A Little Legal Handbook for Data and Computer Scientists

by Rafael "RAFA" Baca, Esquire & MS Data Science
San Francisco, California

FORWARD: Why do this Little Legal Handbook?

At its very heart, the Law provides written coded instructions for human compiling and implementing for what is deemed as innately fair by our human moral compass – our sense of sense of "justice" for settling any conflict or harm that threatens our daily happiness. Societies in the past would use caveman clubs and royal sepulchers but today we use gavels in courtrooms to settle tensions ranging from business fraud to targeted drone strikes.

At this time, data and computer scientists have amassed human-engineered data to a scale that is unprecedented in the entirety of human history. This is indeed a remarkable feat and yet, as a profession, we might possibly have been ignoring a critical need to align such feats with the moral compass within each of us to responsibly address the social challenges in our world today. As data scientists in the 21st century, we are so often busy optimizing machine learning and neural network models, web scraping and spinning up EMR clusters in each digital universe of our own creation, we may tend to pay little attention to some real-world legal issues that often arise within our computer-based and math driven profession.

As a newly-minted data scientist, I too wonder what legal issues might have an affect as I may work on optimizing models for artificial intelligence in business as well as practice law. With my career knowledge as an attorney, it is my hope that this Mini Handbook will give my fellow technical professionals a basic understanding of how the law works here in San Francisco and the rest of the world. Moreover, to do good to the many who will use our data and software, I present the current ethical, moral, and legal challenges that a data and computer scientists are presently encountering in search for a consensus of agreeable conduct that could reliably be enforced online.

The Mini Handbook is organized in three primary sections that all highlight a few instances of how I see the law can most often creep into the life of a data and computer scientist, and would thus be good to know to avoid creating some unintended human chaos. A fourth, bonus section strives unravel the often-intertwined concepts in ethics, law, and morality for developers, data, and computer scientists. So this Little Handbook is divided in these sections:

1. **Legal Stuff 101**
2. **Data Science Careers & Gigs in the Legal World**
3. **Freelance Consulting Contract Breakdown**
4. **Ethics Sidebar**

And remember - *good data comes from good people*.

Rafael Victor Baca, Esquire, M.S. Data Science

CONTENTS:

LAW 101

_____

A. What is Law All About and Why Should Data & Computer Scientists Care:

The law is simply a system of enforceable rules implemented by a formal body to govern a particular community or state. There are many different types of formal legal systems throughout the world as each system is a direct reflection of the culture that is governs. In the sovereign states of the Western world, there are two basic types of legal systems: 1) Roman civil (aka "Justinian", "Codified") law and 2) English common law that includes the United States. Beyond Europe itself, these two legal systems were introduced across the world with the expansion of Western colonialism.

Most all legal systems of the European continent and Latin America are based on Civil law. Civil law is principally predicated on classified or "codified" legal categories for statutory governance, such as from Assault to Zoning. In a codified courtroom, notary publics are the civil law counterpart to a U.S. lawyer that play the role of presenting their client's case by directly referencing applicable civil code. Civil law judges thus interpret the facts based on the given code to arrive at a judgment. Law is changed by governing bodies that edit the existing code and not typically by the civil law courts as is occasionally done with common law courts. Intuitively, Civil law is similar to a programmer's flow diagram given the input steps a fairly consisted legal ruling becomes the output. Much of civil law is interpreted from the outcome of past circumstances or forecasted events. It typically takes a bit of time for this legal code to reflect new, never before social phenomena brought by new technology and social situations, such as social network trolling and possible culpability arising from development in artificial intelligence.

Many of the world's wealthiest nations are based on a flexible common law system that complements and governs capitalist enterprise. The English colonial system of common law is based on a concept called "*stare decisis*" which means that court decisions should be guided by legal precedence but not steadfastly adhering to the exact same past instances of facts and legal interpretations. Similar to developing a machine learning model, common law courts are free to rationally improvise given the most persuasive (or "weighed" in our algorithm analogy) arguments made by the advocating lawyers from the opposing parties in a dispute where such arguments are based on recorded case decisions, codified regulations, and differences in fact.

Each common law case is similar in format to a scientific journal article and is written by the ruling judge. Each legal case is a written discussion of the legal issues that were decided, the factual context of the case, and the laws used to support the winning and, optionally, losing decision. A legal case will be sure to discuss the outcome of proceedings by a jury trial if such a jury option was evoked. A case may reference other previous common law cases, government regulations and statutes, and treaties that support the reasoning behind the legal outcome. In programming, a common law case best resembles the outcome of object oriented programming - as opposed to functional programming - as a legal case is best implemented given certain facts (similar to continuous and categorical variable inputs) where the legal

outcome (aka "legal principle" or "point-of-law") is often characteristically similar to a mutable object.

The first common law case that cites a computer program was a patent law case (*Gottschalk v. Benson*, 409 U.S. 63 (1972)) where the U.S. Supreme Court ruled that a process claim directed to a numerical algorithm is not patentable. Any person can access most all court cases online, although the real magic trick is threading the cases and other laws to support each position in your particular argument so as to persuade a judge and, optionally, jury to rule in your favor.

As common law is often the preferred legal system for global business, either a data or computer scientist it is highly likely to encounter various aspects of this formal legal paradigm in their professional careers. This Handbook strives to address what are the most likely aspects of the common law that a data or computer scientist will likely encounter.

B. Oh no! I have to Deal with Lawyers - What do I do Now!!

This section provides some basic information about some typical real-world legal scenarios that a data or computer scientist is very likely to encounter while working in the business world. This information is not legal advice, as that should be discussed with and received from your lawyer. This information is to give you the likely next steps to follow after you have been legally bombarded while acting in your capacity as a tech worker:

1. NDA & PROPRIETARY INFORMATION AGREEMENTS:

Signing Non-Disclosure Agreements (NDAs) and Proprietary Information Agreements are a very standard business practice not only for data and computer scientists but all parties who either access data or another form of property owned by the provider of the NDA. The degree and time period for legal enforceability typically varies by U.S. state but rarely do the respective state courts grant the effectiveness of such agreements to be more than one year. Enforcing these agreements permits the owner to use a court order stop other party (aka "injunction") from using or even selling the disclosed information – that usually includes asking the court for any monetary and punitive damages as well as attorney fees to be provided to the owner from the disclosed party.

Generally, NDAs tend to prevent dissemination of information outside of the defined boundaries described in the NDA document, such as information disclosed during a startup pitch or business demo. By distinction, confidentiality and proprietary agreements tend to be at a deeper level of specificity of describing the actual sensitive information disclosed. It is common to see confidentially and proprietary information agreements in the context of employment matters and corporate valuations as well.

As your professional reputation will often precede you, it always best to respect the wishes of the data and information owner when they request your signature to one of these types of agreements. These agreements are a very common business practice and are usually executed during the brief period before you access the owner's data and/or are at the owner's premises. Again, NDAs and proprietary agreements should basically be interpreted as "a promise not to tell"

along the basic school yard intuition we all remember but set to grow-up's businesses under the principles of contract law.

2. NON-COMPETE AGREEMENTS:

Non-compete agreements are a somewhat extension of an NDA or Confidentiality Agreement, but often implemented as part of a tech worker's written employment agreement executed during a company's initial employee orientation. Non-compete agreements acknowledge that as you gain access to company data, trade secrets, and other information during the course of your full-time or contract employment and will restrict you for a period of time after your employment not to use that knowledge while subsequently staring your own spin-off company or immediately working for a direct competitor.

Non-compete agreements are state-enforced agreements and the term of non-competition typically lasts for not more than a year or, like in California, not at all (but for a few exceptions). Generally, many state courts tend to become concerned about enforcing these agreements, as they can be too restrictive on an employee's individual livelihood as well as prohibiting competitive enterprise. However, it is likely that many formal employment agreements for tech workers will have such built-in language – but these agreements may be successfully challenged according to applicable State Constitutions.

3. CEASE & DESIST LETTER:

Cease and desist letters are typically sent from a lawyer on behalf of their client informing the recipient to refrain from conducting activity that the lawyer's client is claiming some rights to. Effectively, these Cease & Desist letters act as a firm warning of a lawsuit if the recipient does not amend their ways. Alternatively, if you receive a written Complaint and Court Summons by registered mail or special human delivery (aka a "process server") having special stamps or seals on the document from the court clerk - that is the official start of a law suit. It will then be time to lawyer-up.

In either event, being the recipient of such a warning often brings immediate and genuine fear to most anyone, including computer and data scientists. After receiving such a letter, the "next-step" often depends on the recipient's tolerance of fear. Most recipients stop the action of concern described in the letter. Thereafter, some rights owners never do anything more after the recipient refrains from the concerning actions. Alternatively, other recipients will bring the letter to their employer's attorney and/or their personal attorney to strategize the next possible move with respect to the rights owner. Notably, it is possible that your legal interests may differ from that of your employer, in which case you must find your own personal legal representative and be prepared for a lawsuit.

USE CASE EXAMPLE - Webscraping:

Your company's marketing department asks you to create a user distribution list to send spam emails about your new product. As a data scientist, you then webscrape from WebsiteX which provides an updated membership list to create your requested spam list. Soon after, the owner of the membership list at WebsiteX

has their lawyer send you a Cease & Desist letter stating a violation of WebsiteX's Terms of Use Agreement.

What should you do next? There are a few choices, and here are the most commonly observed actions: Generally, on receipt of the Cease & Desist letter, its is best to immediately stop the webscraping activity just as prescribed by the WebsiteX letter. Thereafter, 1) pass letter along to your lawyer or company's lawyer, and 2) wait and see if WebsiteX will then expense-out a law firm to initiate a law suit against you – depending on how aggravated the folks at WebsiteX really are.

Further, here are some continuing thoughts on websraping: Suppose WebsiteX posts information that is arguably considered as "public information", like reporting the local weather. An immediate notion from a Data Scientist, Data Engineer, Computer Scientist, Software Engineer would be: "if this weather information is public and by reasonable extension 'open source', then why am I served with a stinkin' Cease & Desist letter?"

The simple answer can be found to what degree was this weather information transformed so as become the legally valid property of WebsiteX? The concern derives from basic property rights and who can claim such ownership. Intellectual Property legal concepts typically in copyright may be layered on this rudimentary notion. Often a concerned "owner" in this webscraping scenario could claim that the accompanying metadata for indexing the raw weather data or the graphics for visualizing the raw information is arguably enough of a legitimate legal transformation so as to claim valid property ownership to the data (in its rendered form). The data owner thus asserts a legally potent, valid Cease & Desist letter against webscraping you and your colleagues.

As will be discussed in greater detail further below, another use case factor refers to open source property. Where although it is free to use such software, the original owner and/or creator of this open source property conditionally grants a license to use the software property in a certain way. Although no money is exchanged, open source users must carefully adhere to the terms and conditions stipulated by the owner while using the software and related information. By analogy consider the terms and conditions from a software license agreement to that of an apartment rental agreement, whereby a landlord requires a tenant to use a kitchen oven in a rented apartment for cooking only but not for drying stinky wet socks. As contracts will be discussed in a subsequent chapter of this handbook, a license is a written contractual agreement that, by analogy, is similar to an apartment rental agreement. Contrastingly, given the same analogy, an assignment agreement transfers ownership rights of software property similar to real estate purchase contract.

Another use case factor that will be discussed in greater detail below, is that some webscraped information could very well be subject to US, European Union or other international individual privacy laws. The U.S. Federal Trade Commission (FTC) requires giving written usage notice to the privacy owner as well as secured storage of such private information among many other complicating factors that come with the simple routine act of webscraping.

4. WITNESS - SUBPOENA (Not to be confused with a SUMMONS):

While consistently interacting with potentially sensitive business information on a daily basis, it is likely that a data or computer scientist will encounter a subpoena at least once in their career. A SUBPOENA is a court or legislative order requiring that the recipient appear at a legal proceeding and to provide some specific written evidence, appear at a deposition, or potentially appear at a trial or hearing as a witness. So what can be expected? The subpoena is a written order that will provide in great detail the underlying legal reasons for the recipient's appearance, what information is sought, and instructions for the appearance that includes a time, date, location to appear.

