DHBW
Duale Hochschule
Baden-Württemberg

Katja Wengler, Dirk Reichardt, Dennis Pfisterer (Eds.)

DHBW AI Transfer Congress 2024

Proceedings

Imprint

Stuttgart, Juli 2024

DHBW AI Transfer Congress

Published by:
Duale Hochschule Baden-Württemberg
Baden-Wuerttemberg Cooperative State University
Friedrichstraße 14, 70174 Stuttgart
represented by: Prof. Dr. Martina Klärle

Editors

Prof. Dr. Dennis Pfisterer
Program Chair
DHBW AI Transfer Congress

Prof. Dr. Dirk Reichardt
Program Chair
DHBW AI Transfer Congress

Prof. Dr. Katja Wengler
Program Chair
DHBW AI Transfer Congress

This volume presents the conference proceedings of the 3rd AI Transfer Congress, organized by the Cooperative State University Baden-Württemberg (DHBW), held on July 11th, 2024, in Stuttgart, Germany.

Inaugurated in 2022, the conference aims to establish a transfer platform fostering exchange between academia and practical applications regarding the field of Artificial Intelligence (AI). The event features a conference track that explores current research in applied AI, and a workshop track that facilitates trainings and discussions on applications, tools, and the implementation of AI in economics as well as society. This structure fosters an exchange between academic and industrial researchers and users and practitioners interested in AI.

AI applications have recently permeated numerous domains and research areas. The conference addresses a broad spectrum of topics, including unveiling bias, future skills, AI literacy, industrial applications, recommender systems, autonomous driving, stock market prediction, and the interdisciplinary field of sustainability, alongside educational aspects of AI.

The workshop track serves as a forum for training and discussions on selected application areas, particularly with practitioners, e.g. machine learning, embedded AI, generative AI, empowering leaders, retrieval augmented generation or AI in marketing.

The review process was overseen by an internationally diverse program committee of experienced practitioners and scientists. We extend our gratitude to the authors and reviewers for their outstanding contributions. Lastly, we thank the research officers, the members of the DHBW university administration and communication departments for their dedicated efforts in making the AI Transfer Congress successful.

Prof. Dr. Martine Klärle
President DHBW

Welcome to the third AI Transfer Congress, or AITC, here at the Baden-Württemberg Cooperative State University! It is my great pleasure to greet you all on this special occasion.

As a transfer university, DHBW offers a unique platform for the exchange between theory and practice. With AITC, we continue this tradition and focus on one of the most exciting and forward-looking topics of our time: Artificial Intelligence. We are particularly proud that Artificial Intelligence is a central topic for DHBW, and we successfully practice this not only in teaching but also beyond.

This year is especially significant for us as we celebrate the 50th anniversary of our university. For five decades, DHBW has been at the forefront of integrating academic knowledge with practical application, and we are thrilled to mark this milestone with you all.

It is impressive that we have gathered around 200 guests from academia and industry from six different countries here today.

Your dedication and expertise are crucial in achieving progress in the research and application of Artificial Intelligence.

I am especially pleased that, as part of the EU4DUAL project, we are networking European universities. This

project promotes cross-border exchange and strengthens cooperation in dual education. Our partner universities have actively contributed to the program committee and provided valuable input.

I am convinced that today's congress, through your contributions and the intensive dialogue between theory and practice, will lead to new insights and innovative solutions. Let us use this opportunity to learn from each other, exchange ideas, and initiate new collaborations.

In this spirit, I wish us all an inspiring and successful congress!

Prof. Dr. Beate Sieger-Hanus
President DHBW Stuttgart

The DHBW AI Transfer Congress is an outstanding opportunity for networking and exchanging theoretical knowledge and practical experience on the topic of artificial intelligence.
Here we can jointly strengthen the innovative power in the research and application area of artificial intelligence and develop sustainable solutions for the future.
Immerse yourself in an exciting discourse on the technical, economic, social and also ethical aspects of AI. Inspire others with your ideas and be inspired by the ideas of others.

Keynote

Unveiling Bias: Navigating the Complexities of Artificial Intelligence in Society

Dr. Alina Gales

Technische Universität München

Artificial Intelligence is evolving at an unprecedented pace, driving innovations such as autonomous drones, medical diagnosis systems, and recommendation algorithms that shape our online experiences.

While these advancements promise transformative benefits for society, they also raise concerns about bias and misinformation. In this keynote address, Dr. Alina Gales explores how Artificial Intelligence reflects and amplifies societal structures, potentially reinforcing biases against marginalized communities.

