Proposal - NFDI4Ing Seed Fund

Year	**Institution**	**Research area**	**Date of application**
2023	Helmholtz-Zentrum hereon GmbH	Digital Materials-, Surface- and Corrosion Engineering	30-9-2022
Funding amount	65,588.00 €		
Term	01.01.2023 to 31.12.2023		
Contact	Dr. Daniel Höche, daniel.hoeche@hereon.de, +49(0) 4152-87 1914		
Project title	**CorWiz -** Corrosion-data wizard for smart selection of corrosion models for novel sustainable engineering concepts with steels		
Member of the NFDI e.V.	<Are you or your institution members of the NFDI e.V.>	Yes x	No ☐

Have you worked with people from the NFDI4Ing steering committee (link to overview) in the last five years?

Prof. Dr. A. Streit

Project outline (<1 page)

Damage caused by corrosion has a big impact on today's society, especially damage to critical infrastructures, for example, corrosive damage to bridges often make the news and cause high economic or even ecological damage. As such, corrosion, corrosion prevention, and corrosion data utilisation are of great importance to engineers.

Corrosion protection concepts are crucial for successful, economical, and sustainable component design and construction. They follow defined protocols and standards in safety and quality assessment design, e.g., ISO 9223 "Corrosivity of metals under atmospheric corrosion". Digital twins as part of corrosion engineering conceptualization are not established yet at high maturity level. Calculations to predict corrosion and its propagation are inherently difficult as all types of corrosion are caused by a multitude of different mechanisms, which depend on the application (in-service), environmental and material factors, and the surface finish according to its application. Furthermore, physical-based simulations of corrosion are often prohibitively expensive, because of the involved time- and length scales and complex interactions [Mir20], e.g., electrochemical systems requiring large-scale DFT simulations of defects towards its effect on microscale corrosion mechanism (e.g., pit growth). For example, for steels, heuristic models for weight loss are well known [Gie19]. They can (if they are applied correctly) assist engineers to conceptualise the corrosion protection strategy.

Hence, we propose the NFDI co-development of CorWiz, a web tool based on heuristic / data-based models [Seg2022] to give insights into the corrosion process and providing both a first assessment on the impact of corrosion on the constructive design as well as an appropriate model for the provided material, corrosion conditions, data-availability, and engineers choice performance criteria. To facilitate such a web tool (Betty), data from different sources (Doris), such as microstructural information, microscope images, composition vs. processing, data on the protective coatings, corrosion information from electrochemical experiments and the field (Golo), and data from existing simulation models are collected in a database [Höc20] and aggregated (Caden) into CorWiz (Kadi4Mat support). An algorithm is then developed to estimate the corrosion impact for a predefined steel in a component system exposed to chloride containing environment (recommended by the community clusters), as well as to select the best available (heuristic) corrosion model to simulate the problem under the defined conditions (immersed or atmospheric).

This model and its prediction are then made accessible via a web tool, allowing the easy and broad use of the results, and collected as reusable data. Since the model will be available as a web-service, it can be utilised by a wide array of users (engineers), who might not have a strong background in corrosion engineering. The reference database will give access to a wide range of data for the training of neural network-based models, as well as complex future model workflows, provide validation data for physical-based simulations and allow better design of corrosion experiments (lab and field) and steel-based constructions.

Brief description of the proposed project (~3-5 pages)

Introduction

CorWiz will implement a web service by collecting available data on steel corrosion by the applicants and project partners (incl. NFDI and external) to get sufficient FAIR data for the model utilisation and subsequent corrosion engineering processes. Based on this knowledge (database) a model will be generated (calibration to existing models) that contains an extensive range of mathematical corrosion description as shown in Fig. 1 and currently available. The adaptation and extension capabilities are considered too, as well as data tricking aspects.

This knowledge primarily exists in the form of even unstructured data like research articles and is neither directly available to engineers nor directly useful for computations. Selecting and getting an overview of all available knowledge in the form of data and models thus is paramount for an efficient modelling workflow. Getting this information is complicated as the nature of corrosion processes is complex [Zan16, Sedr96], and available digital solutions do not possess the domain knowledge required to select appropriate models and available data. Thus, a tool to help (corrosion) engineers to select models and to obtain available interpretable data to perform corrosion simulations as part of a construction concept is indispensable.