Immediately, regarding matters relating to data or computer science, you will not be asked to go to court if at all. Lawsuits rarely go to trial as business owners typically find a greater cost benefit to their bottom-line to quickly settle most disputes – unless it gets overly emotional and personally charged between the disputing parties. Thus, in the near term, it is likely that you will be interacting with your company's lawyer or, if an individual data consultant or developer, your hired lawyer. The lawyer will inform you and carefully instruct you on what to do next before you eventually meet the opposing party's representatives. Again, please keep in mind that your company's interests may not always be aligned with your personal legal needs which in this case you will need to seek out your individual attorney to additionally participate in the subpoena request at this time.

The two most common activities requested by a Subpoena are to provide witness testimony at either a deposition and/or at trial. Depositions are witness-interview proceedings conducted during the "discovery period", a timeline set by the court to gather evidence for trial. For a deposition, lawyers on opposing sides meet in an office or hotel conference room along with their subpoenaed witnesses and a court reporter that actively transcribes whle the witness is being interviewed or "deposed" by the lawyers. The lawyers for both parties take turns at asking questions or "interrogatories" from each deposed witness. You and your lawyer will rehearse your interrogatory answers many times beforehand (for at least a day or more), as it is often best not to provide more information than necessary. As a deposed witness, your answers may or may not be admitted as evidence for trial based on a number of legal procedure factors.

On the other hand, trial testimony of a subpoenaed witness may or may not arise from a deposition but the style of interrogatories will be the same as a deposition although you are asked real-time questions in an active courtroom setting. By the time of trial, you would have endlessly rehearsed what to say with your lawyer. Not all trials have a jury and some trials are even conducted remotely online.

It should also be added that a subpoena order should not be confused with a summons. A SUMMONS is the official kick-off court document that sets a lawsuit in motion and instructs when to first appear before to your assigned trial judge. A subpoena is delivered to a witness whereas a summons along with a written legal complaint is served to a party (plaintiff or defendant) to a lawsuit. It is very common for individuals not familiar with legal processes to become confused with the two very different terms.

5. WITNESS – NOTARY:

Generally, in the United States, there is often quite a lot of confusion as to what a notary public can do as a witness. Depending on the seal of the U.S. state used by your signing Notary, there is strictly a combination of two common objectives for seeking notarization: (1) acknowledgement for a written document that the signing person's identity has been authenticated (a similar concept to user authentication) and that the authenticated party freely signed the document without coercion; and (2) acknowledgement the signing party swears (under penalty of perjury) that the information contained in the notarized document is believed by the signing party to be factually true. In the United States, a notary only acts as a witness for just only these two objectives as verified by their state assigned seal or stamp and notary signature. As such, a notary public is not responsible for reviewing the actual substantive validity of the document that receives the notary seal.

If you need an actual legal witness, those individuals who are notary publics will most likely never completely satisfy this need (nor will the urban myth of U.S. postal mailing confirming documentation to yourself). As a viable solution, ideally consider two impartial, healthy witnesses (one typically for redundancy) that will be able to provide a live answer to a judge's questions as a legal witness in about 20 years time or more. Impartiality often includes finding individuals that are free of familial or some form of monetary biases relating to you.

C. LEGAL LANDMINES FOR DATA & COMPUTER SCIENTISTS:

This chapter includes is a brief survey of specific of areas law that a data or computer scientist should be familiar with to avoid explosive, costly issues in the future. Although this survey is not comprehensive, this list is something that I personally use while working in the field of Data Science and Software Development in addition to my a practice in Technology & Intellectual Property Law. I am always open to any feedback to improve this list for the developer community as this MiniHandbook is dedicated to you (-we developers and data scientists, that is).

1. LEGAL RESEARCH – AN APPROACH

A suggested method for research and understanding specific laws that a data or computer scientists may encounter during the course of their profession, should be considered as follows:

1) First execute a specific key word search providing critical facts and some narrow legal terminology to an online search engine of your choosing. Deriving the critical facts and legal terminology will iteratively arise from the careful activity of reading each specific case, statute, regulation or treaty retrieved by your search and trying to generally understand as best as you can how similar legal issues are addressed in the research results with the applied law. As the ultimate goal, you hope to gain an intuition of the broader category of law and how this aspect generally applies to facts that are similar to your particular issue(s).

2) Second, to help you to gain a broader understanding of how your applicable category of law is currently applied, consider applying the terms "law journal" or "law review" or "Restatement" to the iterative search to give you a deeper, "next level" of understanding.
3) Third, consider further adding the terms "annotation" or "annotated" to the search terms used in the above second step. Lawyers also refer to annotations as a succinct abstract for interpreting an issue of law – but this deeper-dive is often in lawyerese. You may access annotations online by simply including the terms "annotated", "footnote" to the specific legal statute, regulation or case you are looking for.

It is likely that you may not understand entirely what is being said from your search results, however you are certain to gain very important insights to help you make highly informed decisions as the law relates to your specific situation. Furthermore, similar to having limited online access to IEEE software journal articles, some law journal articles my need to be accessed while at your local law library, a law school or at some court house libraries that are open to the public.

2. INFORMATION AS CURRENCY:

As a data scientist you will quickly realize that most employers are highly protective of their data and often want some legal assurances that their data will be safely and securely managed. In the SOMA district, I have personally observed that data is now far more valuable to this City of San Francisco during this 21st century information revolution than the ore from the gold strike of 1849 ever had. Information has become the currency of a global digital economy.

3. INFORMATION AS INTELLECTUAL PROPERTY:

<u>Trades Secrets & Proprietary Information</u>-

Often data is characterized as a trade secret and so data & computer scientists must be acutely aware of what are trade secrets within a business setting before their professional work begins. From a legal sense, a trade secret is a legal form of intellectual property. A trade secret is information that provides a competitive commercial advantage to its owner, and, would trigger a state civil action against anyone using this information without the owner's permission. A trade secret is any formula, pattern, device, process or compilation of information that provides a competitive commercial advantage in the marketplace. Some examples of a trade secret are a computer program, a specialized customer or marketing distribution list, or data collected by a wearable device.

To gain a further intuition, consider the following factors that differentiate a trade secret from other forms of information. Ask who knows of the information within the company or the outside freely in the marketplace, ex: is the information restricted to upper management? Ask how much effort, resources, and money was expended in creating this information, where an increasing amount of such is a great indicator of a trade secret. Ask how difficult would it be for others to arrive at the same information?

Key to trade secrets is maintaining secrecy at all times. Once this secret is breached, the value of this information to the company is immediately lost. Typically, one distinguishing factor is that loss of information in trade secret is characteristically fatal to the business whereas loss in proprietary or confidential information is often less than fatal -- although severely damaging. It is so critical for data & computer scientists to gain utmost respect for trade secrets – especially when considering undergoing a potential background check before your employer grants you limited access to this information.

It is important for data scientists to further realize that data may not necessarily be a trade secret but nonetheless is distinctively proprietary or confidential information to its owner and must be likewise afforded great respect during access. Typically, the data owner will mark such data as "proprietary" or "confidential" and will often make you sign a legal document, such as a Proprietary or Confidential Information Agreement, which confirms that you understand this designated status before you are even given access to such data.

Data and computer scientists should also be aware that information including trade secrets, proprietary, and confidential information sent across U.S. borders may require extra consideration relating to U.S. Department of Commerce export control regulations. For example, the online export of software or storing of data outside U.S. borders are quite common concerns to the U.S. government. To avoid losing their work status, data and computer scientists who may be foreign nationals working in the United States under H-1B need to be extra vigilant about performing technical services or accessing sensitive company information that may also be regulated under Export Administration Regulations (EAR, 15 C.F.R. Parts 730-774 (2017)[https://www.ecfr.gov/cgi-bin/text-idx?sid=d41c11b9e64e0ef23cd750f3f3e6deed&c=ecfr&tpl=/ecfrbrowse/Title15/15cfrv2_02.tpl], see eg. confidential business information, telecom and information security, electronics, computers, as well as lasers & sensors.)

Copyright Law -

Copyrights protect any combination of literal and non-literal elements of computer programs from being reproduced (ie "copied"), used, or distributed without permission from the copyright owner (until 70 years after the owner's death). Literal elements of computer software that can be copyrighted are source and object code. Non-literal elements of software affording copyright are user interface; screen displays; menus (structure & hierarchy); flow charts, umls, & DAGs; and various means (substantially similar structure, sequence, and organization) chosen to achieve the desired purpose of the program. By analogy, literal elements refer to verbatim aspects whereas non-literal copying refers to paraphrasing those aspects of a computer program. It should be added that, under the U.S. Supreme Court ruling of *Feist Publications v. Rural Tel. Serv. Co.,* 499 U.S. 340 (1991), creative aspects associated with compilations of data *aka* "databases" (including original, non-obvious aspects of database design) may also be afforded copyright protection.

Copyright registration is simple, such that it is often a fast, self-service procedure - in a manner of minutes to hours. Additional legal benefits can be obtained by officially registering your copyright works (software) with the US Copyright Office as an optional, affordable measure of good practice: https://www.copyright.gov.

So, although optional, why is it best to register my software with the U.S. Copyright Office, and what are these additional legal benefits? The simple answer is that it is an affordable process that officially date-stamps your work with federal government. Copyright registration is a strict requirement before starting a federal copyright infringement lawsuit.

Only with prior Copyright registration will you be entitled to the opportunity to pursue additional legal damages against a copycat in the event of a subsequent copyright infringement lawsuit. Specifically, if you promptly register your software with the copyright office no more than three months after creation, the added benefits arising from U.S. Copyright Office registration include recovery of your attorney fees and court costs from the infringing party as well as an added recovery of statutory damages with a successful trial victory. Statutory damages permit a favorably greater calculation for monetary damage awards at a successful trial ruling against a copycat than an actual damage compensation afforded with a not-so-earlybird registration. It should be added that early registration acts as a deterrent to would be infringers knowing the financially detrimental consequences of added damages arising from an early-bird registration at the U.S. Copyright Office.

Data & Computer Scientists may also become introduced to copyrights when encountering pirated software. Piracy describes the act of reproducing copyrighted works without permission from the copyright owner. On the other hand, by distinction, counterfeit further includes the element of an intentional act of deceptively producing copyrighted software that is not the genuine article. The Federal Bureau of Investigation (FBI) and Customs & Border Patrol (CBP) investigate copyright counterfeit and piracy within the United States whereas International Trade Commission (ITC) investigates such activity between the United States and a limited number of foreign nations.

Ownership of software typically refers to whether the software is owned under the laws required in copyright. Copyright ownership should be, but practically not always, indicated by some combination of ©, the word "copyright", the owner's name, dates of creative development, and possibly the registration(s) of a combination of literal and nonliteral aspects of the software with the the US Copyright Office repository. Specifically, all creative works that qualify for copyright under the U.S. adoption of the Berne Convention (after March 01, 1989) do not require the above written notice for a copyright to be displayed – but such a notice will overcome a copycat's defense of Innocent Infringement. Generally, all that is required for a software copyright is that the elements be original and fixed in some written format. Specifically, all the handy details of Copyright subject matter as related to software can either be found in Frequently Asked Questions (FAQ) style from the US Copyright Office's copyright registration portal listed above or comprehensively found in Title 17 of the U.S. Code (https://en.wikisource.org/wiki/United_States_Code/Title_17 (The Computer

Software Act of 1980 reflected in Amendments to Chapter 1, at 17 U.S.C. §§ 101-117 (1982 – present)); *see also*: https://www.copyright.gov/title17/) aka - the US adoption of the Berne Convention).

Like most intellectual property rights, exact procedure for copyright registration and enforcement varies between nations. However, copyright law arguably provides the most internationally uniform set of laws thanks to the 1994 TRIPS treaty, 1996 WIPO Copyright Treaty (with respect to the Berne convention)- which should be the first stop for data and computer scientists requiring research on international copyright issues.

