Advances in Reliability, Safety and Security, Part 7 – Kołowrocki, Magryta-Mut (eds.)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Application Of Differentially Private Algorithms To Reliability Data Evaluation In Industrial Internet Of Things Context

Andreas Joanni

Siemens Technology, Siemens AG, Munich, Germany

Abstract

Suitable methods and computing infrastructures for reliability data collection and evaluation are relevant aspects for the Industrial Internet of Things (IIoT). However, reliability data such as failure times or down times is often considered confidential by companies that contribute to the IIoT. This may lead to less amount of available data and, therefore, increased difficulties in taking advantage of it. The application of differentially private algorithms to evaluate and publish reliability data in the context of IIoT is, therefore, useful by ensuring that potentially sensitive reliability data is sufficiently protected by adding only as much randomness to the data or published results as necessary, such that useful evaluations are still possible. Differential privacy is a rather recent approach that has a rigorous theoretical framework to quantify the required degree of randomness to limit the greatest possible information gain by a hypothetical attacker, while offering several advantages as detailed below. This paper gives a brief introduction to the concept of differential privacy and outlines its application to reliability data evaluation in the context of Industrial IoT, including discussion of the specific benefits in this context as well as possible drawbacks. Two numerical examples are given.

Keywords: reliability data, statistical methods, differential privacy, internet of things, asset management

1. Introduction

Suitable methods and computing infrastructures for data collection and evaluation are relevant aspects for the Industrial Internet of Things (IIoT), see e.g. (Compare et al., 2019). Frequently, the collected and evaluated data are related to reliability in some sense: it could be sensor information such as temperature & vibration signals for tracking the health state of monitored components (Hong et al., 2018), or it could be information directly on failure times or down times of components. However, reliability data such as failure times or down times is often considered confidential by companies or other parties contributing to the IIoT, in the sense that they are the owners or operators of the monitored components. There are various reasons for that, such as (i) sensor information like temperature & vibration signals may reveal that components were operated outside of the specifications stated by the manufacturer of the components, so owners or operators of the monitored components fear that they might lose warranty rights; and (ii) failure times or down times of components may, for instance, reveal information on the cost structure of a manufacturer, or details on the operating and environmental conditions, or information on the plant utilization.

This may cause companies to share less available reliability data and, therefore, this leads to increased difficulties in taking advantage of it (Lazarova-Molnar and Mohamed, 2019). For instance, these kinds of reliability data could be analyzed and used for optimizing the design of the system that comprises the monitored components (such as production system where it is inferred from the reliability data that a higher level of redundancy is needed to reach given availability targets); or, the analyzed data could be used for improvement of the logistics of maintenance operations, such as setting a just-in-time logistic support that reduces the required number of stored spares. In addition, the data could be used for learning fault models, including causality among faults and failures, which may be followed by advanced simulation and data analytics. The results of simulation and data analytics may be used for decision support on improved system configuration and generation of preventive maintenance schedules for increased reliability of the system (Lazarova-Molnar and Mohamed, 2019).

Therefore, the application of differentially private algorithms to evaluate and publish reliability data in the context of IIoT is useful by ensuring that potentially sensitive reliability data is sufficiently protected. This is done by adding only as much randomness to the data as necessary, while useful evaluations are still possible, and while offering several advantages as detailed below.

The paper is organized as follows. The next Section 2 briefly introduces some concepts of differential privacy to the extent needed for understanding the remainder of the paper. This is, of course, not in too much detail and far from being exhaustive. Sections 3 and 4 outline its application to reliability data evaluation in the context of Industrial IoT, including the specific benefits in this context and possible drawbacks. Two numerical examples are given for a simple mean value estimation, and estimation of parameters for a Weibull distribution. Finally, the paper concludes with a summary and outlook.

2. Some concepts of differential privacy

Traditional methods for publishing sensitive data while protecting it from privacy-oriented attacks include data anonymization techniques, which may be susceptible to reconstruction attacks especially with more and more computing power available (Wood et al., 2020). Others, like cryptographic approaches, have other drawbacks in an Industrial IoT context such as increased computational complexities and computational overhead, see e.g. (Husnoo et al., 2021). Differential privacy is a rather recent, state-of-the-art approach that ensures that sensitive information in a statistical database is sufficiently protected by adding only as much randomness to the data or published results as necessary, such that useful evaluations are still possible, but an attacker cannot infer any information about a particular record in the database with high confidence. Hence, a processing (or evaluation) step satisfies differential privacy if its output is relatively insensitive to any change of a single record in the original database. This is done by adding randomness (noise) to the processing result or to the individual reliability data, depending on the chosen model (central or local). For an accessible description and details on the differential privacy concept, see e.g. (Wood et al., 2020) and (Desfontaines, 2023). The suitable degree of randomness depends on the processing step and the specified limit on the greatest possible information gain by the attacker. The interesting feature of the differential privacy concept is that this information gain can be quantified based on a rigorous theoretical framework.