She highlights impacted areas of biased AI that particularly affect women and people of color. Furthermore, she will emphasize the critical importance of fostering discussions on responsible AI, posing the question: How can we ensure that the potential benefits of AI are realized equitably and ethically**?**

Table of Contents

Graph-Based Automated Styling and Recommendation System for Personal Shopper Services in the Fashion Industry

Jauder Uribarri Sánchez
Mondragon Unibertsitatea - Business Faculty
jauder.uribarri@alumni.mondragon.edu

Maialen Aguiriano Beristain
Mondragon Unibertsitatea - Business Faculty
maialen.aguiriano@alumni.mondragon.edu

Nerea Zuaznabar Montoya
Mondragon Unibertsitatea - Business Faculty
nerea.zuaznabar@alumni.mondragon.edu

Sara Crego Herrero
Mondragon Unibertsitatea - Business Faculty
sara.crego@alumni.mondragon.edu

Nahia Uncetabarrenechea Rodríguez
Mondragon Unibertsitatea - Business Faculty
nahia.uncetabarrenec@alumni.mondragon.edu

Unai Garcia Garcia
Mondragon Unibertsitatea - Business Faculty
unai.garciag@alumni.mondragon.edu

***Abstract*—The adoption of machine learning (ML) by industries has paved the way for disruptive companies and methodologies. In particular, in the fashion industry, personal shopper services such as Lookiero have begun to integrate ML to improve their style recommendations. In this paper, we present a dual-purpose framework designed to automate the generation of fashion outfits and to tailor these recommendations to customers' tastes. The first component of our framework involves the construction of a graph, where each node represents a garment from the available inventory, complete with detailed attributes. Edges between garment nodes are established based on predefined compatibility rules, facilitating the assembly of coordinated looks. This graph serves as the training ground for a graph neural network (GNN), which learns to integrate new garments into the network by predicting their compatibility. To achieve the second goal - personalised recommendations - a bipartite undirected graph is constructed. One set of nodes represents customer clusters, while the other corresponds to the looks generated by the first graph. Customers are grouped into clusters, and recommendations are made by analysing the strength of connections between customer clusters and fashion looks. The results of the GNN are promising and indicate potentially significant benefits for the company. However, the recommendation network did not fully meet expectations due to the limitations of the tools used. Despite these challenges, our tests yielded encouraging results with representative samples, suggesting that with further refinement, the framework could be effectively scaled and implemented within the company, revolutionising the personal shopping experience.**

***Index Terms*—Personal shopper, fashion look, graph theory, generative neural network, clustering, recommendation system**

I. Introduction

Graph Neural Networks (GNNs) have revolutionised the way we interpret complex data which contains rich relation information among elements, transforming industries with their predictive power and adaptability [1] [2].This type of neural networks have their origins traced back to the late 1990s where some papers showed that neural networks can be adapted to process and classify structured information by using generalized recursive neurons and directed acyclic graphs, extending traditional sequence-based models to handle complex data structures [3] [4]. Recurrent Neural Networks (RNNs) were subsequently extended to accommodate a broader range of graph structures, leading to the development of GNNs for addressing node-focused problems [4]. Advancements such as feedforward architecture and mapping graphs into n-dimensional Euclidean space have enhanced the technology [5] [6], increasing its popularity and paving the way for the development of various GNN variants. In 2015, convolutional networks where adapted to non-euclidean data resulting in Graph Convolutional Networks (GNNs) and their posterior variants [7] [8]. The good results obtained from these models boosted the interest in GNNs sparking new variations like:

- *GraphSage*: Aggregates neighborhood information using sampled nodes to improve representation learning [9].
- *Graph Attention Networks (GATs)*: Uses attention mechanisms to weigh the importance of neighbors for each node in the graph [10].
- *Graph Isomorphism Networks (GINs)*: Embeds[1] nodes by aggregating information from their neighborhoods, invariant to graph isomorphism [11].

GNNs have revolutionised what is possible in several domains, including fashion. In recent years, several innovative approaches have been proposed to improve fashion recommendation systems by using graph-based neural networks to model the complex relationships between fashion items. One

[1] Embedding is a means of representing objects like text, images and audio as points in a continuous vector space where the locations of those points in space are semantically meaningful to machine learning (ML) algorithms. [12]

such approach is Neural Graph Filtering, which enhances the flexibility and diversity of recommendations by modelling garments as nodes in a graph neural network, leading to significant improvements in user preference and performance metrics across different datasets [13]. Another notable method is the Node-wise Graph Neural Network (NGNN), which represents outfits as graphs to better capture compatibility between fashion items, demonstrating superior performance in item suggestion and compatibility prediction tasks [14]. This method was compared in a study against Hypergraph Neural Network (HGNN) on the Polyvore dataset [15]. The paper shows that HGNN slightly outperforms NGNN on fill-in-the-blank and compatibility prediction tasks, with further accuracy improvements achieved using vision transformer embeddings [16].

Deep Relational Embedding Propagation (DREP) further advances the field by incorporating extra-connectivity between items and user interactions into the compatibility modelling process. This graph-based framework significantly improves pairwise compatibility modelling and outperforms state-of-the-art methods [17]. Complementing these efforts, an approach that learns image embeddings to capture both item similarity and compatibility for fashion outfit construction has shown a 3-5% improvement over existing methods in compatibility prediction and fill-in-the-blank tasks using datasets from Polyvore [18]. Finally, the Attention-based Dataset Distillation Graph Neural Network (ADD-GNN) uses designer-generated data to improve feature representation and outfit compatibility assessment. This method outperforms several competitive baselines on real-world fashion datasets [19].

Our work advances the current state of the art in fashion recommendation systems by implementing theoretical models in a real-world business context. Using anonymized data samples provided by the company, we aim to demonstrate the feasibility of integrating these innovative methods into the business processes, effectively bridging the gap between theoretical advances and practical implementation challenges in real-world settings.