Open Source & Public Domain in Copyright – *The Daily Grind*

In the context of copyright licensing, as an alternative to a purchased end-user copyright license agreement, the concept of open source software copyright licensing plays a critical role in software development and will be encountered by computer and data scientists on a daily basis. Open source development is decentralized and encourages collaboration and interoperability through the no-cost use of software provided under an open source license. Use of open source software permits direct access to source code and of its modifications. However, failure to abide by the terms of use under an open source license could mean not only revocation of further use but also a possible copyright infringement lawsuit that includes those infamous statutory damages calculations.

Many businesses prefer developing their own software applications using at least some elements of open source software licensing with the new code ultimately developed and copyrighted by a business. By contrast to proprietary, custom-made software, open source software is generally cheaper, high quality, robustly tested & secure, and interoperable with other business-to-business applications.

Generally it is best to navigate the concepts of open source software by first determining whether software source code is copyrighted, if the copyright owner places use restrictions on that software, and whether there is some language indicating an applied "Open Source" software license.

Look to the actual use restrictions placed by the copyright owner to distinguish an open source license from a standard for-profit copyright license. For-profit copyright licenses (such as purchased end user license agreements) are more restrictive (especially toward using the source code if at all provided), are more likely to include royalty payments, and agreement by signature such as either opening a "shrink-wrapped" package or affirmatively checking a box. By contrast, some open source licenses (such as "permissive" open source licenses) guarantee the right to use, modify, and redistribute the associated software.

Before considering open source software for your company's development efforts, first see if the software best matches the needs of the company and meets the open source codes' licensed use-obligations. After receiving approval to use open source software, your company must keep an accurate manifest of the open source software used and ensure the software is from a trustworthy source. It is likely that the use license may require you reciprocally dedicate your programmatic derivations from the original code as part of future open source licenses as well.

As an alternative to either open source or copyright licensing, many developers do not claim software ownership rights so that there are no licensing restrictions on using the desired software. Accordingly, by definition, there is no claim to copyright ownership to software that is in the "public domain". Software in the public domain, however, often does not permit user access to the owner's source code. In sum, no restrictions and no claim of copyright ownership are the key distinguishing characteristics of public domain software apart from either open source or paid end-user software.

Digital Millennium Copyright Act (DMCA) –

Copyright as a tool for Censoring Online Content

The US Constitution's 1st Amendment provides for free speech. Under the DMCA, social media websites encourage posting of user-generated content but provide a procedure for copyright owners to remove any online user postings without the owner's permission, aka "a takedown notice". (See https://www.gpo.gov/fdsys/pkg/USCODE-2011-title17/pdf/USCODE-2011-title17-chap5-sec512.pdf). Under the "Safe Harbor" provisions of the DMCA, Internet Service Providers (ISPs) are legally protected from any harms arising from copyright infringement or other liabilities from the content postings made by online users. To maintain safe harbor protection, ISPs cannot substantially change the user-generated content to ensure the copyright remains with the owner (i.e. -so as to avoid creating a new copyrighted work made by the ISP and/or posting user).

ISPs and social media companies have arguably used to DMCA takedown censorship to positively mitigate concerning user activities such as revenge porn, unwanted posted images and language, commercial defamation, and ensuring licensed audio & visual media goes through proper distribution channels and not through unauthorized peer-to-peer postings. On the negative side, to the developer community, the DMCA has been arguably used as an anti-hacking weapon to stifle further development on copyrighted software apart from the established authorized dealer or licensor business model. Specifically, DMCA is used against hackers who typically unlock proprietary software (usually firmware or embedded software) and post the hacked code online, typically to some community message board. Hackers feel this use of DMCA violates free speech, freedom to operate, compete as well as apply "white hat" product testing and research.

Information Subject to U.S. Constitutional & Human Rights -

At this juncture, it is quite important for computer and data scientists to recognize how constitutionally guaranteed human rights affect everyday professional activities. The Bill of Rights (Amendments 1-10) is the first document in human history to constitutionally guarantee individual rights to its citizens. As opposed to public, "free-speech zones" where citizens can exercise their individual inalienable 1st Amendment rights, the DMCA leverages the fact that ISPs' and business owners' with private online venues are commercial property and thus individual online content on these venues can be legally censored by relying on property rights, namely in copyright. Contrastingly, as stated in greater detail below, the individual right to Privacy is never specifically stated in the U.S. Constitution, which legally supplants all other laws as "the highest law of the land."

Actually, the U.S. Supreme Court by interpreting a succession of court cases invented the first ever notion of the individual right to privacy as arising from the due process clause of the 14th Amendment to the U.S. Constitution along with a combination of the 1st Amendment (Privacy to gather & profess a faith), 3rd Amendment (Privacy of residence), 4th Amendment (individual and property), and the 9th Amendment (catch-all or "enumerations" clause).

Unlike 1st Amendment free speech rights that are asserted from the top (US Constitution) down (local laws) to broadly apply to non-commercial as well as commercial individual rights, the "constitutionally implied" right to Privacy in the U.S. has grown organically from local, state laws on up to federally unifying privacy legislation presently proposed by the U.S. Congress. The bottom-up assertion of individual privacy currently results in the narrow application of privacy rights through statutory regulations on just only some information of individuals used for commercial purposes (such as the financial and healthcare industries) as differentiated from the blanket application of 1st Amendment free speech rights to all individual citizens in public places. Arguably, expanding privacy in the United States as an individual right is thus legally more difficult than the narrowing of broader constitutionally enumerated, free speech rights with the DMCA.

This understanding is critical in providing software developers and engineers a general intuition for determining what information may be subject to individual privacy, freedom of speech or censorship takedown laws. A further realization must be considered that individual rights may apply to collected data depending on how the purpose for collecting digital information and even how the data is stored.

Starting with each written constitution from across the world, privacy and free speech rights of citizens varies. Those privacy rights specifically mentioned in a constitution, such as that of the European Union, are legally the strongest individual right conveyed by that government and must be respected as such. Alternatively, in some non-EU nations — such as the U.S.— interpretations of privacy are much more difficult as various legal jurisdictions and type of individual activities do not legally guarantee an individual's right to privacy.

Software Patents -

As discussed, copyright is one of a variety of intellectual property rights for legally protecting aspects of software that further includes patents and trade secrets. Particular software embodiments may also be eligible for patent rights. A U.S. patent is a limited monopoly granted by the federal government to the owner for about a twenty-year period. Arguably, when compared with copyrights in practice, a patent is a much broader legal right that permits the owner to prevent others from making, using, or selling their invention within the United States - although the patent holder need not partake in any of these actual activities legally forbidden to others. Patents refer to inventions that at least in part use software components whereas copyrights refer to literal and nonliteral elemental aspects relating to actual software development.

Unfortunately, the U.S. government application, review, and approval process for a patent is typically long, expensive, and arduous and can take several years as

opposed to a just few hours for a registered software copyright that affords narrow legal rights. Software startup lifecycles in Silicon Valley are far shorter than the typical time to have received a patent grant. However, seeking patent protection depends on what your company's needs are. For example, despite the slow process for approving a software patents, patents often exponentially enhance the valuation of a startup software company in the eyes of investors and are often a "must-do" legal activity for business purposes.

What sort of software can be patented is currently in a state of flux given a recent Supreme Court ruling called *Alice Corp. Pty. Ltd. v. CLS Bank Int'l*, 134 U.S. ___ (2014). In the months after the *Alice* ruling, the number of issued software patents with the eCommerce group of the U.S. Patent Office fell by up to 95% of pre-*Alice* patent allowances. It will take a few years of court rulings to fully unravel the meaning of the *Alice* decision to reliably understand what software aspects will be patentable in this post-*Alice* era. Successful software inventions that have recently received a U.S. patent registration are often described to the Patent Office from the perspective of how a computer (compiler) operates (and, at times, in conjunction with a highly specific UI/UX perspective). Embedded software is also more likely to receive patent registration at the present time.

4. INFORMATION AS SECURE & PRIVATE – A DEEPER PERSPECTIVE

A. Privacy of Individual Electronic Information – A Legal History:

Legal privacy protection of electronic information first appeared in the United States as computers became a viable business tool with the invention of the integrated circuit. Specifically, *Griswold v. Connecticut*, 381 US 479 (1965), is a modern landmark case where the U.S. Supreme Court ruled that the Due Process clause of the Constitution *implies* a fundamental right to individual privacy. The fact that individual privacy was never explicitly mentioned in the Constitution's Bill of Rights is quite an important philosophical distinction from the European Union's Constitution that a data scientist must pay very close attention to.

Unlike the United States, Chapter II, Article 8 of the European Union (E.U.) Constitution (http://www.europarl.europa.eu/charter/pdf/text_en.pdf) states that everyone has an inalienable right to personal privacy protection – including an E.U. citizen's right to delete their personal data. These distinctive approaches to privacy between the U.S. and E.U. play an everyday role in how a data scientist should legally adhere to privacy rules while engaging in their professional activities – and culminate in the E.U. General Data Protection Regulation (G.D.P.R.)-mandated protocols discussed below.

Generally, E.U.-based laws are very strict to ensure that privacy remains a fundamental constitutional right in all circumstances. In the U.S., enforcing individual privacy laws is a patchwork of federal regulations and state laws that apply to only a few circumstances, such as health information, financial data, and geolocation among others. Only the State of California leads the way with privacy law in the U.S. as an inalienable, state constitutional right like that of the multinational E.U. As many California-based tech businesses reside within the E.U., the following amendment to the California State Constitution provides for the individual constitutional right to privacy to all Californians, and even provides for

the right to delete a minor's or student's personal data at the state level: (https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=CONS&sectionNum=SECTION%201.&article=I).

Similar to that of the E.U., in light of many data breeches in the recent news, the U.S. Congress recently introduced legislation to provide a more unified privacy protection within the context of commercial data regulated at the federal level. Despite multiple Congressional votes toward passage, none of these recent bills have been enacted into law. Two of the most well known pieces of US federal legislation are S.1995 Personal Data Protection & Breach Accountability Act (2014) (see https://www.congress.gov/bill/113th-congress/senate-bill/1995 , see also https://www.us-cert.gov/ for US gov't enforcement of breaches) as well as S.547 Commercial Privacy Bills of Rights Act (2015-2016) (see https://www.congress.gov/bill/114th-congress/senate-bill/547/text?q=%7B%22search%22:%5B%22data+breach%22%5D%7D) would provide access to and right to correct one's personal data.

Again it should be emphasized that because of the inalienable rights afforded by the E.U. and, to an extent, California Constitutions, citizens are afforded broad, comprehensive (non-commercial & commercial) privacy rights and protections that are the most favorable for individual data. Although Californias and EU citizens enjoy the greatest rights to individual privacy, all other governments within U.S. state and federal jurisdictions provide some reasonable expectations of individual data privacy that are however not uniform and, more restricted than the E.U. citizen rights, as such U.S. laws usually apply in just commercial settings, for example health care and financial data.

Data and computer scientists will however encounter the greatest difficulty when their professional work is applied within national legal systems other than Common and Civil law systems that do not legally provide actual human rights enforcement to its citizens. Professionally, for example, an ethical dilemma may quickly arise as developers may believe that disrespecting individual human privacy rights is morally wrong while they must adhere to the potentially repressive human rights laws of that nation as a matter of good business judgment. Accordingly, I have included an APPENDIX to this Handbook entitled "*ETHICS SIDE BAR"* to introduce some concepts in ethics, law, and morality for developers, data and computer scientists.

B. US-related Legal Privacy Issues –

In addition to case law, such as *Griswold v. Connecticut* that inferred information privacy as Constitutional with the due process clause, the United States Health & Human Services (HHS) Agency implemented Fair Information Practice statutory regulations for "Records, Computers & the Rights of Citizens in 1973. This was the first regulatory approach to privacy in the U.S. and the world for individual health records and was succeeded by the HHS's Health Insurance Portability and Accountability Act of 1996 (HIPPA) for healthcare informatics.