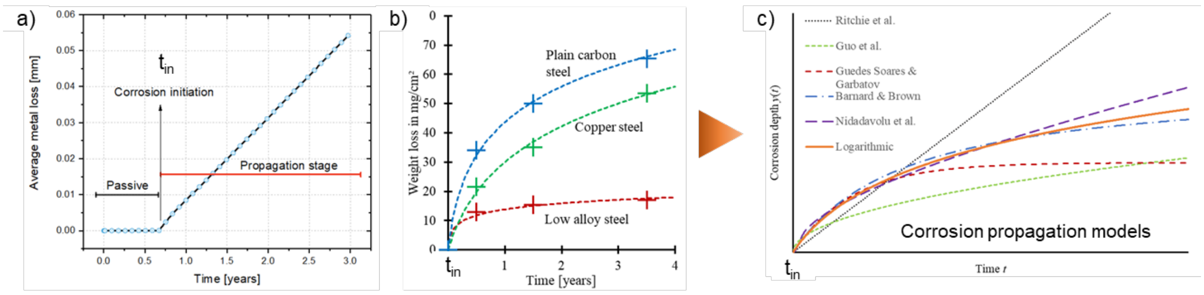


Fig. 1: P*rinciples of mass loss corrosion data a), its material dependency (e.g., steels) b), and its propagation modelling via heuristic approaches c) [Gie19]*

Thus, we want to develop CorWiz, a hands-on web tool for engineers to give access to the broad corrosion knowledge in the form of available (meta-)data and models (according to the archetypes). As well as to give end-users (from different stakeholder groups) first insight into the corrosion challenges they would face in a certain construction application (in this project one example for salt induced steel corrosion), which will accelerate the design process as well as bring unforeseen corrosion issues (electrolyte sinks, galvanic couples, crevices, etc.) into their mind during the design process. This will consequently lead to more economical and ecological design by providing a longer lifetime or preventing unnecessary hazardous or even toxic corrosion protection measures.

An end-user will be able to select, for example, substrate material, application, and (simple) environmental conditions, from which CorWiz will provide all relevant data in the form of a corrosion grade, relevant literature, suitable heuristic model, experimental and simulation results in the form of structured data. This allows the end-user to spare valuable time searching and processing data or creating their own experiments to use for design validation. The development of such a tool requires data scraping [Ryan18] from all available sources to collect all relevant data, as well as processing and quality assessment of the data (support by NFDI4Ing base services). These steps will be performed by developing data scraping and processing tools for Kadi4Mat [BGH21]. Kadi4Mat is a research data management platform developed in part for the NFDI4Ing consortium. Allowing the efficient storage and provenance tracking required for this endeavour. This also links the applicants' activities to at least one of the NFDI4Ing partners and provides access to the data in a widely accepted format.

Implementation

After the identification of the types of corrosion, the next step involves the collection of data (Fig. 2). Corrosion related data must be collected from all possible sources, including the internal consortium partners, the NFDI network and literature data. This data then needs to be carefully segregated according to the material, type of coating and corrosion. Different methods of data collection leads to different kinds of data collected for the same type of corrosion. Thus, the parameters in the data samples collected must be identified and precisely collaborated with the type of corrosion.

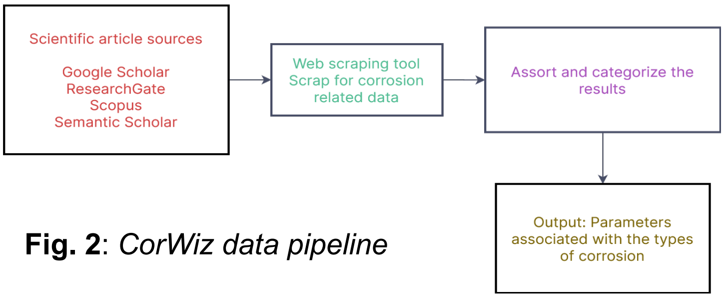


Fig. 2: *CorWiz data pipeline*

Microstructural data associated with steels would also have to be collected to gain insight about the microstructural features, grain sizes, presence of carbides, hydrides and the grain boundary network. EBSD data and SEM images of steels would provide information on the microstructure and the phase distribution of the steels. This information could be used to classify the steels according to the microstructure. A database holding information on the microstructure of steels would be valuable not just to the current project, but to the whole corrosion community as well. Since EBSD and SEM, information about a variety of steels has already been collected by many researchers, but this information is not readily available for others to use. Collecting and making them accessible to the whole community would cut down considerable amounts of man-hours and resources spent on procuring this information again. This database would also be very convenient for modellers looking for microstructural data.

