apply the judgment needed to utilize ethical guidelines.

As defined by Briggie & Mitcham (2012), the ability to explain/justify why one or another ethical guideline or principle pertains, or should be prioritized over other considerations, constitutes the target level of engagement with what constitutes “ethical data science practice”; handing out (or testing on memorization of) the ethical guidelines themselves is clearly not sufficient.

Engagement with case studies, which itself requires a set of *qualitative* skills that many (if not all) quantitative students need to develop, should help students to progress to B4 (at a minimum) in terms of ethical practice in data science over the curriculum. Students can engage with similar cases, at increasingly more complex and sophisticated levels, as they progress through the curriculum (Tractenberg, 2016-a p. 125). Assignments in earlier courses/earlier in their programs can focus on one, and then the other, set of ethical guidelines from the ASA and ACM. As learners approach their program or degree completion, they should be prepared by the prior work in the curriculum for comparing and contrasting the two practice standards.

Discussion

The National Academies of Sciences, Engineering, and Medicine (NAS, 2018), Committee on Envisioning the Data Science Discipline “... underscores the centrality of studying the many ethical considerations that arise as workers engage in data science. (p. 2) and two key recommendations are that “*Academic institutions should ensure that ethics is woven into the data science curriculum from the beginning and throughout*” (Recommendation 2.4) and “*The data science community should adopt a code of ethics; such a code should be affirmed by members of professional societies, included in professional development programs and curricula, and conveyed through educational programs*” (Recommendation 2.5).

These NAS Recommendations are aspirational in the sense that academic institutions cannot ensure that a curricular focus on ethics will support ethical data science practice if there is no agreement on “normative ethical data science”. This is due to another aspirational aspect of the Recommendations: there is no “data science community”; The domain is too new and practitioners operate, and join the practice, from too many other domains for there to be a single community. While, in 2015 it may have been true that, “The development of a discipline or profession tends to proceed in tandem with a normative agreement on a code of professional ethics or conduct for the practitioner in the community” (Tractenberg et al. 2015; p. 1487), the domain of data science has not followed the same path.

The development of professional identity is supported by instruction and practice in justifying professional decisions: “...students need to learn to articulate the reasons behind their actions.” (Trede & McEwen, 2012:10), and learning to utilize the ethical practice guidelines of the ASA and ACM in the practice of data science, at whatever level students are practicing (creators, consumers, innovators) is a pedagogically sound way for them to develop professional identities that include confidence about ethical data science. “(S)tudents could learn more from their experiences if they were more explicitly guided to look out for certain aspects of professionalism and given further opportunities to discuss and critique their observations and experiences.” (Grace & Trede, 2011:12; see also Trede et al. 2011) Comparing and contrasting the guidance that the ASA and ACM practice standards offer on any given professional decision is a concrete way to both direct student attention to these standards and to enable the development and demonstration of the minimum Bloom’s level of complexity (B4) required for recognizing and prioritizing applicable ethical guidance.

Whether the data scientist’s job is primarily computational or primarily statistical, training, and practice, with the ASA and ACM ethical practice standards will help any data scientist to be a *more professional* -i.e., more ethical - practitioner in the quantitative sciences. The purpose of these practice standards is not to correct unethical habits, but to inculcate a respect for, and awareness of, the complexities of ethical statistical practice in the modern data intensive landscape. The Code (ACM) and Guidelines (ASA) represent the ‘habits of mind’ that characterize the experienced quantitative practitioner (Tractenberg, 2015). Thus, they can be useful in and across academic curricula in developing professionalism and professional identity. They can also promote *discussion* and *explication* of expectations & standards of practice – within & outside of the discipline. While of essential importance for data science students and practitioners, these practice standards also represent important information for those who use data driven approaches (or the results therefrom) in decision making across science, business, and policy.

“...(E)thics is not a vaccine that can be administered in one dose and have long lasting effects no matter how often, or in what conditions, the subject is exposed to the disease agent”. (National Academy of Engineering, 2009:34).

It is very difficult (maybe impossible) to enumerate all the things that *are* ethical, and so most ethics education focuses on identifying just what is NOT ethical. Instead, achieving NAS Recommendation 2.4 requires academic institutions to acknowledge that - and act as if - ethical practice is something worth working for. Instructors, mentors, and supervisors also need to be

familiar with ethical practice standards like the Code and Guidelines, and effectively convey to newer/more junior practitioners that they are essential for those who want to be considered, and recognized, to be a “professional” or “ethical professional”. Simply articulating Guidelines is not enough – and standard approaches to dissemination of the articulated practice standards may also not be sufficient (Tractenberg & DiMartino, 2017) to truly promote engagement with what constitutes ethical professional practice. This is partly due to the requisite level of Bloom’s cognitive complexity needed to exercise real professional judgment: memorizing the ethical principles, and their constituent elements, can only contribute foundational knowledge (since it is at Bloom’s level 1, not judgement which requires Bloom’s levels 4-6) to professional, ethical, decision making. Introducing and giving practice with these principles and their constituent elements is best suited to the academic setting - *before* quantitative scientists are focusing on their “day job”. However, these ethical practice guidelines and their use must also be accessible to practitioners, as well as instructors (who are often times practitioners). All relevant professional organizations with applicable ethical guidelines need to convey the importance of the guidelines - and their importance - for those who use statistical or computational methods or their outputs to make decisions.

Training and continuing professional development for the analyst and the decision-maker should promote/create opportunities to learn & practice prioritizing the ethical practice principles. Achieving NAS Recommendation 2.4 requires structuring courses, programs, & curricula to promote ongoing, *deepening sophistication* of thinking about what constitutes ethical practice and the ASA and ACM standards are a natural starting point. This essential training should accompany all efforts towards “data literacy”, “data acumen”, and also “data science”. Given their clear foundational relationships to the domain of data science, engagement with the ASA and ACM practice standards could be added to *every course that includes quantitative content* (Tractenberg, 2016-a; 2016-b).

References

American Statistical Association (ASA) *ASA Ethical Guidelines for Statistical Practice-revised* (2018) downloaded from <https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx> on 30 April 2018.

Association for Computing Machinery (ACM). *Code of Ethics* (2018) downloaded from <https://www.acm.org/about-acm/code-of-ethics> on 12 October 2018.

ACM/IEEE-CS Joint Task Force. Computer Science Curricula 2013 (2013). Downloaded from <http://www.acm.org/education/CS2013-final-report.pdf> 20 January 2015. (available in 2020 from <https://dl.acm.org/doi/book/10.1145/2534860>)

Baumer BS, Garcia RL, Kim AY, Kinnaird KM, Ott MQ. (2020, 21 January). *Preprint*. Integrating data science ethics into an undergraduate major. Published to the Open Statistics Archive, Stat arXiv arXiv:2001.07649v1 [stat.OT]

BS Bloom (Ed.), with Engelhart MD, Furst EJ, Hill WH & Krathwohl DR. (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York, NY: David McKay.

Briggle, A., & Mitcham, C. (2012). *Ethics and science: An introduction*. Cambridge, UK: Cambridge University Press.

Dow MJ, Boettcher CA, Diego JF, Karch ME, Todd-Diaz A, Woods KM. (2015). Case-based learning as pedagogy for teaching information ethics based on the Dervin sense-making methodology. *Journal of Education for Library and Information Science* 56(2-Spring): 141-157.

Fiesler C, Garrett N, Beard N. (2020). What Do We Teach When We Teach Tech Ethics? A Syllabi Analysis. In, *The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20), March 11–14, 2020, Portland, OR, USA*. ACM, New York, NY, USA, 7 pages. Downloaded from https://cmci.colorado.edu/~cafi5706/SIGCSE2020_EthicsSyllabi.pdf on 14 Feb 2020.

Golde CM, & Walker GE. (Eds.). (2006). *Envisioning the future of doctoral education: Preparing stewards of the discipline-Carnegie essays on the doctorate*. San Francisco: Jossey-Bass.