Similarly, the Federal Trade Commission (FTC) applies statutory regulations to sensitive data associated with consumer goods and services that are actively transferred across state and international borders. Some examples of FTC regulated

non-public "sensitive data" include financial and banking data, health information, individual's geographic information, online children's privacy protection, and social security numbers. As a specific example, regarding financial and banking data, the FTC enforces the Gramm-Leach-Bliley (GLB) Act with fines up to $100,000 per data breech violation and up to five (5) years imprisonment for corporate activities of banks, credit unions, financial aid lenders, insurance companies, auto lenders, real estate, securities brokers, and retail store credit cards (https://www.ftc.gov/tips-advice/business-center/guidance/how-comply-privacy-consumer-financial-information-rule-gramm). In a further example, the FTC broadly applies the Controlling the Assault of Non-Solicited Pornography and Marketing (CAN-SPAM) Act to commercial emails when harvesting valid email addresses with random send-blasts (a criminal act, with $16k penalty for each email). Accordingly, such canned emails must clearly state spam activities are either an ad or solicitation, a non-deceptive subject line, and clearly provide a mechanism for consumer opt-out for at least 30 days.

1. Daily Basis for Data Scientists: FTC Law --

On a daily basis, a data scientist will most commonly encounter the FTC's regulations applications to "sensitive data" relating to online behavioral advertising (FTC – OBA) for eCommerce. Data Scientists are encouraged to read the following US government publications to give an idea of what is expected of data privacy for information related to U.S. businesses under the FTC -OBA's regulatory jurisdiction: 1. FTC primer (https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-staff-report-self-regulatory-principles-online-behavioral-advertising/p085400behavadreport.pdf); 2. FTC Update (https://www.ftc.gov/news-events/press-releases/2009/02/ftc-staff-revises-online-behavioral-advertising-principles).

A general breakdown of the FTC online behavioral advertising principles may be appreciated as follows:

(1). US Data Privacy:
a) A privacy notice under the GLB, should describe categories of information collected and disclosed, categories of affiliated and non-affiliated entitles that the collected information is shared with, and notice of how a consumer can opt-out.
b) For data collection, under FTC-OBA, website administrators should receive written confirmation of consent before collecting sensitive data (data specifically listed above for FTC purposes), disclose the data collection practice of cookies, and providing an opportunity and mechanism to the consumer for opting out from the data collection. Data holders must inform the consumer how data is collected, stored, used, and disclosed.

(2). US Data Storage: reasonable security:

FTC-OBA highly recommends that those who collect and/or store consumer data for eCommerce provide reasonable security for stored data. Data should only

be retained only while fulfilling a valid business or law enforcement needs. Express consent from the consumer is required, before a data holder uses such stored data in a manner materially different from the promises made during original data collection. Security protection should be based on sensitivity of the data, nature of commercial activity, risks involved, and all potential reasonable protections available. Further, FTC's Safeguards Rule (https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/safeguards-rule) requires that all other business partners of the data holder should also safeguard such information. Specifically, FTC Red Flag rule requires that data owners implement routine measures to deter identity theft (https://www.ftc.gov/tips-advice/business-center/guidance/fighting-identity-theft-red-flags-rule-how-guide-business).

(3). US Data Transfer – *post-EU Safe Harbor*:

After 9/11 the U.S. Congress enacted the Patriot Act which removed past individual privacy rights bestowed to all U.S. citizens. As the E.U. Constitution grants greater, comprehensive privacy rights to its citizens than the U.S. (http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679), the U.S. FTC created a voluntary "Safe Harbor" certifications for US companies receiving data from the EU. Safe Harbor allowed US companies to remain legally complaint with the stricter EU privacy laws as well as enabling the US surveillance agencies access to transferred EU citizen data, such as monitoring internet traffic via the National Security Agency (NSA)'s PRISIM program.

Recently, many EU citizens formed and won many class action lawsuits before the European Court of Justice (ECJ) claiming breach of their constitutional rights to privacy while their individual information was accessed from the E.U. for eCommerce activities within and subject to U.S. jurisdiction (http://www.europe-v-facebook.org/EN/Complaints/complaints.html). Thereafter, U.S. companies must voluntarily comply with the current provisions under the Privacy Shield through the FTC (https://www.privacyshield.gov/PS-Application). As long as the discrepancy of privacy between U.S./E.U. citizens remains and 9/11 style US surveillance remains, data transfer between the U.S. & E.U. will remain tenuous for some time. I would encourage all developers to research the many E.U. governmental websites on this matter, see for example http://ec.europa.eu/justice/data-protection/international-transfers/transfer/index_en.htm . As discussed below, the GDPR makes this once voluntary compliance, a legal requirement for most U.S. business to accommodate the individual privacy rights of the data derived from E.U. citizens if such U.S. businesses wish to continue to sell to E.U. citizens.

(4). US Data Disposal:

As previously discussed, until a unifying data privacy law is passed by the U.S. Congress, data disposal laws continue to remain a hodge-podge soup of state and federal laws. Data disposal under federal jurisdiction requires the destruction of (financial) consumer reports so that the information can no longer be reconstructed for future use. (see https://www.ftc.gov/news-events/blogs/business-blog/2016/09/disposal-proposal-ftc-reviews-disposal-

rule). Moreover, about 30 states require personal information be rendered unreadable, undecipherable or destroyed.

California State Law

Developers working within or with California-based companies should be generally aware of state privacy laws. For example, state laws provide quite detailed instructions on how and when a business must disclose use of commercial information under civil penalty. State law further requires reporting data breaches to any California resident reasonably believed to be acquired by an unauthorized party. The State of California Privacy Enforcement and Protection Unit provides great online resources that are truly a recommended read if conducting professional data and computer science based professional activities relating to the Golden State: https://oag.ca.gov/privacy .

C. European Union GDPR Protocols for U.S. Data Professionals --

1. GDPR – A General Notion of Why this Matters

Online commerce is predicated on an information supply chain where goods and services are sold to consumers through an intermediary, a big data infrastructure for collecting, processing, and storing historically unprecedented amounts of data – including personal data – "on the cloud". Recently, through the General Data Protection Regulation (GDPR - http://www.eugdpr.org/), the European Union has made pioneering efforts to permit E.U. government regulators to hold all companies accountable for inappropriately protecting personal data – especially from the recent surge in data breaches (where such companies have failed to be held accountable for their past actions). Effective May 25, 2018, any company (inside or outside the E.U.) that either directly or indirectly works with information relating to E.U. citizens must either comply with the GDPR or face monetary fines from E.U. privacy regulators - newly appointed by the GDPR.

In short, the GDPR is the first globally-scaled data governance law where all participants in the information supply chain are now held accountable if they touch data of E.U. citizens even if geographically located outside of the E.U. Notably, the GDPR also holds "data processors" accountable (including data and computer scientists who may be third party contractors) to mitigate risks from inappropriate use and possible breach of personal information relating to E.U. citizens. Intermediary companies in the information supply chain, such as cloud computing platforms, UI/UX designers and marketers, and data analytics consulting firms, must also now either comply with the GDPR or face legal liability.

Under the GDPR, the European Union expands the scope of personal data that is currently used by United States Federal Trade Commission regulators to now include protecting an individual's social, cultural and economic information, browsing history, and even physical device information. Most all personal data will subject to the GDPRs regulations and monetary penalties.

2. Hefty fines! Are we awake now?

Under the GDPR, the European Data Protection Supervisor (https://edps.europa.eu/) is the sole authority for levying significant fines against EU and non-EU violators of the GDPR. A fine for a regulatory violation under the GDPR is €20 million (approx. $24 Million (USD)) or 4% of an entity's annual turnover, whichever is greater.

To date, U.S. and E.U. representatives have not formalized the final mechanisms for enforcing GDPR fines to non-E.U. entities. As the U.S. and E.U. remain historically and economically tied, it is highly likely that U.S. authorities will assist with enforcing the GDPR against U.S. violators.

As a good rule of thumb, the sustained, direct or indirect processing activities of U.S. businesses relating to the offer of goods or services to E.U. citizens are subject to GDPR regulations. Ultimately, the E.U. regulatory courts have the discretionary ability to determine if such activities are in the spirit of compliance or in disregard of the GDPR.

3. 72 Hour Mandatory Reporting of Data Breach to GDPR Supervisory Authority

Under Article 33 of the GDPR, notification must be provided to at least the Supervisory Authority of the E.U. government (Article 55) within 72 hours of becoming aware of a personal data breach of an E.U. citizen. Under some conditions, as discussed further below, notification to the affected data subject(s) is further required. Inasmuch, all data and computer science professionals should be prepared to implement or asked to develop standard operating procedures for facilitating 72 hour reporting of a breach of personal data under the GDPR. If such a notification requirement is ignored, please be aware that any E.U. citizen may complain directly to the E.U. government of a GDPR violation: https://edps.europa.eu/data-protection/our-role-supervisor/complaints_en

4. PRIVACY BY DESIGN

The GDPR is a new approach to globally governing data through economic means arising from European Union enforcement activities. "Privacy by Design" is a GDPR concept that enables a company to strategically develop new standard operating procedures, impact assessments, approaches to data security, means for operational transparency by which personal data is collected, used, and terminated as well as mechanisms for mitigating personal data breaches. Professionally, to accommodate the GDPR, data and computer scientists will need to holistically adopt the concept of "privacy by design" though constant training and regulatory awareness while developing software as well as new architectures for processing and destroying data along the information supply chain. New management techniques will now be designed to accommodate GDPR regulations that mitigate the risk of personal data breaches whereby, as discussed below, some companies will need appoint a Data Protection Officer, DPO.

5. PIAs or DPIAs - Mandatory Roadmap for GDPR compliance

As part of the underlying "Privacy by Design" approach of the GDPR, a Privacy Impact Assessment (PIAs) (also called a Data Protection Impact Assessment

(DPIAs)) seeks to identify and implement a plan that provides a sustained, internal infrastructure for documentation and implementation of data protection and GDPR compliance within a company to mitigate risk of data breach. PIAs are a mandatory requirement of the GDPR and must be conducted before any personal data is ever handled. PIAs are documentary evidence to be supplied to E.U. and U.S. privacy regulators when requested by GDPR enforcement officers.

6. What must be either Deleted or Exempted by Psuedoanomization:

Each PIA must sufficiently outline all activities associated with the data lifecycle, including the deletion of data. Under the GDPR, data must be deleted for two principle reasons:
(1) the data is longer used for the purpose it was originally collected; or
(2) at E.U. user's request, under the "Right to be Forgotten" principle. Namely, E.U. citizens have the constitutional right to directly demand that their personal data be erased and, optionally, sue entities for compensation if distressed by acts of GDPR non-compliance. Most notably for data scientists, an E.U. citizen has the right to receive an explanation and challenge an automated decision made from a predictive, machine or deep learning algorithm under the GDPR in some circumstances. Thus, a PIA should not only include processes for deleting data but also means for obtaining new user consent for every instance the data is used differently.

Under Recital 26 of the GDPR, Pseudonomization is a means for rendering data such that a user's specific identity is made anonymous to all, including the data processors. Pseudonomization provides immediate practical and regulatory benefits. Practically, pseudonomization transforms the data to accommodate existing IT infrastructure as a cost-effective measure that averts reconfiguring existing company infrastructure. Moreover, pseudonomization is selectively reversible for instances where full data access is needed. As discussed below, according to GDPR regulation, reporting a data breach to each affected user is not required if such data was initially pseudonomized.

Two primary means for pseudonomization are either Format Preserving Encryption (FPE) or Tokenization, a process similar to "one-time pad" encryption. FPE uses referential integrity encryption methods from the U.S. National Institute of Standards and Technology (NIST), for example NIST FF1 AES – SP800-38G, that does not alter the data format to ensure application and data store consistency. NIST FPE encryption is widely used by government and commercial institutions. In practice, be careful when selecting a specific NIST algorithm as some formats have a history of commercial breaches, namely NIST FF3. Tokenization substitutes sensitive data with an undecipherable token. Tokenization cannot be decrypted in a mathematical sense *per se* but de-tokenization may occur by swapping out the assigned tokens by strictly using the original tokenization system.

