Catastrophic Failures, by Data Design (*)
by Kevin Loney, April 2019

When the shoe factory's old boiler exploded, it turned into a rocket — tearing upward through four wooden floors and destroying their support beams before crashing through the roof and landing down the street. The old boiler had been pressed into service that day during maintenance on the primary boiler system that heated the Grover Shoe Factory in Brockton, Massachusetts. With its structural supports destroyed by the blast, the building collapsed - floors pancaked onto each other, walls caved in, burning coals left behind by the boiler lit fires, and solvents fueled them. By the end of the day two nearby factories, seven residences and a tavern had burned along with the shoe factory, with the death toll reaching 58 and another 150 injured. The shoe factory became a scar on the ground in a single day, March 20, 1905.

Boiler explosions weren't rare a century ago — steamboat boiler explosions in the 1800s claimed the lives of many, including over a thousand paroled Union soldiers in a single incident in 1865. But with the Grover Shoe Factory deaths in Brockton and a fatal December 1906 factory boiler explosion in Lynn, Massachusetts, the state government created a board to write boiler safety rules. The board's standards for boiler design, manufacture, and testing became state law in 1909. In writing their standards, the board leaned heavily on the Boiler and Pressure Vessel Code published in 1884 by ASME, the American Society of Mechanical Engineers. The resulting boiler testing code was written into the regulations of most states, territories, and provinces throughout the Americas, and what had been common disasters became rare anomalies. (*see note 1*)

The founders of ASME had started their work together because catastrophic failures in boiler design, maintenance, and use had direct impact on human lives; the engineers took it upon themselves to create a professional society and solve the core problems. At the heart of their professional society is a code of ethics that calls for the advancement of human welfare written directly into its constitution. Their approach to problem-solving and their commitment to the public would begin from a common set of principles. Given those principles as engineering project requirements, the outcomes of any development effort would be expected to be aligned to that code of ethics and any failures and their impact would be minimized.

Modern day data-related failures have a titanic breadth in terms of the number of people impacted by each data breach reported. In light of repeated catastrophic failures in the data field, legislators are stepping in to drive principles and regulations into businesses' handling of data where the businesses — and the engineers they employ — failed to adopt such principles as an industry. The California Consumer Protection Act (CCPA) is the most recent legislation, and its introduction explicitly calls out a 2016 case as part of its motivation:

```
(g) In March 2018, it came to light that tens of millions of people had their
personal data misused by a data mining firm called Cambridge Analytica. A
series of congressional hearings highlighted that our personal information
may be vulnerable to misuse when shared on the Internet. As a result, our
desire for privacy controls and transparency in data practices is heightened.
```

Given publicly disclosed data failures, the CCPA starts with a set of foundational principles and from those principles develops regulations and aligned standards. Like the boiler safety codes, the CCPA regulations enforce principles that advocates argue should have been part of the

design process all along. In the case of data, international guidelines for ownership and privacy of an individual's data have been published since at least 2010. These privacy guidelines direct the handling of privacy data and the design of applications that access it.

Design is often an exercise in weighing the costs and benefits of alternatives. It's critical for the architects and engineers reviewing designs to use common accepted strategic principles when evaluating those alternatives so risks are properly quantified and understood. How are the catastrophic risks (such as data breaches) weighed against tactical risks (including project delays impacting quarterly earnings) and strategic benefits? And who is representing the customer's data privacy rights in those discussions?

For design requirements concerning data privacy, the Privacy by Design (PbD) principles issued by the Canadian Information & Privacy Commissioner in 2009 and 2010 provide foundational guidance. The PbD principles are a core part of the GDPR (General Data Protection Regulation) implemented by the European Union in 2018 and to a less stringent extent the CCPA. PbD's guidance, for example, is the basis for the GDPR opt-in approach to data collection consent. The PbD guidance requires the user actively consent to the collection of data (opt-in):

Consent – The individual's free and specific consent is required for the collection, use or disclosure of personal information, except where otherwise permitted by law. The greater the sensitivity of the data, the clearer and more specific the quality of the consent required. Consent may be withdrawn at a later date.

while clause 32 (and later articles) of GDPR also requires a proactive consent from the user with an explicit assumption of non-consent (highlighted in bold), again requiring the consent be freely given, specific, and informed:

32) Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. This could include ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. **Silence, pre-ticked boxes or inactivity should not therefore constitute consent.** Consent should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them. If the data subject's consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided.

When its guidance was first published, PbD was critized for being vague and unwieldy, and the full PbD approach was not widely implemented as a set of business requirements. A decade later, as a remedy to catastrophic failures in the marketplace, PbD is being instituted via regulation to protect the privacy rights of individuals. For example, PbD's "Access" requirement is a broader version of the CCPA data disclosure/deletion sections:

Access – Individuals shall be provided access to their personal information and informed of its uses and disclosures. Individuals shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.

