Quality Improvement Mechanism for Cyber Physical Systems – An Evaluation

Gilbert Regan[1], Fergal Mc Caffery[1], Pangkaj Chandra Paul[1], Jan Reich[2], Eric Armengaud[3], Cem Kaypmaz[3], Marc Zeller[4], Joe Zhensheng Guo[4], Simone Longo[5], Eoin O Carroll[6] and Ioannis Sorokos[7]

[1] Lero @DKIT, Dundalk, Ireland
Gilbert.regan,fergal.mccaffery and Pangkajchandra.paul @dkit.ie
[2] Fraunhofer IESE, Kaiserslautern, Germany
jan.reich@iese.fraunhofer.de
[3] AVL List GmbH, Austria, Turkey
eric.armengaud and cem.kaypmaz@avl.com
[4] Siemens, Munich, Germany
joe.guo@siemens.com
Marc.zeller@siemens.com
[5] General Motors, Turin, Italy
simone.lomgo@gm.com
[6] Portable Medical Technology, Kerry, Ireland
eoin@portablemedicaltechnology.com
[7] University of Hull, Hull, UK
I.Sorokos@hull.ac.uk

Abstract. The future will encompass heavily interconnected, distributed, heterogeneous and intelligent systems which are bound to have a significant economic and social impact. Cyber Physical Systems (CPS) such as autonomous cars, smart electric grid, implanted medical devices and smart manufacturing are some practical examples of these intelligent systems. However, due to the open and cooperative nature of CPS, assuring their dependability is a challenge. The DEIS project addresses this important and unsolved challenge by developing the concept of a **D**igital **D**ependability **I**dentity (DDI). A DDI contains all the information that uniquely describes the dependability characteristics of a CPS or CPS component. DDIs are synthesised at development time and are the basis for the (semi-)automated integration of components into systems during development, as well as for the fully automated dynamic integration of systems into systems of systems in the field.

In this paper we present an overview of the DDI. We provide the metric selection process for evaluating the DDI's impact on CPS dependability. The results of an evaluation of the DDI's impact on dependability in four CPS industrial systems are provided, both for design time and runtime. These results demonstrate the positive impact of the DDI on the dependability of CPS.

Keywords: dependability, cyber physical system, evaluation, cyber security.

1 Introduction

Cyber-Physical Systems (CPS) harbor the potential for vast economic and societal impact in domains such as mobility, home automation and delivery of health. At the same time, if such systems fail they may harm people and lead to temporary collapse of important infrastructures with catastrophic results for industry and society. There are two core challenges while assessing the dependability of a CPS. First, the inherent complexity of modern CPS and the resulting complex market organisation requiring the tight cooperation between different teams, expertise, and institutions, while managing confidentiality issues. The second challenge is related to the increase of connectivity, e.g., through machine to machine cooperation enabled by Internet of Things, which introduces a new dynamic in system operation. As a result, Cyber-Physical Systems of Systems (CPSoS) come together as temporary configurations of CPS, and which dissolve and give place to other configurations. This leads to a potentially infinite number of variants, with cooperation between systems potentially not analysed during design time.

The world is changing fast with a wave of digitisation and disruptive innovation across industry and society, exploiting artificial intelligence (AI), low-power computing, IoT and edge computing platforms underpinned by developments in advanced semiconductors including mixed signal, sensor, and power technologies. Notably, artificial intelligence, digital security and connectivity are areas that have also been identified as strategic technologies by China in its Made in China 2025 strategy[1] , by South Korea under a USD 1.5 billion initiative[2], and by the US as part of a strategic programme run by the US National Science Foundation[3]. This led to a new generation of trusted, collaborative systems of systems, implementing reasoning capabilities at the edge and able to take a safety-critical decision relying on the information received from their environment. Such a concrete example is connected and automated mobility (CAM) for the automotive domain. With the roll-out of 5G[4][5], CAM is a step closer towards supporting a solution for the many challenges faced by today's transport sector. CAM is expected to offer significant societal benefits– ranging from enhanced safety (reducing accidents caused by humans), increased energy efficiency (smoother traffic) to greater comfort (non-driving tasks while travelling), social inclusion (personal mobility for all, including elderly and impaired users) and accessibility (facilitated access to city centres). A list of 30 use cases and potential service requirements have been specified by 3GPP[6].

The DEIS project[7] addresses these important and unsolved challenges by developing technologies that form a science of dependable system integration. In the core of these technologies lies the concept of a Digital Dependability Identity (DDI) of a component or system. The DDI targets (1) improving the efficiency of generating consistent dependability argumentation over the supply chain during design time, and (2) laying the foundation for runtime certification of ad-hoc networks of embedded-systems. During the DEIS project, four industrial systems are provided to evaluate the performances of the DDI. The target is to evaluate the impact of the proposed methodology for process improvement during product development, and to support the emergence of new functions, e.g., by higher degree of collaboration. The core challenge for the evaluation of the project relies on two levels of innovation: first the dependability engineering approach shall be enhanced, second this shall be deployed to improve the industrial product with new solutions.

Assuring dependability of CPS is the core challenge of the DEIS project. Dependability is qualitatively defined as *'the ability to deliver service that can justifiably be trusted*', and quantitatively defined as '*the ability to avoid service failures that are more frequent and more severe than is acceptable to its user(s)*[8]*'*. Dependability encompasses the following attributes: availability; reliability; safety; confidentiality; integrity; maintainability. The security attribute is considered a triage of confidentiality, integrity, and availability. These 'primary' attributes may contain 'secondary' attributes e.g. accountability, authenticity, and non-reputability can be considered secondary attributes of security[8].