One of its advantages is that there is no need for attack modeling because it does not matter what the attacker knows about the data. As a worst-case assumption, it is supposed that all the sensitive information in the database is known to the attacker except the data of one individual record that is the target of the attack. Further, there is also a rigorous quantification of the information gain if the results of several processing steps are published and known to the attacker (so-called composed algorithms).

A basic definition of differential privacy, the so-called ε-DP (short for differential privacy) introduced by (Dwork et al., 2006), is as follows. Suppose that an attacker tries to distinguish between two adjacent databases D_1 and D_2 that differ by only one record (i.e., the record could have different data, or it could exist in one database but not in the other). Then, if an algorithm A is ε-DP, then $A(D_1)$ and $A(D_2)$ will return output O with similar probability:

$$P[A(D_1) = O] \leq e^{\varepsilon} \cdot P[A(D_2) = O] \tag{1}$$

or

$$e^{-\varepsilon} \cdot P[A(D_2) = O] \leq P[A(D_1) = O] \leq e^{\varepsilon} \cdot P[A(D_2) = O] \tag{2}$$

because D_1 and D_2 are interchangeable.

The parameter ε is the so-called privacy budget. In case of $\varepsilon = 0$, the probabilities are identical, which means that the degree of added randomness is so large that the two databases are completely indistinguishable. Hence, to be meaningful, we always have $\varepsilon > 0$. The added noise (or randomness) may be sampled from a variety of distributions, such as the Laplace or Gaussian distribution. For the required amount of randomness in case of the Laplace distribution, which will be demonstrated by means of the numerical example in the next section, the so-called sensitivity Δf of the function f in the following form is important:

$$\Delta f = \max_{D_1,\, D_2} \|f(D_1) - f(D_2)\|_1 \,. \tag{3}$$

As before, the two databases D_1 and D_2 differ by only one record. The function f stands for the processing (or evaluation) step before adding randomness, for instance the calculation of the mean value or other statistics from a database. Hence, the result of Eq. (3) is the maximum difference in the output of f when applied to any two adjacent databases.

3. Differentially private reliability data evaluation

There are many examples of processing of records from a statistical database in the field of reliability. For instance, the records in the database could be failure times collected from several devices in the field, and the processing result could be the maximum-likelihood estimator of the mean operating time between failures (MTBF or MOTBF, see 192-05-13 in IEC 60050-192:2015), or corresponding upper and lower confidence bounds. An attacker might wonder if the database contains the failure time of an individual device, or what the failure time would be if it is contained in the database. After performing the usual calculation (dividing the number of failures by the total time on test), one would have to add a sufficient degree of randomness to the result.

There is one important thing to observe in this case, however. Since an individual failure time could, in principle, be arbitrarily large, an infinitely large degree of randomness would have to be added to the database to protect an individual record. Therefore, to limit the degree of randomness to be added, the values of the individual records must be restricted to a specified allowable range, which depends on the specific application. This is called clamping and, for reliability data such as failure times, this is not too much of a sacrifice, because outliers are usually excluded from a statistical analysis anyway.

As a concrete example, consider a database of complete uncensored samples of $n = 500$ failure times from a Weibull distribution with scale parameter of 24 and a shape parameter of 2. The regular maximum likelihood estimator of the mean time between failures equals to 21.6. Figure 1 shows two density plots of 4096 published values, each after applying a differentially private calculation of the mean time between failures using the Laplace mechanism.