Grace S & Trede F. (2011). Developing professionalism in physiotherapy and dietetics students in professional entry courses. *Studies in Higher Education*. DOI: 10.1080/03075079.2011.603410.

Hogan H, Tractenberg RE, Elliot AC. (2017, July). Ethics and Big Data: Perspective of the American Statistical Association Committee on Professional Ethics. Presented at the 61st International Statistics Institute World Statistics Congress, Marrakesh, Morocco.

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2018). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent*

Systems, First Edition. Downloaded from <https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html> on 27 June 2019.

International Statistics Institute. (2010). *Declaration on Professional Ethics*. Downloaded from <https://isi-web.org/images/about/Declaration-EN2010.pdf> on 28 January 2015.

Loukides M, Mason H, Patil DJ. (2018). *Ethics and data science*. O'Reilly. Downloaded from <https://learning.oreilly.com/library/view/ethics-and-data/9781492043898/?ar> on 2 August 2018.

National Academies of Sciences, Engineering, and Medicine. (2018). *Data Science for Undergraduates: Opportunities and Options*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25104>.

National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/21896>.

National Academy of Engineering and National Research Council. (2009). *Ethics Education and Scientific and Engineering Research: What's Been Learned? What Should Be Done?* Washington, DC: The National Academies Press.

National Academy of Engineering (2013). *Practical Guidance on Science and Engineering Ethics Education for Instructors and Administrators: Papers and Summary from a Workshop December 12, 2012*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/18519>.

Paterson, M., J. Higgs, S. Wilcox, and M. Villeneuve. 2002. Clinical reasoning and self-directed learning: Key dimensions in professional education and professional socialisation. *Focus on Health Professional Education* 4, no. 2: 5–21.

Rios CR, Golde CM, Tractenberg RE. (2019). The preparation of stewards with the Mastery Rubric for Stewardship: Re-envisioning the formation of scholars *and practitioners*.

Royal Statistical Society (2014). Code of Conduct. Downloaded from <https://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf> on 28 January 2015.

Tractenberg, RE. (2015). Professionalism, professional identity, and ethics education in statistics: Discussion. *Presentation at the 2015 Joint Statistical Meetings*, Seattle, WA. Slides published in the *Open Science Foundation*, osf.io/chqj9 DOI 10.17605/OSF.IO/BQ8U5

Tractenberg RE, Russell AL, Morgan G, FitzGerald KT, Collmann J, Vinsel L, Steinmann M, Dolling LM. (2015) Amplifying the reach and resonance of ethical codes of conduct through ethical reasoning: preparation of Big Data users for professional practice. *Science and Engineering Ethics* 21(6): 1485-1507. DOI 10.1007/s11948-014-9613-1

Tractenberg RE. (2016-a). Why and How the ASA Ethical Guidelines should be integrated into every quantitative course. *Proceedings of the 2016 Joint Statistical Meetings*, Chicago, IL. Pp. 517-535.

Tractenberg RE. (2016-b). Institutionalizing ethical reasoning: integrating the ASA's Ethical Guidelines for Professional Practice into course, program, and curriculum. In, J. Collmann & S. Matei (Eds.), *Ethical Reasoning in Big Data: An Exploratory Analysis*. New York: Springer. 115-139.

Tractenberg RE, DeMartino G. (2017, December). What guidance is provided by ethical guidelines for professional practice in economics and statistics & data science? *Amstat News* <http://stattrak.amstat.org/2017/11/01/ethical-guidelines/>

Tractenberg RE. (2019-a, April 30). Preprint. Strengthening the practice and profession of statistics and data science using ethical guidelines. Published in the *Open Archive of the Social Sciences* (SocArXiv), <https://doi.org/10.31235/osf.io/58umw>

Tractenberg RE. (2019-b, April 23). Preprint. Teaching and Learning about ethical practice: The case analysis. Published in the *Open Archive of the Social Sciences* (SocArXiv), <https://doi.org/10.31235/osf.io/58umw>

Trede F. (2012). Role of work-integrated learning in developing professionalism and professional identity. *Asia-Pacific Journal of Cooperative Education* 13(3): 159-67.

Trede F & McEwen C. (2012). Developing a critical professional identity: Engaging self in practice. In J Higgs, R Barnett, S Billett, M Hutchings, F Trede. (Eds.) *Practice-based education: perspectives and strategies*. Rotterdam, The Netherlands: Sense Publishers. Pp. 27-40.

Trede F, Macklin R & Bridges D. (2011). Professional Identity development: a review of the higher education literature. *Studies in Higher Education* 37(3): 365-384.

Wang MQ, Yan AF & Katz RV. (2018). Researcher requests for inappropriate analysis and reporting: A US survey of consulting biostatisticians. *Annals of Internal Medicine* 169(8), 554–558.

Wolpe PR. (2006). Reasons scientists avoid thinking about ethics. *Cell* 125 (June 16): 1023-1025.

APPENDIX. Alignment of ASA with ISI and RSS Ethical Guidance

The Table below examines the alignment of the ASA GL Principles (2018) with the Declaration on Professional Ethics of the International Statistics Institute (2010).

International Statistical Institute

From the preamble: “For the purposes of this document, the definition of who is a statistician goes well beyond those with formal degrees in the field, to **include a wide array of creators and users of statistical data and tools.**” (Emphasis added)

And

“The aim of this declaration is to enable the statistician's individual ethical judgments and decisions to be informed by shared values and experience, rather than by rigid rules imposed by the profession. The declaration seeks to document widely held principles of the statistics profession and to identify the factors that obstruct their implementation. It recognizes that, the operation of one principle may impede the operation of another, that statisticians – in common with other occupational groups – have competing obligations not all of which can be fulfilled simultaneously. Thus, statisticians will sometimes have to make choices between principles. The declaration does not attempt to resolve these choices or to establish priorities among the principles. Instead it offers a framework within which the conscientious statistician should be able to work comfortably. It is urged that departures from the framework of principles be the result of deliberation rather than of ignorance.”

And

“Although not explicitly stated, the Principles inherently reflect the obligations and responsibilities of – as well as the resulting conflicts faced by – statisticians to forces and pressures outside of their own performance, namely to and from:

- Society
- Employers, Clients, and Funders
- Colleagues
- Subjects

In carrying out his/her responsibilities, each statistician must be sensitive to the need to ensure that his/her actions are, first, consistent with the best interests of each group and, second, do not favor any group at the expense of any other, or conflict with any of the Principles.”

Table 1. Alignment of the ASA Ethical Guidelines for Statistical Practice (2018) with ISI (2010) 15 items

Mashup, ASA (2018, columns) x ISI (2010, rows).

			RESPONSIBILITIES TO/REGARDING:					
ASA:	A Professional Integrity and Accountability	B Integrity of data and methods	C Science/ Public/ Funder/ Client	D Research Subjects	E Research Team Colleagues	F Other Statisticians or Statistics Practitioners	G Allegations of Misconduct	H Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners
ISI:								
SHARED PROFESSIONAL VALUES Our shared professional values are respect, professionalism, truthfulness and integrity.								
1. Respect We respect the privacy of others and the promises of confidentiality given to them. We respect the communities where data is collected and guard against harm coming to them by misuse of the results. We should not suppress or improperly detract from the work of others.			X	X				
2. Professionalism The value Professionalism implies Responsibility, Competence and Expert Knowledge, and Informed Judgment. We work to understand our users' needs. We use our statistical knowledge, data, and analyses for the Common Good to serve the society. We strive to collect and analyze data of the highest quality possible. We are responsible for the fitness	X	X	X		X	X		X