A. Objectives

This study aims to use GNN technology to provide a solution to a problem faced by Lookiero, a Spanish fashion company. Lookiero is a platform aimed at women that selects 5 items of clothing based on the user's personal tastes and preferences, and sends them to an address in a unique and personalised box. The company wants to automate the creation of the "looks" it recommends to its customers, in order to speed up the process of recommending clothes and classifying new products as they arrive. To this end, two objectives were set: to automate the process of creating looks by associating garments in a graph; and to recommend these looks to its customer base based on their characteristics.

II. Data processing and analysis

While all data used in this study is proprietary, it has been anonymized to protect the privacy and confidentiality of the individuals involved. Two different datasets are used: one containing clothing data where names and brands have been replaced with generic identifiers to protect proprietary inventory information. This approach protects supplier relationships and minimises brand perception bias while maintaining the utility of the data. In the client data set, all personal information was systematically removed, leaving only essential attributes linked to an encrypted primary key for analysis. This measure is critical for compliance with data protection and privacy laws, and effectively prevents unauthorised access, identity theft, or misuse of sensitive personal information.

A. Garment data

The data of the different products are stored in a relational database, where the garments are divided into 3 different seasons, 19 garment families and 163 brands. For this paper, data from 2932 products were used. Thus, each product has different characteristics that define it and must be taken into account to develop the links between garments. Data processing is performed to clean and prepare the data for further development. Then a descriptive analysis is performed to understand the data in a statistical and visual way. The garments in each season can be divided into different categories according to their characteristics, but an important distinction is the "category". Each garment falls under a category of clothing (T-shirts, sweaters, pants, bags, jackets...) that are at the same time divided into 3 different levels:

1) *Base layer*: T-shirts, dresses, trousers...
2) *Second layer*: Sweaters, hoodies, Jackets...
3) *Accessories*: Scarfs, Bags, Hats...

The idea, as shown in III-A, is that by differentiating clothes by levels, recommendations are diverse and can assemble an outfit.

For instance, figure 1 shows the distribution by colour and size of the garments.

Other types of visualisation (see figure 2), allow the visualisation of both the different values that a feature can take and the features that a garment has. The features can be Boolean, numerical or categorical, and with this type of aid it was possible to understand the meaning and implications of each. The image below shows the characteristics of a particular garment. Although each garment has dozens of features, only 6 are shown in this visualisation to avoid clutter.

B. Client data

A sample of real, anonymised customer data is used for the recommendation system. When creating an account in Lookiero, and in similar personal shopping companies, a survey must be filled with information about your sizes, style preferences, parts of the body to show or cover, types of clothes to avoid and more. With this valuable information the personal shoppers at the company decide which clothes to send to each customer. In this work an algorithm is implemented using the data from surveys of approximately 2000 customers, to create custom recommendations.

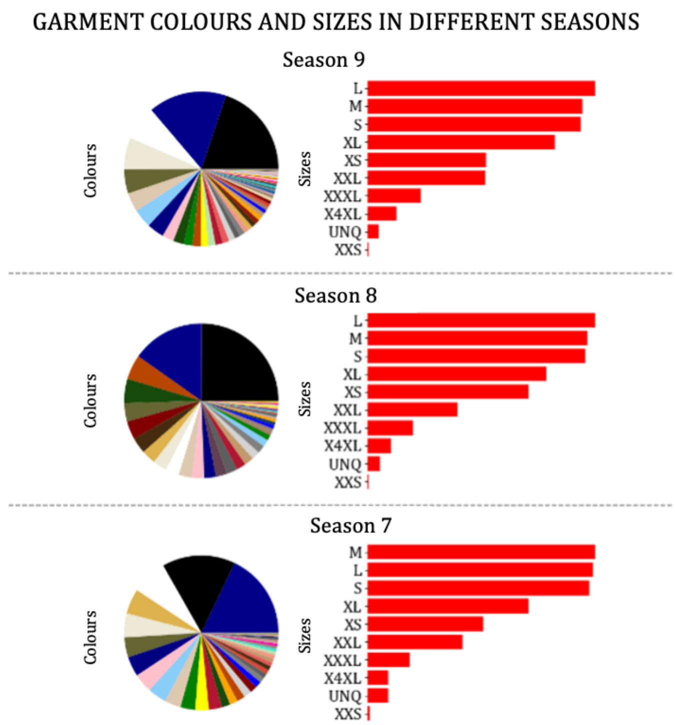


Fig. 1: Visualisation of the importance of colours (pie charts) and sizes (horizontal bar charts) in the three seasons used in the analysis.

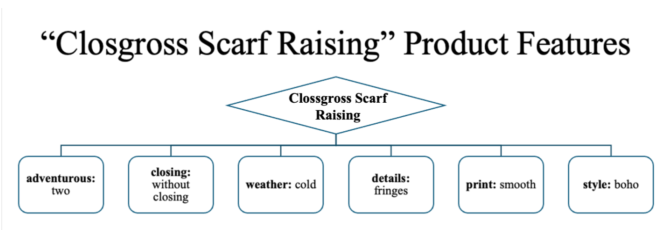


Fig. 2: Tree graph showing the feature values assigned to a garment.