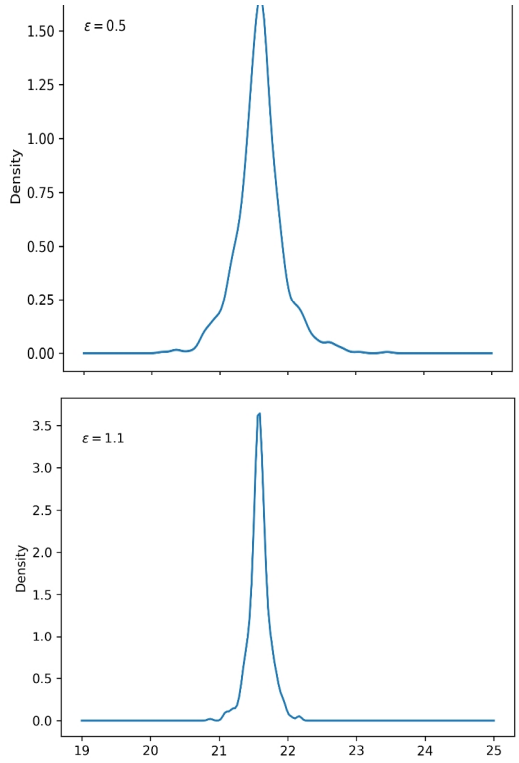

Fig. 1. Density plot of 4096 published values from a differentially private calculation of the MTBF ($\varepsilon = 0.5$ in the upper plot, $\varepsilon = 1.1$ in the lower plot) using the Laplace mechanism.

The privacy budget in the upper plot equals to $\varepsilon = 0.5$, and in the lower plot $\varepsilon = 1.1$, and the individual failure times are clamped at a value of $M = 60$. It can be seen by comparing the two plots in Figure 1 that a smaller value for the privacy budget ε leads more spread in the values. That is, it is harder for a hypothetical attacker to infer the value of one failure time in the database with confidence, even if he knew all other values in the database.

Nevertheless, in both cases the published values are in a range that makes them useful. In the present example, applying the Laplace mechanism that satisfies ε-DP means that a random variable with probability density function

$$f(x|\mu, b) = \frac{1}{2b} e^{-\left(\frac{|x-\mu|}{b}\right)} \tag{4}$$

with location parameter $\mu = 0$ and scale parameter $b = M/(n\,\varepsilon)$ has been added to the mean value calculated using the usual formula. The scale parameter is determined by the sensitivity according to Eq. (3), which is a measure of the maximum difference in the output when (in this case) the mean value is calculated for two adjacent databases that differ by only one record. This, in turn, is influenced by the clamping value M, because the maximum difference is equal to M/n for the mean value function.

To continue with the example and to illustrate the meaning of the parameter ε, consider Figure 2 showing the densities from the Laplace distribution added to the mean time between failures calculated from the database above (mean value equal to 21.6), as well as from the database with one additional record included, resulting in a mean value equal to $21.6 + 60/500 = 21.72$. Here, the maximum failure time after clamping is assumed.

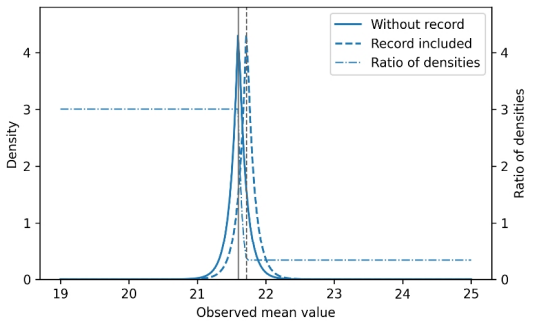

Fig. 2. Comparison of densities of observed mean values for database with and without an additional record, based on a mean time between failures equals to 21.6 ($\varepsilon = 1.1$).

Hence, for any observed (or published) value for the mean time between failures, resulting from a differentially private algorithm using $\varepsilon = 1.1$, it is hard to tell for a hypothetical attacker if the additional record is included in the database or not, even if they knew all other records in the database. In other words, differential privacy is totally independent of the prior knowledge of the attacker. This is evident here because the densities in both cases are not too different. More precisely, Figure 2 shows that the ratio of the two densities varies between about $1/3$ and 3, which is exactly the requirement for ε-DP as stated by Eqs. (1) or (2) since $e^{1.1} \approx 3$.

4. Application to an Industrial IoT context

Some of the more widely adopted IIoT infrastructure variants are described in the following (see also Husnoo et al., 2021) and depicted in Figure 3. The commonality is that sensing and actuating are carried out at the lowest layer, also referred to as the device layer. The next layer up, the edge layer, provides the connectivity between the devices and the application or cloud layer. Sometimes, this connectivity is provided by semi-capable devices behaving as gateways, collecting data from the sensors, and transmitting them to the cloud layer. The cloud layer is responsible for everything related to IIoT applications, such as device and system management, for data processing and for storage.