7. Data Collection under the GDPR: OP-IN Collection & Scope

"Large Scale" data collection under the GDPR requires companies to switch to an "op-in" data approach. An "op-in" user selection requires a user's actual consent to participate in data collection after the collecting entity explains exactly what personal data will be collected, processed and used. A user's silence or inaction will

not be considered consent for "op-in" purposes.

Article 37 of the GDPR does not quantitatively specify what "large scale" means although examples are given such as personal data for behavioral advertising, patient data collected by a hospital, insurance company data, and ISP & telecom data. By contrast, Recital 91 of the GDPR provides examples of what is not considered "large scale" personal data processing – namely, data collected by an individual physician, lawyer, or other health care professional.

For children under thirteen (13) years old who are citizens of the E.U., the GDPR strictly requires that businesses collecting data (i.e. Data Controllers) must demonstrate affirmative parental consent for each child. Otherwise, like other existing international privacy laws, the GDPR forbids the collection, storing, and processing of personal data for minors under age of thirteen (13).

The scope of private information under GDPR has notably increased to include almost all related personal data. For example, GDPR information includes physical device information, browser history and cookies, religious and political beliefs, organization memberships, sexuality, biometric data, and online financial transaction histories.

8. Data Protection Officer, Data Controller, and Data Processor – Requirements:

Under the GDPR, a dedicated Data Protection Officer (DPO) is required for each organization (i.e. company) with more than ten (10) employees and the organization's business depends on processing personal information of at least one E.U. citizen. The DPO is a legal extension of the GDPR supervisory authority that promotes privacy by design while ensuring compliance with the GDPR for personal data collection, storage, and processing activities. Apart from this, all other entities that collect, store or process data from any EU citizen are also subject to the GDPR requirements but without a formal DPO representative.

Therefore, in the event of a data breach, the DPO or informal representative for the breached organization must notify: (1) the Supervisory Authority of the member E.U. nation where the organization's principle operational activities take place and of the locale where E.U. users with the breached data reside. Furthermore, the organization must also notify (2) the affected users only if the breached organization did not apply pseudonomization techniques, failed to apply its PIA to the personal data, and take corrective measures after the breach to ensure greater privacy protection in the future.

The GDPR distinguishes a data organization (or individual) as either a "Data Controller "who provides the business reasons for collecting, storing and processing personal data or a "Data Processor" who actually gets the job done regarding collecting, storing, and processing data. The GDPR requires that Data Controllers appoint the DPO, establish PIAs, and notify the supervisory authority of breach whereas Data Processors are required to notify the supervisory authority of an established PIA before data is processed, ensure data security is maintained, and comply with the GDPR principles while collecting, storing, and processing data. Arguably, as I data scientist, I recognize that the labels between Data Controller and Data Processor can become easily blurred in actual practice. As it will take years for the GDPR to become fully understood, I would strongly recommend that data and

computer science professionals safely assume both GDPR responsibilities as a Data Controller & Processor rather than to receive a fine or reprimand from the GDPR authorities against your means and reputation as you professionally make a livelihood from the data. Thus, for this reason, I stated above that any informal representative of an organization must report a data breach within 72 hours.

2. Careers & Gigs in the Legal World:

Data and Computer Scientists are often informally introduced to the law from the movies, detective novels, and daily newsfeeds containing lawsuits. Many lawyers will tell you they are the happiest when they take-on new case matters as each case often presents interesting backstories and new human narratives in the epic struggle for justice. The daily energy level at a law practice does make for good movie entertainment and often feels like a sporting competition between at least two opposing parties, the plaintiff (complaining party) and defendant (defending party) to a lawsuit.

Legal culture is quite new to the tech world as the rate of legal tech startups receiving significant funding in the Bay Area continues to significantly increase. In the near term, many tech workers such as data and computer scientists will now quickly become new, established members of a 21st century legal team and provide a variety of supporting tasks, such as legal analytics, eDiscovery, Data Mining, Forensics as well as scaled network and security infrastructure support. This new convergence will supplant the historical divisions between the tech industry and legal profession based on inherently different professional cultures as data and computer scientists work with code and succinct algorithms on a daily basis, and lawyers read, write, and communicate with an overabundance of complex and arcane words with vague meanings.

Law suits are truly a team effort all in support of a client's needs. As a new introduction to the legal team, tech workers, such as data and computer scientists, can play a critical role to change the laws that govern our everyday society by using technology to provide a magnified social impact that can instantly affect the lives of many people at a global scale that is unprecedented in human history.

This field is very new and there are many opportunities for entrepreneurial data and computer scientists. Anecdotally, I have heard that legal tech careers pay as much as non-partnered, associate attorneys ($150K on up) and expert or consultant witness gigs with respectable hourly fees ($300/hr on up).

Much of the current rendering of "legal tech 1.0" works to provide rapid, online workflow assistance to an attorney's legal team to ultimately make a compelling story with supporting arguments (based on law and facts) with the hope of successfully persuading a judge to rule in your clients favor. I would suggest try a gig or two in the following areas - before committing to a full-time career, and here are a few helpful observations :

A. LEGAL ANALYTICS -

The same skills for gaining insights from data in other industries are applied to either the legal-side or business-side of law. On the legal-side, an analyst will most likely gain statistical and predictive insights on similar cases, judges, attorneys or other information to help your client win their legal dispute. As lawyers must uphold a unique lawyer-client level of confidentiality, a data or computer scientist must also be extra respectful of many regulations that come along with using legal-side data – similar to healthcare patient privacy.

For business-side data, an analyst will most likely perform standard finance analytics for either private law or corporate law practices. For example, obtaining

business insights as to how profitable an individual lawyer may be or whether to forego hiring a new full-time attorney.

I recommend the American Bar Association (ABA) website to gain ideas on current market trends. Often books, blogs, and non-lawyer resources are available on this subject matter.
https://apps.americanbar.org/dch/committee.cfm?com=CL570000

B. eDISCOVERY (ELECTRONIC DISCOVERY) -

Discovery is a term that refers to the activity where the opposing parties to a lawsuit gather evidence in preparation for trial according to a single timeline set by the presiding judge. For data and computer scientists, eDiscovery work may come in two common varieties: data mining or forensics.

C. DATA MINING

As shown frequently in lawyer-drama movies, one often encounters the infamous tactic of a "document dump" where lawyers for the opposing side provide requested evidence in typically truckloads of paper file boxes. As lawyers in lawsuits are paid hourly, the strategic tactic is to quickly burn through the document dump recipient's resources on the opposing client's side while they look for that "needle in a haystack" nugget of evidence.

For a computer and data scientists and data engineers, the process of Optical Character Recognition or other structuring of the document data and subsequent data mining a document dump for one legal case could easily keep one tech worker employed on a full-time for a few years. From a business standpoint, law firms relatively pay far more to hire teams of people including junior lawyers to manually comb through the dump for "smoking gun" type evidence.

D. DATA (or DIGITAL or COMPUTER) FORENSICS

Forensics is an evidentiary approach that uses scientific principles to gather evidence in preparation for court. Forensics in a digital context is applying Hollywood-style detective work toward gathering electronic evidence for either criminal or civil trial. Forensic procedures are an especially ideal standard for gathering indirect or "circumstantial" evidence. Forensics applies a very methodical scientific procedure that in practice should increase chances that such forensically rendered evidence will be deemed accepted or "admissible" by presiding court judges for official use at trial. Thus, admissible evidence is used by a combination of either a judge or jury to form a legal basis for a court ruling. A great college text on digital forensics for beginners is: *Digital Evidence & Computer Crime: Forensic Science, Computers and the Internet* by Eoghan Casey.

As everyday life events are becoming digitized, there is an ever-growing need for forensic-based evidentiary retrieval from the digital world. Data and Computer Scientists can easily satisfy the exploding demand digital investigators provided that they pass a couple of certification exams. To become a digital forensic investigator there is no formal professional certification exam that is analogous to a lawyer passing a bar exam before practicing their profession. However, some digital

instigators optionally take the Certified Computer Examiner "CCE" certification exam (https://www.isfce.com/) sometimes along with EnCE software certification (and a minimum of three months experience with the software) (https://www.guidancesoftware.com/training/certifications?cmpid=nav_r#EnCE) to informally gain a recognized "expert" status with the digital investigation industry. The State of California has further popular certification requirements worth mentioning as well (http://sidebars.cdaa.org/new-item4). On average eDiscovery certified specialists make salaries between $100k-$250k by gathering legally valid evidence within the digital realm for court cases, government matters as well as for corporate business intelligence.

E. EXPERT & CONSULTING WITNESSES

Either a gig as an expert witness or consulting witness is a wonderful, highly paid introduction for tech workers to the legal world. A legal team retains you for your professional experience as a data scientist, computer scientist, data engineer, developer and so on. Traditionally, you earn a few hundred dollars per hour such a temp gig and many referrals are easily made if you do a great job.

Specifically, for the legal team, a consultant witness acts as an interpreter and educator for a facet of technology that you are professionally familiar with while the legal team figures out a compelling case story for their client. While acting as a consultant witness you pretty much hang around that team's cushy law office during the gig, as there is no expectation, by definition, that you will be subpoenaed as a witness in the courtroom.

On the other hand, an expert witness is expected to testify in court by providing either their professional technical knowledge or observations from the presented evidence. As the case develops over time, it should be added that is likely that an individual tech worker can be both consultant and expert witness.

There is no single path for obtaining such a well paying legal witness gig but for networking. It is better to try many different approaches to get noticed as a technical expert. One method is contacting and frequently visiting only those law firms that you reasonably believe can use your specific data, computer, and software expertise in the future. Also consider going to your local and state bar associations to see if directories of expert witnesses or advertising is available to the local attorney membership. Frequent personal visits still go a long way when it comes to referrals.

Otherwise, you can be sure to create an online profile. Typically, your profile in an established directory for expert witnesses may cost some money. You will need to experiment what directories are most likely to get you those quick, high paying gigs. Well-known Professional Associations are one great way to get the word out to legal teams who may be searching online. For example the IEEE Consultants Network Membership is highly well known by the legal world for all things tech. https://www.ieee.org/membership-catalog/productdetail/showProductDetailPage.html?product=UH3001 . Additionally, there are many online directories for legal and insurance expert witnesses such as Thompson Reuters (https://www.trexpertwitness.com/expert-witness-service/) & TASA (https://www.trexpertwitness.com/expert-witness-

service/) as well as HG (http://www.hgexperts.com/consultants-expert-witnesses.asp) As there are many options for online advertising, these are exam

3. Freelance Consulting – A Work Contract Breakdown:

Quick, temporary gigs are often the stock and trade of tech developers. As you may take-on freelance consulting contracts either as a full-time endeavor or for some extra pocket money, understanding the basic concepts of a legally binding contract is a critical business function that freelance developers must assume.

A well-executed consulting contract will mitigate many miscommunication pitfalls that lead to lawsuits as well as gain a well-deserved reputation as a software consultant with exceptional business acumen. With a clear, contractual understanding of each party's roles and obligations early on, a consultant's work efficiencies will profitably increase.

A. For freelance developers, a *Basic Contractual Equation* is as follows -

K = Offer + Acceptance

A binding contract refers to at least one legally enforceable promise between two parties, an advancer of a conditional promise or "promisor" as well as one individual who acts to fulfill and benefit from the promise or "promisee". Optionally within a consultancy, a promisee may "subcontract-out" their obligations to other parties to assist with carrying out the promisee's conditional promise – typically with the Promisor's permission. Further, the benefit of a promise could be assigned by the promisee to a third party beneficiary.

Specifically, as a traditional approach to contract law, Offer and Acceptance are key elements for identifying a legally binding contract.