III. Data Modelling

A. Garment graph creation

Three combined garment networks are created, one for each season, because garments from a particular season are only purchased in conjunction with garments from the same season. The process for creating the three networks is identical and is as follows:

First, a node is created for each garment, with the designated characteristics assigned as attributes. Variables containing more than one value for a garment are stored as a list. For example, since the original value of 'size' is 'XS—S—XXL', it is stored as ['XS', 'S', 'XXL']. In terms of creating relationships, bi-directional links are created as clothes go together in both ways; thus creating a non-directional network. This approach makes it possible to connect the different nodes to create outfits, once the links between the nodes have been established according to certain rules.

The criteria for a 'combines' relationship between two garments are outlined below. It is important to note that the nodes are compared in pairs and a scoring system is designed to determine whether they match. According to a predefined set of rules, each node is assigned a score which quantifies the strength of the connection between the nodes in question. The resultant graph is a "weighted graph," in which each edge is allocated a specific weight or cost that reflects the relationship's intensity. Here are the rules to score this connections:

- *Level rule*: it is fulfilled by pairs of nodes that are level 1 and 2, 1 and 3, 2 and 2, 2 and 3 or 3 and 3. That is, all pairs of nodes except when they are both level 1:
  - One garment of level 1, one of level 2 and one of level 3.
  - One garment of level 1 and two of level 2.
  - One garment of level 1 and two of level 3. One point is awarded for pairs of nodes that meet this rule.
- *Colour rule*: This rule takes the colour feature of the clothes and assigns points based on if they match or not.

  The valid colour combinations have been defined in a table following the guidelines of colour theory and with the help of a stylist. While neutral colours like white and black combine with nearly every other colour, with shiny colours some general rules can be applied. For example while a yellow garment can combine with something in navy, it is not recommended to style it with green.
- *Category rule*: This rule is fulfilled and a point is added if the two nodes do not belong to the same category (as explained in II-A). The aim of this rule is that a look does not contain two garments of the same type if they are of the same level. In other words, if it consists of one item of level 1 and two items of level 2, this will prevent the last two items from being, for example, a T-shirt and a top, rather than a T-shirt and a pullover. Also, if one of the nodes is a dress and the other is a T-shirt, they do not match because they are incompatible in a look.
- *Usage rule*: The usage in the clothes is specified as the environment the garment is supposed to be worn. With this rule points will be scored if the two products match in their usage ('special-occasion', 'night', 'work' ...). If one garment is 'special-occasion' and the other is 'night', they also meet the rule, but at a lower level, so they get half a point. The aim is to combine garments that can be used in the same environments.
- *Print rule*: : Clothes in the dataset can either have a print (floral, geometric, animal...) or a smooth print (does not have any pattern). The pairs of nodes that satisfy this rule will get a point if:
  - The two have a smooth print.
  - One has a smooth print and the other a different print. Combining different prints in an outfit is not recommended and this rule tries to avoid it.
- *Style rule*: The clothes have been divided into one or more style categories (street, classic, boho...). A full point will

be granted if the two garments are of the same style. If the garments do not fall in the same style, but in a similar one, half point will be granted. For example if one garment is a street style and the other is a casual style. As with the previous rules, the idea is to output coherent outfits.
- *Weather rule*: : the pair of garments that have the same value in the variable 'weather' fulfil the weather rule and get one point. That is, they must both be from the warm season or both be from the cold season.

After establishing the criteria, a connection is formed between two nodes if they achieve a score of 1 point in level, colour, and category, and exceed 0 points in application, pattern, style, and time. The edge's weight is determined by the total points accrued. The attributes of the resultant graphs for seasons 7, 8, and 9 are detailed below:

TABLE I: Number of nodes, connections and isolated nodes of the networks of seasons 7, 8 and 9.

Season	Number of Nodes	Number of Connections	Number of Isolated Nodes	Avg.grade of nodes
Season 7	2033	109045	327	148
Season 8	1929	89278	444	174
Season 9	1732	80460	328	134

After observing the results, it was considered positive that the nodes did not have too many connections, as this meant that they were more precise (more robust relationships). As a result, there were nodes that were not connected to any other node, so they were assigned connections in the following way:

1) In order to identify the most similar node to the isolated node in the whole network, a similarity value was calculated by adding the number of attributes with an identical value to that of the isolated node for each non-isolated node. It has been defined that the most similar node must be of the same season, the same colour, the same level and the same category, so that it does not violate the main rules of connection between nodes.
2) After ranking the nodes from most similar to least similar to the isolated node, the node with the highest similarity was chosen and its neighbouring nodes were extracted and assigned to the isolated node. The weight of the edges has been divided by two so that these relationships do not have the same relevance as those of an originally non-isolated node.
3) The same process was repeated for each isolated node.

These are the resulting graphs after isolating the nodes:

Once the combined garment graphs are complete, they are stored in Neo4j [20], an open-source graph database management system. This Graph database is chosen for its simplicity, powerful Cypher query language, flexible schema, and the ability to choose between embedded and server mode architectures, which facilitate fast and specific information retrieval within graph databases [21]. The next step is to create the graph of looks formed by 3 related garments.

TABLE II: Number of nodes, connections and isolated nodes of the networks of seasons 7, 8 and 9 after deisolating nodes.

