33rd Annual INCOSE international symposium

hybrid event

Honolulu, HI, USA
July 15 - 20, 2023

Modeling & Simulation SPICE

Assessing the Capability of Credible Simulation Processes

Frank Eichenseer
Systems Engineer & Researcher | SETLabs Research GmbH, Munich, Germany

About me…

Frank Eichenseer (* 1992, Ingolstadt, Bavaria, Germany)

- 2019: **Master's Thesis – FUSE@AUDI**
  AUDI AG, Ingolstadt, Germany
- 2020: **M. Eng. "Systems Engineering"**
  University of Applied Sciences Landshut, Germany
- '20-'22: **Consultant – Systems Engineering, Processes & Methods**
  iCONDU GmbH, Ingolstadt, Germany
- '21-'23: **Systems Engineer & Researcher**
  SETLabs Research GmbH, Munich, Germany

Certificates:

- **INCOSE ASEP™** (2019 - 2024)
- **Certified Systems Engineer (GfSE)® - Level C** (2019 - 2024)

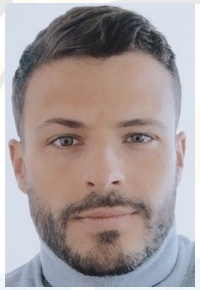
Co-Authors

Hans-Martin Heinkel
Robert Bosch GmbH
Stuttgart, Germany

Dr. Michael Holzner
iCONDU GmbH
Ingolstadt, Germany

Dr. Martin Benedikt
Virtual Vehicle Research GmbH
Graz, Austria

Christoph Stadler
AUDI AG
Ingolstadt, Germany



Maurizio Ahmann
SETLabs Research GmbH
Munich, Germany

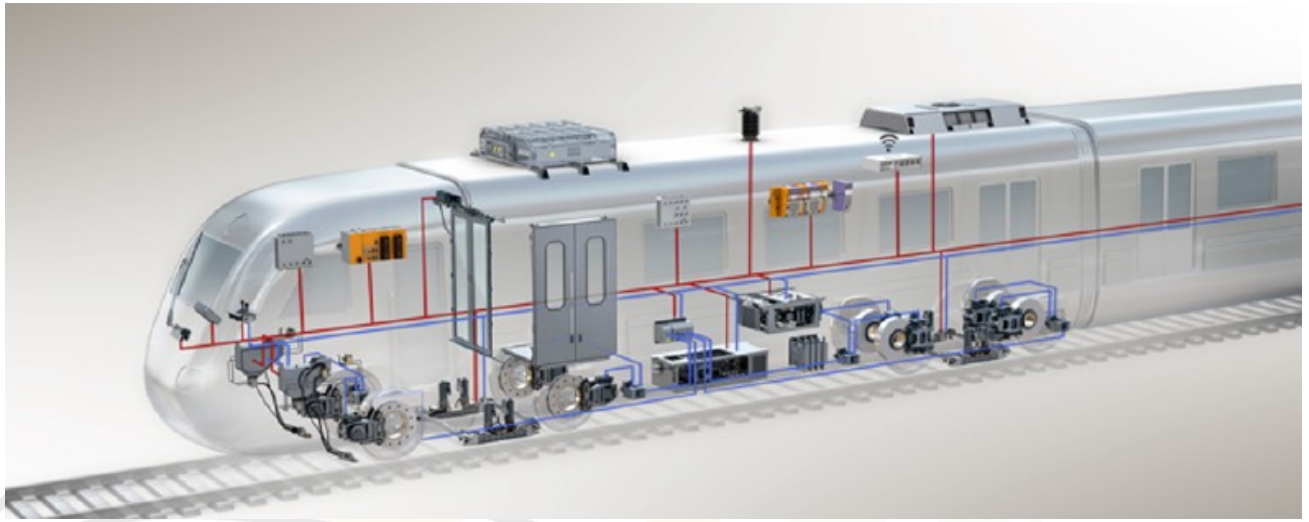
Modeling & Simulation on the Rise

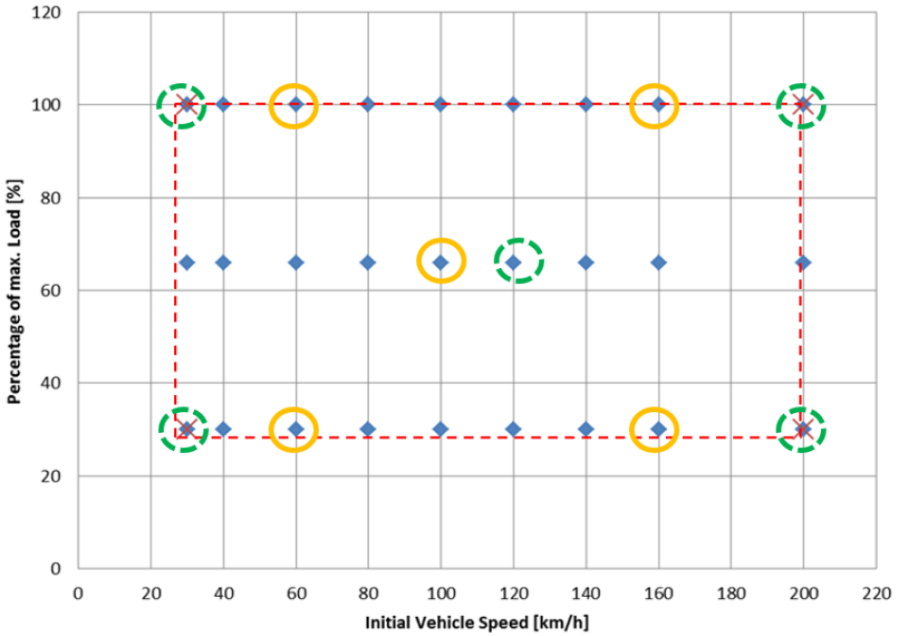
Industry Potentials and Regulations

Rail Domain

Example: **Virtual Validation of Railway Vehicle Braking Performance**

- Approval regarding **EN16185** (test specification) and **EN16834** (brake performance evaluation)





- Real tests reduced by **63%** (!) from 27 to 10
- Regulative requirements on use of M&S incoming (**prEN17833**)

Source: Krammer, M.: Virtual Validation for Certification of Railway Vehicle Braking Performance, EuroSPI publication 08/2022; https://link.springer.com/chapter/10.1007/978-3-031-15559-8_10

Automotive Domain

Example: **Virtual Testing for Automotive Vehicle Safety**

To deliver on these objectives, Euro NCAP must also innovate the way testing is performed. It will further step up the use of **virtual testing,** complementary to crash tests in the laboratory, tests on the track and on the road. Subsystem testing, for instance using a body-in-white on a sled, can also provide additional insights in performance of restraint systems under more variable conditions. While these tests add value and will help keep the programme feasible and manageable, they require trust and cooperation with the vehicle manufacturer, and their outcome and application in the rating must be carefully weighed.

Source: Euro NCAP: Euro NCAP Vision 2030 – A Safer Future for Mobility; https://cdn.euroncap.com/media/74468/euro-ncap-roadmap-vision-2030.pdf

Automotive Domain

Example: **Type Approval for Automated Driving Systems - EU Regulation**

COMMISSION IMPLEMENTING REGULATION (EU) 2022/1426

PART 4

PRINCIPLES FOR CREDIBILITY ASSESSMENT FOR USING VIRTUAL TOOLCHAIN IN ADS VALIDATION

General

The credibility can be achieved by investigating and assessing five properties of Modelling and Simulation (M&S):

(a) capability – what can the M&S do, and what the risks are associated with it;

(b) accuracy – how well does M&S reproduce the target data;

(c) correctness – how sound & robust are M&S data and algorithms;

(d) usability – what training and experience is needed.