As scientists, we pursue promising new ideas and discard those demonstrated to be invalid. We work towards the logical coherence and empirical adequacy of our data and conclusions. We value well-established objective criteria of assessment.								
ETHICAL PRINCIPLES								
1. Pursuing Objectivity Statisticians should pursue objectivity without fear or favor, only selecting and using methods designed to produce the most accurate results. They should present all findings openly, completely, and in a transparent manner regardless of the outcomes. Statisticians should be particularly sensitive to the need to present findings when they challenge a preferred outcome. The statistician should guard against predictable misinterpretation or misuse. If such misinterpretation or misuse occurs, steps should be taken to inform potential users. Findings should be communicated for the benefit of the widest possible community, yet attempt to ensure no harm to any population group.	x	x	x		x	x	x	x
2. Clarifying Obligations and Roles The respective	x							x

obligations of employer, client, or funder and statistician in regard to their roles and responsibility that might raise ethical issues should be spelled out and fully understood. In providing advice or guidance, statisticians should take care to stay within their area of competence, and seek advice, as appropriate, from others with the relevant expertise.								
3. Assessing Alternatives Impartially Available methods and procedures should be considered and an impartial assessment provided to the employer, client, or funder of the respective merits and limitations of alternatives, along with the proposed method.		X	X					X
4. Conflicting Interests Statisticians avoid assignments where they have a financial or personal conflict of interest in the outcome of the work. The likely consequences of collecting and disseminating various types of data and the results of their analysis should be considered and explored.	X			X				
5. Avoiding Preempted Outcomes Any attempt to establish a predetermined outcome from a proposed statistical inquiry should	X	X	X		X		X	X

be rejected, as should contractual conditions contingent upon such a requirement.								
6. Guarding Privileged Information Privileged information is to be kept confidential. This prohibition is not to be extended to statistical methods and procedures utilized to conduct the inquiry or produce published data.		X	X	X	X	X		
7. Exhibiting Professional Competence Statisticians shall seek to upgrade their professional knowledge and skills, and shall maintain awareness of technological developments, procedures, and standards which are relevant to their field, and shall encourage others to do the same.								
8. Maintaining Confidence in Statistics In order to promote and preserve the confidence of the public, statisticians should ensure that they accurately and correctly describe their results, including the explanatory power of their data. It is incumbent upon statisticians to alert potential users of the results to the limits of their reliability and applicability.	X	X			X	X	X	X
9. Exposing and Reviewing Methods and	X	X	X		X	X		X

Findings Adequate information should be provided to the public to permit the methods, procedures, techniques, and findings to be assessed independently.								
10. Communicating Ethical Principles In collaborating with colleagues and others in the same or other disciplines, it is necessary and important to ensure that the ethical principles of all participants are clear, understood, respected, and reflected in the undertaking.	X			X	X	X	X	
11. Bearing Responsibility for the Integrity of the Discipline Statisticians are subject to the general moral rules of scientific and scholarly conduct: they should not deceive or knowingly misrepresent or attempt to prevent reporting of misconduct or obstruct the scientific/scholarly research of others.	X	X		X	X		X	X
12. Protecting the Interests of Subjects Statisticians are obligated to protect subjects, individually and collectively, insofar as possible, against potentially harmful effects of participating. This responsibility is not absolved by consent or by		X	X	X	X		X	

the legal requirement to participate. The intrusive potential of some forms of statistical inquiry requires that they be undertaken only with great care, full justification of need, and notification of those involved. These inquiries should be based, as far as practicable, on the subjects' freely given, informed consent. The identities and records of all subjects or respondents should be kept confidential. Appropriate measures should be utilized to prevent data from being released in a form that would allow a subject's or respondent's identity to be disclosed or inferred.								
--	--	--	--	--	--	--	--	--

Working to change laws that practitioners perceive as unjust (ISI Shared Professional Values, item #2) and seeking to upgrade professional skills (ISI Ethical Principles, item #7) are not aligned with any of the ASA Ethical Guideline principles or elements. However, “exhibiting professional competence” (ISI #7 theme title) *is* aligned with ASA Principles A and G.1; the constituent elements of ISI #7 are not.

Royal Statistical Society (RSS)

From the RSS Preamble: “In every civilised society rules of conduct exist for the benefit of society at large and in order to give freedom for individual members to go about their legitimate business within bounds of behaviour which are accepted and observed by others. In common with professional bodies in other fields, the Royal Statistical Society (RSS) has formulated its own rules as a Code of Conduct to define the actions and behaviour expected of RSS Fellows practising in everyday professional life. This code of conduct has been drawn up to reflect the standards of conduct and work expected of all practising statisticians. It is commended of all Fellows of the Society and is mandatory on all Professionally Qualified Fellows as defined in paragraph 1(s) of the Society's Bye-Laws.

Constitutional Authority

The Royal Statistical Society (referred to as ‘the Society’) is a professional and learned society which, through its members, has an obligation in the public interest to provide the best possible statistical service and advice. In general, the public has no ready means of judging the quality of professional service except from the reputation of the provider. Thus it is essential that the highest standards are maintained by all Fellows whenever they are acting professionally and whatever their level of qualification. Professional membership of the Society is an assurance of ability and integrity.”

<http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf>

ASA (2018, columns) x RSS (2014, rows).

			RESPONSIBILITIES TO/REGARDING:					
ASA:	A Professional Integrity and Accountability	B Integrity of data and methods	C Science/ Public/ Funder/ Client	D Research Subjects	E Research Team Colleagues	F Other Statisticians or Statistics Practitioners	G Allegations of Misconduct	H Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners
RSS:								
Public Interest								
1. Fellows should always be aware of	*Preamble.							

their overriding responsibility to the public good; including public health, safety and environment.								
2. Fellows shall in their professional practice have regard to basic human rights and shall avoid any actions that adversely affect such rights.	*Preamble			x				
Obligation to Employers and Clients								
3. Fellows shall carry out work with due care and diligence in accordance with the requirements of the employer or client.	x	x	(**)	(**)				
4. Fellows shall respect any agreements of confidentiality entered into with an employer or client.		x	x					
5. Fellows should not allow their name to be attributed to work that they have either not	X						X g1	

contributed to or which presents their contribution in a misleading way.								
Obligation to the Profession and the Society								
6. Fellows shall uphold the reputation of the Profession and the Society								
7. Fellows shall seek to advance knowledge and understanding of statistical science and advocate its use	*Preamble	X			x			
8. Fellows shall act with integrity towards fellow statisticians and to members of other professions with whom they collaborate	(x)				x	x	(x)	
9. Fellows shall take personal responsibility for work bearing their name.	x							
Professional Competence and Integrity								
10. Fellows shall strive to act with	*preamble						X (G1)	

honesty and integrity in all aspects of their professional life.	x							
11. Fellows shall undertake continuing professional development (CPD)... and shall encourage others to do likewise.								
12. Fellows shall seek to conform to recognised good practice including quality standards which are in their judgement relevant, and shall encourage others to do likewise.	x	x	x	x	x	x	x	x
13. Fellows shall report to the Society any criminal convictions against them in respect of violence, dishonesty or professional misconduct; or upon becoming bankrupt or disqualified as Company Director.								

Notes: There are three (of 13) RSS Principles that have no alignment with any ASA Ethical Guideline Principle: 6, 11, and 13.

** The ASA Ethical Guidelines for Statistical Practice articulate the opposite of RSS Public Interest item 3: “Fellows shall carry out work with due care and diligence in accordance with the requirements of the employer or client.” Instead, ASA Guideline Principle C specifies that “the ethical statistician supports valid inferences, transparency, and good science in general, keeping the interests of the public, funder, client, or customer in mind (as well as professional colleagues, patients, the public, and the scientific community)”, but also that, “Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations. In some situations, Guideline principles may conflict, requiring individuals to prioritize principles according to context. However, in all cases, stakeholders have an obligation to act in good faith, to act in a manner that is consistent with these Guidelines, and to encourage others to do the same. Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical.” (ASA Preamble). The RSS item 3 specification suggests that work should be consistent with *employer or client requirements*, while the ASA Principles suggest that the statistician should not work in cases where employer or client requirements are inconsistent with the public good, good science (Principle C), or professional considerations (e.g., if a client “requires” that a statistically significant result be obtained when the data do not support such a result or when such a result would be harmful to the public/science/other collaborators or practitioners), then the ethical statistician does not carry out such work. Also, ASA Principles D5 and D6 describe situations where the manner in which data were obtained specifically could preclude the statistician’s engagement, irrespective of the employer or client:

D5. Considers whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations, before analyzing data from such a study, and while reviewing manuscripts for publication or internal use. The statistician considers the treatment of research subjects (e.g., confidentiality agreements, expectations of privacy, notification, consent, etc.) when evaluating the appropriateness of the data source(s).