Season	Number of Nodes	Number of Connections	Number of Isolated Nodes
Season 7	2033	150300	0
Season 8	1929	166974	0
Season 9	1732	114602	0

B. Looks graph creation

Using related garment graphs, groups of three connected nodes are obtained to create looks, forming a connected subgraph. Each look becomes a virtual node combining the attributes of the three garments it contains. To start the process, all level 1 nodes are identified. For each, all existing pairs of neighbours are found to check for connections. If two neighbours are connected, they form a triangle with the level 1 node and this configuration is called a look. The total number of looks is 4,184,529. This process is repeated for all Level 1 garments across all seasons, and the garment attributes are incorporated to determine the look attributes. Nodes are then created for each look with its attributes. This forms a non-directional graph, as future relationships will be bidirectional, i.e. a look will be linked to a client and vice versa.

C. Modelling the Looks Graph

With the graphs in place, algorithms are implemented to predict potential new connections, or axes, within the graphs. The aim of these algorithms is to identify possible connections for a new garment entering the database by finding the most similar node in the graph and assigning it the edges of that node, a process known as 'axis prediction'. Graph Neural Networks (GNNs) are used in this modelling phase. These algorithms differ from traditional neural networks in that they operate within a graph structure, as opposed to the vector-based processing of conventional networks. GNNs have gained popularity in various domains, including social networks, knowledge graphs, recommender systems and bio informatics, due to their ability to extract knowledge from complex data, taking into account interactions rather than just individual values. GNNs can perform node-level, edge-level and graph-level tasks.

To work with this graph format, embeddings are required to reduce nodes and axes to a dimensional vector space, consisting of an encoder and a similarity function. The encoder transforms the dimensions of the inputs, while the similarity function defines the similarity between the vector space and the neural network. In convolutional neural networks, nodes gather information from their neighbours through embeddings, with the number of layers of the network corresponding to the number of hops considered when extracting neighbourhood information.

Choosing an appropriate loss function is crucial before building the network; for this project, Binary Cross Entropy with linear sigmoid activation[2]was chosen. Since we want to determine if one garment fits with another, the task is binary classification. Each layer contains a message and an aggregation function that allows communication between layers. This mecanism combines feature information from a node's neighbors to update the node's representationv [23]. This is crucial for learning since it allows the propagation and integration of features across the graph, enabling the capture of local and global graph structure. The message mathematically represents the information from the nodes, while the aggregation function combines multiple messages at a node into a single value. There is also a matrix of weights, denoted W_l, where l is the layer number, which is shared within each layer but varies between layers. This approach may result in some loss of information, but it helps to prevent overfitting.

In this project, we employ three distinct aggregation functions: Graph Convolutional Networks (GCN), GraphSAGE, and Graph Attention Networks (GAT). To introduce nonlinearity into our model, we utilise an activation function. The ReLU (rectified linear unit) [24] is selected due to its exemplary performance in preliminary tests with our dataset, outperforming alternatives such as ELU [25] and Leaky ReLU [26]. Below, we provide a concise overview of each aggregation function:

- *GCN*: The main idea behind GCNs is that nodes in a graph can influence other nodes through their connections, so GCNs use a filtering operation that considers information from a node's neighbours to update its representation. This operation is performed by applying a convolutional layer to the node representation. Each previous embedding is divided by the number of neighbours to relativise, then multiplied by the layer's weight matrix and applied to the activation function. In addition, only the neighbours of v are found in the summation and not v itself, so the node being analysed itself is not taken into account.

$$h_v^{(l)} = \sigma\left(\sum_{u\epsilon N(v)} W^{(l)} \frac{h_u^{(l)}}{|N(v)|}\right)$$

- *GraphSage*: is a generalisation of the GCN architecture that allows deep learning to be applied to graphs in a scalable and efficient way. Instead of applying convolutions to nodes and their neighbours, GraphSAGE uses a neighbour aggregation technique to summarise the information of a node's neighbours and update its representation. The first part of the formula would be the embedding part of creating the neighbourhood embedding of v, which is then concatenated with v's own embedding in the previous layer, multiplied by the weight matrix and activation function applied.

$$\mathrm{h}_v^{(l)} = \sigma\Big(W^{(l)} * CONCAT\Big(h_v^{(l-1)}, AGG\big(\{h_u^{(l-1)}, \forall u \epsilon N(v)\}\big)\Big)$$

- *GAT*:is a neural network that specialises in learning representations of nodes in graphs. The idea behind GAT is to use attention weights (α_{vu}) to assign weights to the neighbours of each node, rather than simply aggregating or averaging their features. Attention is based on the structural properties of the node. In this aggregation function, not all neighbours are equally important, and the calculation of α takes into account not only the degree of the node, but also other structural properties of the network. Then, as with the other two, the activation function is applied to calculate the new embedding of the node.

$$h_v^{(l)} = \sigma\left(\sum_{u\epsilon N(v)} \alpha_{vu} W^{(l)} h_u^{(l-1)}\right)$$

To model the data, it is divided into 3 subsets: training, validation and test. The training set is used to optimise the GNN parameters, the validation set is used to develop the model and hyperparameters, while the test set is kept separate and reports the performance of the network on unknown data. It is worth noting that data in networks have a structure in which the relationships between them provide information. If the data is not independent and there is a relationship between them, the dependency of the data cannot be ignored. For example, one of the nodes can be in the training set and another with which it has a dependency can be in the validation set. To solve this problem, the separation of the data has been done in a transductive way. This means that the whole network can be observed in all divisions of the dataset by simply splitting the node labels. The embeddings are computed using the whole graph to preserve the structural and relational information and then, only the embeddings of the training set are fed to the model. The model's performance will be evaluated on the embeddings of the validation and test sets. This transductive approach ensures that the embeddings reflect the global structure of the graph and the dependencies among nodes, even if the labels are only known for a subset of the nodes during training, providing better results in certain cases [27].