Contribution of this paper is to present a systematic approach for the evaluation of dependability methodologies for CPS, and to apply this method for the evaluation of the DDIs in the four industrial systems of the DEIS project. The paper is organized as follow: Section 2 presents related works on quality assessment. An overview of the DDI is presented in Section 3, and the research methodology as well as the systems are introduced in Section 4. In Section 5, the tailoring of the standards used for evaluation are presented. Section 6 provides the design time evaluation results, while Section 7 provides the runtime evaluation results. Finally, Section 8 concludes this work.

2 Related Work

A software quality model can be defined as ‘a model that describes, assesses and/or predicts quality’[9], or as ‘a set of factors, criteria and metrics (characteristics) and the relationship between them. These relations provide the basis for specifying quality requirements and evaluating quality’[10].

Several models of software quality factors and their categorisation have been suggested over the years. The first software quality models were published in the mid 1970’s by Boehm et al[11] and Mc Call et al.[12]. Mc Call identified three main perspectives for characterising the quality attributes of a software product i.e. a product’s ability to change, adaptability to new environments, and basic operational characteristics. From these three perspectives Mc Call identified eleven characteristics. The major contribution of the McCall method was to consider relationships between quality characteristics and metrics. This model was used as a base for the creation of others quality models[13]. The main drawback of the Mc Call model is the accuracy in the measurement of quality, as it is based on just positive/negative responses (Yes/No). Furthermore, the model does not consider the functionality so that the user's vision is diminished[14]. Boehm’s model constitutes an improvement on Mc Call’s model because it is based on a wider range of characteristics and because it adds factors at different levels.

The FURPS quality model[15], which was proposed by Robert Grady from Hewlett Packard in 1992, takes into account the following five characteristics: Functionality, Usability, Reliability, Performance, and Supportability. A main drawback of this model is that it does not consider some important characteristics such as portability, which may be an important criterion for application development[16]. In 1995 Robert Dromey proposed a product based quality model[17] based on the idea that a more dynamic way of modelling process was needed. This was due to the fact that quality evaluation differs between products and the model needed to be wide enough to apply to different systems.

In order to standardise quality assessment, the International Organisation for Standardisation (ISO) developed ISO 9126 [18] in 1991 and revised it in 2001. This standard is an extension of previous models as defined above, and is divided into four parts which address the following subjects: quality model; external metrics; internal metrics; and quality in use metrics. The quality model is divided into the following six characteristics: Functionality; Reliability; Usability; Efficiency; Maintainability; and Portability. The internal metrics are static metrics that do not rely on software execution, whereas the external metrics rely on running software. Quality in use metrics can be measured only when the final product is used in a real environment with real conditions.

The ISO 9126 model was updated in 2005 and evolved to become part of the ISO 25000:2005 series[19], and which has further been revised with ISO 25000: 2011[20]. Studies conducted by[21][22][23] indicate that the ISO/IEC 25010 model[24] is the most comprehensive quality model available because it covers the most quality characteristics and sub-characteristics. It achieves this by adding new characteristics such as security and compatibility.

3 Overview of DDI

Assurance cases represent the backbone of modern dependability assurance processes. A given assurance case captures the underlying argument of how the subject system meets its dependability requirements. The assurance case takes into account the subject system's intended operational environment together with the evidence that supports requirement validity in the finally implemented system. In practical terms, producing, maintaining and reviewing an assurance case is a process that aims to increase confidence in the quality of the subject system's dependability properties, as well as its development process.

Since there is an interrelation between the system, its dependability claims, and the supporting evidence artifacts that exist in the real world, we claim this should also be the case for the system's model-based safety reflection, i.e. its DDI (see Fig. 1). DDIs represent an integrated set of dependability data models that may be (semi-)automatically analysed, or generated during the execution of safety engineering processes.

A DDI contains information that uniquely describes all dependability characteristics of a system required for certifying the system's dependability. DDIs are formed as modular assurance cases and their composability allows for the (semi-)automatic synthesis of system DDIs from the DDIs of the subcomponents. The DDI of a system contains a) claims about the dependability guarantees given by a system to other systems and derived system dependability requirements and b) supporting evidence for those claims in the form of various models and analyses. For security assurance, it contains e.g. threat and risk analyses (TARA) and attack trees, while for safety assurance, hazard and risk analyses (HARA), architecture modeling and failure propagation modeling such as fault trees, FMEA or Markov chains are supported.