(e) fit for purpose – how suitable is the M&S for the ODD and ADS assessment.

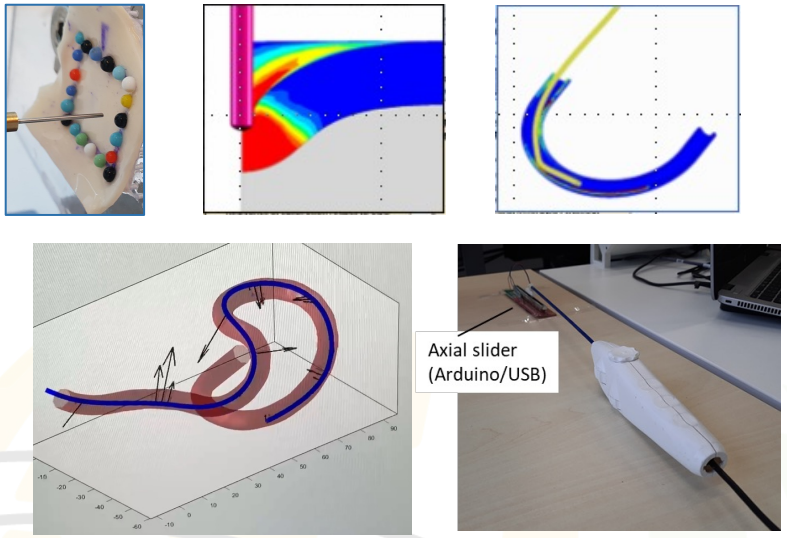
At the same time, the credibility assessment framework shall be general enough to be used for different M&S types and applications. However, the goal is complicated by the broad differences between ADS features and the variety of M&S types and applications. These considerations require a (risk-based/informed) credibility assessment framework relevant and appropriate to all M&S applications.

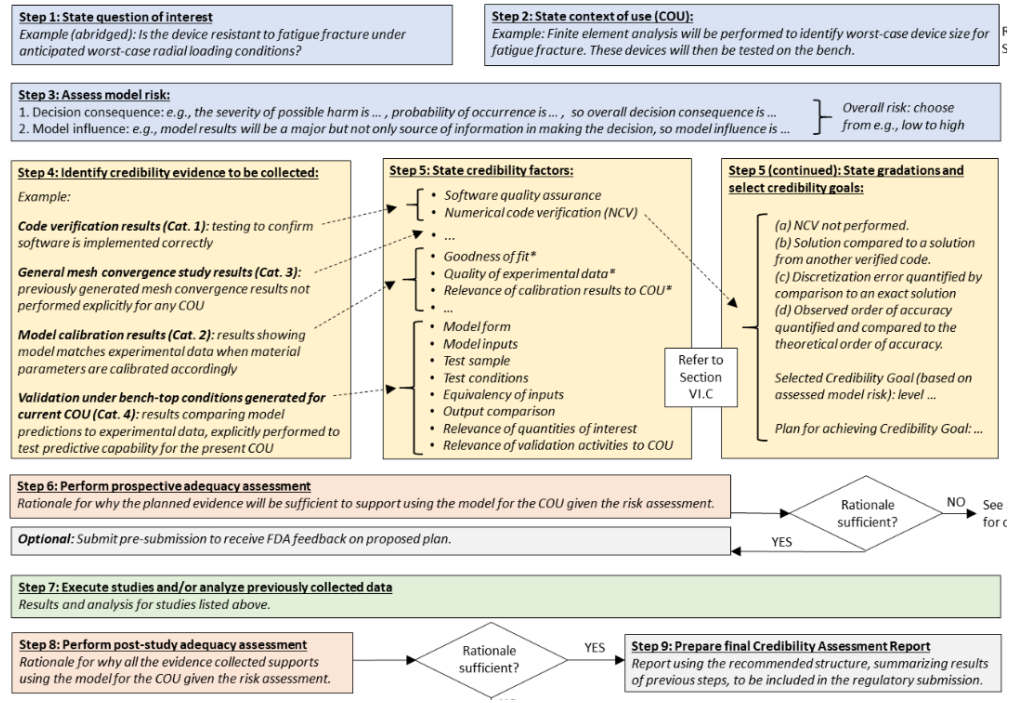
Source: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1426

Healthcare Domain

Example: **Virtual Verification of Medical Devices for Regulatory Approval**

- Regulatory requirements on use of M&S (**FDA-2021-D-0980**)





- Accelerate pre-market device evaluation (<< 8 years)
- Identify adverse events often hardly detectable

Source: FDA: Draft Guidance, Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions, 12/2021; https://www.fda.gov/media/154985/download

M&S - A Sleeping Giant

Obstacles & Barriers w.r.t. Market Uptake

Why M&S does not unfold its Potential…

In a Nutshell

- **(Too) High-level Guidelines & Regulations**
  - Mostly high-level credibility requirements for simulation use & quality defined
  - NASA, FDA, EuroNCAP, EU, etc.
  - Company-specific implementation is difficult
- **Focus on Artifact Quality Approaches**
  - Most available approaches focus more on the V&V of the M&S results than on the procedure
- **Missing Frameworks for Virtual Development**
  - Development processes are still tailored for real testing; simulation often only by-product
  - Often punctual, department-specific use of M&S; integration is challenging
  - Lack of M&S performance control & improvements
- **Missing Trust in Modeling & Simulation**
  - Industry prefers to play it safe
  - As a result, mostly expert-driven M&S assessment, according to real measurements

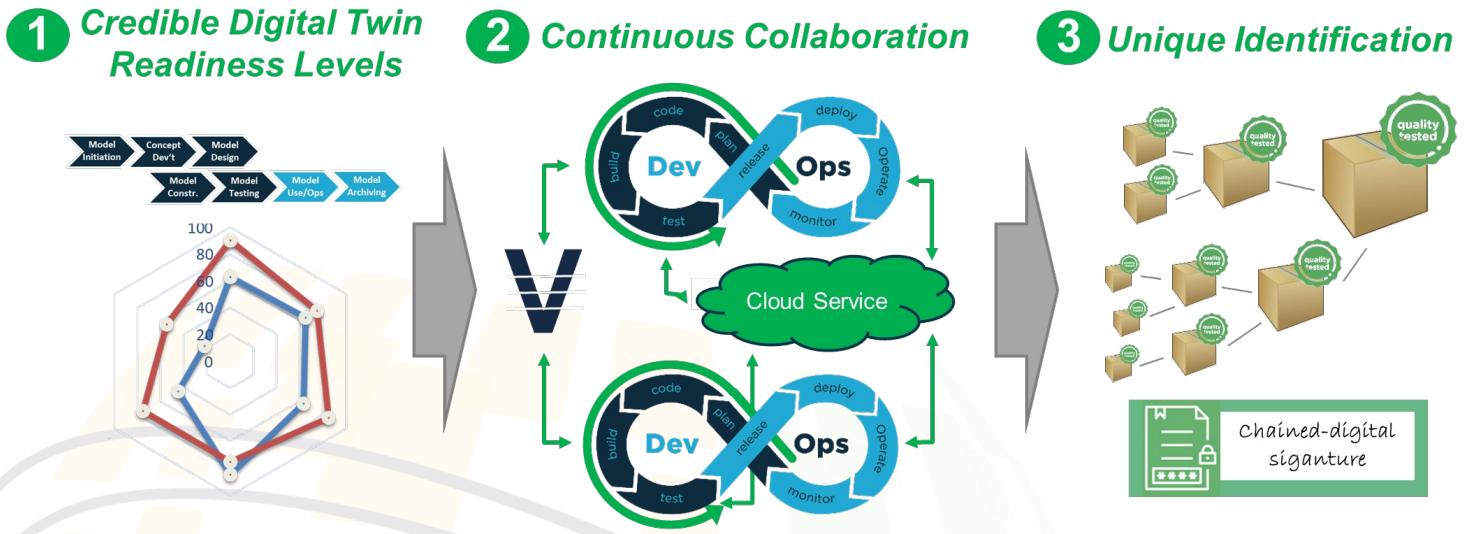
How to unleash M&S Potentials

Credibility through Consistency

The ITEA3 UPSIM Project

Overview

Solution Approach:



Expected Results:

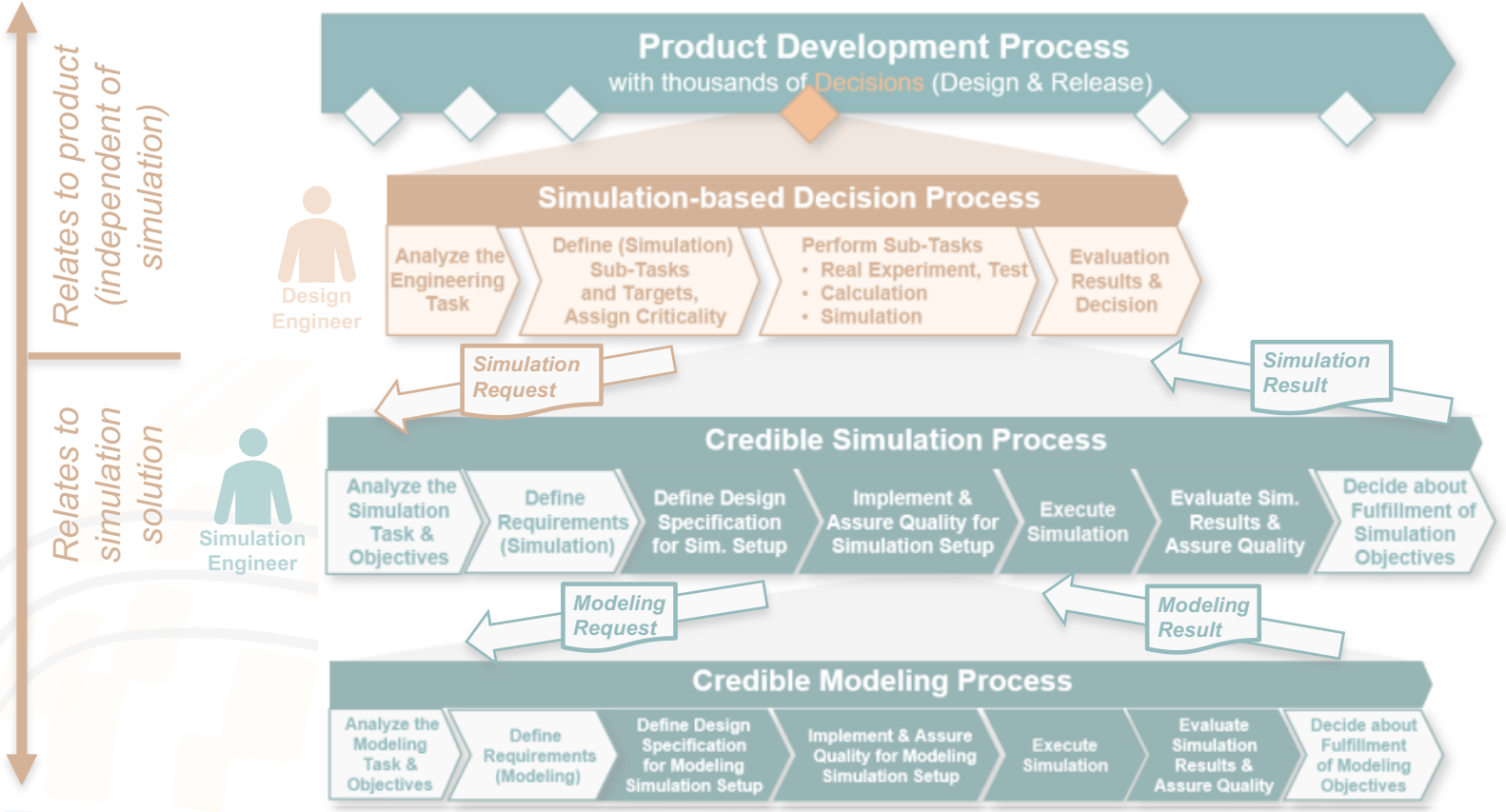
Open Access **Standards & Tooling** for the quantification of **Simulation Credibility** via a reasonable **Credibility Assessment Framework** and application-specific **Best Practices.**

Facts

Framework:	EUREKA / ITEA3 – Call 6
Duration:	10/2020 – (09)12/2023
Total Budget:	14,268 M€
Partners:	29
Countries:	AT, DE, DK, NL, RO, UK
Coordinator:	VIRTUAL VEHICLE
Website:	www.upsim-project.eu

Credible Simulation Process Framework

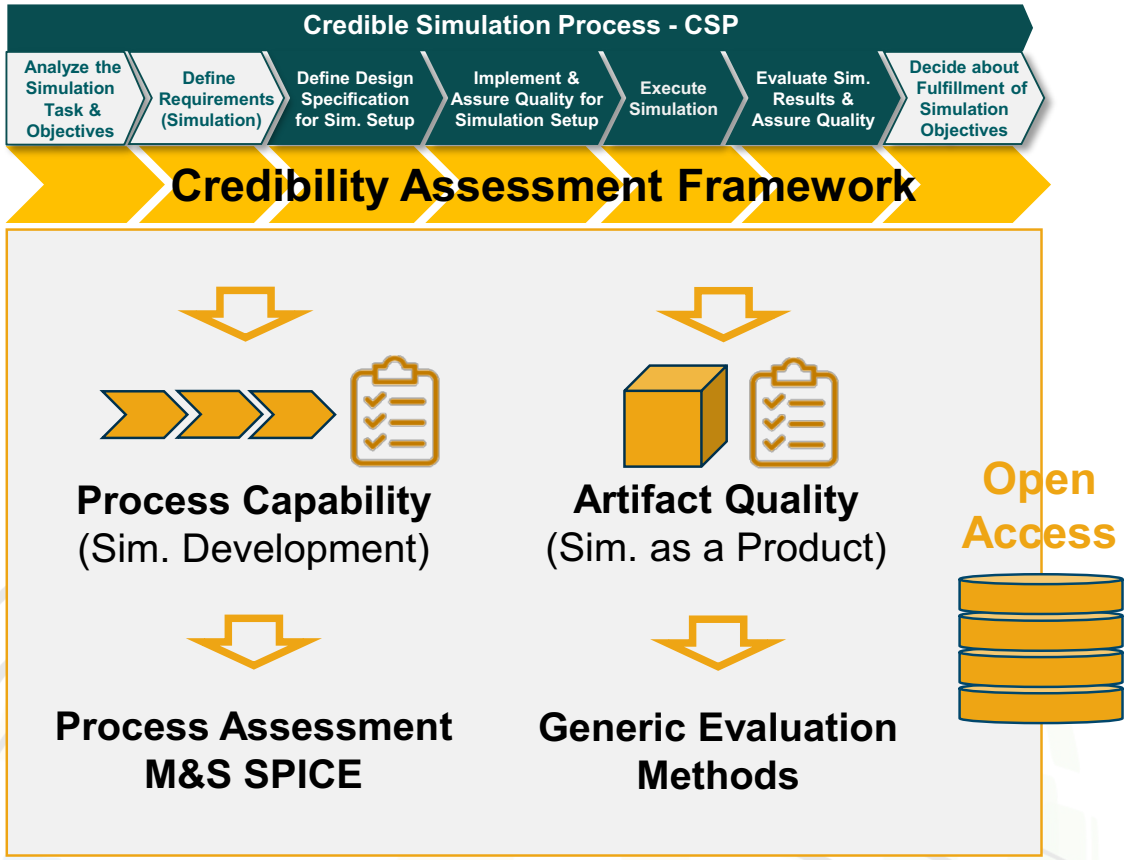
Reference Model



Source: Heinkel, HM et al.: Building Blocks for Simulation-Based Cooperation between Partners, Presentation, prostep ivip Symposium 2023;
© 2023, prostep ivip e. V.