Non-atomistic corrosion models range from phase field models [WSB16], finite element models [CK19], mechanistic [Ziru07], stochastic models [VCARH07] to very simple models [CMBG08], which capture corrosion with a set of equations. There are complex phase field models that capture the intricate nature of the corrosion process and provide insight into the corrosion process itself, by considering the microstructure, the corrosion current, evolution of damage due to corrosion and much more. Though these models provide very detailed results, they are computationally very expensive and might not be feasible for quick results. While there are simple models that capture corrosion with just a set of equations and provide very quick and targeted results, they might not provide sophisticated results, especially not for complicated geometry. Depending on the requirements of the end user, a wide variety of corrosion models are available in the literature. In tasks T1.3 and T1.4, the corrosion models associated with the steels will be evaluated and categorised according to complexity and the input parameters.

A database will be developed using the Kadi4Mat with the data that has been collected using the data scraping tools developed in T1.1. The collected data will be categorised according to the type of corrosion, the parameters involved, the results provided and for the models, the input parameters, complexity of implementation and the results provided by these models. Task T2.1 will deal with the development of the said database and aggregation of the collected data. With the database prepared, a web tool shall be developed to access the database. The web tool will provide the users an intuitive way to access the database and look up models to simulate their steel corrosion problems. The users shall provide their base metal, type of corrosion, information on the corroding environment and the parameters available to them as an input and the web tool will provide the users with a list of possible models.

Current state of corrosion research of construction steels

Construction steels are used for a multitude of purposes, ranging from culinary products, surgical instruments, architecture to automobile and aeronautical applications [Zan16]. Although construction steels are exclusively coated against corrosion, they are not immune to corroding effects, especially in the long term. Corrosion of steel can lead to rupture of pipes and tanks, causing leaks, reduction in load bearing capacity, leading to fracture and failure, production of rust and other harmful contaminating products.

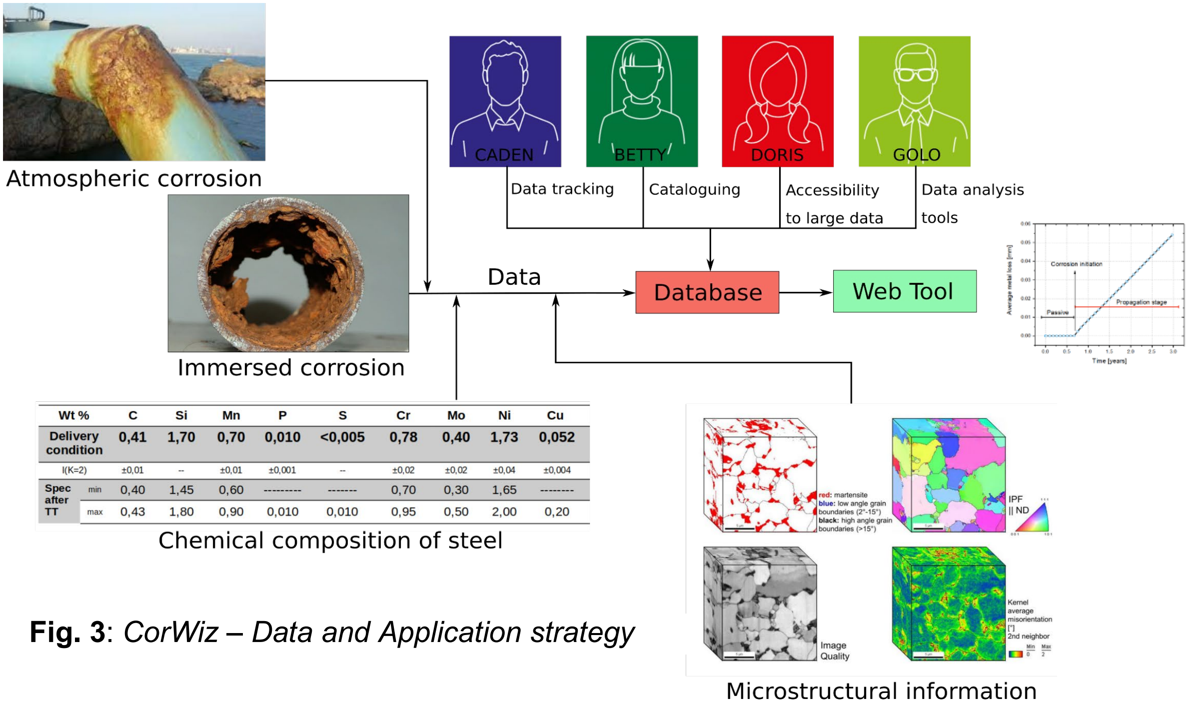
Dissolution is observed when construction steel parts are exposed to a chemical media, which leads to a continual removal of the bulk metal. Ample amount of research has been done on dissolution corrosion and data with tables on resistance against various chemicals are already available for a wide selection of construction steels [Kadry08]. Dissolution driven corrosion is primarily characterised by the amount of weight loss. Dissolution can be mainly classified into two types, atmospheric corrosion, and immersed corrosion. Complex corrosion models can provide highly detailed results at the cost of high computational cost. To provide engineers with highly targeted results, we are interested in heuristic models that can predict the loss of weight with the evaluation of algebraic equations. These models also allow the implementation in a web tool because they are computationally inexpensive and can be solved on web-based applications. A list of heuristic models is highlighted in Momber et al. [Mom22].