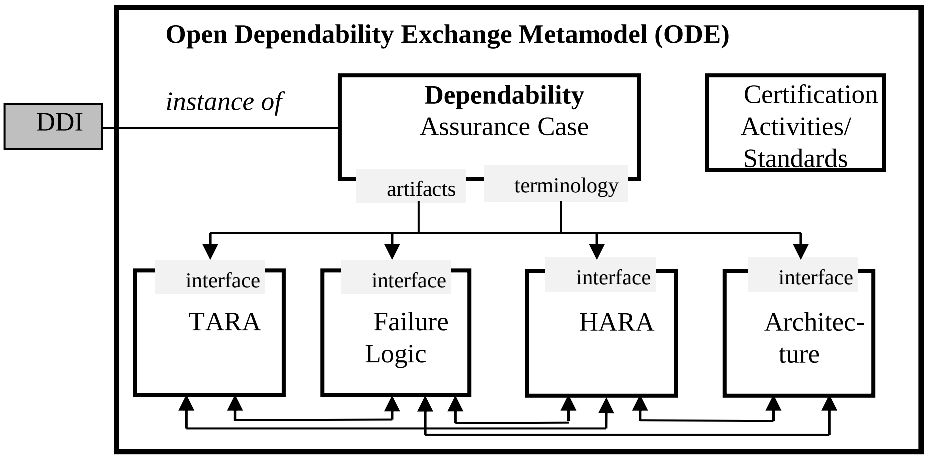
Open Dependability Exchange Metamodel (ODE)
DDI
instance of
Dependability Assurance Case
Certification Activities/ Standards
artifacts
terminology
interface
TARA
interface
Failure Logic
interface
HARA
interface
Architecture

Fig. 1. The Open Dependability Exchange Metamodel (ODE)

Thus, a DDI is powered by a set of dependability claims, for which the models produced during dependability activities shall provide convincing evidence of satisfaction for the developed system. For both risk management planning and dependability assessment purposes, an explicit argument is indispensable inductively relating the created evidence to the top-level claim through several step-wise layers of argumentation. DDIs can deal with safety and security risks, thus the set of currently supported assurance activities focus on industrially well-established methods. These activities proved sufficient over the last decades in demonstrating the dependability of embedded systems.

The advantage of continuous traceability in DDIs between a safety argument expressed in the Goal Structuring Notation (GSN) and safety-related evidence models is enabled by an integrated meta-model, the Open Dependability Exchange (ODE) meta-model v2.0. Since it is very likely that new dependability standards will come up in the future, the ODE has been designed to be extendable through a modular package concept. Around this technical DDI backbone, an automation framework was built to support automated change impact or argument validity analyses on the DDI data contents[25]. Details on the DDI framework as well as an open-source version of the ODE meta-model can be found at Github[26].

4 Research Methodology

Section 4.1 provides an overview of the methodology employed to select the quality metrics used to evaluate the impact of the DDI, while section 4.2 provides a brief description of the four systems used for this same evaluation.

4.1 Methodology

The methodology used to conduct this research comprised the following main stages: Select metrics; Select systems; Evaluate DDI impact in systems; and Report findings.

As stated in section 2, the ISO 25010 quality model is the most comprehensive quality model available and so this model, specifically metrics from the following standards within the ISO 25000 series were selected to assess the impact of the DDI: ISO 25022[27]; ISO 25023[28]; and ISO 25024[29]. Details of the metrics and their selection are provided in section 5.

Four industrial partners on the DEIS project each put forward a system for assessing the impact of the DDI. Two systems are embedded in the automotive domain while the remaining ones are embedded in the railway and healthcare domains. A short description of these systems is provided in section 4.2. For each system a team of people from within each system's organisation conducted the evaluation. Each system was evaluated both before and after application of the DDI. The make-up of the teams was decided upon by the organisation themselves. For example, the Siemens team included 1 model-based safety and reliability engineer, 1 model-based safety and reliability consultant, 1 safety engineer, and 1 reliability engineer. Through expert judgement and consensus, and with the use of the measurement formulae within the standards, each team determined values for the selected characteristics. The results for each organisation's quality characteristic assessment are provided in section 6.

4.2 Systems

Portable Medical Technology (PMT): PMT's ONCOassist is a clinical decision support app for oncology professionals. It contains all the key oncology decision support tools oncology professionals need and makes them available in an easy to access and interactive format at point of care. ONCOassist is an application which aims to engage and connect with EHRs. This will allow for automatically reading patient specific data from EHRs in a safe and secure manner, in order to perform calculations that would not be possible on the hospital system. However, with the variety of EHR vendors it is very difficult to create consistent interoperability between EHRs and third party medical devices. This raises combined concerns of safety, security and privacy. Patient data must be secure and kept private. If the system's operation is maliciously attacked successfully or there are errors in the system's nominal functions, patient safety may also be compromised due to mistakes introduced to their medical information. As health practitioners must also use their credentials to access the associated service, those credentials must also be secured. It is envisioned that the DDI will help alleviate these concerns.

General Motors (GM): The Dependable Physiological Monitor System (DPMS) use case describes the sensing environment inside the vehicle that monitors the health condition of drivers and passengers in order to improve the safety of the driver. The DPMS aids prevention of accidents in cases where the occupant's health condition deteriorates. A dedicated sensing solution based on a Single Photon Avalanche Diode camera and Physio hardware board have been developed. In the case of the drivers' health deteriorating, a high level emergency manager feature makes mitigating decisions depending on the severity of the deterioration in the drivers health, for example taking control of the car and parking in a lay-by, or notifying emergency responders. The DPMS will apply the DDI at both development time and at runtime. During the development phase a reduction of time-to-market is expected as a consequence of the usage of DDI methodology. At runtime, DDI is evaluated against the overall dependability of this system, dealing with security and privacy aspects in V2V and V2C communication. For this system, safety of the passenger must be assured as part of the nominal behavior of the system. Additionally, as the system is designed to communicate over open channels and manage sensitive and personally identifiable information, security and privacy must also be assured.

