Decentralisation of reliability data through blockchain

G. Willoughby*, T. Ingram**, T. Byrne***, D. Smith, N Rahman

* Brew House, Wilderspool Business Park, Warrington, WA4 6HL: gwilloughby@reliabilityblockchain.com
** Brew House, Wilderspool Business Park, Warrington, WA4 6HL: tingram@mmiengineeirng.com
*** Apollo House, Eboracum Way, York, YO31 7RE: tbyrne@geosyntec.com

Abstract. The intent of this paper is to illustrate how the current paradigm of reliability information held by a small selection of companies who sell the information could soon be outdated. We believe that Reliability data ownership could be transferred by a collaborative, decentralized approach, with data sourced from the communities that use the data stored in a blockchain. The advent of blockchain as a shared, immutable ledger, has the potential to be the technology that provides the platform for this. Such a ledger can provide a new way of thinking about reliability data potential: Original Equipment Manufacturers (OEM's) and users contributing data in a transparent manner, clear visibility of underpinning data continually evolving in a live database, self-checking across the community, data beyond the base of the bathtub, quality certification, storage of source information such as Failure Modes Effects Criticality Analysis (FMECAs), direct linkage into the Industrial Internet of Things. This fundamental shift in approach could see a focus of practitioner held information, with trust held within the community. We will present some potential methodologies, and how we are proposing to establish collaborative working groups to establish a standard on this technology. The objective is to provide a platform for the international reliability community that allows for the development of a truly free, trustworthy, reliability resource.

1. Background

The intent of this paper is to present:

- Current Reliability data: considering approaches of how data is used; data sources; and the resulting issues and benefits of these approaches;
- Future of decentralised reliability data repository: presenting a description of decentralisation; benefits of decentralisation; and tools to support decentralisation;
- Implementation of decentralised repository: considerations for implementation of a block chain; challenges of implementation; a forward-looking roadmap from conception;
- Future potential of a decentralized approach: presenting the opportunities from a new way of thinking
- Conclusions: presenting recommendations for the next steps.

2. Current Reliability data

2.1 *How data is used*

Data is collected for a variety of uses but is most often used to prove or evidence something, OEMs might use data to advertise the capability of their products. End users might use data to evidence the performance of their equipment to regulators or data might be used internally to justify investment or aid decision making. Whatever the reason, reliability data has an integral part to play in almost all elements of the asset life cycle.

This reliability data is typically used to underpin frequency of failures and resulting events, whether that be Availability Analysis, Quantitative Risk Assessment (QRA), Failure Modes Effects Analysis, Failure Analysis, Spares Ranging & Scaling, Multi Criteria Decision Analysis, Accelerated Life Testing, with many more potential uses available. The output of such analyses focusses on a single point result when the reality of operations is much more complicated.

Reliability of equipment is typically categorized into 3 periods: infant mortality in which issues that have not been identified during testing and commissioning result in a high initial failure rate which decreases over time; operational life in which the failure rate remains reasonably constant, due to increased duration and increased data population; end of life phase results in an increasing failure rate as the component approaches or exceeds its intended life. The majority of data sources focus on the portion of "useful life" which is represented by a constant failure rate, or simplified wear out patterns characterised through distributions such as Weibull.

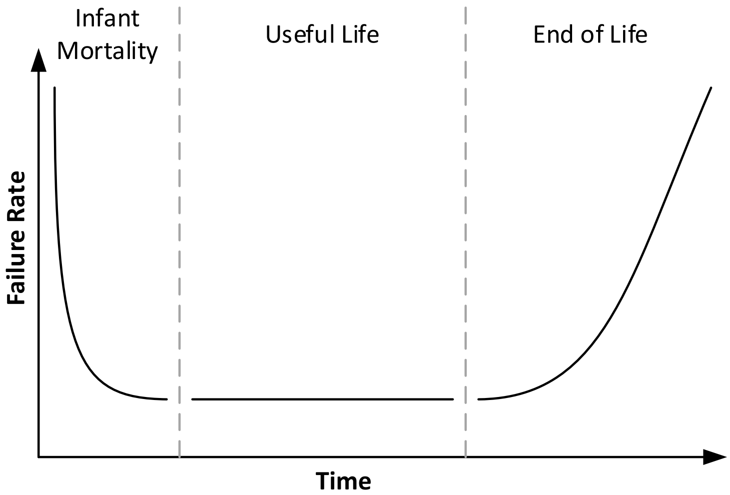

Figure 1. Example "Bathtub curve" with split into 3 typical life stages.