In contrast to the basic IIoT connectivity provided by IoT gateways, which often collect and transmit data from legacy devices, the characteristic of edge computing is that data processing occurs directly on the smart IoT device or a smart edge device physically close to the field device level. This allows for real-time processing and decision-making at the source of data, reducing latency and bandwidth usage. As another variant, fog computing serves as a distinct layer between edge layer and cloud layer with intermediate level of computing power, and implies distribution of the communication, computation, storage resources, and services on or close to the field devices.

Keeping these IIoT infrastructure variants in mind, it becomes clear that processing reliability data in way that satisfies differential privacy is, naturally, tied to devices or levels where data processing can be carried out. For instance, the failure data could have been collected by a smart edge device, by which also the differentially private processing as described in the previous section is performed. This approach is the so-called central model (as

opposed to the local model). In principle, both the central model and the local model of differential privacy are applicable for reliability data evaluation in the context Industrial IoT applications. Mixtures of the local and global model also exist, see (Husnoo et al., 2021).

In the central model, a central processing device, such as a smart edge device (see Figure 3), has access to the actual and raw reliability data. The appropriate degree of randomness, as prescribed by the chosen differentially private algorithm, is only applied once after the evaluation step, and the result can then be published. This central model, therefore, has the advantage of higher accuracy of the result because a lesser degree of randomness is required to achieve sufficient protection. The drawback is that the processing device must be trustworthy, because all the real data is collected in one place. That means that the central processing device must be sufficiently protected from hacking attacks and, in addition, the party that has access to the device must be trustworthy. Alternatively, the processing could be performed by an IIoT application for the fog computing variant where, again, sufficient trustworthiness must be ensured.

For the local model of differential privacy, the central processing device, such as a smart edge device or an IIoT application, does not have access to the real data. Instead, the appropriate degree of randomness as prescribed by the differentially private algorithm is applied by every individual data source of the reliability data, such as a smart IoT device, an IoT gateway, or any other lower or intermediate level as in the case of fog computing (see Figure 3), before sending the data to the downstream processing device or IIoT application for evaluation. This way, the processing device does not have to be trustworthy and there is no need for protection against hacking attacks. The drawback is that the total degree of randomness is larger (although still only as large as required as the basic premise of differential privacy), because each individual source must add randomness to the data. As a consequence, the result of the processing step is less accurate.

As a practical prerequisite for a wider application of differential privacy in an IIoT context, the availability of suitable state-of-the-art libraries and tools must be ensured. Fortunately, there are more and more implementations of differentially private algorithms published by corporations such as Google and IBM as open-source projects, and also the number of practical tools offered by smaller companies is growing. In addition, experts with knowledge about differential privacy as well as domain knowledge are needed who decide if, in which form, and where differentially private processing can be integrated into realistic IIoT setups.

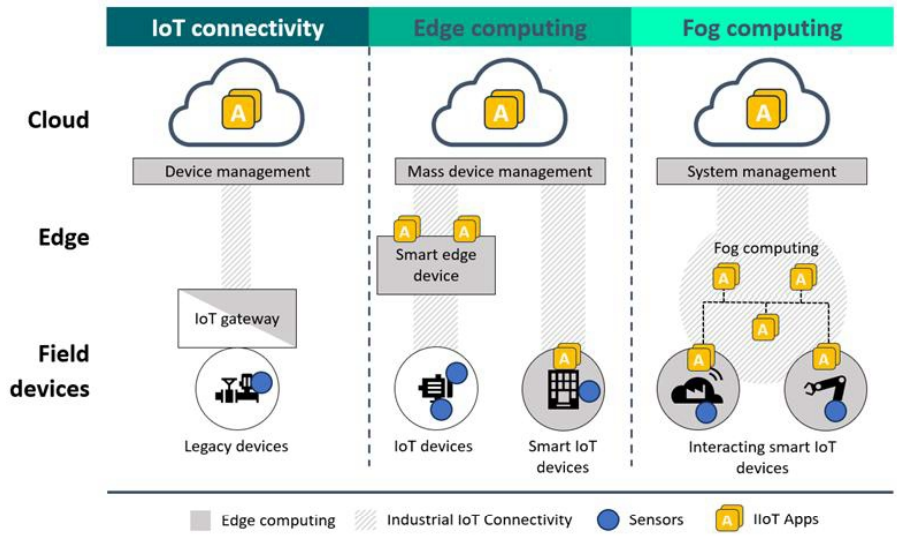

Fig. 3. Different variants of computing infrastructure in an Industrial IoT context.