Siemens (SAG): The European Train Control System (ETCS) provides standardised train control in Europe and eases travelling with trains crossing the borders of all countries. The ETCS consists of an on-board and a trackside part. Both sub-systems must fulfil the safety requirement as defined in Subset-091 (Safety Requirements for the Technical Interoperability of ETCS in Levels 1 & 2) [30]of the ERTMS/ETCS specification. Systems in the railway domain are also produced by various stakeholders in the value chain (such as railway undertaking, OEMs, suppliers, etc.) and, therefore, safety information about components and subsystems (rolling stock, trackside and interlocking systems) need to be interoperable and exchangeable. In this use case a generalized ETCS is used. The generalized ETCS is realized as an Emergency Brake Command (EBC) functionality of the Train Interface Unit (TIU) of the ETCS On-Board sub-system. This EBC system is used for among others to illustrate strengths and drawbacks of DDI in several engineering phases, such as architecture, qualitative and quantitative dependability analysis, and Goal Structural Notation (GSN) based dependability assurance case development. It is envisioned that DDI usage in railway domain will assist in achieving safe and interoperable railway systems.

AVL: Heavy-duty trucks create a platoon to reduce time gaps between the trucks, to increase energy efficiency, improve safety, and to reduce truck driver loads. In the SAE L4 platoon function, platoon level decisions are taken by the platoon leader and broadcasted to the follower vehicles and executed by each member without any need of a human driver or operator in a constrained operation boundary. Two-way information flow between vehicles and different communication topologies between members creates a wide range of dependability. The communication between the trucks for increased efficiency and safety comes with increased security threats, and guaranteeing the systems' dependability requirements poses new challenges. For this use case, safety of the passengers and surrounding traffic must be assured; due to communication over open channels being required, security is also an applicable issue, with safety implications should successful attacks induce dangerous platoon behavior.

5 Metric Selection

The quality characteristics, and the metrics for these characteristics, were selected from the following standards: ISO 25022; ISO 25023; and ISO 25024. ISO 25022 defines characteristics and measures for evaluating quality in use characteristics i.e. quality from the end user's perspective, ISO 25023 defines characteristics and measures for quantitatively evaluating system and software product quality characteristics, while ISO 25024 defines data quality characteristics and measures for quantitatively evaluating the quality of the data within the system.

These standards contained more quality characteristics than were relevant for measuring the impact of the DDI, therefore the first task was to select a subset of those quality characteristics which were relevant to the study. A focus group, containing members from the DEIS project partners, was formed to conduct this task. The members of this focus group are listed as authors of this paper. The process of selecting the relevant metrics included five 1.5 to 2 hour on-line meetings. At these meetings, each metric within the three standards was discussed in detail with two main considerations in mind: 1) its relevance to assessing the impact of the DDI; and 2) effort required for both data collection from the underlying system models and calculation of the metric. Decisions on whether to include a metric, or not, were based on a general consensus which was largely unanimous in each case. The number of characteristics and measures within each standard, along with the number of characteristics and measures selected from the standards, is displayed in Table 1.

Table 1. Number of characteristics and measures within standards versus number of characteristics and measures selected for measuring DDI impact

Number of characteristics and measures within standards			**Selected for measuring impact of DDI**		**Selection versus total %**	
Standard	Characteristics	Measures	Characteristics	Measures	Characteristics	Measures
ISO 25022	5	36	4	10	80	28
ISO 25023	8	86	7	29	87.5	33.7
ISO 25024	15	63	10	19	66.7	30.2
	28	**185**	**21**	**58**	**75**	**31.4**

In total, the team selected 21 quality characteristics (from 28 within the standards). From these 21 characteristics, 42 sub-characteristics were selected as shown in section 6 below. A total of 58 separate measures (from 185 within the standards) were employed in order to determine values for the quality characteristics.

Due to the large number of total measures (and submeasures) discussed, a complete account of the considerations for each measure could not be included in this publication due to space limitations. Instead, presentation of the partner views regarding a few measures follows, as examples indicative of the overall process.

The first example concerns the ‘Functional appropriateness (sub) measures’, of the ‘Functional suitability measures’ ISO 25023. These (sub) measures were viewed as only partially relevant to one of the use cases and not the remaining ones. Additionally, the DEIS partners noted that the question of (nominal) function appropriateness fell outside the scope of the DDI. This is the case, as the DDI instead deals with a given system’s dependability requirements and their assurance. On a similar note, the second example pertains to the ‘User interface aesthetic measures’ of ISO 25023, which are a quality not affected by the presence or absence of the DDI. A more meaningful example is that of the ‘Data accuracy measures’ of ISO 25024, most of which were considered for evaluation of the DDI. For instance, the data accuracy assurance measure was found to be appropriate for DDI evaluation. This sub measure evaluates the ratio of data items measured for accuracy over the data items whose measurement is required for accuracy. Effectively, this is a requirement coverage evaluation, which is in line with the DDI’s role in supporting requirement assurance.

6 Design Time DDI Evaluation Results

The results from evaluating the DDI in the four industrial systems at design time are now presented in the following three subsections.

6.1 ISO 25022 Quality in Use Results

Table 2 presents the results from evaluating the four ‘Quality in Use’ characteristics which have been selected from ISO 25022. The selected sub-characteristics (three in total) are listed in column 2. The ‘Effectiveness' and 'Efficiency' characteristics have no sub-characteristics. The results for each system is presented in 2 columns with the first column presenting the result without the DDI applied, and the second column (in italics) presenting the result with the DDI applied at design time. Each individual result can vary from 0 to 1, with 1 being equivalent in percentage terms to 100. The last column in the table presents the average improvement, in percentage terms, across the four systems, so for example 'Effectiveness' increased by an average of 14.2% when DDI was applied.