UPSIM Credibility Assessment Framework

Big Picture

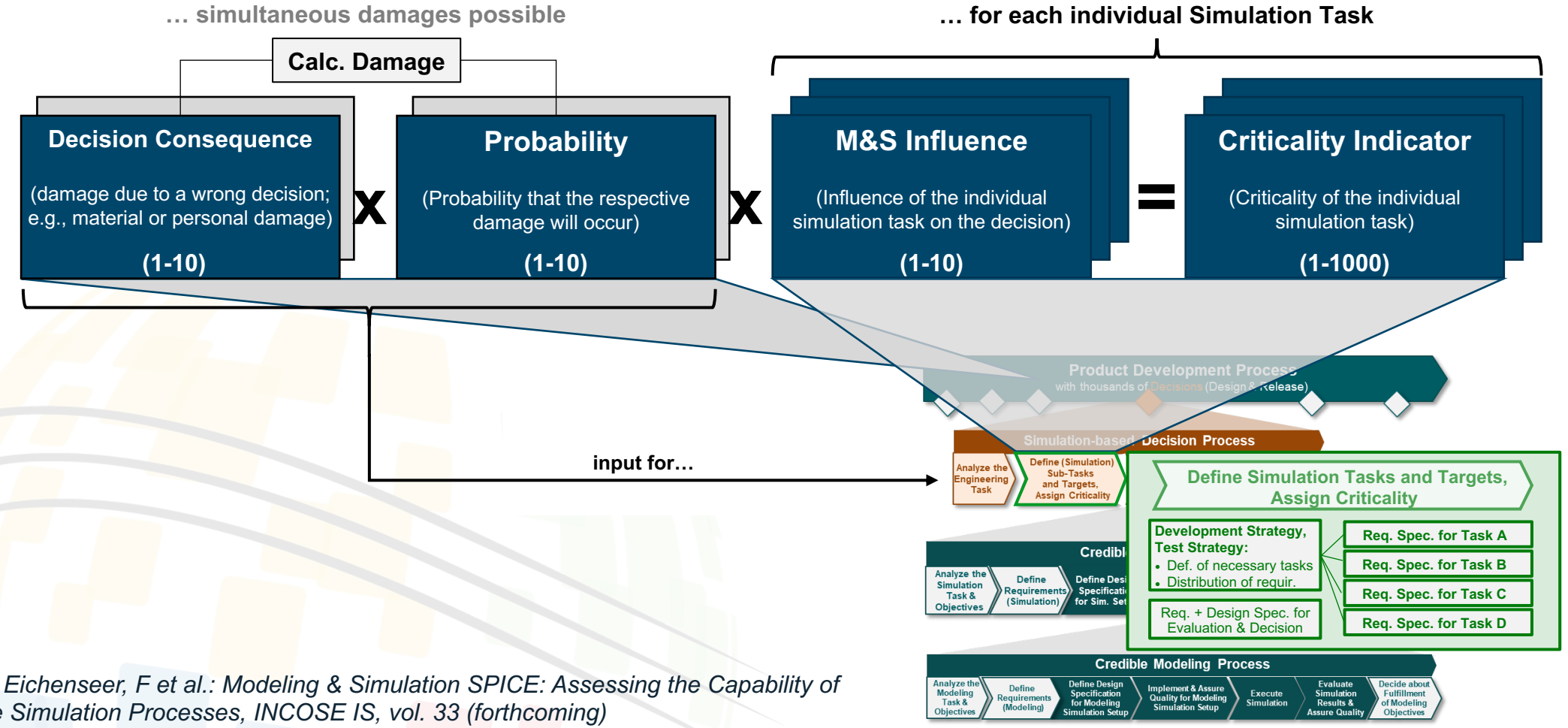


- **The CSP-Framework as Foundation**
  UPSIM has chosen the CSP-Framework as a suitable foundation for the UPSIM CAF
- **Intentional Inclusion of the Process Dimension**
  Additional focus on the M&S procedure w.r.t. Model & Simulation development processes
- **Holistic Framework for Virtual Development**
  Consideration of both **implicit** (process) and **explicit** (artifact) credibility criteria

Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

UPSIM Criticality Indicator Method

A Risk-based Approach



Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

UPSIM Criticality Indicator Scales (1/3)

Decision Consequence

Level		Label	Consequence for Product and/or Humans and/or Business in case of failure
Safety	10	Fatal	**Product:** Fatal impact on user and uninvolved persons. Causes the operational safety to be violated without a prior warning, so that government regulations will be violated. 100% scrap, no rework possible. **Human**: Fatal injury that leads to permanent disability or death. **Business**: Fatal detriment to organizational business. Causes major project cancellation and/or has significant influence on the overall organizational business.
	9	Severe	**Product:** Severe impact on user and uninvolved persons. Causes the operational safety to be endangered with a prior warning, so that government regulations will be violated if the bad event cannot be prevented after the warning. 100 % scrap, no rework possible if bad event will not be prevented. **Human**: Severe injury that may lead to temporal disability or death. **Business**: Severe detriment to project and overall organizational business. Causes major project reductions.
	8	Serious	**Product**: Serious deterioration of product safety. Even may call for total rework, up to 100% scrap possible. However, no violation of regulation or government norms. **Human**: Serious injury or severe occupational illness leading to hospitalization and typically associated with claims for compensation. **Business**: Serious detriment to full completion of project. Typically, a significant scope of the project will be cancelled for being able to deploy a distinctive sub-scope of the project.
Functionality	7	Significant	**Product**: Significant deterioration of product functionality, including non-functioning or inoperable. Serious disruption of work. Some part may completely scrap. **Human**: Significant injury or significant occupational illness that makes a doctor's visit necessary immediately after the event. **Business**: Significant detriment to project execution. Requires cancellation of smaller sub-projects or partial scopes of the project.
	6	Moderate	**Product**: Moderate deterioration of product functionality. May be noticed by nearly all users (more than 95%) with annoyance and discomfort. Some part may call for rework, scrap possible. **Human:** Moderate injury or little occupational illness that makes a doctor's visit necessary within one week after the event. **Business:** Moderate business detriment. Requires the project's business plan to be refined.
	5	Minor	**Product**: Minor deterioration of product functionality. May be noticed by most of users (around 80%). Unscheduled rework. **Human:** Minor injury or minor occupational illness without the need for a doctor's visit. **Business**: Minor business detriment, but outside calculations.
Performance	4	Notable	**Product:** Notable faults to detriment of product performance. May be noticed by around the half of all users. Also, minor rework may be called for. **Human:** Notable human detriment like significant discomfort. **Business:** Notable business detriment, but within calculations.
	3	Inconsequential	**Product**: Inconsequential faults to detriment of product performance. May be noticed by some users (around 20%). **Human**: Inconsequential human detriment like significant annoyance and/or little discomfort. **Business**: Inconsequential business detriment.
	2	Negligible	**Product**: Negligible faults to detriment of product performance. May only be noticed by attentive users (less than 5%). **Human**: Negligible human detriment like little annoyance. **Business**: Negligible business detriment.
	1	Nonexistent	No **human**, **product** or **business** detriment.

Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

UPSIM Criticality Indicator Scales (2/3)

Probability

Level		Label	Probability that event from decision consequence happens at least once during lifetime
Likely	10	**Certain**	Event occurrence is almost certain (p ≥ 1:2)
	9	**Frequently**	Very high number of event likely (p ≥ 1:3)
	8	**Repeatedly**	High number of event likely (p ≥ 1:8)
Occasionally	7	**Periodically**	Moderately high number of event likely (p ≥ 1:20)
	6	**Notable**	Medium number of event occurrence likely (p ≥ 1:80)
	5	**Erratically**	Occasional event likely (p ≥ 1:400)
Unlikely	4	**Scarce**	Very low likelihood of event (p ≥ 1:2K)
	3	**Rarely**	Rare likelihood of event (p ≥ 1:15K)
	2	**Negligible**	Extremely low chances of event occurrence (p ≥ 1:150K)
	1	**Impossible**	Event will practically not occur or will be occuring very rarely (p ≤ 1:1.5M)

Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

UPSIM Criticality Indicator Scales (3/3)

M&S Influence

Level		Label	Influence on decision of higher-level engineering task
Controlling	10	Critical	The simulation task is critical for the decision, as it has prevailing influence, meaning that it can veto the contribution of other sub-tasks for the decision.
	9	Dominating	The decision is highly dependent on the simulation task, meaning that it is one of the dominating sub-tasks for the decision, but without the ability to veto the decision.
	8	Governing	The decision is dependent on the result of the simulation task, meaning that it is among the most important sub-tasks for the decision, but without dominating character.
Tantamount	7	Important	The simulation task will have important influence, meaning that it is among the important sub-tasks for the decision, but not among the most-important sub-tasks.
	6	Significant	The simulation task will have significant influence, meaning that it will be taken into account for the decision with significant importance compared to other sub-tasks in the decision-making process, without belonging to the important sub-tasks.
	5	Evident	The simulation task will have evident influence, meaning that it will be taken into account for the decision without significant importance compared to other sub-tasks in the decision-making process.
Supporting	4	Considerable	The simulation task will only have supporting character, but has considerable supporting influence, meaning that it can support the other sub-tasks in the decision-making process as one of the most important supporting tasks.
	3	Notable	The simulation task will only have supporting character and has notable supporting influence, meaning that it can support the other sub-tasks in the decision-making process as one of the moderate important supporting tasks.
	2	Negligible	The simulation task will only have supporting character and it has negligible supporting influence, meaning that it can support the other sub-tasks in the decision-making process as one of the less important supporting tasks.
	1	Insignificant	The decision will be made exclusively based on results from other tasks. The simulation task will only have confirming character.

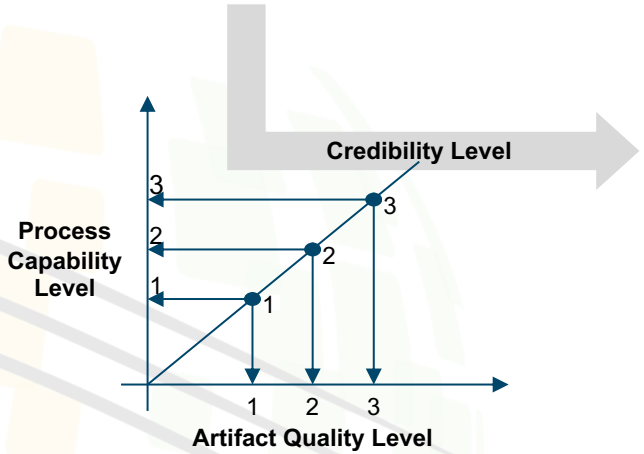
Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

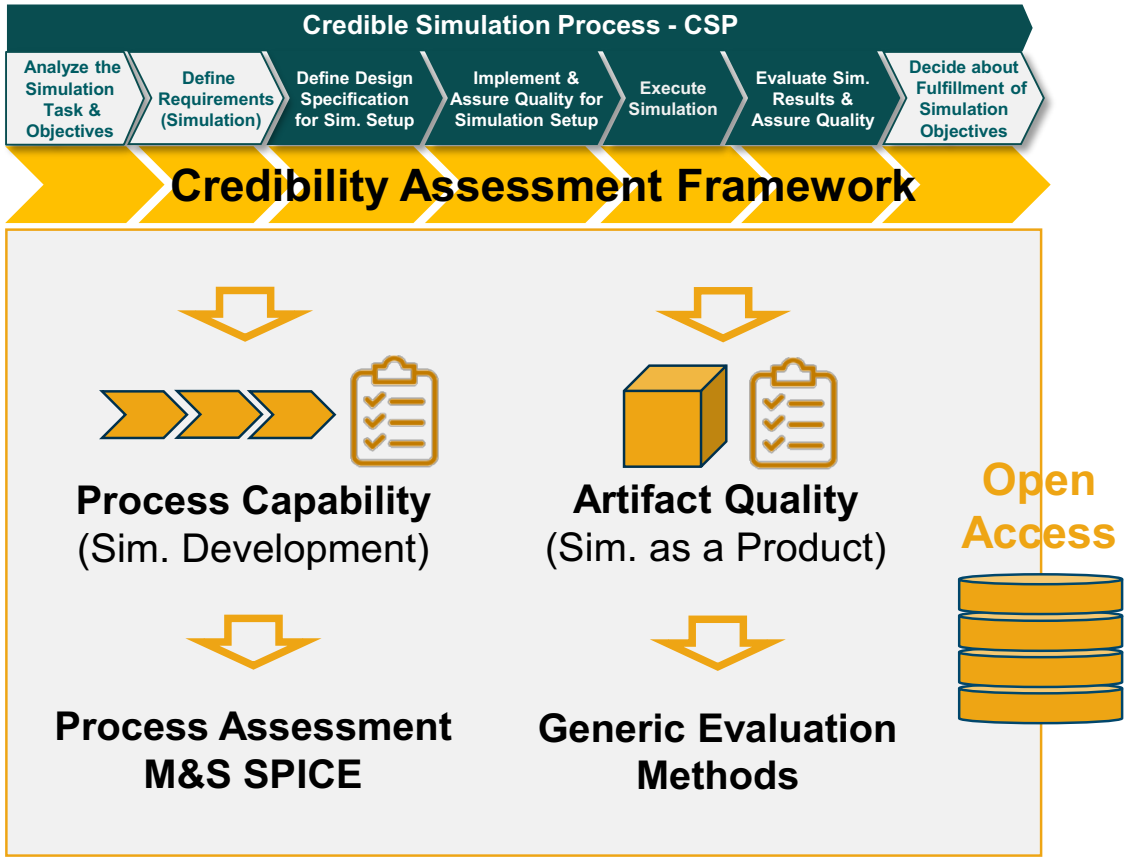
From Criticality to M&S Credibility…

Ensuring Adequacy of Credibility Assurance Measures

Criticality Indicator	Credibility Level
≤1000	3 – High Credibility
<250	2 – Medium Credibility
<50	1 – Low Credibility
<10	0 – No Credibility