At the Institute of Surface Science at Helmholtz-Zentrum Hereon, we work on developing state-of-the-art corrosion models. We have already published many scientific articles that deal with corrosion modelling, e.g. [SHLM19]. We are also involved in developing phase field corrosion codes and models for hydrogen embrittlement of steel with metallic coatings. Since we primarily work on the modelling side of corrosion and are involved with several experimental and industrial partners, we can predict the needs of an engineer in an ever-evolving field and provide the best solution for corrosion model selection.

Contact areas with NFDI4Ing

Associated archetypes - Caden, Betty, Doris and Golo

The data that needs to be collected for the development of CorWiz involves the collection of data sets (Fig. 3) that includes both experimental (lab + field) data and metadata. This collected data must be stored in a structured manner so that they can be easily accessed. Data is also collected from various institutions, whose data storage solutions are expected to be different. Thus, consolidation of data into a structured format is an issue that needs to be addressed as well. The collection and storage of literature data from the data scraping process will also pose significant challenges in terms of consolidation. Archetype Caden offers approaches and insights to overcome these issues. Furthermore, the Kadi4Mat will be used to establish a structured format to store the collected experimental data, simulation data and metadata.



Wt %		C	Si	Mn	P	S	Cr	Mo	Ni	Cu
Delivery condition		**0,41**	**1,70**	**0,70**	**0,010**	**<0,005**	**0,78**	**0,40**	**1,73**	**0,052**
I(K=2)		±0,01	--	±0,01	±0,001	--	±0,02	±0,02	±0,04	±0,004
Spec after TT	min	0,40	1,45	0,60	---------	-------	0,70	0,30	1,65	--------
	max	0,43	1,80	0,90	0,010	0,010	0,95	0,50	2,00	0,20

Fig. 3: *CorWiz – Data and Application strategy*

The common gripe with the development of research software and scientific methods is, they are not readily available for the end-user. Archetype Betty of NFDI4Ing addresses this issue and provides some insights on how to overcome this issue. In this project, we want to implement the measure **B-3** of archetype Betty to develop a web tool to act as a frontend to solve an engineer's corrosion modelling needs. This approach would containerize the complexity of the models and would provide a simple and elegant solution. The availability of research software is always in doubt, since they are not always published due to confidentiality and even when it is available, their usability is not guaranteed due to a number of reasons, such a lack of documentation, proprietary data storage and handling systems and dependencies. To deal with this issue, the models researched in this project will be subjected to cataloguing, which exactly lines up with the **B-5** measure of archetype Betty.