Table 2. 'Quality in Use' characteristic values across four systems at design time

Characteristic (4)	**Sub characteristic (3)**	**GM**	***GM***	**AVL**	***AVL***	**PMT**	***PMT***	**SAG**	***SAG***	**AVG % imp.**
Effectiveness	n/a	0.39	*0.56*	0.5	*0.63*	0.5	*0.58*	0.64	*0.83*	14.2
Efficiency	n/a	0.44	*0.62*	0.25	*0.44*	0.53	*0.82*	0.05	*0.95*	39.0
Freedom from Risk	Economic risk mitigation	0.61	*0.73*	0.65	*0.85*	0.33	*0.68*	0.73	*0.87*	20.3
Context Coverage	Context completeness, and Flexibility measures	0.2	*0.2*	0.42	*0.67*	0.5	*1.0*	0.35	*0.69*	27.3

The last column above indicates that the application of the DDI resulted in significant improvement in each of the 'Quality in Use' metrics. The 'Efficiency' metric, at 39% improvement, is particularly influenced by the SAG results who state that '*We are expecting a significant increase of the number of the objectives achieved for the same period of time by introducing DDI. Furthermore, we are expecting a significant decrease in the cost for carrying out the task for the same amount of objects in ETCS use case'*.

Another interesting observation from Table 2 is the GM result for 'Context Coverage'. Context coverage assesses the degree to which a product or system can be used with effectiveness, efficiency, satisfaction, and freedom from risk in both specified contexts of use and in contexts beyond those initially explicitly identified. GM results indicate no improvement as '*no other scenario has been evaluated for DPMS usage'*.

The average improvement, with the introduction of design time DDI's, for all 'Quality in Use' quality metrics is calculated to be 25.2 %.

6.2 ISO 25023 System and Software Quality Results

Table 3 presents the results from evaluating the seven ‘System and Software Quality’ characteristics which have been selected from ISO 25023, and is structured the same as Table 2. The seven characteristics contain twenty sub-characteristics.

While all characteristics indicate an average improvement across the four systems, ‘Performance efficiency’ is the lowest at 4.5%. This is due to most of the systems reporting a very small improvement in this characteristic, with Siemens reporting practically no improvement due to such factors as: ‘Resource utilization: *for mean processor utilisation and bandwidth utilisation, we could not observe any improvement by use of DDI. Both mean processor utilisation and bandwidth utilisation remain low for railway safety-critical system’*

Table 3. System and Software characteristics across four systems at design time

Characteristic	**Sub characteristics**	GM	***GM***	AVL	***AVL***	PMT	***PMT***	SAG	***SAG***	AVG % imp
Functional suitability	Functional completeness,and correctness	0.65	*0.7*	0.55	*0.65*	0.67	*0.83*	0.99	*0.99*	7.8
Performance efficiency	Time behavior, and Resource utilisation	0.46	*0.51*	0.39	*0.47*	0.44	*0.48*	0.39	*0.4*	4.5
Compatability	Co-existance, and Interoperability	0.5	*0.5*	0.71	*0.83*	0.43	*0.73*	0.79	*0.83*	11.5
Reliability	Maturity, Availability, and Fault tolerance	0.71	*0.77*	0.66	*0.78*	0.66	*0.70*	0.75	*0.63*	7.0
Security	Confidentiality, Integrity, Authenticity, Accountability, and Non-repudiation	0.6	*0.68*	0	*0.2*	0.2	*0.38*	0	*0.2*	16.5
Maintainability	Reusability, Analysability, Modifiability, Testability	0.45	*0.52*	0.39	*0.6*	0.44	*0.78*	0.51	*0.55*	16.5
Portability	Adaptability, Replaceability	0.3	*0.3*	0	*0*	0.33	*0.83*	0	*0*	12.5

The ‘Portability’ characteristic, which assesses the degree of effectiveness and efficiency with which a system, product or component can be transferred from one hardware, software or other operational or usage environment to another, indicated no improvement across three systems (GM, AVL, and SAG). This was mainly due to factors such as no portability is conducted in the system: GM stated '*No improvement here, considering that no other scenario has been evaluated outside the GM architecture ecosystem',* while AVL stated *'No portability related implementation has been done'*. For SAG the reason was somewhat different in that they reported that '*the DDI does not offer additional data or handling in case of porting the ETCS system onto another environment'*.

For the 'Security' characteristic, which assess the degree to which a system protects data so that persons or other systems have the degree of data access appropriate to their types and levels of authorization, two systems reported a score of zero for the security sub-characteristics listed in Table 3, however with application of the DDI their security score improved to 0.2. Both of the systems reported an improvement due to implementation of authentication rules.

For 'Functional suitability' SAG reported no improvement due to the fact that they consider their system to be practically functionally complete and correct. They stated that their ETCS products *have 1% of missing intended usage of the system without DDI (99% of usage completeness)…….This estimation is also true for the correctness of functions'*.

While the majority of the selected system and software quality metrics are applicable to most of the systems, according to the industry partners there are occasions where some metrics may not apply to some systems. For example the 'portability' metric only showed improvement in the PMT system.