D6. In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.

APPENDIX. Alignment between ASA Ethical Guideline Principles (Table A) and the core principles of the IEEE Ethical AD (2019 1e) <https://ethicsinaction.ieee.org/>

Table A. ASA (2018, columns) x IEEE EAD (2019, rows).

			RESPONSIBILITIES TO/REGARDING:					
ASA:	A Professional Integrity and Accountability	B Integrity of data and methods	C Science/ Public/ Funder/ Client	D Research Subjects	E Research Team Colleagues	F Other Statisticians or Statistics Practitioners	G Allegations of Misconduct	H Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners
IEEE EAD:								
Human Rights	*Preamble							
A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.	A7		(C)	D; D1, D3, D5, D6		(F4)	(G; G6)	ALL H
Well-being								
A/IS creators shall adopt increased human well-being as a primary success criterion for development								
Data Agency								
A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have			(C5)	(D4, D5, D6)			(G5)	

control over their identity.								
Effectiveness								
A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.	(A; A2)	ALL B	(C1, C2, C3)	(D2; D7)	E2, E3, E4	F1	(G1)	(H4, H5, H7, H8)
Transparency								
The basis of a particular A/IS decision should always be discoverable.	(A; A2)	ALL B	C		E2, E4	F1	(G1)	(H4, H5, H7)
Accountability								
A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.	(A; A1,A2, A4)	ALL B	C; C1, C2, C3		E2, E4	(F4)		(H4, H5, H7)
Awareness of Misuse								
A/IS creators shall guard against all potential misuses and risks of A/IS in operation.	(A; A1; A2)	ALL B	(C3)		ALL E	(F1)	G1	(H2, H3, H4, H5, H7)
Competence								
A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.		(B)					G1	(H3, H7, H8)

Table B. ACM (2018, columns) x IEEE EAD (2019, rows).

ACM: IEEE EAD:	1. General Ethical Principles	2. Professional Responsibilities	3. Professional Leadership Principles	4. Compliance with the Code
Human Rights	*Preamble			
A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.	1.1	2.3	3.1	
Well-being				
A/IS creators shall adopt increased human well-being as a primary success criterion for development.	1.1, 1.2		3.1	
Data Agency				
A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity.	1.6			
Effectiveness				
A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.		(*preamble; 2.1; 2.5)	3.1	
Transparency				
The basis of a particular A/IS decision should always be discoverable.				
Accountability				
A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.		(2.5)		
Awareness of Misuse				
A/IS creators shall guard against all potential misuses and risks of A/IS in operation.	(1.2)	(2.9)	(3.6)	
Competence		*2.2		
A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.		2.6 (2.1, 2.2)		

Not surprisingly, the ACM Code element “compliance” (ACM #4) has no representation in the IEEE ethically aligned design document, since it is not member/membership delimited.

APPENDIX: Ethical Guideline Documents: ASA (2018), ACM (2018), ISI (2010), RSS (2014), IEEE EAD (1e, 2019)

The ASA Ethical Guidelines for Statistical practice were prepared by the Committee on Professional Ethics (chaired by the author) and approved by the Board of Directors of the American Statistical Association in April 2018. The Guidelines appear here verbatim, and are also available at <https://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf>

Purpose of the Guidelines

The American Statistical Association's Ethical Guidelines for Statistical Practice are intended to help statistics practitioners make decisions ethically. Additionally, the Ethical Guidelines aim to promote accountability by informing those who rely on statistical analysis of the standards that they should expect. The discipline of statistics links the capacity to observe with the ability to gather evidence and make decisions, providing a foundation for building a more informed society. Because society depends on informed judgments supported by statistical methods, all practitioners of statistics, regardless of training and occupation or job title, have an obligation to work in a professional, competent, respectful, and ethical manner.

Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations. In some situations, Guideline principles may conflict, requiring individuals to prioritize principles according to context. However, in all cases, stakeholders have an obligation to act in good faith, to act in a manner that is consistent with these Guidelines, and to encourage others to do the same. Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical.

Ethical statistical practice does not include, promote, or tolerate any type of professional or scientific misconduct, including, but not limited to, bullying; sexual or other harassment; discrimination based on personal characteristics; or other forms of intimidation.

The principles expressed here should guide both those whose primary occupation is statistics and those in all other disciplines who use statistical methods in their professional work. Therefore, throughout these Guidelines, the term "statistician" includes all practitioners of statistics and quantitative sciences, regardless of job title or field of degree, comprising statisticians at all levels of the profession and members of other professions who utilize and report statistical analyses and their implications.

A. Professional Integrity and Accountability

The ethical statistician uses methodology and data that are relevant and appropriate, without favoritism or prejudice, and in a manner intended to produce valid, interpretable, and reproducible results. The ethical statistician does not knowingly accept work for which he/she is not sufficiently qualified, is honest with the client about any limitation of expertise, and consults other statisticians when necessary or in doubt. It is essential that statisticians treat others with respect.

The ethical statistician:

1. Identifies and mitigates any preferences on the part of the investigators or data providers that might predetermine or influence the analyses/results.
2. Employs selection or sampling methods and analytic approaches appropriate and valid for the specific question to be addressed, so that results extend beyond the sample to a population relevant to the objectives with minimal error under reasonable assumptions.
3. Respects and acknowledges the contributions and intellectual property of others.
4. When establishing authorship order for posters, papers, and other scholarship, strives to make clear the basis for this order, if determined on grounds other than intellectual contribution.
5. Discloses conflicts of interest, financial and otherwise, and manages or resolves them according to established (institutional/regional/local) rules and laws.
6. Accepts full responsibility for his/her professional performance. Provides only expert testimony, written work, and oral presentations that he/she would be willing to have peer reviewed.
7. Exhibits respect for others and, thus, neither engages in nor condones discrimination based on personal characteristics; bullying; unwelcome physical, including sexual, contact; or other forms of harassment or intimidation, and takes appropriate action when aware of such unethical practices by others.

B. Integrity of data and methods

The ethical statistician is candid about any known or suspected limitations, defects, or biases in the data that may impact the integrity or reliability of the statistical analysis. Objective and valid interpretation of the results requires that the underlying analysis recognizes and acknowledges the degree of reliability and integrity of the data.

The ethical statistician:

1. Acknowledges statistical and substantive assumptions made in the execution and interpretation of any analysis. When reporting on the validity of data used, acknowledges data editing procedures, including any imputation and missing data mechanisms.
2. Reports the limitations of statistical inference and possible sources of error.
3. In publications, reports, or testimony, identifies who is responsible for the statistical work if it would not otherwise be apparent.
4. Reports the sources and assessed adequacy of the data; accounts for all data considered in a study and explains the sample(s) actually used.
5. Clearly and fully reports the steps taken to preserve data integrity and valid results.
6. Where appropriate, addresses potential confounding variables not included in the study.

7. In publications and reports, conveys the findings in ways that are both honest and meaningful to the user/reader. This includes tables, models, and graphics.
8. In publications or testimony, identifies the ultimate financial sponsor of the study, the stated purpose, and the intended use of the study results.
9. When reporting analyses of volunteer data or other data that may not be representative of a defined population, includes appropriate disclaimers and, if used, appropriate weighting.
10. To aid peer review and replication, shares the data used in the analyses whenever possible/allowable, and exercises due caution to protect proprietary and confidential data, including all data that might inappropriately reveal respondent identities.
11. Strives to promptly correct any errors discovered while producing the final report or after publication. As appropriate, disseminates the correction publicly or to others relying on the results.

C. Responsibilities to Science/Public/Funder/Client

The ethical statistician supports valid inferences, transparency, and good science in general, keeping the interests of the public, funder, client, or customer in mind (as well as professional colleagues, patients, the public, and the scientific community).