(1) **OFFER**: An *Offer* is a statement of terms that defines how the promisor will remain obligated to the promisee if the promisee accepts such terms of the offer. For example, an offer may be: "I promise to pay DevelopR Corp. $50,000 if DevelopR creates working software that accurately predicts consumer online grocery delivery spend in Palo Alto for the next six months".

(2) **ACCEPTANCE**: An *Acceptance* is a willingness by the promisee to take action (or refrain from action - called a "forebearance") up on and be bound to the conditions set forth in the *Offer* also called willingness to take "performance (of the Offer)".

(3) "**+**" refers to a Mutual Understanding, Mutual Assent or "Meeting of the Minds" between the two parties: each party has an complete, identical understanding of the terms of the offer and acceptance performed as that of the other party. The plus symbol is an additional requirement needed for the promise to become a legally binding contract. Surprisingly, a complete and identical understanding of the terms of the agreement is more difficult than it would seem – even if the Agreement were in written form. Many expensive law suits require a drilling down to a deeper level of what each party understood the contract to be in their "minds' eye".

(4) "**=**" refers to the complete execution of the performance required of the promisor under the agreement. Only on completing all the terms of the agreement,

will the Promisee receive the offered "consideration", i.e. valued payment of some item or action by the Promisor.

(5) Lawyers typically use the symbol "**K**" to refer to a contract.

Unilateral vs. Bilateral Flavors of Contracts - It should be added that there are a few basic flavors of contracts based on the promises made between parties. Unilateral and Bilateral Contracts are the two most commonly seen by the tech community.

B. Unilateral Contracts: A unilateral contract occurs when only one party (usually the Offeror) makes a specific promise and is obligated to perform on that promise if another party agrees or "assents". Characteristically, unilateral promises are often broadcasted out to anybody who accepts the terms of the agreement and are typically not made to just specific individuals. Distinctive to unilateral contracts, one individual party (the Offeror) makes both a promise and performs on that promise – whereas the other party to the agreement basically just agrees to that offer.

"Clickthrough" or "Clickwrap" license agreements are one of the most common types of unilateral contracts experienced in the tech field. A contract is thus formed as a user selects an agreement button - at the user's option. Specifically, these unilateral agreements typically provide a user with the option to click a button or take some other affirmative form of action to agree to the terms and conditions for using a website, application or software download. In this case, the Offeror makes a unilateral or "one-way" promise. So long as a user agrees to accept the "terms of use" by clicking some provided agreement button on the screen of an electronic device, the Offeror must then contractually allow those clicking individuals use of the version of software described under the Offeror's terms.

C. Bilateral Contracts: The above-presented "Basic Contractual Equation" assumes a bilateral relationship. Bilateral contracts require that mutual promises be exchanged between two parities, whereby each party receives a benefit. By contrast, only one party is making a promise in a unilateral situation.

Bilateral contracts are the most commonly used legal instruments for business transactions. In bilateral situations, both parties to either side of an Agreement make promises and must each take further action (or inaction) to complete their obligations.

A contract is broken or "breached" by a party who first fails to perform their obligations on the terms of the contract. However, the degree of breach matters for determining whether the other party must continue to fulfill their end of the bargain.

Specifically, for a minor breach, both parties can still work to fulfill the original obligations under the contract. An example of a minor breach: software was set to be delivered on January 1st but not realizing that such a date is a holiday and thus actually delivered on the next business day.

A material breach is where the other party provides something significantly or "materially" different from the agreed terms –that includes failing to act altogether. A material breach consequently removes contractual obligations from the other party that was harmed by the breach.

Thereafter, breach of contact disputes may be resolved in court typically in actual economic terms of what was lost from not making good on contractual promises plus reasonable ancillary economic damages arising from the breach. It should be added that a valid, legally enforceable contract could be in either oral or written form. Oral contracts are more costly to enforce in court as witnesses are often needed to validate the contractual terms *in lieu* of a written agreement.

D. *CONSULTING CONTRACT TEMPLATE: HOW TOs -*

Below, let us discuss the basic elements of a bilateral contract in the form of a Consultant Agreement. Illustratively, consider a Software Consultant Agreement as a practical example to help computer and data scientists gain some background business knowledge while pursuing their interests in acquiring lucrative and interesting consulting gigs.

There are many online consultant contracts out there, I will attempt to give you general understanding of the basic terms and conditions. As a rule of thumb, the very best contract is one that memorializes the obligations and agreeable understandings of both parties in writing as completely and accurately as possible.

All other terms to the contract outline what the parties should do in the event of a dispute or if some legal friction happens in the future. Lawyers are trained to be very specific when adding beforehand what steps the parties must take in the event of a dispute to avoid a drawn out and costly ordeal between the parties. This back and forth of written contract corrections between lawyers is often a sidebar ritual that can be distracting at times from quickly furthering the real business agreement between the parties to the actual contract. Thus, much of this back and forth between lawyers can be mitigated if the actual parties to a contract are as clear and specific from the very start regarding what to do in the event of a dispute.

The below template will provide numbered sections that are blocked off in red color. After introducing the template, I will discuss the numbered sections:

1 [

INDEPENDENT [software/data science] CONSULTING AGREEMENT

This Agreement is made this _____ day of _____, 20__, between _____ _____ (the "Client"), a _____(Name of State) corporation with a principal place of business at the following address _____ _____, and _____ (the "Consultant"), a sole proprietor and an individual resident of the County of _____ in the State of _____ .]

2 [WHEREAS, the Consultant, is an independent consultant in the field of [Data Science / Computer Science] and desires to perform business consulting services for legal business entities, subject to the terms of this Agreement;

WHEREAS, the Client desires that the Consultant provide advice and assistance to the Client in his or her area of expertise;

WHEREAS, the Client and the Consultant agree the Effective Date for beginning this Agreement is _____ and consulting services shall continue at least until this expected time (that is subject to change) _____;

WHEREAS the Client desires the Consultant to perform the following services as understood by the Consultant and the Client to be the following in summation and where the entirety of services is set forth in EXHIBIT A that is appended to this Agreement:

_____

_____

_____,

; and

WHEREAS, the Consultant desires to provide such advice and assistance to the Client under the terms and conditions of this Agreement;

NOW, THEREFORE, the Client and the Consultant hereby freely agree as follows:]

3 [1. Consulting Services

(a) Subject to the terms and conditions of this Agreement, the Client hereby retains Consultant as a technical advisor to perform the consulting services specifically set out in Exhibit A and made a part hereof (hereafter referred to as the "Consultant Services"). Such Consultant Services shall be limited to the area of expertise described in Exhibit A. The parties agree that this Agreement creates an independent contractor relationship. The parties further acknowledge that neither party has the authority to bind the other party during the rendering of the Consulting Services. Consulting Services under this Agreement may not be assigned or subcontracted by the Client to any third party without express permission from the Consultant. This Agreement does not create a partnership relationship between the parties.

(b) During the rendering of Consulting Services, the Client shall provide the Consultant with sufficient information relevant to forming any pertinent conclusions by the Consultant. It is well understood that no fiduciary obligation to the Client arises from the Consulting Services. The Consulting Services are further understood to provide independent conclusions. Therefore the Client and the Consultant both understand that under no circumstance is the Consultant obligated to be an advocate for the Client in any forum, public or private. The Client further agrees that under no circumstances will these Consulting Services be compromised in terms of independent conclusions or inaccurately represented.

2. Compensation and reimbursement.

In consideration of the services to be provided by the Consultant to the Client hereunder, the Client shall pay to the Consultant $ _____.
The Consultant shall render services at times and places mutually agreed by the Client and the Consultant. Expenses incurred during the course of the Consulting Services shall be reimbursed by the Client for the following mutually agreed aspects: _____. Invoices for all Consulting Services shall be rendered in a timely manner and, optionally, at an agreed upon periodic basis. Payment shall be provided within _____ business days after receipt of such Invoice. Failure to pay the Consultant in a reasonable time shall constitute a material breach of this Agreement.]

4 [3. Indemnification

The Client shall indemnify, defend, and hold harmless the Consultant and its employees against any claim, liability, cost, damage, deficiency, loss, expense or obligation of any kind (including without limitation reasonable attorneys' fees and other costs and expenses of litigation) incurred by any claims arising out of this Agreement. The Client will provide notice to the Consultant of any such claim and will assist the Consultant in defending such claim.]

5 [4. Intellectual Property

[CHOOSE OPTION 1 – YOU KEEP THE INTELLECTUAL PROPERTY DERIVED FROM THE GIG i.e. “CONSULTING SERVICES” and thus might charge less for consulting fees]
“Work Product” derived from Consulting Services under this Agreement includes, but is not limited to, use cases, specifications, visualizations, flow and uml diagrams, dashboards, content, object and source codes, data, all ideas and other materials, in whatever form, developed during the course of the Consulting Services for the Client.

Therefore, the Client agrees that the Client will retain all rights the Consultant may have in the Work Product.

Therefore, under the Agreement, the Consultant grants the Client an unrestricted, nonexclusive, perpetual, fully paid-up, worldwide license to use the Work Product for the purposes of developing and selling the Clients products and services but not for the purpose of selling the Work Product as a separate body of work that is not combined with the Client’s products and services.

[CHOOSE OPTION 2 – YOU GIVE AWAY ALL THE INTELLECTUAL PROPERTY DERVIDED FROM THE GIG i.e. “CONSULTING SERVICES” and thus typically charge more]
“Work Product” derived from Consulting Services under this Agreement includes, but is not limited to, use cases, specifications, visualizations, flow and uml diagrams, dashboards, content, object and source codes, data, all ideas and other materials, in whatever form, developed during the course of the Consulting Services for the Client.

The Consultant hereby assigns to the Client entire right, title, and interest including all patent, copyright, trade secret, trademark, and all other residual rights arising from the Work Product developed during the Consulting Services.

At no cost, the Consultant will assist and execute in the preparation of any papers that the Client considers as necessary to obtain or maintain any patents, patent, copyright, trade secret, trademark, and all other residual rights arising from the Work Product. The Client assumes all expenses for obtaining and maintaining the assigned rights under the Agreement and shall reimburse the Consultant for reasonable out of pocket expenses incurred in pursuit of obtaining and maintaining such assigned rights derived from Consulting Services.

5. Consultant Tools
Consultant tools refer, but are not limited, to software, programs, documentations, and visualizations that include among others object code, source code, utilities, software tools, utilities, and copyrightable material that does not comprise the Work Product and are owned or licensed solely by the Consultant with a right to sublicense.

The Consultant’s tools may include the following _____

_____.

6. Confidential Information

(a) The parties acknowledge that during the Consultant Services, the Client may disclose to the Consultant and the Consultant's employees and agents confidential and proprietary information and trade secrets of the Client. Moreover, the Consultant may also create such information within the scope and in the course of performing the Consultant Services. The Consultant will not use or disclose such information to third parties without the Client's written consent. The Consultant as well as the Consultants employees and agents shall each be individually subject to all Confidentiality provisions under this Paragraph.

Such information includes but is not limited to: Written or electronically recorded materials furnished by the Client for the Consultant's use. The Client's customer lists, trade secrets, operating procedures, know-how, business and marketing activity, data and software conferring a competitive advantage in the market place to the Client. All information marked as "Proprietary", "Confidential" or the like. Oral information conveyed to the Consultant as "Proprietary", "Confidential" or the like. All information indicated by the Client as "Proprietary", "Confidential" or the like. Oral information conveyed to the Consultant as "Proprietary", "Confidential" or the like and then later summarized in written memorandum marked as "Confidential" and legally delivered to the Consultant within thirty (30) calendar days after disclosure.

Such information does not include information that is publically available, independently developed by the Consultant or the Consultant's employees and agents or in the Consultant's possession prior to the Effect Date of this Agreement. The Consultant must provide reasonable evidence of independent development or possession to the Client no later than ninety (90) from the Effective Date of this Agreement.