However, many organisations are planning to operate well beyond their design life to strike a careful balance of cost and performance to maintain cost effective production. In order to support accurate predictions operators need a significant amount of intelligence to characterise operational life, and usage for equipment that may be lost 50 years on.

2.2 *Data sources*

Across industries, approaches differ slightly, but it is typical for data sources to be prioritized, for example: internal operational data; regulatory compiled data (e.g. United Kingdom Health & Safety Executive (HSE)); government sponsored data (e.g. Military Handbooks (Mil-HDK)); group sponsored data (e.g. Non-electronic Parts Reliability Data (NPRD); sector specific data (e.g. Off-shore and onshore Reliability Data (OREDA)).

Organizations constantly look internally to collect and use data in a more effective way, considering the rage and scale of the data they collect through a combination of manual and automated approaches. Such data is directly applicable to the application, and the organization has full visibility of the supporting information. However, internal sources are limited by the available equipment, knowledge of failures, internal confirmation bias, small dataset.

External data sources, such as those mentioned above, provide information from typically a much larger data set, with statistically more confident results. However, the source of the data can raise issues regarding trust and applicability to a given application. For example, government sponsored data such as Mil-Hdbk217F, was last updated in 1995 and data within could be considered out of date when compared to current equipment. Safety related data such as Failure Rate and Event Data for use within Risk Assessments, HSE Planning Case Assessment Guide (PCAG) Chapter 6K (28/06/2012), provides a consistent data source for safety assessments, however its use to underpin operational availability may be questioned as a perception may be that certain failures may reduce operational availability without impacting safety. Paid for standards such as Nonelectronic Parts Reliability Data Publication (NPRD-2016), represent a more varied and independent source. However, the paid for mechanism can also provide uncertainty regarding the validity of the data as new databooks are released, and the provenance of the data can be limited. Sector specific data such as OREDA Handbook 2015, 6th edition – Volume I and II provide useful sector specific data, but may not provide sufficient information or context to enable translation into other sectors or applications.

2.3 *Issues and benefits*

The referenced data sets provide an average mean time between failure, which then needs to be reflected against the condition of in-service assets in-service. Reliability professionals are experts in translation of the current asset performance, or design scope, into a prediction of future performance. As such, within the reliability domain, the confidence level of the data is understood. However, as the resulting analyses are used to underpin wider business decisions the understanding of the supporting data can be lost. Inherent weaknesses within the sources such as, dataset size, validity against current & existing equipment, cross sector applicability, and underpinning resolution can limit the confidence and trust outside of the reliability community.

Even if we view these data sources as high integrity, the use of the data may be different to the source, for example risk of a complete failure vs a partial degradation are likely to have different failure rates. Publications such as Failure Mode/Mechanism Distributions FMD-2016 fill some of the gaps through articulation of the distribution of failure modes within the data set. However, there can be a question on the integrity of the data and whether that the usage/environment can translate.

Within the right context each of the data sources mentioned in Section 2.2 can be a valuable asset to support the reliability predictions.

3. Future of decentralised reliability data repository

3.1 *What is decentralisation of data?*

A comparable example of data decentralisation is the evolution of encyclopedias, which once were limited to printed books, moving onto electronic media. Such publications were a “trusted source” because there was no other information available and had inherent invested value as they were paid for publications. The notion and validity of a crowdsourced data set within the 1990’s would have raised that decentralising such information would result in a lack of trust that may render the data useless. Moving forward to the present day, Wikipedia has made that transition, and disrupted the paid for model, but it is not the only data source. Also the perception of the data is to utilize it with caution, rather than with implicit confidence. Since its creation a significantly information resource has developed within www.wikipedia.com sourced from volunteers with a reasonable level of accuracy.

In a similar manner we propose the transition of reliability data into a shared data set, that self regulates, and provides sufficient granularity, traceability and lineage to make it useful to a wide array of applications and end users.