The optimiser chosen for the model is ADAM (Adaptive Moment Estimation) [28], which is a variation of the stochastic gradient descent method. Unlike traditional stochastic gradient descent, ADAM retains information about the current state and history of the model parameters as they are updated. This approach allows for an adaptive learning rate, providing a more efficient and effective optimisation process.

The Pytorch libraries [29] and their derivatives are used to create the networks. Once the graphs of seasons 7, 8 and 9 are loaded, certain characteristics of the network are extracted to ensure that they are correctly implemented.

[2]Binary Cross-Entropy Loss measures the performance of a binary classification model by quantifying the difference between the predicted probabilities and the actual binary labels. It is used alongside a sigmoid activation because the sigmoid function maps predictions to a probability distribution between 0 and 1, aligning perfectly with the binary nature of the labels [22]

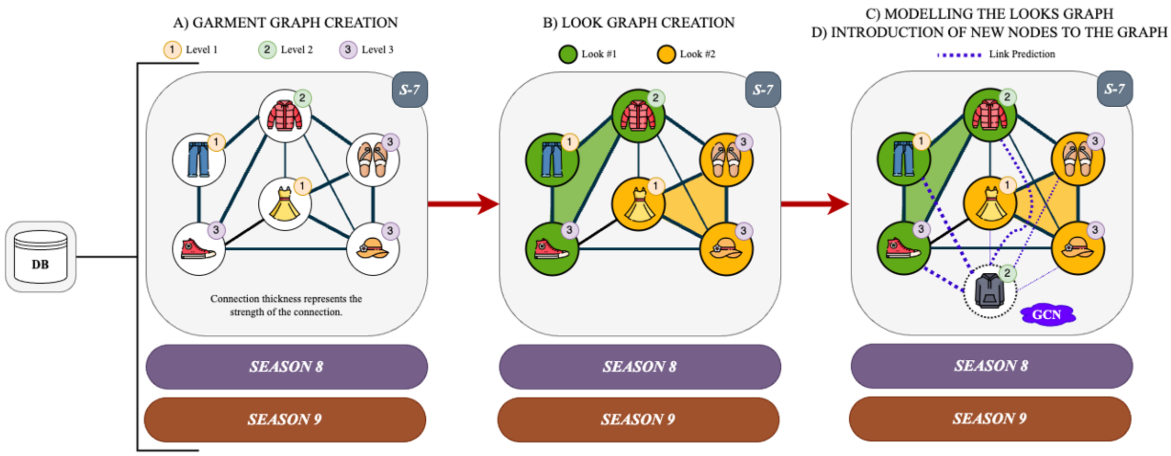


Fig. 3: Diagram showing sections A, B, C & D

This diagram has been designed using images from Flaticon.com

When modelling, it is essential that the data is in numerical format, specifically 'float 32', so a function is developed to convert the variables needed for the prediction into this format by enumerating the different classes of variables. The characteristics of the garments that we will use for prediction are: level, colour, category, application, print, style and weather. After the format change, each of the graphs is converted to the pytorch tensor[3]format and the features of each are assigned in the correct format. The data from each network is then divided into 3 sets as mentioned above and each of the GNNs is run with the different aggregation functions used. The number of channels used in the networks is 128 in the first layer and 64 in the output layer. This number is chosen because testing with a lower number does not allow the network to reach its full potential, and a higher number results in overfitting. The need for a dropout has been considered, but when implemented little change was observed, so it has been decided to not implement it to also reduce complexity in the network. Two layers are used because if neighbours greater than order 2 are taken into account, the nodes would be so similar that there would be little difference between them. Finally, Adam is used as the optimiser, as it is one of the best performers using gradient descent. In order to prevent the networks from running unnecessarily long, an early termination technique is implemented, whose function is to terminate the execution if there is no improvement in the last 20 epochs. A table of the validation results for each model is shown below:

TABLE III: GNN accuracy results.

	GCN	GraphSage	CAT
Season 7	0.8409	0.5065	0.6662
Season 8	0.8323	0.5545	0.6244
Season 9	0.8121	0.5012	0.6468

TABLE IV: GNN number of epochs.

	GCN	GraphSage	CAT
Season 7	25	100	32
Season 8	24	26	21
Season 9	23	23	100

It can be seen that the network with the best results in the 3 seasons is the GCN, so it will be the one used for the predictions.