The microstructural data, namely the EBSD and SEM data collected to categorise the steels, must be collected from a variety of sources. The collected data must also be archived and made available to access. Archetype Doris deals with the problems associated with storage and sharing of large data. We hope to implement the measure **Storage & archive for very large data** in archetype Doris to make the collected microstructural data (e.g., with MatWerk) available to end-users via the frontend and making them searchable and accessible. A database of microstructural data would be very significant for engineers and researchers alike. Corrosion engineers collect large amounts of field data from sensors. These data can be collected and used as either input parameters for the heuristic models or also used as calibration data for the various models. The collected field data must be categorised and associated with the different processes. We believe archetype Golo will help us to prepare for the problems associated with the handling of field data. With the development of CorWiz, we plan to use, as well as contribute to the base services S3 and S4. With the database developed, CorWiz will provide the community clusters with access to large amounts of data on corrosion and microstructural data. This will enable the sharing and use of data and implementation into workflows. The base service S4 will particularly be useful for the development of the database, since it provides tools, standards and protocols for the collection and storage of (corrosion-type) large data.

The archetypes deal with the collection, consolidation, and storage of the data, but the identification of the various sources of data also pose a challenge. We believe the community cluster 43 of NFDI, Materials science and engineering, would prove very significant in overcoming this challenge. With the ample knowledge and experience behind the members of CC43 and the pre-existing work, vast amounts of data must already be available. Through the community cluster, establishing contact between the different institutes should be relatively uncomplicated. Community cluster CC45 (civil engineers) is a strong community and CorWiz user and profiteer too.

Community extension and networking

Corrosion is a bad taste problem in a wide range of engineering disciplines with huge economic relevance. More than 3% of GDP is lost every year due to corrosion. A multidisciplinary community approach is required to combine data from materials to weather towards the selection of the most promising corrosion protection concept. NFDI provides a perfect platform for that as it might be a neuron between:

1. the NFDI consortia with facing corrosion problems, and with data exchange in CorWiz,
2. our simulation platform activities as within DGM where we are co-chairing, as well as on EU level where we deal with currently three platform project and are co-chairing the European Materials Modelling Council Focus Area – Industrial impact, and
3. the native CorWiz related activities within the Helmholtz association on artificial intelligence, research software engineering, metadata or digital (corrosion) twins

Based on recent community feedback (GfKorr, at EuroCorr, Eurostars programme, etc.) and based on our own experience, a corrosion data infrastructure would be highly appreciated.

D. Höche

References

[Mir20] Mir, Z. M., Friis, J., Hagelien, T. F., Svenum, I. H., Ringdalen, I. G., Konchakova, N., Höche, D. (2020). Interoperability architecture for bridging computational tools: Application to steel corrosion in concrete. Modelling & Simulation in Mat. Sci. and Eng., 28(2), 025003.

[Gie19] Gießgen, T., Mittelbach, A., Höche, D., Zheludkevich, M., & Kainer, K. U. (2019). Enhanced predictive corrosion modelling with implicit corrosion products. Materials and Corrosion, 70(12), 2247-2255.

[Seg22] Seghier, M. E. A. B., Höche, D., & Zheludkevich, M. (2022). Prediction of the internal corrosion rate for oil and gas pipeline: Implementation of ensemble learning techniques. Journal of Natural Gas Science and Engineering, 99, 104425.

[Höc20] Höche, D., Konchakova, N., Zheludkevich, M., Hagelien, T. F., & Friis, J. Ontology Assisted Modelling of Galvanic Corrosion of Magnesium. In 14th WCCM-ECCOMAS Congress 2020 (Vol. 2100).

[WSB16] Mai, Weijie, Soheil Soghrati, and Rudolph G. Buchheit. "A phase field model for simulating the pitting corrosion." *Corrosion Science* 110 (2016): 157-166

[CK19] Liu, Chao, and Robert G. Kelly. "A review of the application of finite element method (FEM) to localized corrosion modeling." *Corrosion* 75.11 (2019): 1285-1299

[Ziru07] Zhang, Ziru, et al. "A mechanistic model of top-of-the-line corrosion." *Corrosion* 63.11 (2007): 1051-1062

[VCARH07] Valor, A., et al. "Stochastic modeling of pitting corrosion: a new model for initiation and growth of multiple corrosion pits." *Corrosion science* 49.2 (2007): 559-579

[CMBG08] da Costa-Mattos, H. S., I. N. Bastos, and J. A. C. P. Gomes. "A simple model for slow strain rate and constant load corrosion tests of austenitic stainless steel in acid aqueous solution containing sodium chloride." *Corrosion Science* 50.10 (2008): 2858-2866

[SHLM19] Snihirova, Darya, et al. "Galvanic corrosion of Ti6Al4V-AA2024 joints in aircraft environment: Modelling and experimental validation." *Corrosion Science* 157 (2019): 70-78

[Zan16] Zander, D., Höche, D., Deconinck, J., & Hack, T. "Corrosion and its context in service-life", ICME Handbook (2016): 227-245. Weinheim: Wiley VCH.