As already mentioned above, there are other approaches such as homomorphic encryption that are useful for protecting sensitive data in an IIoT context. For the latter approach, this means that the data is encrypted and transmitted before performing the evaluation on a processing device. The algorithms for the evaluation must be adapted to work on encrypted data. Then, result is transmitted again in encrypted form and can decrypted it to obtain the result. This homomorphic encryption technique, however, typically has an extremely large performance overhead.

5. Differentially private estimation of parameters for a Weibull distribution

In order to present a slightly more elaborate example of a differentially private processing of reliability data, the estimation of the parameters of a Weibull distribution is considered based on a statistical database of complete failure times (i.e., without censoring). The parameters can be determined by means of linear regression, which becomes clear if the cumulative distribution function $F(\cdot)$ of the Weibull distribution with scale parameter α and shape parameter β

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \tag{5}$$

is reformulated as

$$\ln\left(-\ln\left(1 - F(t)\right)\right) = \beta \ln t - \beta \ln \alpha \,. \tag{6}$$

Eq. (6) can be expressed as the linear equation

$$y = \beta t' + a \tag{7}$$

where $y = \ln\left(-\ln\left(1 - F(t)\right)\right)$, $t' = \ln t$ and $a = \beta \ln \alpha$. The $F(t)$ are obtained from the ordered failure times and can be approximated by

$$p_i = \frac{i - 0.5}{n} \tag{8}$$

with the i-th largest failure time in a sample of size n, see (Meeker et al., 2022).

Figure 4 below shows the results based on $n = 1000$ samples from a Weibull distribution with a scale parameter of 24 and a shape parameter of 2. Two fitted lines are also shown; a solid line using ordinary linear regression, and a dotted line using a differentially private linear regression algorithm with $\varepsilon = 1.0$. Expressed in terms of the values for the shape and scale parameters, which can be obtained from the intercept and slope of the fitted lines, for ordinary linear regression they are 1.99 and 23.8 respectively (close to the true parameters), whereas for differentially private linear regression they amount to 2.69 and 22.5, respectively.

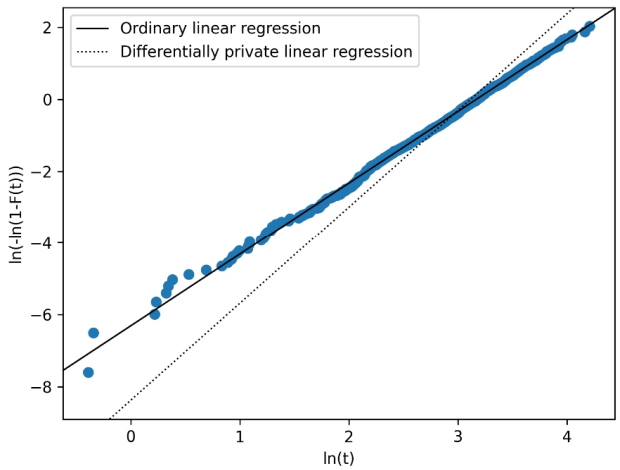

Fig. 4. Linear regression based on samples from a Weibull distribution using ordinary linear regression (solid line) and a differentially private version (dotted line, $\varepsilon = 1.0$)

It should be noted that ε-DP for linear regression is more difficult than ε-DP for evaluating and publishing simple statistics such as mean values or counts. The reason is that regression involves solving an optimization problem, and it is harder to infer from the original dataset to the results of the optimization task, which is necessary to determine the minimum amount of randomness to obtain differentially private optimization results. The differentially private linear regression algorithm used here uses a so-called functional mechanism that adds randomness to the objective function of the optimization problem instead of its results (Zhang et al., 2012).

A more recent work (Alabi et al., 2020) investigates several differentially private simple linear regression algorithms where the added randomness is less than the standard even for small datasets. In order to achieve this, however, it was found that a different, robust linear regression estimator, Theil-Sen, had to be used instead of ordinary linear regression. Moreover, most of the investigated methods come with hyperparameters that govern their behavior. It is still and open problem how to optimally choose these parameters, and this may also hinder practical applicability. In addition, the focus of the referenced paper is on outputting accurate point estimates, rather than confidence intervals.