Such information shall be returned to the Client, except the Consultant may keep one file copy of all documents (which copy shall be subject to the confidentiality and non-use requirements set out in this Agreement). Confidentiality shall remain in effect after the termination date of this Agreement for five (5) Calendar Years or another term mutually agreed by the parties in writing. The Client further agrees that the Consultant shall not be liable to the Client or to any third party claiming by or through the Client for any unauthorized disclosure or use of such information that occurs despite Consultant's compliance with the Consultant's obligations under this Agreement.]

6 [7. Term and Termination of the Agreement

The Term of the Agreement was set forth above and starting with the Effective Date. This Agreement will remain in effect during the term of Consulting Services. Alternatively, this Agreement may be terminated by either party under the following procedures:

(a) Terminated without cause (for any reason), by fifteen (15) business days after receipt of a written termination notice by either party; or
(b) Terminated with cause, immediately upon material breach of any term of the Agreement by either party.

At termination, the Client is responsible for paying the Consultant up to the day of receipt of written termination notice or material breach; and the Client shall pay no later ten (10) business days from the day of receipt of written termination notice.

At termination without cause, the Consultant shall be entitled to compensation and reimbursement accrued under the terms of this Agreement. In addition, the Consultant shall be reimbursed for any noncancellable obligations and penalties arising from such obligations, unless the Consultant terminates the agreement without cause.

8. Warranties

(a) The Consultant and the Client each warrant that they respectively have authority to enter into this Agreement and perform the arising obligations thereto.

(b) Work Product is created by the Consultant or under the Consultant's supervision by the Consultant's employees or authorized agents.

(c) For a period of _____business days the Work Product is warranted to be substantially in conformance under the terms of the Consultation Services set forth in Exhibit A.

9. Other Agreements

(a) This Agreement with the appended Exhibits constitutes the entire Agreement between the Consultant and the Client. No modification of this Agreement and the appended Exhibits, shall be valid unless mutually agreeable, made in writing, and executed by the Consultant and the Client.

(b) All notices and communication between the Client and the Consultant pursuant to this Agreement shall be in writing with some means of providing a validating date stamp.

(c) The Agreement is governed by either the laws of the State of _____ and federal laws of the United States relating to intellectual property.]

IN WITNESS WHEREOF, the parties are legally authorized to have executed this Agreement on the dates indicated below.

_____ _____
[Consultant's Signature] [Date]

[CLIENT]

By: _____
[Client Representative Signature]

Business Title: _____

Date: _____

----------NEW PAGE----------

7 [Exhibit A- Comprehensive Description of the "Consulting Services"

Task at hand requiring Consultant Services:

Field of expertise Consultant provides to the Consultant Services (list all technical skills):

]

E. DISCUSSION to the (Contract Template) NUMBERED SECTIONS:

1. This section formally and legally introduces the parties to the Agreement. For parties that are legal business entities, such as a corporation, information should closely reflect the information provided during the formal process of business

incorporation. If a party is unincorporated or not a formal business entity, then indicating "sole proprietor" with information of residency is sufficient.

2. This section very much provides information like a technical or scientific abstract. This section captures the essence of what the contract is all about and what was negotiated. Lawyers often term this as the "Recitals" section.

3. This section encapsulates the heart of the Consulting Services provided in terms of what the consulting gig is all about, what information will be supplied to assist with the consulting services, and payment terms for the gig. This section makes reference to Exhibit A, which comes at the end of the Agreement. Exhibit A should concisely provide the mutually understood job description (i.e. task to be resolved and expected deliverables) and all the technical skills that the consultant is bringing to and will use for the contracted gig.

4. An indemnification clause basically says that if the consultant is sued based on this Agreement, the Client will step in and take the responsibility (i.e. "legal liability") away from the Consultant in the pending lawsuit. This is a very nice clause for a Consultant to have – but may not always be accepted by a potential Client who is often hesitant for making good on the Consultant's misdeeds.

5. This section attempts to clear-up ownership and stewardship of each party's data during the gig. If some valuable information arises from the project, the Agreement may either assign all intellectual property rights to the Client or the Consultant can keep them. Moreover, the Agreement discusses how to respectfully care for confidential information provided by the Client for the contracted gig.

6. This section provides nice, extra "bells and whistles" features to avoid any contention between the parties arising from the executed Agreement. For example, this section provides formal procedure for terminating the contract, communication between the parties, and allows for a time after the agreement to fix any bugs on deliverables to the Client. Although this sample Agreement provides clauses requiring court venues to settle a dispute, this Agreement does not provide an arbitration clause (-- of which you can easily find online). Although arbitration clauses are a cheap and quick way to potentially settle disputes, they often will restrict or preclude you from your full legal rights to settle your dispute in court (i.e. the court option). It is thus highly recommended that you carefully research and consult a legal professional to help with an overall approach is ideal for your situation.

7. Again, Exhibit A should concisely set forth the mutually understood consulting job description, expected deliverables, and all the technical skills that the consultant will bring to the gig. Ideally, you want to capture all that was negotiated regarding the objectives in this section so it is likely that both parties should work on the draft together before signing.

APPENDIX – *AN ETHICS SIDEBAR*

Data and computer scientists often encounter great difficulties when their professional work is applied within national legal systems, other than common and civil law systems, that do not guarantee human rights protections to its citizens. Further concerns might arise as current employers may require some professional activity that is legal but may be objectionable as being personally immoral. As an APPENDIX to this Lit'l Legal Handbook, the following "*ETHICS SIDEBAR"* strives to unravel the often-intertwined concepts in ethics, law, and morality for developers, data, and computer scientists.

FOO GOOD:

As a profession, we programmers strive to do good. "Real World" motivations are often murky and, comparably, not as clear as our programming which reliably provides well-structured output in bits and bytes. In the end, we should always strive to hack the real world based on our personal moral compass to do good to the many who will use our data and software. *Coding foo should never bar our personal and professional obligations to foo good.*

So how do we know our own personal brand of foo is good? Let's first gather a distinguishing understanding of law, ethics and morality . . . .

Digital Ethics, Morality, and the Law:

Introduction: The Digital Realm -

As technology within the digital domain continues to rapidly expand and affect our everyday lives, building resonate and inspirational communities on Internet frontier is arguably one of the most noteworthy efforts in recent human history. By its very nature, creation is a messy process where a sense of wonder, lawlessness, amusement, love, and anger can all be seen all at once during this Internet "Wild West" period.

Presently, many communities across the Internet are struggling for a shared moral, ethical, and legal ethos for a fair, consensus of agreeable conduct that can be reliably enforced - a search for an innate digital sense of right and wrong. All online communities tend to agree that there should be baseline understanding of respect and accountability, although there are many approaches for how this understanding should enforced.

Illustratively, groups of entities, namely individuals, corporations, and nations, have recently charted quite different paths toward this unified goal for cyberspace. In many instances, online communities of individuals develop guidelines for conduct within their corresponding fields of interest. For example, in making a small, solitary pledge at Softwarethics.org,[1] individuals working within their professional occupations are in the process of actively developing codes of ethics for software and social engineering on the Internet. Other individuals are compelled toward online activism and even vigilantism to address many digital and physical world objectives, through the actions of such groups as Anonymous; Never

Again's "Tech Pledge" (http://neveragain.tech/), and with social engineering certification training programs for hackers.[2]

With the absence conduct in the digital realm that consistently advocates respect for basic human liberties as well as the peaceful, well-being of citizens in other nations, companies across the globe are taking a collective initiative toward creating online communities for promoting fundamental online norms while addressing a variety of issues. In response to the San Bernardino shooting, Apple and a community of companies, collectively defended their position that defends writing software code, even for screen-locks, as protected free speech and for the right to resist the orders of the government to assist in potentially incriminating a U.S. citizen in the name of counter-terrorism. (https://www.apple.com/newsroom/2016/03/03Amicus-Briefs-in-Support-of-Apple/). In light of the Russians interfering with US presidential elections, Facebook, Twitter, and a community of companies have actively and favorably responded in support of "electoral integrity", the democratic right of free and fair elections.[3] As a consequence of Volkswagen software engineers programming a hack around mandatory government auto carbon emissions testing, the S&P Dow Jones Indices, the London Stock Exchange, and RobecoSAM permanently removed, for social and ethical reasons, Volkswagen AG stock from the Sustainability Indices for globally responsible companies.[4]

From a sense of right and wrong embedded deep down within our shared human experience, all these participants no mater how big, small, or sophisticated that they may be ultimately know that many of the above acts just "feel wrong". The issue is how can that individual feeling for determining "wrongness" be applied to something as abstract and wild, at times, as the Internet? What behaviors would be acceptable to you, and how can these behaviors be fairly applied to activity on the Internet in terms of morality, ethics or laws? First, let us address the basic differences between law, ethics, and morality. Second, we see how various communities are struggling to apply various laws, ethics, and morality to the digital world at this time.

I. Morality, Laws & Ethics 101: The Basic Differences between Law, Ethics, & Morality

Ultimately, I believe that a reputation of trust is the currency that each individual actively gains while on the Internet and information or "data" is the fuel that drives one's perceived reputation. Formulaically, as expressed in terms of statistical machine learning, one's reputation is based on how likely others will have a "good feeling" that an individual will act favorably within accepted norms based on information of validated past events. As fake information is often purposely juxtaposed next to the truth, the Internet is a murky space to make a "good feeling" judgment from one's perceived prior behaviors. So how do we go about navigating or perhaps socially designing some basic level of accepted norm for trust as humans increasingly connect digitally and interact with one another on the Internet?

As an overview to the following discussion, consider that each of us have personal set of values of what innately feels right or wrong. In my opinion, this innateness tends to come from a user manual for all humanity. Therefrom, morals

tend to become a validated set of values when recognized by at least one other individual as fundamentally agreeable. Ethics tends to add some civic structure to some select moral values, when becoming a majority consensus of what particular behaviors within a community are agreeable and tolerable. Notably, ethics and laws are applied differently in practice – although laws can directly arise from ethical standards that a government decides to apply to preserve peace and order. Ethics are an active means for dynamic engagement within a community whereas laws tend to be static rules by which to govern effectively from. However, great social tensions happen within a nation when some ethical principles are incorrectly made law.

Let's first gain a deeper understanding about the concept of morality, stemming from the Latin word *mores* referring to "folkways"[5] (or values) of agreeable significance shared within a community.[6] Our personal liberty interests are a good example of morals and are also referred to as our fundamental personal rights[7] that, in part, comprise the Bill of Rights of the U.S. Constitution -- such as freedom of religion, right to peacefully assemble as well as the right to self determination and movement. Morals are self-understood from ones' personal opinions of right and wrong.[8] By extension, a moral community thus develops a core set of guiding principles that two or more individuals can tolerably agree.[9] Thus, moral values form the foundational layer of right and wrong that corresponding moral communities ultimately ascribe to.[10]

"Having a moral duty" is a popular phrase and suggests that morals occupy the same social space within our greater community just as our laws do.[11] In reality, as morals are a personal experience shared by individuals, morals tend not to be written or clearly defined as our laws, which are systematically written and codified by the government. However, individuals often mistakenly (and dangerously) confuse morals with the law when justice is called for. Many legal actions are taken to what is actually considered a moral wrong, such as many disputes between one's religious values and many other personal ideologies that are not collectively shared by the tolerant majority.[12] Further, to avoid the tragedies of genocide, most modern nations strive to be respectful of diverse moral interests to protect the interests of all citizens.[13]

"The wellspring of ethical change is personal morality"[14] as individuals expand their shared values toward a consensus that forms a broader community's sense of right and wrong. As illustrated by many groups of players within a single online multiplayer gamer community, the variety online sub-communities and their specific customs may be incomprehensively vast such that ethics dynamically strives to underscore those acceptable customary behaviors shared by the majority within the broader community for that particular game, such as the perennial Warcraft series.

Ethics derives from the Greek word *ethos* meaning "customary".[15] Ethical behaviors derive from a dominant moral community that is respectfully tolerant of other moral views in the minority. Ethical principles remain an open-ended narrative within a specific community as well as in a cluster of communities alike. In short, ethics is principally conveyed as shared moral values in action.[16] Unlike the law, ethics cannot easily be generalized into a defined set of written rules.