Initial proposal; to be clarified/specified by collaborating partners





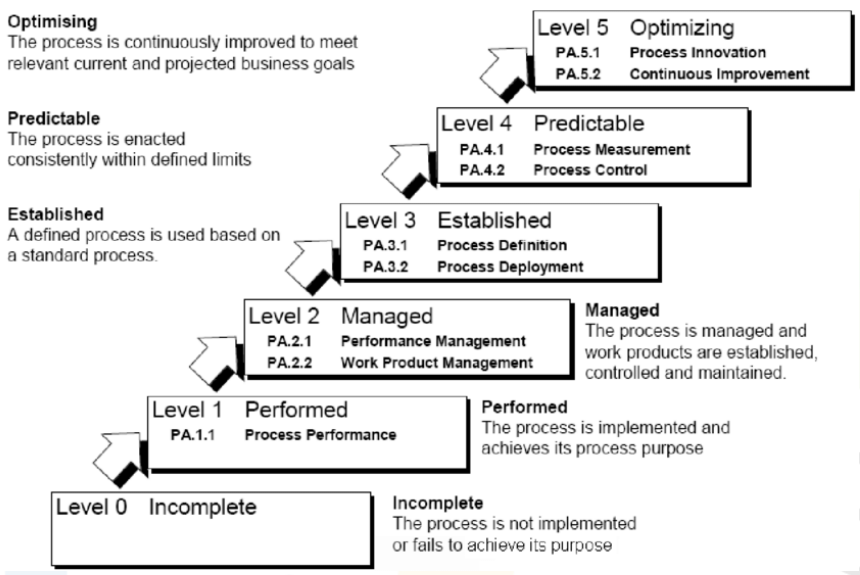
Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

Automotive SPICE®

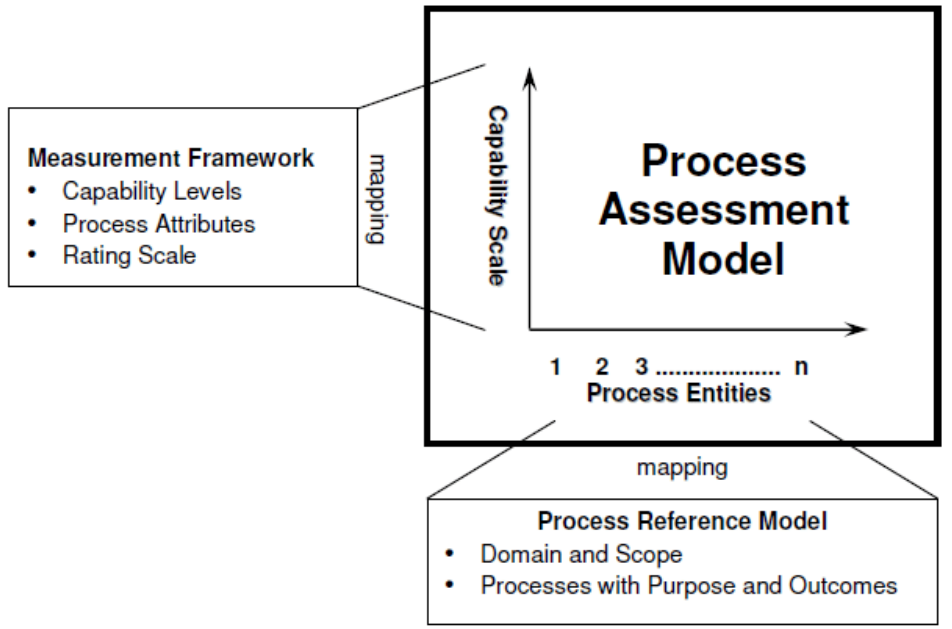
Process Capability Assessment for Automotive Systems & Software

Automotive SPICE® is based on the ISO 330xx Series

- ISO 330xx revises the ISO/IEC 15504 (SPICE) series, initiated by an EU-funded project.
  - SPICE: **S**oftware **P**rocess **I**mprovement and **C**apability d**E**termination
- provides a framework for the assessment of Process Capability characteristics and organizational maturity.



Source: ISO/IEC 15504

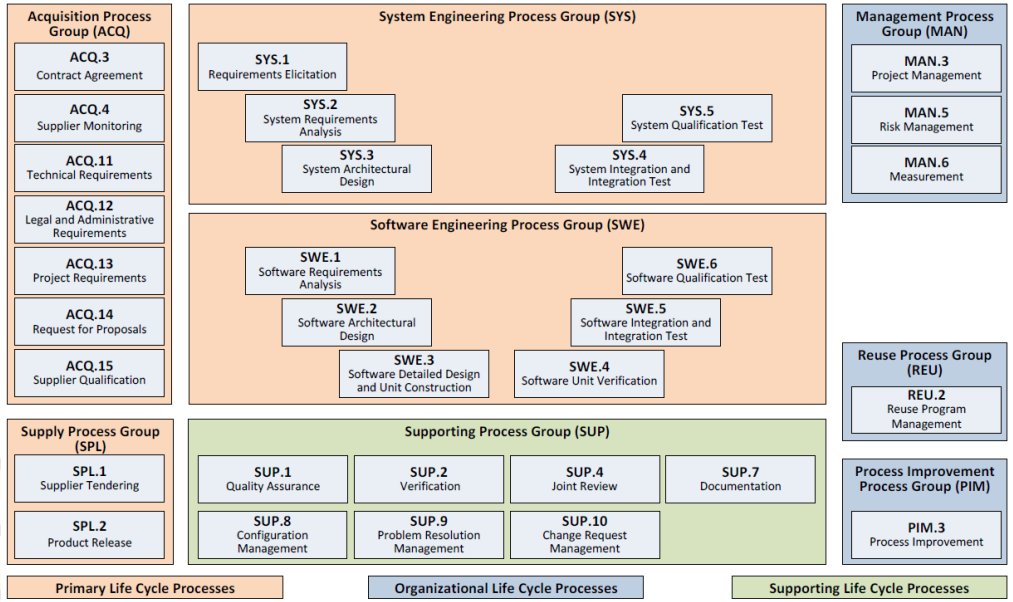


Source: ISO/IEC 15504

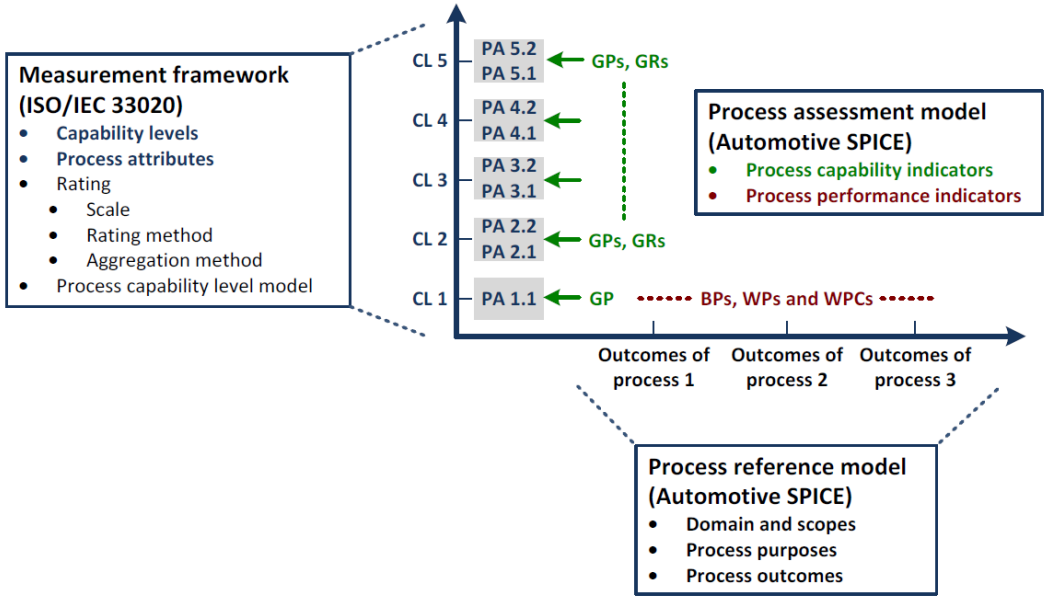
Automotive SPICE®

Process Capability Assessment for Automotive Systems & Software

Process Reference Model (PRM)