The average improvement, with the introduction of design time DDI's, for all 'System and Software' quality metrics is calculated to be 10.9 %.

6.3 ISO 25024 Data Quality Results

Table 4 presents the results from evaluating the ten 'Data Quality' characteristics which have been selected from ISO 25024, and is structured the same as Table 3. The ten characteristics contain nineteen sub-characteristics.

Table 4 indicates that all the selected data quality characteristics show an average improvement across the 4 systems, ranging from 10.3% to 26.8%. However for one system, seven of the characteristics show no improvement. For the data completeness, data credibility, data precision and data compliance, SAG state that their ETCS system has to be certified according to relevant safety standards and that these values are at 100% regardless of whether the DDI is applied or not.

For data confidentiality, SAG state that this metric is not applicable but give no reason for this. With regards to data understandability SAG report no improvement, stating that ‘the *semantic understandability will not be changed by introducing DDI’*, and for data portability SAG state that ‘*the DDI does not influence the portability in the sense of operation environment adaptability and data reusability/import capability’*. The fact that three of the metrics have scored zero with and without DDI implementation clearly indicates that SAG do not think that these metrics can be improved in their system.

Table 4. Data quality characteristics across four systems at design time

Characteristic	**Sub characteristics**	GM	***GM***	AVL	***AVL***	PMT	***PMT***	SAG	***SAG***	**AVG % imp**
Data accuracy	Syntactic accuracy	0.6	*0.68*	0.5	*0.63*	0.5	*0.63*	0.88	*1*	11.5
Completeness	Record, and Attribute completeness	0.55	*0.61*	0.38	*0.63*	0.25	*0.88*	1	*1*	23.5
Consistency	Data format and Architecture consistency, and Risk of data inconsistency	0.38	*0.48*	0.33	*0.58*	0.33	*0.53*	0.61	*0.66*	15.0
Credibility	Values, Source, and Data model credibility	0.6	*0.64*	0.5	*0.67*	0.47	*0.87*	1	*1*	15.3
Compliance	Regulatory compliance of value, and Technology	0.54	*0.65*	0.44	*0.75*	0.3	*0.8*	1	*1*	23.0
Confidentiality	Encryption, and Non-vulnerability	0.66	*0.66*	0.38	*0.63*	0.55	*0.71*	0	*0*	10.3
Precision	Data values, and Format precision	0.48	*0.51*	0.38	*0.63*	0.6	*0.8*	1	*1*	12.0
Traceability	Traceability of data values, Data value Traceability	0.5	*0.64*	0.38	*0.5*	0.19	*0.75*	0.7	*0.95*	26.8
Understandability	Semantic undstandability	0.55	*0.68*	0.25	*0.63*	0.6	*0.8*	0	*0*	17.8
Portability	Data portability ratio	0.48	*0.63*	0.38	*0.63*	0.4	*0.8*	0	*0*	20.0

For GM, the data confidentiality metric indicated no improvement. The reason for this according to GM is that applying the DDI guarantees the same level of data confidentiality as against not applying the DDI. However GM further state that the DDI can help in selecting at design time the best security solution to satisfy confidentiality requirements.

While the majority of the selected data quality metrics are applicable to most of the use cases, there were occasions where some metrics may not apply to some systems.

The average improvement, with the introduction of design time DDI's, for all 'Data' quality metrics is calculated to be 17.5 %.

7 Runtime DDI Evaluation Results

The results from evaluating the runtime DDI in the GM and the AVL use cases are now presented in the following three subsections. Only the GM and AVL use cases have been evaluated at runtime because the PMT and Siemens use cases are not related to the runtime DDI concept. The same metrics that were used for the evaluation of the design time DDI in Section 6 are used here for runtime DDI evaluation.

7.1 ISO 25022 Quality in Use Results

Table 5 presents the results from evaluating the four 'Quality in Use' characteristics which have been selected from ISO 25022. The results for each system is presented in 2 columns with the first column presenting the result without the DDI applied, and the second column (in italics) presenting the result with the DDI applied at runtime. The last column in the table presents the average improvement, in percentage terms, across the two systems, so for example 'Effectiveness' increased by an average of 14.5 % when DDI was applied at runtime.

Table 5. 'Quality in Use' characteristic values across two systems at runtime

Characteristic	**Sub characteristic**	**GM**	***GM***	**AVL**	***AVL***	**AVG% imp.**
Effectiveness	n/a	0.39	*0.56*	0.83	*0.95*	14.5
Efficiency	n/a	0.44	*0.62*	0.75	*0.90*	16.5
Freedom from Risk	Economic risk mitigation	0.61	*0.81*	0.5	*0.9*	30.0
Context Coverage	Context completeness,	0.20	*0.20*	0.50	*0.55*	2.5

	and Flexibility measures					

The results show an improvement across all metrics with the introduction of DDI’s. The ‘Freedom of risk’ measures have the highest improvement with an increase of 30.0 % while the ‘Context coverage’ measures have the lowest increase of 2.5 %. The GM results for ‘Context Coverage’ indicates no improvement. This is the same as for the design time DDI evaluation (see Section 6.1).

The evaluation shows that in both cases it is possible to complete all tasks in less time (effectiveness and efficiency) with the introduction of runtime DDI’s. This results in a reduction of costs in implementation (freedom from risk).

The average improvement, with the introduction of runtime DDI’s, for all ‘Quality in Use’ metrics is calculated to be 15.9%%.