The ethical statistician:

1. To the extent possible, presents a client or employer with choices among valid alternative statistical approaches that may vary in scope, cost, or precision.
2. Strives to explain any expected adverse consequences of failure to follow through on an agreed-upon sampling or analytic plan.
3. Applies statistical sampling and analysis procedures scientifically, without predetermining the outcome.
4. Strives to make new statistical knowledge widely available to provide benefits to society at large and beyond his/her own scope of applications.
5. Understands and conforms to confidentiality requirements of data collection, release, and dissemination and any restrictions on its use established by the data provider (to the extent legally required), and protects use and disclosure of data accordingly. Guards privileged information of the employer, client, or funder.

D. Responsibilities to Research Subjects

The ethical statistician protects and respects the rights and interests of human and animal subjects at all stages of their involvement in a project. This includes respondents to the census or to surveys, those whose data are contained in administrative records, and subjects of physically or psychologically invasive research.

The ethical statistician:

1. Keeps informed about and adheres to applicable rules, approvals, and guidelines for the protection and welfare of human and animal subjects.

2. Strives to avoid the use of excessive or inadequate numbers of research subjects, and excessive risk to research subjects (in terms of health, welfare, privacy, and ownership of their own data), by making informed recommendations for study size.
3. Protects the privacy and confidentiality of research subjects and data concerning them, whether obtained from the subjects directly, other persons, or existing records. Anticipates and solicits approval for secondary and indirect uses of the data, including linkage to other data sets, when obtaining approvals from research subjects, and obtains approvals appropriate to allow for peer review and independent replication of analyses.
4. Knows the legal limitations on privacy and confidentiality assurances and does not over-promise or assume legal privacy and confidentiality protections where they may not apply.
5. Considers whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations, before analyzing data from such a study, and while reviewing manuscripts for publication or internal use. The statistician considers the treatment of research subjects (e.g., confidentiality agreements, expectations of privacy, notification, consent, etc.) when evaluating the appropriateness of the data source(s).
6. In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.
7. Recognizes that any statistical descriptions of groups may carry risks of stereotypes and stigmatization. Statisticians should contemplate, and be sensitive to, the manner in which information is framed so as to avoid disproportionate harms to vulnerable groups.

E. Responsibilities to Research Team Colleagues

Science and statistical practice are often conducted in teams made up of professionals with different professional standards. The statistician must know how to work ethically in this environment.

The ethical statistician:

1. Recognizes that other professions have standards and obligations, that research practices and standards can differ across disciplines, and that statisticians do not have obligations to standards of other professions that conflict with these Guidelines.
2. Ensures that all discussion and reporting of statistical design and analysis is consistent with these Guidelines.
3. Avoids compromising scientific validity for expediency.
4. Strives to promote transparency in design, execution, and reporting or presenting of all analyses.

F. Responsibilities to Other Statisticians or Statistics Practitioners

The practice of statistics requires consideration of the entire range of possible explanations for observed phenomena, and distinct observers drawing on their own unique sets of experiences can arrive at different and potentially diverging judgments about the plausibility of different explanations. Even in adversarial settings, discourse tends to be most successful when statisticians treat one another with mutual respect and focus on scientific principles, methodology and the substance of data interpretations. Out of respect for fellow statistical practitioners, the ethical statistician:

1. Promotes sharing of data and methods as much as possible and as appropriate without compromising propriety. Makes documentation suitable for replicate analyses, metadata studies, and other research by qualified investigators.
2. Helps strengthen the work of others through appropriate peer review; in peer review, respects differences of opinion and assesses methods, not individuals. Strives to complete review assignments thoroughly, thoughtfully, and promptly.
3. Instills in students and non-statisticians an appreciation for the practical value of the concepts and methods they are learning or using.
4. Uses professional qualifications and contributions as the basis for decisions regarding statistical practitioners' hiring, firing, promotion, work assignments, publications and presentations, candidacy for offices and awards, funding or approval of research, and other professional matters.

G. Responsibilities Regarding Allegations of Misconduct

The ethical statistician understands the differences between questionable statistical, scientific, or professional practices and practices that constitute misconduct. The ethical statistician avoids all of the above and knows how each should be handled.

The ethical statistician:

1. Avoids condoning or appearing to condone statistical, scientific, or professional misconduct.
2. Recognizes that differences of opinion and honest error do not constitute misconduct; they warrant discussion, but not accusation.
3. Knows the definitions of, and procedures relating to, misconduct. If involved in a misconduct investigation, follows prescribed procedures.
4. Maintains confidentiality during an investigation, but discloses the investigation results honestly to appropriate parties and stakeholders once they are available.
5. Following an investigation of misconduct, supports the appropriate efforts of all involved, including those reporting the possible scientific error or misconduct, to resume their careers in as normal a manner as possible.
6. Avoids, and acts to discourage, retaliation against or damage to the employability of those who responsibly call attention to possible scientific error or to scientific or other professional misconduct.

H. Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners

Those employing any person to analyze data are implicitly relying on the profession's reputation for objectivity. However, this creates an obligation on the part of the employer to understand and respect statisticians' obligation of objectivity.

Those employing statisticians are expected to:

1. Recognize that the Ethical Guidelines exist, and were instituted, for the protection and support of the statistician and the consumer alike.
2. Maintain a working environment free from intimidation, including discrimination based on personal characteristics; bullying; coercion; unwelcome physical (including sexual) contact; and other forms of harassment.
3. Recognize that valid findings result from competent work in a moral environment. Employers, funders, or those who commission statistical analysis have an obligation to rely on the expertise and judgment of qualified statisticians for any data analysis. This obligation may be especially relevant in analyses that are known or anticipated to have tangible physical, financial, or psychological impacts.
4. Recognize that the results of valid statistical studies cannot be guaranteed to conform to the expectations or desires of those commissioning the study or the statistical practitioner(s).
5. Recognize that it is contrary to these Guidelines to report or follow only those results that conform to expectations without explicitly acknowledging competing findings and the basis for choices regarding which results to report, use, and/or cite.
6. Recognize that the inclusion of statistical practitioners as authors, or acknowledgement of their contributions to projects or publications, requires their explicit permission because it implies endorsement of the work.
7. Support sound statistical analysis and expose incompetent or corrupt statistical practice.
8. Strive to protect the professional freedom and responsibility of statistical practitioners who comply with these Guidelines.

The ACM Code of Ethics and Professional Conduct

Adopted by ACM Council 6/22/18.

Four main areas are elaborated on with specific behaviors that “a computing professional should” do. The Code is presented below, and is Copyrighted (c) 2018 by the Association for Computing Machinery.

Preamble

Computing professionals’ actions change the world. To act responsibly, they should reflect upon the wider impacts of their work, consistently supporting the public good. The ACM Code of Ethics and Professional Conduct (“the Code”) expresses the conscience of the profession.

The Code is designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way. Additionally, the Code serves as a basis for remediation when violations occur. The Code includes principles formulated as statements of responsibility, based on the understanding that the public good is always the primary consideration. Each principle is supplemented by guidelines, which provide explanations to assist computing professionals in understanding and applying the principle.

Section 1 outlines fundamental ethical principles that form the basis for the remainder of the Code. Section 2 addresses additional, more specific considerations of professional responsibility. Section 3 guides individuals who have a leadership role, whether in the workplace or in a volunteer professional capacity. Commitment to ethical conduct is required of every ACM member, and principles involving compliance with the Code are given in Section 4.

The Code as a whole is concerned with how fundamental ethical principles apply to a computing professional’s conduct. The Code is not an algorithm for solving ethical problems; rather it serves as a basis for ethical decision-making. When thinking through a particular issue, a computing professional may find that multiple principles should be taken into account, and that different principles will have different relevance to the issue. Questions related to these kinds of issues can best be answered by thoughtful consideration of the fundamental ethical principles, understanding that the public good is the paramount consideration. The entire computing profession benefits when the ethical decision-making process is accountable to and transparent to all stakeholders. Open discussions about ethical issues promote this accountability and transparency.