3.2 *Benefits of decentralisation*

What if we could work together to produce much larger reliability data sets, ones that not only represented the “as is” state for an individual organisation but a data set that could help to evidence the whole spectrum of performance across industries and applications? The potential to work collaboratively and drive up standards across the board is hard to dispute and has potential benefits for all stakeholders involved in the reliability data domain:

- OEM’s – visibility of how their products perform long term in-service, and how they perform against competitors and see how end users are using their products. This will focus investment in value add research & development and develop better products long term;
- End users – informing equipment procurement decisions, benchmarking of their performance within their sector, and beyond. They can use information to help drive investments in new assets, and potentially benefit from best practice and innovation from other sectors;
- Regulators – provision of a simplified holistic view across relevant sectors which can help organisations to drive up performance in targeted areas.

Expectations of customers, users and regulators are increasing all the time. This results in a need to drive up standards, reduce risk and protect profits. Central to this is the ability to make good decisions based on good data. “Lessons learnt” or “best practice” are terms often used to emphasise the virtues of effective information sharing but is something that is often difficult to achieve even on a small scale because of some of the perceived difficulties in this process. What if we could embrace new technologies that would allow us to share and utilise reliability data in a shared, common, de-centralised network; not for profit, simply to improve asset management performance.

3.3 *Tools to support decentralisation;*

Numerous data storage types exist, and all have varying levels of complexity. A key factor in developing a decentralised set of data relating to equipment failure will be trust and visibility, due to many contributors in many applications and industry sectors. The trust needs to be formed both through an ability to summarise data providing visibility of the supporting data sources, and fundamentally that no single entity controls the whole data set. On this basis we propose that a blockchain infrastructure could be used to provide a decentralised, secure, and reliable platform.

Experience of conventional repository technologies to date have demonstrate an almost Darwinian evolution, as software is upgraded, copied, changed etc, the data within can suffer random mutations. This can be in the form of corrupt, or missing data, which is typically picked up during or after analysis and can costly or devastating results. The advent of blockchain demonstrates an evolutionary leap, which to date through cryptography, and hash values has been robust, in part through inherent error checking from block to block. Also, if a mutation occurs this is stabilised through the peer to peer nature of the system, and the mutation can be readily removed. As technologies improve and cryptography is cracked, this position may change, however there is sufficient financial backing through the likes of Bitcoin to keep investment in cryptography ahead of the game.

4. Implementation of decentralised repository

4.1 *Considerations for implementation of a block chain*

In its simplest form blockchain acts as a large database of information from which data can be traced, verified and stored securely. A blockchain is shared across a peer-to-peer network in which multiple versions of the same data are stored across the network, which adds resilience and security.

The technology itself prevents changes from being made by stringing together a series of transactions into a single chain where each record is linked to the previous record using a hash function of the value of the previous and current transaction, a form of digital signature. If changes are made the chain is broken as the hash will change and not be recognised by any other part of the blockchain. This immutable ledger gives credence to the validity of the record of the change or sharing of data and could be used, for example in the energy sector for a peer-to-peer community-based sharing network, allowing for community auditing and transparent functionality of the industry.

Understanding the concept of decentralization is paramount to understanding the implementation of a blockchain network. Several questions must be answered by an organisation or community of people to decide what type of blockchain to implement. Public blockchains usually consist of a large group of nodes (computers) which are connected to the internet to validate transactions ongoing within the blockchain. As a result of this, consensus transactions are agreed by other ways rather than selective endorsement IE, Mining - computers on the network try to solve a complex cryptographic problem to create a proof of work. In contrast private blockchains can be set to only permit known organisations or users to join, enabling the grounds for a permissioned network in which consensus can be achieved through a process called selective endorsement, where known or selected users can verify changes in data. In essence, private blockchains use identity to confirm membership and access privileges so that participants in the network know who they are dealing with, allowing for enhanced auditing and transparency as noted above.

4.2 *Challenges of implementation*

Information sharing across organisations is notoriously difficult, due to the invested time and money required to collect, verify, manage and analyse data. The inherent requirement to manipulate data into a form that can be shared across electronic platforms causes additional time and financial burden. This is further emphasized when referring to failure related data, as there is additional complexity due to mis representation of failures outside of the organization. However the benefit to all involved (if a common platform can be created) is to drive OEM's to better products, provide more opportunities for informed decision making, ability to truly learn across sectors with trust in a dataset which is large enough to provide statistical confidence in results.

We believe the significantly increase data repository and granularity would outweigh these issues for a significant number of organisations.

A key challenge will be a common understanding of what information (meta data) would be captured in support of the reliability data. For example something as simple as a flag to identify a failure rate that has been influenced in some way by human factors, could provide a new insight into the influence on humans on certain equipment types or industry sectors. This will require a significant amount of cooperation across multiple sectors to develop a set of definitions that work across industries.