With the results, some graphs are created for each of the models, measuring the difference between the loss of the training set and the validation set. This approach allows for the implementation of an early stopping mechanism [31]. As an example, the GCN model for season 7 is shown in figure 4. The chart shows that the validation loss flattens and shows minimal improvement after around epochs 5-7. Beyond this point, the training loss continues to decrease, indicating a potential overfitting scenario. Based on this graph, a good candidate for an early stopping point would be around epoch 7, where the validation loss no longer shows significant improvement [32]. This would help prevent the model from overfitting and ensure that the training process is halted at the optimal point, where the model has achieved a good balance between training and validation performance. Another point of interest in the graph is the abrupt change in direction observed between epochs 2 and 3. This could be due to various factors such as noise in the data, adjustments to the learning rate, or simply the inherent fluctuations that occur during the training process. Given that the overall trend in loss continues to be downward, the behaviour of the loss function appears to be within expected norms. It is important to monitor these changes closely, but they do not necessarily indicate a problem

[3]Tensors are a specialized data structure that are very similar to arrays and matrices. In PyTorch, tensors are used to encode the inputs and outputs of a model, as well as the model's parameters. [30]

unless the upward trend in the loss continues over subsequent epochs.

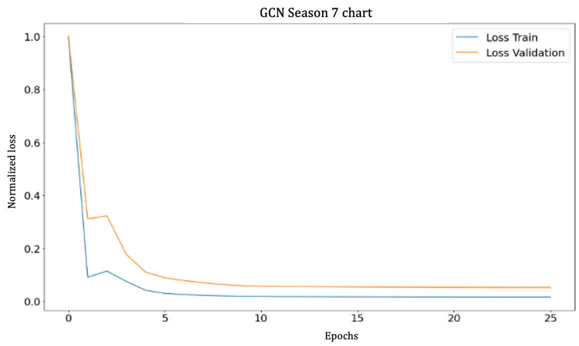


Fig. 4: Difference between train loss and validation loss in the GCN for S7

D. Introduction of new nodes into the graph

The link prediction carried out at the beginning of the project does not use neural networks to link the garments. The whole process is done manually, which can be very computationally intensive and time consuming.

To predict links to possible new garments, GNNs are used. These networks are specifically tailored to handle interconnected data and possess the capability to make predictions at various levels: node, edge, and graph. The relevant features of the aforementioned garments are used to train the networks. As they are based on axis prediction, given a garment, the result of the algorithm will be the axes it should contain based on neighbours with similar characteristics. In this way, new relationships with other garments would be assigned to it, creating new looks.

E. Looks and clients graph creation

In order to simplify the assignment of looks to the 2000 clients, clients with the same values for the selected variables are grouped together. This results in 1820 different groups, which are added as nodes with the following information as attributes

- *Cluster id*: client cluster identifier.
- *Client id*: identifier of the clients belonging to this cluster.
- *Number of clients*: the number of clients in this cluster.
- *Cover*: parts of the body the clients wish to cover.
- *Avoid*: types of clothing that clients want to avoid.
- *Styles*: styles that clients focus on.
- *Adventurous*: how adventurous the customers are.
- *Size bottom*: size of the lower part of the customers.
- *Size top*: size of the upper part of the customers.
- *Type*:type of "customer" node, so that it can be distinguished from nodes of the "look" type.

The network consists of two types of nodes: 'look' and 'client'. As nodes are not connected to those of the same type, the network is bipartite. When creating the links, it is checked that each group of clients is connected to each of the looks, i.e. checks have been made to nodes by pairs. As with the clothing network, a scoring system is designed to decide whether a 'connected' relationship is created between a customer and a look. Once the rules are defined, a link is added between the pair of nodes if they scored one point in size, adventurous, style and avoidance, and more than 0 points when the parts to cover do not match. The weight of the edge is the sum of the value of the 'weight' attribute of the 'look' node and the score obtained in the concealment rule, multiplied by two. As the network consists of 4.184.529 'look' nodes and 1.820 'client' nodes, the comparisons to be made are 7.615.842.780. After many attempts, it is concluded that there is neither the capacity nor the power to perform such a task, so 100 random groups of clients are taken and assigned the relationships.

F. Introduction of new clients to the graph

As already mentioned, a clustering of the customers is created, where each group is comprised of the customers who gave the same answer on all the variables taken into account. These clusters are the nodes in the graph of looks and customers, so they have been assigned looks from different seasons. The procedure followed by the recommendation system when a new customer fills in the form is as follows:

1) Store the answers to the questions that have been taken into account when creating the attributes of the network clusters, i.e. the answers to the questions 'cover', 'avoid', 'styles', 'adventurous', 'size-bottom', 'size-top'.
2) The new client will belong to the cluster that contains those same values in the selected attributes, so that cluster is identified.
3) A certain number of looks related to the cluster she belongs to are recommended. The more weight the look is connected to the client, the more likely she is to be recommended.