1. GENERAL ETHICAL PRINCIPLES.

A computing professional should...

1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.

This principle, which concerns the quality of life of all people, affirms an obligation of computing professionals, both individually and collectively, to use their skills for the benefit of society, its members, and the environment surrounding them. This obligation includes promoting fundamental human rights and protecting each individual’s right to autonomy. An essential aim of computing professionals is to minimize negative consequences of computing, including threats to health, safety, personal security, and

privacy. When the interests of multiple groups conflict, the needs of those less advantaged should be given increased attention and priority.

Computing professionals should consider whether the results of their efforts will respect diversity, will be used in socially responsible ways, will meet social needs, and will be broadly accessible. They are encouraged to actively contribute to society by engaging in pro bono or volunteer work that benefits the public good.

In addition to a safe social environment, human well-being requires a safe natural environment. Therefore, computing professionals should promote environmental sustainability both locally and globally.

1.2 Avoid harm.

In this document, “harm” means negative consequences, especially when those consequences are significant and unjust. Examples of harm include unjustified physical or mental injury, unjustified destruction or disclosure of information, and unjustified damage to property, reputation, and the environment. This list is not exhaustive.

Well-intended actions, including those that accomplish assigned duties, may lead to harm. When that harm is unintended, those responsible are obliged to undo or mitigate the harm as much as possible. Avoiding harm begins with careful consideration of potential impacts on all those affected by decisions. When harm is an intentional part of the system, those responsible are obligated to ensure that the harm is ethically justified. In either case, ensure that all harm is minimized.

To minimize the possibility of indirectly or unintentionally harming others, computing professionals should follow generally accepted best practices unless there is a compelling ethical reason to do otherwise. Additionally, the consequences of data aggregation and emergent properties of systems should be carefully analyzed. Those involved with pervasive or infrastructure systems should also consider Principle 3.7.

A computing professional has an additional obligation to report any signs of system risks that might result in harm. If leaders do not act to curtail or mitigate such risks, it may be necessary to “blow the whistle” to reduce potential harm. However, capricious or misguided reporting of risks can itself be harmful. Before reporting risks, a computing professional should carefully assess relevant aspects of the situation.

1.3 Be honest and trustworthy.

Honesty is an essential component of trustworthiness. A computing professional should be transparent and provide full disclosure of all pertinent system capabilities, limitations, and potential problems to the appropriate parties. Making deliberately false or misleading claims, fabricating or falsifying data, offering or accepting bribes, and other dishonest conduct are violations of the Code.

Computing professionals should be honest about their qualifications, and about any limitations in their competence to complete a task. Computing professionals should be forthright about any circumstances

that might lead to either real or perceived conflicts of interest or otherwise tend to undermine the independence of their judgment. Furthermore, commitments should be honored.

Computing professionals should not misrepresent an organization's policies or procedures, and should not speak on behalf of an organization unless authorized to do so.

1.4 Be fair and take action not to discriminate.

The values of equality, tolerance, respect for others, and justice govern this principle. Fairness requires that even careful decision processes provide some avenue for redress of grievances.

Computing professionals should foster fair participation of all people, including those of underrepresented groups. Prejudicial discrimination on the basis of age, color, disability, ethnicity, family status, gender identity, labor union membership, military status, nationality, race, religion or belief, sex, sexual orientation, or any other inappropriate factor is an explicit violation of the Code. Harassment, including sexual harassment, bullying, and other abuses of power and authority, is a form of discrimination that, amongst other harms, limits fair access to the virtual and physical spaces where such harassment takes place.

The use of information and technology may cause new, or enhance existing, inequities. Technologies and practices should be as inclusive and accessible as possible and computing professionals should take action to avoid creating systems or technologies that disenfranchise or oppress people. Failure to design for inclusiveness and accessibility may constitute unfair discrimination.

1.5 Respect the work required to produce new ideas, inventions, creative works, and computing artifacts.

Developing new ideas, inventions, creative works, and computing artifacts creates value for society, and those who expend this effort should expect to gain value from their work. Computing professionals should therefore credit the creators of ideas, inventions, work, and artifacts, and respect copyrights, patents, trade secrets, license agreements, and other methods of protecting authors' works.

Both custom and the law recognize that some exceptions to a creator's control of a work are necessary for the public good. Computing professionals should not unduly oppose reasonable uses of their intellectual works. Efforts to help others by contributing time and energy to projects that help society illustrate a positive aspect of this principle. Such efforts include free and open source software and work put into the public domain. Computing professionals should not claim private ownership of work that they or others have shared as public resources.

1.6 Respect privacy.

The responsibility of respecting privacy applies to computing professionals in a particularly profound way. Technology enables the collection, monitoring, and exchange of personal information quickly, inexpensively, and often without the knowledge of the people affected. Therefore, a computing professional should become conversant in the various definitions and forms of privacy and should understand the rights and responsibilities associated with the collection and use of personal information.

Computing professionals should only use personal information for legitimate ends and without violating the rights of individuals and groups. This requires taking precautions to prevent re-identification of anonymized data or unauthorized data collection, ensuring the accuracy of data, understanding the provenance of the data, and protecting it from unauthorized access and accidental disclosure. Computing professionals should establish transparent policies and procedures that allow individuals to understand what data is being collected and how it is being used, to give informed consent for automatic data collection, and to review, obtain, correct inaccuracies in, and delete their personal data.

Only the minimum amount of personal information necessary should be collected in a system. The retention and disposal periods for that information should be clearly defined, enforced, and communicated to data subjects. Personal information gathered for a specific purpose should not be used for other purposes without the person's consent. Merged data collections can compromise privacy features present in the original collections. Therefore, computing professionals should take special care for privacy when merging data collections.

1.7 Honor confidentiality.

Computing professionals are often entrusted with confidential information such as trade secrets, client data, nonpublic business strategies, financial information, research data, pre-publication scholarly articles, and patent applications. Computing professionals should protect confidentiality except in cases where it is evidence of the violation of law, of organizational regulations, or of the Code. In these cases, the nature or contents of that information should not be disclosed except to appropriate authorities. A computing professional should consider thoughtfully whether such disclosures are consistent with the Code.

2. PROFESSIONAL RESPONSIBILITIES.

A computing professional should...

2.1 Strive to achieve high quality in both the processes and products of professional work.

Computing professionals should insist on and support high quality work from themselves and from colleagues. The dignity of employers, employees, colleagues, clients, users, and anyone else affected either directly or indirectly by the work should be respected throughout the process. Computing professionals should respect the right of those involved to transparent communication about the project. Professionals should be cognizant of any serious negative consequences affecting any stakeholder that may result from poor quality work and should resist inducements to neglect this responsibility.

2.2 Maintain high standards of professional competence, conduct, and ethical practice.

High quality computing depends on individuals and teams who take personal and group responsibility for acquiring and maintaining professional competence. Professional competence starts with technical knowledge and with awareness of the social context in which their work may be deployed. Professional competence also requires skill in communication, in reflective analysis, and in recognizing and navigating ethical challenges. Upgrading skills should be an ongoing process and might include independent study, attending conferences or seminars, and other informal or formal education. Professional organizations and employers should encourage and facilitate these activities.

2.3 Know and respect existing rules pertaining to professional work.

“Rules” here include local, regional, national, and international laws and regulations, as well as any policies and procedures of the organizations to which the professional belongs. Computing professionals must abide by these rules unless there is a compelling ethical justification to do otherwise. Rules that are judged unethical should be challenged. A rule may be unethical when it has an inadequate moral basis or causes recognizable harm. A computing professional should consider challenging the rule through existing channels before violating the rule. A computing professional who decides to violate a rule because it is unethical, or for any other reason, must consider potential consequences and accept responsibility for that action.

2.4 Accept and provide appropriate professional review.

High quality professional work in computing depends on professional review at all stages. Whenever appropriate, computing professionals should seek and utilize peer and stakeholder review. Computing professionals should also provide constructive, critical reviews of others’ work.