Given PbD's guidance, and with a varied set of regulations (GDPR, HIPAA for health care data, now CCPA for California customers, and potentially additional regulations from other states), what is a reasonable core set of requirements to use when designing a system? How can we work backwards from the legislation to the core principles and guidelines they are built on so the systems we create can be compliant for future use while the regulations are changing?

At their core, each set of regulations is enforcing a code of ethical conduct concerning the use of specific types of data. The mechanical engineering society discussed earlier, ASME, has a code of ethics centered around three Fundamental Principles:
(*see note 2*)

Code of Ethics
ASME requires ethical practice by each of its members and has adopted the following Code of Ethics of Engineers as referenced in the ASME Constitution, Article C2.1.1.

The Fundamental Principles
Engineers uphold and advance the integrity, honor and dignity of the engineering profession by:
1. Using their knowledge and skill for the enhancement of human welfare;
2. Being honest and impartial, and serving with fidelity the public, their employers and clients; and
3. Striving to increase the competence and prestige of the engineering profession

These principles aren't optional for a career as a professional mechanical engineer; they are fundamental to the profession. When an engineer graduates college, she may take the Fundamentals of Engineering exam; after successfully passing that exam, she may work under a licensed Professional Engineer (PE) for four or more years. After that work completes, she can go through the process of testing and certification to become a PE in her state. As noted by the National Society of Professional Engineers (NSPE), "only a licensed PE can prepare, sign, seal, and submit engineering plans and drawings to a public authority for approval, or seal engineering work for public and private clients." The same licensure process does *not* exist in software systems design, so it is critical that those in the software and data professions follow a shared set of principles. And as noted in ASME's second principle above, those that are served are, in order: the public, their employers, and then their clients.

The CCPA offers a striking example of how bad things get in the absence of an enforced ethic. The authors of the legislation felt it necessary to call out places where the data industry failed to behave appropriately. The original version of the bill documenting the issues that drove the creation of this legislation is worth reading to instill where the scorecard stands concerning perceived service to the public good, or the lack thereof:

SEC. 2. The Legislature finds and declares that:
(a) In 1972, California voters amended the California Constitution to include the right of privacy among the "inalienable" rights of all people. The amendment established a legal and enforceable right of privacy for every Californian. Fundamental to this right of privacy is the ability of

individuals to control the use, including the sale, of their personal information.

(b) Since California voters approved the right of privacy, the California Legislature has adopted specific mechanisms to safeguard Californians' privacy, including the Online Privacy Protection Act, the Privacy Rights for California Minors in the Digital World Act, and Shine the Light, a California law intended to give Californians the 'who, what, where, and when' of how businesses handle consumers' personal information.

(c) At the same time, California is one of the world's leaders in the development of new technologies and related industries. Yet the proliferation of personal information has limited Californians' ability to properly protect and safeguard their privacy. It is almost impossible to apply for a job, raise a child, drive a car, or make an appointment without sharing personal information.

(d) As the role of technology and data in the every daily lives of consumers increases, there is an increase in the amount of personal information shared by consumers with businesses. California law has not kept pace with these developments and the personal privacy implications surrounding the collection, use, and protection of personal information.

(e) Many businesses collect personal information from California consumers. They may know where a consumer lives and how many children a consumer has, how fast a consumer drives, a consumer's personality, sleep habits, biometric and health information, financial information, precise geolocation information, and social networks, to name a few categories.

(f) The unauthorized disclosure of personal information and the loss of privacy can have devastating effects for individuals, ranging from financial fraud, identity theft, and unnecessary costs to personal time and finances, to destruction of property, harassment, reputational damage, emotional stress, and even potential physical harm.

(g) In March 2018, it came to light that tens of millions of people had their personal data misused by a data mining firm called Cambridge Analytica. A series of congressional hearings highlighted that our personal information may be vulnerable to misuse when shared on the Internet. As a result, our desire for privacy controls and transparency in data practices is heightened.

(h) People desire privacy and more control over their information. California consumers should be able to exercise control over their personal information, and they want to be certain that there are safeguards against misuse of their personal information. It is possible for businesses both to respect consumers' privacy and provide a high level transparency to their business practices.

The CCPA isn't solely about data breaches. At its core, CCPA is about reclaiming each person's data as private information — private property belonging to the individual. A person's transactions create a digital fingerprint — an identifiable set of data that uniquely identify the individual, and that data fingerprint belongs to the person just as surely as the person's DNA sequence or fingerprints do. The company storing the data as part of a transaction is the steward of the data, and stewardship has always had benefits — and now has a clearly defined set of responsibilities and costs per customer. (*see note 3*)

Every company with customers in California is now under a deadline for compliance with CCPA. Had PbD principles been enforced in requirements for the past decade, the impact of CCPA would be minimal; the regulations would merely be providing guidance on the artifacts needed to show auditors the businesses were acting properly concerning their customers. But the data factory explosions continued, and the California legislature acted; there is no reason to expect it will be the last state to do so.