Process Assessment Model (PAM)

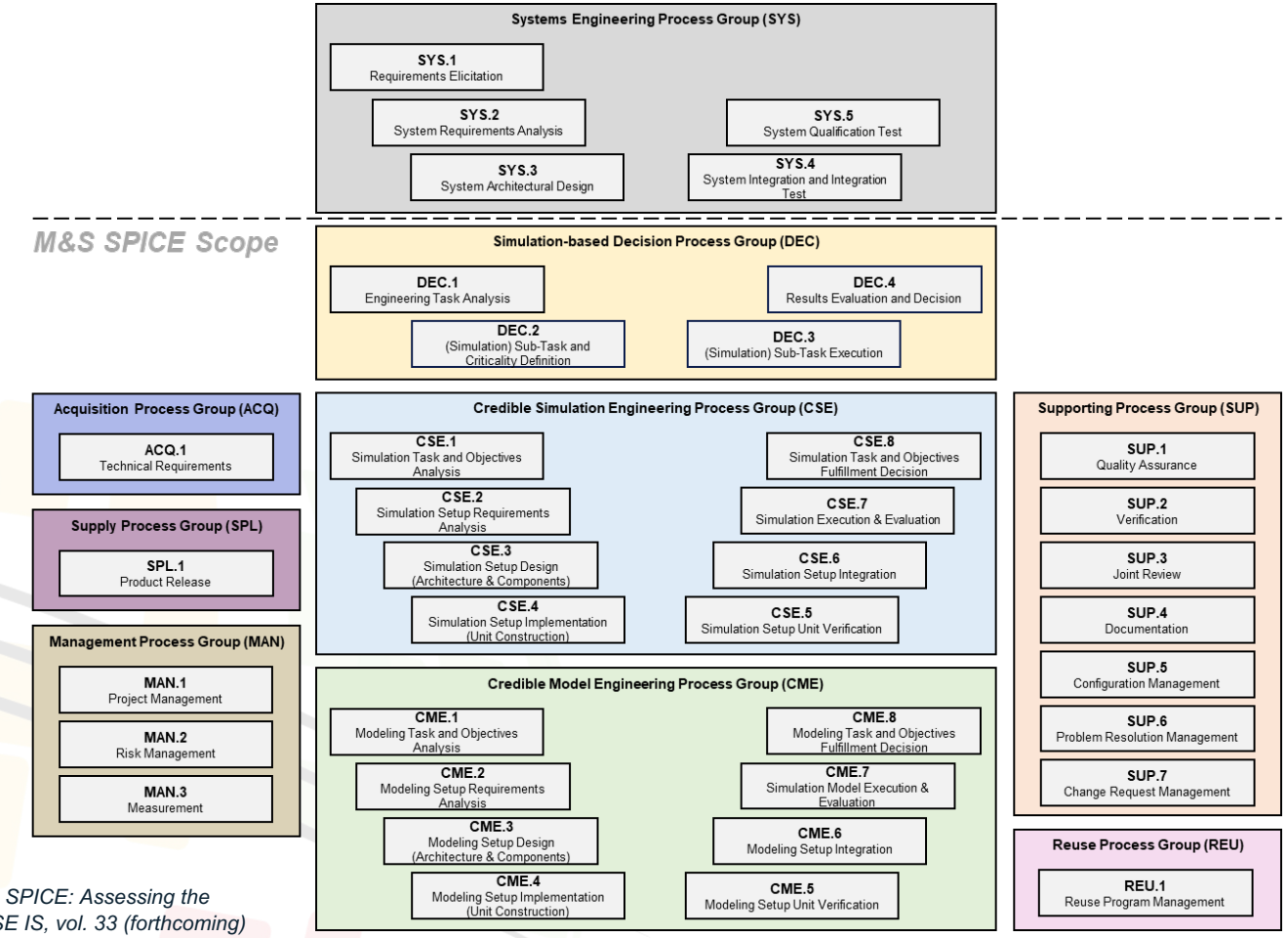


© 2017, VDA QMC | Source: Automotive SPICE® Process Reference and Assessment Model, Version 3.1, https://www.automotivespice.com/download/

ASPICE incorporates the concepts and principles of the ISO 330xx Series and adapts them to address the unique characteristics, requirements, and challenges of developing automotive systems and software.

Modeling & Simulation SPICE

Process Reference Model



Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

Example: Criticality Indicator Determination

Homologation of an Automated Lane Keeping System

Product to be developed: Conditional Automated Vehicle Variant 1

Product Subsystem: Automated Lane Keeping System (ALKS)

Engineering Task: Homologation of the ALKS

Decision to be made: ALKS safety approval, according to UNECE R157

Simulation Task: Generation of simulation results on varied test scenarios, according to UNECE R157, Annex 5, Section 4.1.2.

Excessive Adjustment

Potential Hazard: Excessive lateral adjustment resulting in lateral commuting

Operational Situation: Driving at medium speed (v < 60 kph) on undivided arterial roadway; no pedestrians present

Potential Crash Scenario: No collision, but commuting of vehicle

Criticality Indicator: Decision Consequence (**6**), Probability (**4**), M&S Influence (**10**); **240**

Credibility Level: **Level 2**

Criticality Indicator	Credibility Level
≤1000	3 – High Credibility
<250	2 – Medium Credibility
<50	1 – Low Credibility
<10	0 – No Credibility

Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

Example: Process Assessment – Level 1

PRM & PAM Interplay

DEC.1 – Engineering Task Analysis

The purpose of the Engineering Task Analysis Process is to determine the information required for the execution of the respective higher-level decision or engineering task and to provide this information throughout the subsequent processes. The criticality determination regarding the decision consequence of the engineering task is actually performed at this point according to the workflow of the CSP Framework. In M&S SPICE, however, this part of the criticality determination is allocated to the management process MAN.3 "Risk Management".

Outcomes – as a result of successful implementation of this process

1. the higher-level decision and engineering task are analyzed; and
2. a specific, detailed description of the system under test (product), the relevant environment (of the real-world system), development objectives, requirements, KPI and other relevant criteria is available.

Output work products

DEC-1-1	Decision and engineering task profile	Outcome 1
DEC-1-2	Relevant real-world system (product) specification	Outcome 2

Base practices 1-2

BP1 Analyze the decision and engineering task. Outcome 1

1) *Information and description of the overall development project, the decision to be made and the corresponding engineering task of the higher-level development process as well as the general requirements and objectives, e.g., KPI and other criteria for the engineering task.*
2) *Project Name: Mild Hybrid Variant AAA-55, Project Number: P987658, Version: 2*
3) *Project Leader: J. Miller*
4) *Decision / Engineering Task: A variant of a mild hybrid drive based on the DDC platform is to be chosen and developed. This application is a low-cost variant.*

BP2 Specify and describe the product, environment and requirements in detail. Outcome 2

5) *Product: Mild Hybrid Drive Variant AAA-55*
6) *Environment (of the product): In this simple example, no further details are provided. For a car, the environment would be the road, climate zone, user group, frequency of use, etc.*
7) *Requirements: In this simple example, no further details are provided.*

Process Purpose

Process Goals

Process Evidence

Process Activities

Process Capability Level 1: Performed Process Source: ISO/IEC 33020

The implemented process achieves its process purpose.