[Kadry08] Kadry, Seifedine. "Corrosion analysis of stainless steel." *European Journal of Scientific Research* 22.4 (2008): 508-516

[BGH21] Brandt, Nico, et al. "Kadi4Mat: A research data infrastructure for materials science." Data Science Journal 20.1 (2021)

[Ryan18] Mitchell, Ryan. Web scraping with Python: Collecting more data from the modern web. " O'Reilly Media, Inc.", 2018

[Sedr96] Sedriks, A. John. "Corrosion of stainless steel, 2." (1996)

[Mom22] Momber, A. W., M. Wilms, and D. Brün. "The use of meteorological and oceanographic sensor data in the German offshore territory for the corrosion monitoring of marine structures." *Ocean Engineering* 257 (2022): 110994

NFDI4ing

Timeline, milestones, metrics, KPI (~1-2 pages)

GANTT: CorWiz schedule in 2023

Task	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
T1.1												
T1.2												
T1.3												
T1.4				M1								
T2								M2				
T3.1											M3	
T3.2												
T4												M4

Tasks

T1.1 Develop Data scraping tools

A tool will be developed that facilitates data scraping from literature and other public sources.

T1.2 Data processing tools

Available data will be processed into a common format and augmented with metadata for easier and efficient use in the future.

T1.3 Data quality assessment tools

The quality of available data will be accessed using a tool integrated with Kadi4mat.

T1.4 Collect data in Knowledge base

The developed tools from T1.1-T1.3 will be used to generate a knowledge base.

T2 Develop the model algorithm

A model is developed based on the collected data to allow the assessment of the corrosion model and grade.

T3.1 Web application design and implementation

A web application will be designed and implemented that gives access to the corrosion expert system and the collected data.

T3.2 User feedback adjustments

Based on feedback of test users, the web application usage will be improved.

T4 Report and documentation, incl. Workshop on CorWiz

Write a report on the tools and conclusions from the development, as well as documentation for the usage of the web application and the database.

Milestones

M1 Corrosion Knowledgebase (M4)

T1.1-T1.4 will result in a knowledgebase of corrosion results and models.

M2 Stainless steel corrosion expert (M8)

Corrosion expert is developed that allows the first assessment of corrosion impact in a certain situation. Selection of a corrosion model based on currently available data and performance.

M3 CorWiz web application (M11)

Web application that gives access to the knowledge base and expert system.

M4 Report and Documentation (M12)

Report on capabilities and short-coming of the current approach and future improvement possibilities. And user documentation for the web application and knowledge base

Metrics/KPIs

KPI 1: Quantified improvement of digital corrosion engineering, and human- and computer-readability of corrosion and corrosion testing method specifications and (meta)data

KPI 2: Efficiency, reliability, and user-friendliness of the semantic integration framework with respect to global knowledge sharing by the CorWiz web service (user voting depended)

KPI 3: Reduction (>10%) of the testing and modelling overall effort needed to generate meaningful data on corrosion propagation

KPI 4: Successful seeding of a corrosion database / repository extendable, adaptable, and FAIR, starting with construction steels and chloride induced corrosion

Funding requirements

Description	Count	Sum
Aravinth Ravikumar (Ph.D. Student)	6 PM	26,888.00 EUR
Dr. Sven Berger (Postdoc)	6 PM	38,700.00 EUR
Sum		**65,588.00 EUR**

Project Members

Project-Lead: Dr. Daniel Hoeche

(Institute of Surface Science / Head of Department - Interface modelling)

Staff:

- **Dr. Sven Berger (Postdoc, Git / MOLab System administrator)**
- **Aravinth Ravikumar, M.Sc. (PhD candidate, Hydrogen transport modelling in Zn-Ni plated steel)**

The seed fund would enable direct NFDI activities with respect to the co-funded projects OpenModel: Open Innovative Platform for Integrated Materials Modelling to Bridge the Gap from Translation to Application; H2Free: Investigation and modelling of Hydrogen effusion in electrochemically plated ultra-high strength steels used for landing gear structures; and TAIFUN: Towards Artificial Intelligent Maintenance System (AIMS) via Predictive Failure Modelling and Numerical simulation and link them.