2.5 Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks.

Computing professionals are in a position of trust, and therefore have a special responsibility to provide objective, credible evaluations and testimony to employers, employees, clients, users, and the public. Computing professionals should strive to be perceptive, thorough, and objective when evaluating, recommending, and presenting system descriptions and alternatives. Extraordinary care should be taken to identify and mitigate potential risks in machine learning systems. A system for which future risks cannot be reliably predicted requires frequent reassessment of risk as the system evolves in use, or it should not be deployed. Any issues that might result in major risk must be reported to appropriate parties.

2.6 Perform work only in areas of competence.

A computing professional is responsible for evaluating potential work assignments. This includes evaluating the work’s feasibility and advisability, and making a judgment about whether the work assignment is within the professional’s areas of competence. If at any time before or during the work assignment the professional identifies a lack of a necessary expertise, they must disclose this to the employer or client. The client or employer may decide to pursue the assignment with the professional after additional time to acquire the necessary competencies, to pursue the assignment with someone else who has the required expertise, or to forgo the assignment. A computing professional’s ethical judgment should be the final guide in deciding whether to work on the assignment.

2.7 Foster public awareness and understanding of computing, related technologies, and their consequences.

As appropriate to the context and one’s abilities, computing professionals should share technical knowledge with the public, foster awareness of computing, and encourage understanding of computing. These communications with the public should be clear, respectful, and welcoming. Important issues

include the impacts of computer systems, their limitations, their vulnerabilities, and the opportunities that they present. Additionally, a computing professional should respectfully address inaccurate or misleading information related to computing.

2.8 Access computing and communication resources only when authorized or when compelled by the public good.

Individuals and organizations have the right to restrict access to their systems and data so long as the restrictions are consistent with other principles in the Code. Consequently, computing professionals should not access another's computer system, software, or data without a reasonable belief that such an action would be authorized or a compelling belief that it is consistent with the public good. A system being publicly accessible is not sufficient grounds on its own to imply authorization. Under exceptional circumstances a computing professional may use unauthorized access to disrupt or inhibit the functioning of malicious systems; extraordinary precautions must be taken in these instances to avoid harm to others.

2.9 Design and implement systems that are robustly and usably secure.

Breaches of computer security cause harm. Robust security should be a primary consideration when designing and implementing systems. Computing professionals should perform due diligence to ensure the system functions as intended, and take appropriate action to secure resources against accidental and intentional misuse, modification, and denial of service. As threats can arise and change after a system is deployed, computing professionals should integrate mitigation techniques and policies, such as monitoring, patching, and vulnerability reporting. Computing professionals should also take steps to ensure parties affected by data breaches are notified in a timely and clear manner, providing appropriate guidance and remediation.

To ensure the system achieves its intended purpose, security features should be designed to be as intuitive and easy to use as possible. Computing professionals should discourage security precautions that are too confusing, are situationally inappropriate, or otherwise inhibit legitimate use.

In cases where misuse or harm are predictable or unavoidable, the best option may be to not implement the system.

3. PROFESSIONAL LEADERSHIP PRINCIPLES.

Leadership may either be a formal designation or arise informally from influence over others. In this section, "leader" means any member of an organization or group who has influence, educational responsibilities, or managerial responsibilities. While these principles apply to all computing professionals, leaders bear a heightened responsibility to uphold and promote them, both within and through their organizations.

A computing professional, especially one acting as a leader, should...

3.1 Ensure that the public good is the central concern during all professional computing work.

People—including users, customers, colleagues, and others affected directly or indirectly— should always be the central concern in computing. The public good should always be an explicit consideration when evaluating tasks associated with research, requirements analysis, design, implementation, testing, validation, deployment, maintenance, retirement, and disposal. Computing professionals should keep this focus no matter which methodologies or techniques they use in their practice.

3.2 Articulate, encourage acceptance of, and evaluate fulfillment of social responsibilities by members of the organization or group.

Technical organizations and groups affect broader society, and their leaders should accept the associated responsibilities. Organizations—through procedures and attitudes oriented toward quality, transparency, and the welfare of society—reduce harm to the public and raise awareness of the influence of technology in our lives. Therefore, leaders should encourage full participation of computing professionals in meeting relevant social responsibilities and discourage tendencies to do otherwise.

3.3 Manage personnel and resources to enhance the quality of working life.

Leaders should ensure that they enhance, not degrade, the quality of working life. Leaders should consider the personal and professional development, accessibility requirements, physical safety, psychological well-being, and human dignity of all workers. Appropriate human-computer ergonomic standards should be used in the workplace.

3.4 Articulate, apply, and support policies and processes that reflect the principles of the Code.

Leaders should pursue clearly defined organizational policies that are consistent with the Code and effectively communicate them to relevant stakeholders. In addition, leaders should encourage and reward compliance with those policies, and take appropriate action when policies are violated. Designing or implementing processes that deliberately or negligently violate, or tend to enable the violation of, the Code's principles is ethically unacceptable.

3.5 Create opportunities for members of the organization or group to grow as professionals.

Educational opportunities are essential for all organization and group members. Leaders should ensure that opportunities are available to computing professionals to help them improve their knowledge and skills in professionalism, in the practice of ethics, and in their technical specialties. These opportunities should include experiences that familiarize computing professionals with the consequences and limitations of particular types of systems. Computing professionals should be fully aware of the dangers of oversimplified approaches, the improbability of anticipating every possible operating condition, the inevitability of software errors, the interactions of systems and their contexts, and other issues related to the complexity of their profession—and thus be confident in taking on responsibilities for the work that they do.

3.6 Use care when modifying or retiring systems.

Interface changes, the removal of features, and even software updates have an impact on the productivity of users and the quality of their work. Leaders should take care when changing or discontinuing support for system features on which people still depend. Leaders should thoroughly investigate viable alternatives to removing support for a legacy system. If these alternatives are unacceptably risky or impractical, the developer should assist stakeholders' graceful migration from the system to an alternative. Users should be notified of the risks of continued use of the unsupported system long before support ends. Computing professionals should assist system users in monitoring the operational viability of their computing systems, and help them understand that timely replacement of inappropriate or outdated features or entire systems may be needed.

3.7 Recognize and take special care of systems that become integrated into the infrastructure of society.

Even the simplest computer systems have the potential to impact all aspects of society when integrated with everyday activities such as commerce, travel, government, healthcare, and education. When organizations and groups develop systems that become an important part of the infrastructure of society, their leaders have an added responsibility to be good stewards of these systems. Part of that stewardship requires establishing policies for fair system access, including for those who may have been excluded. That stewardship also requires that computing professionals monitor the level of integration of their systems into the infrastructure of society. As the level of adoption changes, the ethical responsibilities of the organization or group are likely to change as well. Continual monitoring of how society is using a system will allow the organization or group to remain consistent with their ethical obligations outlined in the Code. When appropriate standards of care do not exist, computing professionals have a duty to ensure they are developed.

4. COMPLIANCE WITH THE CODE.

A computing professional should...

4.1 Uphold, promote, and respect the principles of the Code.

The future of computing depends on both technical and ethical excellence. Computing professionals should adhere to the principles of the Code and contribute to improving them. Computing professionals who recognize breaches of the Code should take actions to resolve the ethical issues they recognize, including, when reasonable, expressing their concern to the person or persons thought to be violating the Code.

4.2 Treat violations of the Code as inconsistent with membership in the ACM.

Each ACM member should encourage and support adherence by all computing professionals regardless of ACM membership. ACM members who recognize a breach of the Code should consider reporting the violation to the ACM, which may result in remedial action as specified in the ACM's Code of Ethics and Professional Conduct Enforcement Policy, which can be found here: <https://ethics.acm.org/wp-content/uploads/2018/07/2018-ACM-Code-of-Ethics-Enforcement-Procedure.pdf>

ISI statement on ethics (2010) <https://isi-web.org/images/about/Declaration-EN2010.pdf>

PREAMBLE

The ISI's Declaration on Professional Ethics consists of a statement of Shared Professional Values and a set of Ethical Principles that derive from these values.