7.2 ISO 25023 System and Software Quality Results

Table 6 presents the results for the seven ISO 25023 quality characteristics selected for evaluation.

Table 6. System and Software characteristics across two systems at runtime

Characteristic	**Sub characteristics**	GM	***GM***	AVL	***AVL***	AVG% imp
Functional suitability	Functional completeness,and correctness	0.65	*0.75*	0.4	*0.76*	23.0
Performance efficiency	Time behavior, and Resource utilisation	0.46	*0.51*	0.38	*0.75*	21.0
Compatability	Co-existance, and Interoperability	0.5	*0.56*	0.43	*0.51*	7.0
Reliability	Maturity, Availability, and Fault tolerance	0.71	*0.90*	0.43	*0.99*	37.5
Security	Confidentiality, Integrity, Authenticity, Accountability, and Non-repudiation	0.6	*0.68*	0.50	*0.63*	10.5
Maintainability	Reusability, Analysability, Modifiability, Testability	0.45	*0.53*	0.48	*0.78*	19.0
Portability	Adaptability, Replaceability	0.31	*0.31*	0.90	*0.97*	3.5

All quality metrics show an improvement with the introduction of runtime DDI’s. ‘Reliability’ has the highest increase of 37.5 %, with one partner (AVL) stating: “Reliability measure has improved due to decrease in failure per defined period, and improved failure avoidance. On the other hand, mean notification time has slightly increased."

The 'Portability' measures had the lowest increase of 3.5 % in the evaluations. For the GM system no improvement was observed (this is the same as for the design time DDI, see statement Section 6.2). In the AVL system only a minor improvement (0.07) of the 'Portability' characteristic could be observed. This is as a result of the 'Replaceability' subcharacteristic improving when using the runtime DDI.

The average improvement, with the introduction of runtime DDI's, for all 'System and Software' quality metrics is calculated to be 17.4%.

7.3 ISO 25024 Data Quality Results

Table 7 presents the runtimeevaluation results for the ten ISO 25024 'Data Quality' characteristics. All characteristics show an improvement with the introduction of runtime DDI for the two evaluated systems (GM and AVL). The 'Data consistency' measures indicate the highest improvement of 47.5 % in average. This results mostly from the significant improvement of the 'Data consistency' characteristic in the AVL use case. AVL stated: "The analysis done with the DDI concept has shown that the percentage of consistency of data format is increased, the risk of having inconsistent data itself has decreased and a consistent platooning architecture has been obtained using the DDI concept."

Table 7. Data quality characteristics across two systems at runtime

Characteristic	**Sub characteristics**	GM	***GM***	AVL	***AVL***	**AVG% imp**
Data accuracy	Syntactic accuracy	0.6	*0.85*	0.40	*0.90*	37.5
Completeness	Record, and Attribute completeness	0.55	*0.61*	0.50	*0.95*	25.5
Consistency	Data format and Architecture consistency, and Risk of data inconsistency	0.38	*0.63*	0.30	*1.00*	47.5
Credibility	Values, Source, and Data model credibility	0.6	*0.85*	0.37	*0.92*	40.0
Compliance	Regulatory compliance of value, and Technology	0.54	*0.65*	0.65	*0.975*	21.8
Confidentiality	Encryption, and Non-vulnerability	0.66	*0.66*	0.63	*0.85*	11.0
Precision	Data values, and Format precision	0.48	*0.51*	0.65	*1.00*	19.0

Traceability	Traceability of data values, Data value Traceability	0.5	*0.64*	0.78	*0.98*	17.0
Understandability	Semantic understandability	0.55	*0.83*	0.70	*0.95*	26.5
Portability	Data portability ratio	0.48	*0.63*	0.60	*0.90*	22.5

The ‘Data confidentiality’ metric indicates the lowest improvement of 11.0 % in average. While in the AVL use case a slight improvement in terms of ‘Data confidentiality’ is observed, for the GM system no improved was observed. This is the same as for the design time DDI evaluation (see table 4 in Section 6.3). GM stated: “DDI concept guarantees only the same level of data confidentiality at runtime, but can help selecting at design time the best security solution to satisfy confidentiality requirements.”

The average improvement, with the introduction of runtime DDI’s, for all Data quality metrics is calculated to be 26.8 %.

8 Conclusion

The selected metrics for measuring the impact of the DDI were chosen mainly due to their relevance to assessing the impact of the DDI. The results of the evaluation, at both design time and runtime, indicate that applying the DDI has made significant improvements in the quality of each system, from an end user and from a system and data perspective.

For the ‘Quality in Use’ metrics the average improvement at design time is 25.2% while at runtime it is 15.9%. For the ‘System and Software’ quality metrics the average improvement at design time is 10.9% while at runtime it is 17.4%. For the ‘Data’ quality metrics the average improvement at design time is 17.5% while at runtime it is 26.8%.

These results demonstrate the positive impact of the DDI on the dependability of CPS, with some metrics indicating substantial improvement, for example, the quality in use ‘Efficiency’ metric improved by 39% at design time while the data consistency metric improved by 47.5% at runtime.

However, the results of the evaluation also indicate that not all metrics may apply to all systems, and that not all metrics showed an improvement in all systems. For example, the Systems and Software Quality metric ‘Portability’ only showed improvement in one of the four systems at design time, and in one of the two systems at runtime. While all metrics were applied, the industry partners indicated that in some instances a relatively small number of the metrics did not apply to their system.