The following process attribute demonstrates the achievement of this level:

PA 1.1 Process performance process attribute

The process performance attribute is a measure of the extent to which the process purpose is achieved.

As a result of full achievement of this attribute:

a. the process achieves its defined outcomes.

Generic Practices:

GP 1.1.1 Achieve the process outcomes. Achievement a

Achieve the intent of the base practices. Produce work products that evidence the process outcomes.

Generic Resources:

1. **Resources are used to achieve the intent of process specific base practices.** Achievement a

N	Not achieved	0 … 15%
P	Partially achieved	16 … 50%
L	Largely achieved	51 … 85%
F	Fully achieved	86 … 100%

Source: ISO/IEC 33020

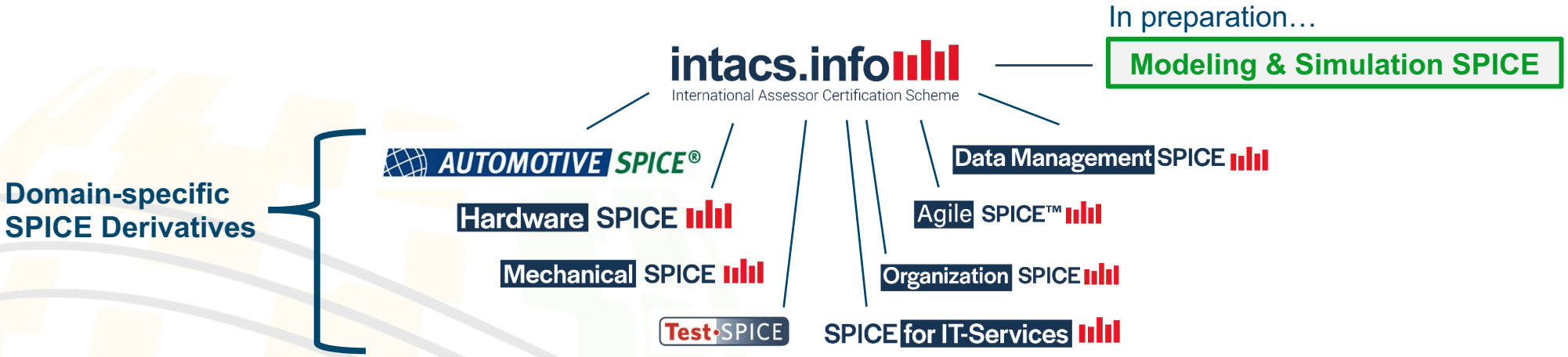
Source: Eichenseer, F et al.: Modeling & Simulation SPICE: Assessing the Capability of Credible Simulation Processes, INCOSE IS, vol. 33 (forthcoming)

Outlook

De Facto Standardization (intacs e. V.)

The intacs association is an independent not-for-profit association, ensuring high-quality assessor qualification for process capability assessment according to ISO/IEC 330xx series:

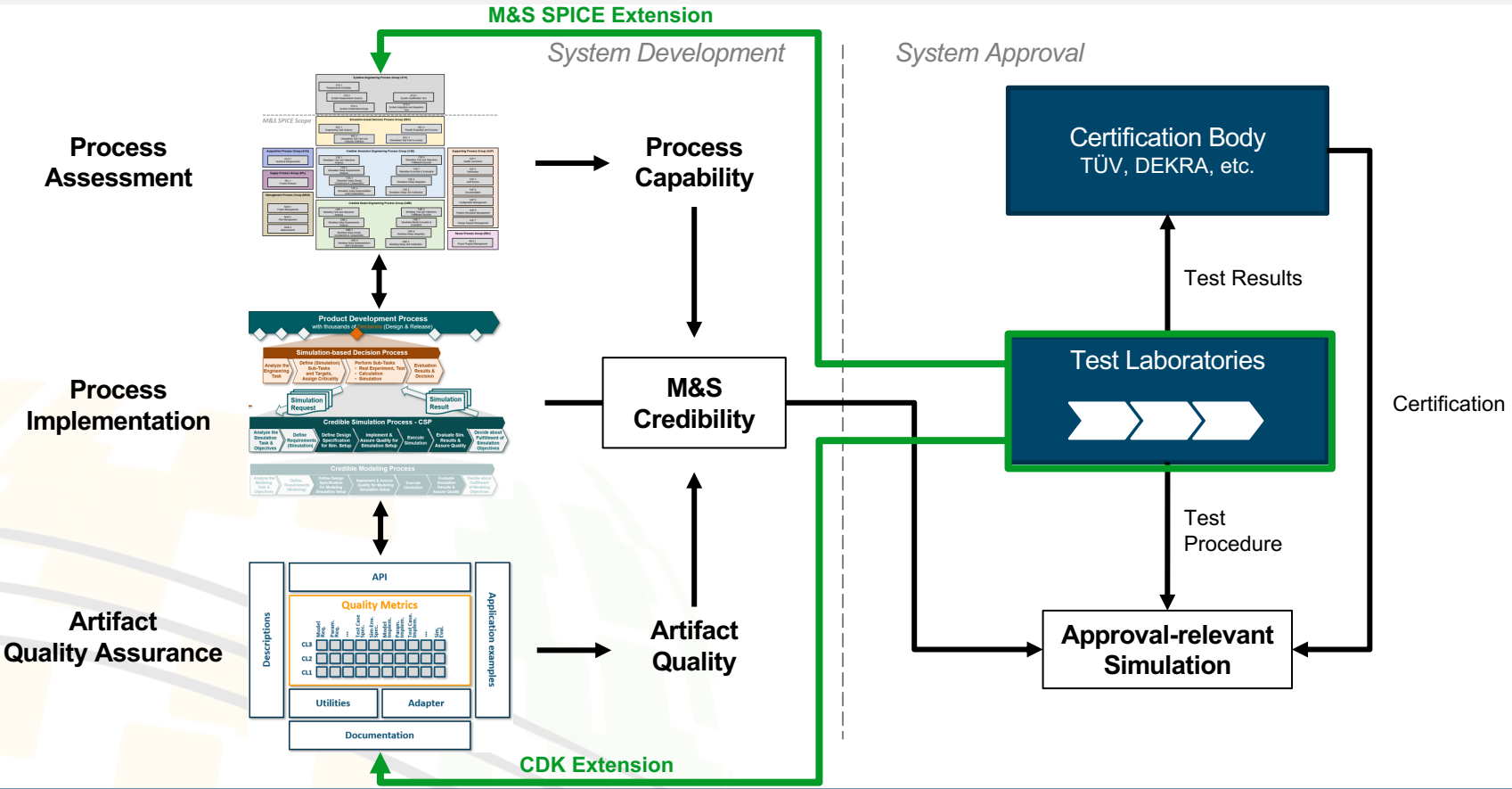
- setting training and certification standards for ISO/IEC 330xx assessors
- setting standards for maintaining assessor competence
- promoting assessment models and community interactions



We are currently in contact with the intacs association regarding a de facto standardization of M&S SPICE (incl. use of the intacs accreditation and certification scheme).

Outlook

M&S SPICE for Virtual Approval (TÜV SÜD)



We are also in contact with TÜV SÜD Germany for a collaboration / joint forces w.r.t. virtual approval, especially for the virtual approval of vehicles (automotive).

33rd Annual INCOSE international symposium

hybrid event

Honolulu, HI, USA
July 15 - 20, 2023

www.incose.org/symp2023
#INCOSEIS

Further Questions? Send me an E-Mail!

frank.eichenseer@setlabs.de