4.3 *Roadmap*

Implementing the blockchain technology requires several steps to be followed as shown below;

- Step 1 – Identify the blockchains purpose;
- Step 2 – Agree on the most suitable consensus mechanism;
- Step 3 – Identify a suitable blockchain platform;
- Step 4 – Understanding the scope of nodes associated with the blockchain;
- Step 5 – Design the blockchain Instance and parameters;
- Step 6 – Build the API's;
- Step 7 - Design the admin and user interface;
- Step 8 – Adding future technology.

5. Future potential of a decentralized approach

There are countless potential benefits of using a blockchain as a decentralised, single source of data for individuals, companies and non-profit organisations operating in multiple industries. The fundamental properties of a blockchain's infrastructure, such as its limitless storage capacity and secure decentralised foundations, make it an ideal solution for the storage of data in a live and manageable repository which can allow for improved data manipulation for the needs of all users.

5.1 Operations Staff and Asset Managers

Operations personnel could be some of the key users of a blockchain based data repository. Access to a blockchain allows operations staff to benchmark and compare performance data internally with their own records, or externally with the data of similar organisations. Comparisons can be based on numerous metrics such as defect rates, production rates, human interaction and other KPI's which will allow for the process optimisation of an organisations operational activities.

5.2 *Procurement and Investment Planning*

As mentioned above, blockchains have the capability to categorise data from a practically infinite number of decentralised locations allowing for effective data comparison. This could be utilised by purchasers to complete a full market comparison and refine their search criteria to find the best suppliers in the market almost instantaneously.

Currently, many organisations struggle to find suppliers which meet their procurement standards and requirements and invest heavily in new market research strategies. If supplier information was held in a blockchain, one could readily complete a full market analysis while ensuring all the data remains secure in decentralised locations.

Blockchain based market research has the potential to save huge amounts of time and effort for procurement functions, being a means for much easier and effective market research and supply chain management. Thus, companies may choose to refresh their procurement strategies on a more regular basis, since the market will be much more transparent allowing purchasers to refine their demands based on the current market offering.

5.3 *Research and Development*

Data comparison using the blockchain could be of significant benefit for R&D organisations when searching for new market opportunities for the development of future products and services. Performance data, asset lifecycle information, new and existing government regulations and policies which are all accessible via one user friendly portal (the blockchain). This could provide R&D organisations with all the necessary information to identify, foresee and exploit opportunities for growth in existing markets. In turn, this could facilitate a rapid growth in research and development and further bridge the gap between R&D and Industry.

5.4 *Regulatory Bodies & Government Departments*
Regulatory bodies could also benefit from having access to industry data via the blockchain. For example, regulatory bodies assessing the general health of equipment used for manufacturing could use the blockchain to gain access to relevant records of the production performance, maintenance history and other statistical information in order to make more informed decisions. Furthermore, regulatory bodies could use this information as a basis to form future regulations based on current or predictable future transitions in industry.

5.5 *Software & hardware developers*
Potentially, company hardware and software could directly interphase with the blockchain, allowing for automated transfer of data directly to the blockchain. This will enhance user experience ensuring that all relevant information is up to date, live and ready for use.

6. Conclusions:

This paper has presented some basic concepts of how reliability data is currently sourced by a small selection of companies who sell the information, and how this paradigm could soon be out dated. We believe that reliability data ownership could be transferred into a collaborative, decentralized approach, sourced from the communities that use the data stored on a blockchain. The advent of blockchain as a shared, immutable ledger, has the potential to be the technology that provides the platform for this.
This fundamental shift in approach could see a focus of practitioner held information, with trust held within the community. We will present some potential methodologies, and how we are proposing to establish collaborative working groups to establish a standard on this technology. To provide a platform for the international reliability community to develop a truly free, trustworthy, reliability resource.
To find out more, register interest or become part of the influencing working groups visit www.reliabilityblockchain.com.

References

HSE Planning Case Assessment Guide (PCAG) Chapter 6K (28/06/2012)
Mil-Hdbk217F (1995) Military Handbook: Reliability Prediction of Electronic Equipment
NPRD (2016) Nonelectronic Parts Reliability Data Publication
OREDA Handbook 2015, 6th edition – Volume I and II
Failure Mode/Mechanism Distributions FMD-2016
www.reliabilityblockchain.com