Acknowledgement This paper is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 732242. It is also supported in part by Science Foundation Ireland grant 13/RC/2094.

References

[1] Institute for Security & Development Policy, "Made in China 2025," 2018. [Online]. Available: https://isdp.eu/content/uploads/2018/06/Made-in-China-Backgrounder.pdf. [Accessed: 18-Feb-2020].

[2] Tim Dutton, Gaga Boskovic, Brent Barron, "Building an AI World: Report on National and Regional AI Strategies," 2018.

[3] S. C. O. A. I. of the N. S. & T. COUNCI, "The National Artificial Intelligence Research And Development Plan:2019 Update," 2019.

[4] E. Commission, "Connecting Europe Facility 2021-2027: Have your say on CEF2 Digital," 2019. [Online]. Available: https://ec.europa.eu/digital-single-market/en/news/connecting-europe-facility-2021-2027-have-your-say-cef2-digital%0D. [Accessed: 19-Feb-2020].

[5] E. Commission, "5G for Europe: An Action Plan," 2016.

[6] T. S. G. S. and S. Aspects, "Study on enhancement of 3GPP Support for 5G V2X Services, 3GPP TR 22.886 V16.2.0," 2018. [Online]. Available: https://www.tech-invite.com/3m22/tinv-3gpp-22-886.html.

[7] "DEIS." [Online]. Available: http://deis-project.eu/. [Accessed: 19-Feb-2020].

[8] A. Avižienis, J.-C. Laprie, B. Randell, and R. Jacquart, "Dependability and Its Threats: A Taxonomy," 2004, pp. 91–120.

[9] S. Wagner, *Software product quality control*, 1st ed. Berlin: Springer-Verlag Berlin Heidelberg, 2013.

[10] R. Lincke and W. Lowe, "Validation of a Standard and Metric Based Software Quality Model : Creating the Prerequisites for Experimentation," The KK foundation, Sweden, 2007.

[11] B. Boehm, J. Brown, and M. Lipow, "Quantitative Evaluation of Software Quality," in *ICSE '76 Proceedings of the 2nd international conference on Software engineering*, 1976, pp. 592–605.

[12] J. A. McCall, P. K. Richards, and G. F. Walters, "Factors in software quality," Griffiths Air Force Base, N.Y. : Rome Air Development Center Air Force Systems Command., 1977.

[13] S. Dubey, G. Soumi, and R. Ajay, “Comparision of Software Quality Models: An Analytical Approach,” *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 2, pp. 111–119, 2012.

[14] J. Miguel, D. Mauricio, and G. Rodriguez, “A Review of Software Quality Models for the Evaluation of Software Products,” *Int. J. Softw. Eng. Appl.*, vol. 5, no. 6, 2014.

[15] R. Grady, *Practical software metrics for project management and process improvement*. Prentice-Hall, 1992.

[16] I. Singh, “Different Software Quality Models,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 1, no. 5, pp. 438–442, 2013.

[17] R. Dromey, “A Model for Software Product Quality,” *IEEE Trans. Softw. Eng.*, vol. 21, no. 2, pp. 146–162, 1995, doi: 10.1109/32.345830.

[18] ISO, “ISO 9126 Software Engineering - Product Quality .” 1991.

[19] ISO, “Software Engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Guide to SQuaRE.” 2005.

[20] ISO, “Ssytem and software engineering - System and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models.” 2011.

[21] G. Samarthyam, G. Suryanarayana, T. Sharma, and S. Gupta, “MIDAS: A Design Quality Assessment Method for Industrial Software,” *ICSE 2013 : Software Engineering in Practice*. San Francisco, pp. 911–920, 2013.

[22] A. Al-Badareen, “Software Quality Evaluation: User’s View,” *Int. J. Appl. Math. Informatics*, vol. 5, no. 3, pp. 200–207, 2011.

[23] S. Dubey, S. Ghosh, and A. Rana, “Comparison of Software Quality Models: An Analytical Approach,” *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 2, pp. 111–119, 2012.

[24] ISO, “ISO/IEC 25010: 2011Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models.” ISo, Geneva, Switzerland, 2011.

[25] J. Reich, M. Zeller, and D. Schneider, “Automated Evidence Analysis of Safety Arguments Using Digital Dependability Identities,” in *Computer Safety, Reliability, and Security. SAFECOMP 2019. Lecture Notes in Computer Science, vol 11698*, 2019.

[26] DEIS Consortium, “DEIS Project,” 2019. [Online]. Available: https://github.com/DEIS-Project-EU.

[27] ISO, “ISO/IEC 25022:2016 Systems and software engineering -- Systems and software quality requirements and evaluation (SQuaRE) -- Measurement of quality in use.” 2016.

[28] ISO, “ISO/IEC 25023:2016 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Measurement of system and software product quality.” 2016.

[29] ISO, “ISO/IEC 25024:2015 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Measurement of data quality.” 2015.

[30] UNISIG, “Safety Requirements for the Technical Interoperability of ETCS in

Levels 1 & 2," 2015. [Online]. Available: https://www.era.europa.eu/sites/default/files/filesystem/ertms/ccs_tsi_annex_a_-_mandatory_specifications/set_of_specifications_2_etcs_b3_mr1_gsm-r_b1/index027_-_subset-091_v340.pdf.