California's state constitution, referenced by the CCPA, documents the core set of rights from which these policies derive. The very first article in the state constitution sets it out boldly:

```
ARTICLE 1   DECLARATION OF RIGHTS.

SECTION 1.  All people are by nature free and independent and have
inalienable rights.  Among these are enjoying and defending life and
liberty, acquiring, possessing, and protecting property, and pursuing
and obtaining safety, happiness, and privacy.
```

When your personal identity is considered your own property and a core component of your privacy, and the data used to verify your identity is part of that property, it follows that no company you interact with could ever be the sole owner of your personal data. Being a consumer does not involve waiving any right guaranteed in your state's constitution; you still own the rights to your digital identity. For companies that have not taken this approach previously, abiding by this principle may involve doing business a different way, starting by working with architects and engineers who operate from the same set of ethical principles businesses publish as their own: that we value our customers and we treat them the best way possible. That we implement data privacy by design, not as an afterthought.

When evaluating a failing system, the first assumption should always be that the system is doing exactly what it has been told to do; this approach holds whether the system is a data warehouse data loader or an application development team. We design and build manufacturing systems to deliver outcomes, and we organize and staff software development teams to deliver code and products. As leaders, we have a unique level of insight and influence over the behaviors of the developers and engineers who design and deliver applications and data products. In handling a customer's personal data, a development team's processes from design through operational support need to be guided by a shared set of ethical principles in order to meet the expectations of the customers, the company, and the public. As leaders, we need to provide direction in the area of ethics and principles — as core components of our relationships with our customers and as part of our guidance for the people who build the systems that serve our customers and process the data they entrust to our stewardship.

-Kevin Loney, April 2019.

ENDNOTES

(*)
As an engineering student I was privileged to take the course "Catastrophic Failures" taught by Dr. McIntyre Louthan, PhD, an expert in the field of failure analysis and an incredible instructor.

See youtube for recordings of his talk called “How Things Fall Apart”
(https://youtu.be/95_-G9Bt0fk) and other works of his.

(1)
The boiler safety issues were so prevalent and devastating that this standard was the first ever issued by ASME. The second, in 1887, was “Standard for the Diameter and Overall Dimensions of Pipe and Its Threaded Ends”, which led to the standardization of pipe dimensions and thread sizes we take for granted.
There are still pressure vessel explosions. An explosion in 2009 threw an 8000-lb vessel fragment through a wall while another piece of steel from the facility was thrown 650 feet, killing a nearby driver. The vessel had not been properly inspected or tested for stress corrosion cracking. See
https://www.csb.gov/csb-releases-safety-video-on-2009-fatal-blast-at-ndk-crystal-animation-depicts-stress-corrosion-cracking-vessels-were-not-inspected-or-tested/

(2)
For similar codes from other engineering societies such as ASCE and IEEE, see
https://www.asce.org/code-of-ethics/
https://www.ieee.org/about/corporate/governance/p7-8.html

(3)
The right to privacy regarding personal data (in CCPA, referenced throughout section 2) is not synonymous with a right to anonymity. When entering into a financial transaction you give up your right to be anonymous in order to be a law-abiding member of society: for fraud agents to verify you are not money laundering, for treasury departments to verify you are not violating sanctions, and for revenue services to verify you are paying taxes on income or sales.

All embedded links:

https://youtu.be/95_-G9Bt0fk

https://chroniclingamerica.loc.gov/lccn/sn88085421/1905-03-20/ed-1/seq-1/#date1=03%2F20%2F1905&sort=date&rows=20&searchType=advanced&language=&sequence=0&index=1&words=Brockton+EXPLOSION+explosion+factories+factory+FACTORY+fire+Fire+FIRE+Grover+Shoe&proxdistance=5&date2=05%2F01%2F1905&ortext=&proxtext=&phrasetext=&andtext=grover+shoe+factory%2C+explosion%2C+fire%2C+Brockton&dateFilterType=range&page=1

https://en.wikipedia.org/wiki/Sultana_(steamboat)

https://www.asme.org/engineering-topics/articles/boilers/the-history-of-asmes-boiler-and-pressure

https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375

https://iapp.org/resources/article/privacy-by-design-the-7-foundational-principles/

https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX%3A32016R0679

https://community.asme.org/colorado_section/w/wiki/8080.code-of-ethics.aspx

https://www.nspe.org/resources/licensure/why-get-licensed

https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375

https://law.justia.com/constitution/california/article_1.html