In the United States, our contemporary understanding of written guiding principles or codes of ethics were first modeled from the Belmont Report of 1978 advocating ethical treatment of human subjects in scientific research.[17] Influenced by the Belmont Report, many communities now tirelessly work to provide updated written codes as ethical behaviors change with time, especially within professional (workplace) communities such as software engineers' codes of ethics under the Institute of Electrical and Electronics Engineers - Computer Science (IEEE-CS) and the Association for Computing Machinery (ACM). Although a noble idea, many professional ethics codes *per se* are not enforced just as regional governments effectively enforce legal codes. Again, ethical principles represent the continuously evolving norms of action and written ethics codes often represent aspirational exemplifications for behavior within a community - as opposed to the laws governed by nation states.[18]

Justice or fairness is a subset of the concept of ethics and not the law. Specifically, a "just" decision is a fair decision to the extent that each individual within a community is treated equally in terms of what they need or deserve.[19] In other words, justice is a principle by which "we render to each what is due and treat like cases alike."[20] Modern democratic legislatures, the courts although to a lesser extent, and similar governing bodies decide what emergent ethical principles, including principles of justice, should be transformed to written law.[21]

Deriving from the Latin word *lex* meaning law or rule, the law creates and maintains an ordered society and ensures safety and welfare of such governed citizens.[22] As discussed above, most constitutionally democratic governments further ensure that their citizens are guaranteed fundamental personal rights.

In contrast to ethics and ethical principles of justice, the law fashions well-defined civic boundaries that are easily generalized, written, and reliably enforced.[23] To this end, the law is often exalted within its own formulaic written terminology that very much differs from our everyday use of language. Further, the law requires well-defined commitment of all those who choose to be within the community of citizens governed by that law as a principal manifestation of a sense of nationalism.[24]

As humanity is never perfect, law and justice very often collides as is manifested in the world's many civil wars and cultural wars, respectively. Laws may not be strictly enforced and give way to the unpredictably turbulent ethical battles, in the name of justice, between various moral communities underneath the canopy of a governed citizenry. Specifically, in the United States, the ethical sensibilities of justice to treat all equally and fairly have episodically intersected with the contemporary written law within the broader discourse of the nation. For example, consider the following equations: the U.S. Slave Codes were inherited from British slavery laws and enforced early on by many of the southern colonial states in United States history = legal but immoral thus unethical. After the written addition of the 13th, 14th, and 15th Amendments to the U.S. Constitution – the supreme law of the land, former slave states enact state and local "Jim Crow" laws to prevent African American and other populations from exercising their legally superior constitutional rights as citizens to vote for well over 100 years after the U.S. Civil war = illegal and immoral thus unethical – but yet wedges one community against another to the

present day. There are many other unspeakable examples within this blurry war of collisions, such as the mandatory sterilization of the mentally disabled and vulnerable ethnic minority populations in the 20th Century within the United States and its protectorates[25] to illegalization of marijuana by Federal Bureau of Narcotics Chief Harry Anslinger[26] and its subsequent legalization in the 21st century (http://fortune.com/2017/09/11/california-marijuana-drone-delivery-ban/). Society can remain conflicted in its legal and ethical codes but each individual's sense of moral fairness toward human rights tends to ultimately bubble-up and prevail as the eventual societal norm (- with much hopeful optimism!).

II. Got Lovn Feeln? A Survey of Getting Along in the Digital Wild West

With a better understanding of morality, ethics, and the law, let us now turn to what might work best in curtailing that growing feeling of "wrongness" on the Internet today as shared by a consensus of many communities of individuals, corporations, and governments.

A. Imparting Morals to Machines: 21st Century Morality Transplant Surgery -

In the field of artificial intelligence, there is a concept of supervised machine learning where you train a computer to statistically draw inferences based on past information when given a description or "label". In particular, to establish an initial point of reference to the learning process of a machine, humans typically provide labeled descriptions. Thereafter, for unseen instances, the machine draws on those labeled past experiences to guess the answer of greatest probability.

As such, Massachusetts Institute of Technology (MIT)'s Moral Machine project (http://moralmachine.mit.edu/) currently crowd sources human labeling in the context of asking human's to provide information to assist computers to innately decide on "what is the moral thing to do?" The project prompts thousands of humans for their moral value intuition so that each dilemma shown on the website is statistically synthesized as a moral value. Those labeled moral values will be "transplanted" to machines through algorithms as if given an innate sense of what an artificial intelligent being will personally value. I suppose a community of robots that share the same values would eventually form the first artificial moral community rooted in from the many human contributors to this project.

Similarly, the U.S. Government's Intelligence Advance Research Projects Activity (IARPA)'s CREATE project[27] applies traditional philosophy to artificial intelligence that assists humans to make enhanced, "human augmented" judgments to best mitigate the harms associated with human conflict. Collectively, a sourced crowd of thousands of people is submitting speech and debate arguments that assist machines to identify the good and bad parts of such arguments, even those arguments based on unreliable pretenses. Ultimately, an artificial intelligence can be developed for application in analytical toolkits to assist humans from making poorly biased judgments as well as promote better communication that yields enhanced human reasoning and conclusions.

B. Imparting Ethics to Coders -

Communities of like-valued coders are currently doing amazing things on the Internet to standardize their respective codes of ethics. Ethical Hacking certifications and professional codes for developers are the most predominant.

Many communities of hackers, computer scientists, software engineers, and digital security experts collectively agree that a coder's social ("ethical") responsibility is to make the public aware of software vulnerabilities; this is called "Responsible Disclosure". In support of the "responsible disclosure" movement, some tech companies will pay individuals who find such software vulnerabilities monetary rewards, aka "bug bounties".[28] Moreover, there are widely available ethical hacker certifications[29] as well as ethical hacking and social engineering initiatives. Such certifications are quite popular in the digital security industry.

Professional codes of ethics have been introduced to software programming in the 1940s by MIT professor Norbert Wiener.[30] Since then, the predominantly recited ethics codes in the digital field are the above mentioned Software Engineering Code of Ethics and Professional Practice, IEEE (Principle 1: Public 1.00 *et seq.*): [31]

" • ~~**Approve software only if they have a well-founded belief it is safe and meets specifications.**~~
- ~~**Accept full responsibility for their own work.**~~
- **Not knowingly use software that is obtained or retained either illegally or unethically.** – IP (copyrights)
- **Identify, define, and address ethical, economic, cultural, legal and environmental issues related to work projects.** – business & employment law, environmental law
- **Ensure that specifications for software on which they work satisfy the users' requirements and they have the appropriate approvals.** – contract law
- **Ensure adequate testing, debugging, and review of software. –** contract law
- **Not engage in deceptive financial practices such as bribery, double billing, or other improper financial practices.** – criminal law ",

and the ACM Code of Ethics & Professional Conduct, § 1 (General Moral Imperatives): [32]

" • ~~**Contribute to society and human well-being**~~.
- **Avoid harm to others.** – criminal law (& civil penalties in California)
- ~~**Be honest and trustworthy.**~~
- **Give proper credit for intellectual property.** IP law
- **Respect the privacy of others.** – due process, US Constitution
- **Honor confidentiality.** – IP law (trade secrets) ".

Notably, I have crossed out the above provisions of ethical code with no foundation in U.S. Law to find that the majority of items in these two ethical codes for professional conduct are generalizations that are duplicated from existing law. In short, the majority of the IEEE and ACM ethical codes also may be separately enforced in both the criminal and civil courts by at least the underlying U.S. laws.

C. Imparting laws (and order?) to the Digital Wild West -

Hacker Laws:

Currently, the Computer Fraud and Abuse Act "CFAA" (18 U.S.C. § 1030 *et seq*.) is the most infamous federal law applied to Internet hacking. The CFAA prohibits accessing or conspiring to access a computer without authorization and leads the way to individuals being prosecuted in both criminal and civil court systems. In practice, the CFAA is applicable to computers greater than ten (> 10)

that are related to the federal government, across interstate boundaries, and for losses exceeding USD$5,000.

Many other U.S. states apply much of the same provisions of the federal CFAA law but in an equivalent state law context. California, for example, subsequently enacted the Computer Data Access And Fraud Act, Cal. Pen. Code, §502 ("CDAFA") that does not require unauthorized breaking into a computer for access as with the federal CFAA but merely requires logging into a database with a valid password - but without permission - to subsequently take data. Accordingly, California's CDAFA has expanded the prosecutorial reach of the federal CFAA where common hacking techniques, such as webscraping, can arguably be punished under the state's CDAFA statute after receiving notice, often the click-thru-notice from a website's Terms & Conditions clause. At this present time, the U.S. Supreme Court is (Oct. 10, 2017) denied review of this very matter in the case *Facebook, Inc. v. Power Ventures, Inc.,* No. C-08-05780 (N.D. Cal. July 20, 2010) where California might have overstepped its prosecutorial authority.

Whistleblower Statutes:

Federal whistleblower statues (5 U.S.C. §1201 *et seq.*) theoretically shield and potentially, financially reward employees that witness some illegal activity during the course of their employment. Whistleblower statutes generally refer to federal protections afforded to government workers and some contractors in areas typically designated by the statute that include digital intelligence (arguably: such as contractor Edward Snowden), military (arguably: U.S. Army veteran Chelsea Manning through the whistleblower repository, WikiLeaks), securities and banking regulation, and labor and federal employment matters among others. Employees of private companies and non-profits may also be eligible for federal whistleblower protection under the law if their work is related to Environmental, Occupational, and Health reporting. There are some states that also enact their own state whistleblower statutes based on the federal law as a guideline.

Whistleblowing is another example of an unsettled cultural war where opposing ethical communities struggle to define what a whistleblower actually is in the digital realm. The overarching whistleblower laws that give protection to just one community strictly falling under the static written law but controversially denies other potential whistleblowers from those legal benefits further conflate this murky confusion. Again, the static whistleblower statues attempt to cover a dynamically unsettled issue between ethical communities at this time. Moreover, from my personal observations, employees actually attempting to evoke whistleblower statutes have found the process disappointingly ineffective, as the U.S. Department of Justice will only accept quick, easy "open-and-shut" cases with little evidentiary holes. However, if you are professionally active in the digital security and intelligence communities, these whistleblower statutes are helpful to keep in mind as they were initially enacted by our legislatures to help those who come forward with the truth, which along with justice must be at the heart of an online democratic society.

FOO ALARM FIRE – SOUNDING THE ALARM:

The law may be regarded as a formal governmental confirmation of what is already generally acceptable by the community with the largest ethical consensus. Given new social situations such as the Internet, the path from ethical principles to laws is fluid, often indirect, and unsettled. Some codes of conduct may be arguably unethical but legally valid and vice versa. In my opinion, justice is often obtained at the junction of being both clearly legal and ethical and with minimal murkiness among most individuals. It should be said that the human narrative continues to include those who struggle for justice.

As professional, individual, and business communities take positive steps to apply codes of digital ethical conduct or become signatories to online petitions, such ethical actions may not be entirely enforceable under the vast, expansive skies and natural landscapes of the lawless Digital Wild West. In my humble opinion, just as many flags where set on the Antarctic and Lunar territories under respective Treaties, I believe to provide effective enforcement a multinational legal agreement, such as an Accord (enacted by the U.S. executive branch) or, better, a treaty (enacted by U.S. Congress) would need to be in place and coordinated by some international body like the United Nations. As a smaller scale exemplary roadmap, to combat recurring digital data breeches by corporations, similar global efforts have been recently made to ensure commercial regulatory compliance with the European Union's General Data Protection Regulation (GDPR – Regulations (EU) 2016/672) for privacy and protection of EU citizens' personal data.

In short, with the concerted approval from our U.S. representative government, a body of international law advocating a *Digital Bill of Rights* for all users remains a valid option. Our journey should begin with each of us as individual citizens and tech companies commanding our U.S. Congress to take legal action based on growing shared ethical concerns for the Internet, and directly contacting our representative through the help of the Electronic Frontier Foundation: (https://democracy.io/#!/).