For the purposes of this document, the definition of who is a statistician goes well beyond those with formal degrees in the field, to include a wide array of creators and users of statistical data and tools. Statisticians work within a variety of economic, cultural, legal and political settings, each of which influences the emphasis and focus of statistical inquiry. They also work within one of several different branches of their discipline, each involving its own techniques and procedures and, possibly, its own ethical approach.

Statisticians work in diverse fields such as economics, psychology, sociology, medicine, whose practitioners have ethical conventions that may influence their conduct. Even within the same setting and branch of statistics, individuals may face various situations and constraints in which ethical questions arise.

The aim of this declaration is to enable the statistician's individual ethical judgments and decisions to be informed by shared values and experience, rather than by rigid rules imposed by the profession. The declaration seeks to document widely held principles of the statistics profession and to identify the factors that obstruct their implementation. It recognizes that, the operation of one principle may impede the operation of another, that statisticians – in common with other occupational groups – have competing obligations not all of which can be fulfilled simultaneously. Thus, statisticians will sometimes have to make choices between principles. The declaration does not attempt to resolve these choices or to establish priorities among the principles. Instead it offers a framework within which the conscientious statistician should be able to work comfortably. It is urged that departures from the framework of principles be the result of deliberation rather than of ignorance.

The declaration's first intention is to be informative and descriptive rather than authoritarian or prescriptive. Second, it is designed to be applicable as far as possible to the wide and changing areas of statistical methodology and application. For this reason, its provisions are drawn quite broadly. Third, although the principles are framed so as to have wider application to decisions than to the issues it specifically mentions, the declaration is by no means exhaustive. It is designed in the knowledge that it will require periodic updating and amendment, reflecting on the one hand developments in the generation of information and technical tools utilized by statisticians and, on the other hand, in the uses (and, consequently, misuses) of statistical outputs. Fourth, the values, principles, and the commentaries which follow acknowledge with the general written or unwritten rules or norms, such as compliance with the law or the need for probity. However, the declaration restricts itself insofar as possible to matters of specific concern to statistical inquiry.

Although not explicitly stated, the Principles inherently reflect the obligations and responsibilities of – as well as the resulting conflicts faced by – statisticians to forces and pressures outside of their own performance, namely to and from:

- Society
- Employers, Clients, and Funders
- Colleagues
- Subjects

In carrying out his/her responsibilities, each statistician must be sensitive to the need to ensure that his/her actions are, first, consistent with the best interests of each group and, second, do not favor any group at the expense of any other, or conflict with any of the Principles.

The Principles are followed by short commentaries on the conflicts and difficulties inherent in their application. A link is provided for each ethical principle for those who wish to pursue the issues. Similarly, a limited annotated bibliography is provided after the commentaries for those who wish to pursue the issues or consult more detailed texts.

SHARED PROFESSIONAL VALUES Our shared professional values are respect, professionalism, truthfulness and integrity.

1. **Respect** We respect the privacy of others and the promises of confidentiality given to them. We respect the communities where data is collected and guard against harm coming to them by misuse of the results. We should not suppress or improperly detract from the work of others.

2. **Professionalism** The value Professionalism implies Responsibility, Competence and Expert Knowledge, and Informed Judgment. We work to understand our users' needs. We use our statistical knowledge, data, and analyses for the Common Good to serve the society. We strive to collect and analyze data of the highest quality possible. We are responsible for the fitness of data and of methods for the purpose at hand. We discuss issues objectively and strive to contribute to the resolution of problems. We obey the law and work to change laws we believe impede good statistical practice. We are continuously learning both about our own field as well as those to which we apply our methods. We develop new methods as appropriate. We do not take assignments in which we have a clear conflict of interest. We act responsibly with our employers.

3. **Truthfulness and Integrity** By **Truthfulness and Integrity**, we mean **Independence, Objectivity and Transparency**. We produce statistical results using our science and are not influenced by pressure from politicians or funders. We are transparent about the statistical methodologies used and make these methodologies public. We strive to produce results that reflect the observed phenomena in an impartial manner. We present data and analyses honestly and openly. We are accountable for our actions. We have respect for intellectual property. As scientists, we pursue promising new ideas and discard those demonstrated to be invalid. We work towards the logical coherence and empirical adequacy of our data and conclusions. We value well-established objective criteria of assessment.

ETHICAL PRINCIPLES

1. **Pursuing Objectivity** Statisticians should pursue objectivity without fear or favor, only selecting and using methods designed to produce the most accurate results. They should present all findings openly, completely, and in a transparent manner regardless of the outcomes. Statisticians should be particularly sensitive to the need to present findings when they challenge a preferred outcome. The statistician should guard against predictable misinterpretation or misuse. If such misinterpretation or misuse occurs, steps should be taken to inform potential users. Findings should be communicated for the benefit of the widest possible community, yet attempt to ensure no harm to any population group.

2. **Clarifying Obligations and Roles** The respective obligations of employer, client, or funder and statistician in regard to their roles and responsibility that might raise ethical issues should be spelled out and fully understood. In providing advice or guidance, statisticians should take care to stay within their area of competence, and seek advice, as appropriate, from others with the relevant expertise.

3. **Assessing Alternatives Impartially** Available methods and procedures should be considered and an impartial assessment provided to the employer, client, or funder of the respective merits and limitations of alternatives, along with the proposed method.

4. **Conflicting Interests** Statisticians avoid assignments where they have a financial or personal conflict of interest in the outcome of the work. The likely consequences of collecting and disseminating various types of data and the results of their analysis should be considered and explored.

5. Avoiding Preempted Outcomes Any attempt to establish a predetermined outcome from a proposed statistical inquiry should be rejected, as should contractual conditions contingent upon such a requirement.

6. Guarding Privileged Information Privileged information is to be kept confidential. This prohibition is not to be extended to statistical methods and procedures utilized to conduct the inquiry or produce published data.

7. Exhibiting Professional Competence Statisticians shall seek to upgrade their professional knowledge and skills, and shall maintain awareness of technological developments, procedures, and standards which are relevant to their field, and shall encourage others to do the same.

8. Maintaining Confidence in Statistics In order to promote and preserve the confidence of the public, statisticians should ensure that they accurately and correctly describe their results, including the explanatory power of their data. It is incumbent upon statisticians to alert potential users of the results to the limits of their reliability and applicability.

9. Exposing and Reviewing Methods and Findings Adequate information should be provided to the public to permit the methods, procedures, techniques, and findings to be assessed independently.

10. Communicating Ethical Principles In collaborating with colleagues and others in the same or other disciplines, it is necessary and important to ensure that the ethical principles of all participants are clear, understood, respected, and reflected in the undertaking.

11. Bearing Responsibility for the Integrity of the Discipline Statisticians are subject to the general moral rules of scientific and scholarly conduct: they should not deceive or knowingly misrepresent or attempt to prevent reporting of misconduct or obstruct the scientific/scholarly research of others.

12. Protecting the Interests of Subjects Statisticians are obligated to protect subjects, individually and collectively, insofar as possible, against potentially harmful effects of participating. This responsibility is not absolved by consent or by the legal requirement to participate. The intrusive potential of some forms of statistical inquiry requires that they be undertaken only with great care, full justification of need, and notification of those involved. These inquiries should be based, as far as practicable, on the subjects' freely given, informed consent. The identities and records of all subjects or respondents should be kept confidential. Appropriate measures should be utilized to prevent data from being released in a form that would allow a subject's or respondent's identity to be disclosed or inferred.

IEEE: Ethically aligned design (2019) 8 elements

1. Human Rights

A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.

2. Well-being

A/IS creators shall adopt increased human well-being as a primary success criterion for development.

3. Data Agency

A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity.

4. Effectiveness

A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.

5. Transparency

The basis of a particular A/IS decision should always be discoverable.

6. Accountability

A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.

7. Awareness of Misuse

A/IS creators shall guard against all potential misuses and risks of A/IS in operation.

8. Competence

A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.