6. Summary and outlook

The paper demonstrates how differentially private algorithms can be applied to reliability data evaluation in general, as well as in an Industrial IoT context. When publishing results of evaluations of sensitive data that requires protection, the question whether differential privacy can be reasonably applied depends mainly on the answer to the following question: is the evaluation method robust, i.e., does the result of the evaluation not depend too much on small changes in the database? This is, fortunately, typically the case for evaluations of reliability data and it means that differential privacy may be applied. This comes with several advantages such as (see e.g. Wood et al., 2020)

- relatively low computational cost, so the differentially private processing steps can be implemented on IIoT devices with limited computing power and memory,
- a rigorous theoretical framework that provides provable privacy guarantees. This is also the case for the cumulative risk from successive publishing of processing results, which has not been dealt with in the present paper,
- there is no need for elaborate attack modeling because the differential privacy guarantees are totally independent of the prior knowledge of the attacker. As a worst case, it can be assumed that the attacker knows the complete database except for the record that is the target of the attack. Finally,
- it is not necessary to ensure secrecy around the differentially private algorithm or its parameters, which is another distinguishing feature from traditional methods for publishing sensitive data.

The drawback is some loss of accuracy, which of course depends on the specific situation but what we consider usually acceptable for reliability applications. Regarding the implementation aspect, there are an increasing number of implementations of differentially private algorithms published by corporations such as Google and IBM, and the number of practical tools is growing.

Currently, we are investigating the application of differential privacy to a wider class of reliability data algorithms, such as maximum-likelihood estimators for failure data with arbitrary censoring mechanisms including confidence bounds, as well as aggregating reliability data on various levels in the IIoT infrastructure.

Acknowledgements

The numerical examples have been produced with the IBM differential privacy library diffprivlib, which is a general-purpose library for experimenting with, investigating, and developing applications in differential privacy (see Holohan et al., 2019).

References

Alabi, D., McMillan, A., Sarathy, J., Smith, A., Vadhan, S. 2020. Differentially private simple linear regression. ArXiv preprint arXiv:2007.05157.

Compare, M., Baraldi, P., and Zio, E. 2019. Challenges to IoT-enabled predictive maintenance for industry 4.0. IEEE Internet of Things Journal, 7(5), pp. 4585-4597.

Desfontaines, D. 2023. A friendly, non-technical introduction to differential privacy. Personal blog, https://desfontain.es/blog/friendly-intro-to-differential-privacy.html

Dwork, C., McSherry, F., Nissim, K., Smith, A. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In: Halevi, S., Rabin, T. (eds) Theory of Cryptography. TCC 2006. Lecture Notes in Computer Science, vol. 3876, pp. 265-284. Springer, Berlin, Heidelberg. doi: 10.1007/11681878_14

Holohan, N., Braghin, S., Mac Aonghusa, P., and Levacher, K. 2019. Diffprivlib: the IBM differential privacy library. ArXiv e-prints arXiv: 1907.02444

Hong, Y., Zhang, M., Meeker, W. Q. 2018. Big data and reliability applications: The complexity dimension. Journal of Quality Technology, 50(2), 135-149.

Husnoo, M. A., Anwar, A., Chakrabortty, R. K., Doss, R. and Ryan, M. J. 2021. Differential Privacy for IoT-Enabled Critical Infrastructure: A Comprehensive Survey, in IEEE Access, vol. 9, pp. 153276-153304, doi: 10.1109/ACCESS.2021.3124309
IEC 60050-192:2015, International Electrotechnical Vocabulary (IEV) - Part 192: Dependability, Ed. 1.0 (2015-02-26)
Lazarova-Molnar, S., Mohamed, N., 2019. Reliability Assessment in the Context of Industry 4.0: Data as a Game Changer. Procedia Computer Science,vol. 151, pp. 691-698, doi: 10.1016/j.procs.2019.04.092
Meeker, W.Q., Escobar, L. A., and Pascual, F. G. 2022. Statistical methods for reliability data. John Wiley & Sons.
Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., Nissim, K., O'Brien, D. R., Steinke, T., and Vadhan, S. 2020. Differential Privacy: A Primer for a Non-Technical Audience. 21 Vanderbilt Journal of Entertainment and Technology Law 209-276
Zhang, J., Zhenjie, Z., Xiaokui, X., Yin, Y., and Winslett, M 2012. Functional mechanism: regression analysis under differential privacy. arXiv preprint arXiv:1