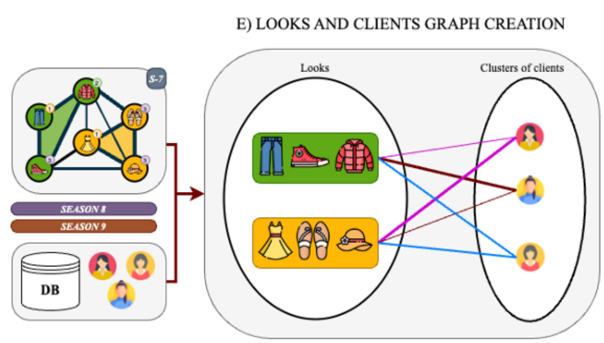


Fig. 5: Diagram showing the looks and clients graph section

This diagram has been designed using images from Flaticon.com

IV. Results

A. Looks graph

The creation of graphs of related garments has been considered satisfactory, since, by means of the rules defined for the connection of these, precise relationships have been obtained. The models have given very good results, exceeding 80%

in each season, as in the case of those using GCN as the aggregation function. Such good results are to be expected because, thanks to the reduction of variables and the correct choice of variables, the model is not excessively complex and can learn adequately. The functionality obtained from these algorithms is the generation of looks on a new product. This is very useful for the company as it allows them to incorporate garments into the seasons in due time instead of having the season totally structured at the beginning. It also brings value to the customer thanks to the functionality of the web mockup and helps the customer to choose looks with their own garments.

B. Looks and clients graph

The structure and process used to set up the customer referral system was costly and therefore did not achieve the desired reality due to the capacity and power of the tools used. Nevertheless, the expected result was achieved with a representative sample of all types of clients. The looks assigned to the clients are correct in terms of the various aspects they have defined and are also different from each other. This means that a client does not get looks that are too similar and has room for exploration. The different rules defined are precise but not too strict, so that no client is left without a look. The result was therefore considered to be very good and applicable to creating value in the company.

C. Practical examples

To test in a practical manner the accuracy of the results with respect to the defined rules has been checked by means of some examples. The first example shown is that of the client cluster no. 242, which contains the following main attributes:

- **Cover**: Arms, waist
- **Avoid**: Skirt, bag
- **Styles**: Classic
- **Adventurous**: Yes
- **Size bottom**: Medium
- **Size top**: Large

According to the information, the customers want to conceal their arms and waist, they don't want skirts or bags, the style they are looking for is classic, they are adventurous and their sizes are 'Medium' bottom and 'Large' top. This group has been assigned 72 different looks, from which 2 random looks have been extracted: look no. 875852 and look no. 1248606.

As it can be seen in figure 6 that the link with 'look_875852' has a weight of 10, while the link with 'look_1248606' has a weight of 14. In this case, it is due to the fact that the garments combine better. Below, the features of the aforementioned looks:

Both look contain classic style garments from S7. In the first look a pair of white trousers, a black top and a burgundy jackets is proposed, while in the second outfit is composed by a pair of dark brown trousers, a black top and a beige jacket. Both looks obey all the rules:

- The trousers are high-waisted to hide the waist.

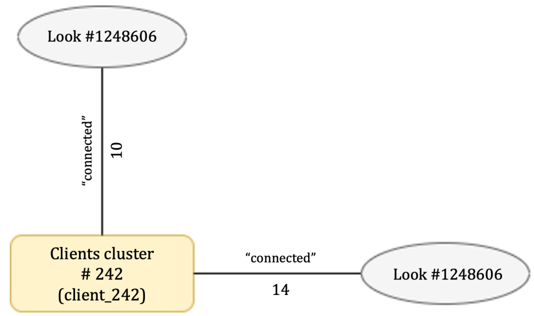


Fig. 6: Examples of looks #1469859 and #1109169) assigned to customer group nº1432.

TABLE V: Client cluster no. 242 attributes.

	Look #1469859	Look #1109169
Product names	Polina pant comfort (Pant) Mejhd top lace (Top) Cressida jacket film (Jacket)	Leopard pant print (Pant) Mejh top lace (Top) French Jacket Lava (Jacket)
Season	7	7
Weight	5	8
Colours	White Black Burgundy	Cream Black Dark Brown
Styles	Classic	Classic
Sleeves	Adjustable sleeve Basic sleeve	Adjustable sleeve Basic sleeve
Necklines	V neckline	V neckline
Shot	High	High
Sizes	Top: S, M, L, XL Bottom: L, XL	S, M, L L, XL
Adventurous	Yes	Yes

- The trousers are available in a medium and the top and jacket are available in a large (sizes defined by the customer).
- They are adventurous, just like the customers.
- They are in a customer defined style; classic.
- They do not contain the garments to be avoided (skirt and bag).

This example exemplifies the verification of the correct assignment of looks to clients.

V. Conclusions

In our work with Lookiero, we have successfully implemented a two-step approach to improving the personal shopping process. First, we created a comprehensive inventory of available garments and assessed their compatibility. Second, we used this inventory to determine the most appropriate outfit combinations for each customer. To facilitate this, we have stored garment data in a graph-based structures where the nodes are garments and the edges represent the compatibility between them. This storage model has significantly streamlined the matching process for workers, allowing them to effortlessly pair garments based on their attributes. In addition, we developed algorithms capable of predicting potential combinations for new garments as they are introduced to the graph. This predictive capability improves the integration of new inventory items and ensures seamless compatibility checks with existing garments. We have also refined the process of personalising outfit combinations for customers using bipartite graphs. One partition of the graph represents the garment sets, while the other partition represents the customers. Through sophisticated algorithms trained on specific variables, we can now provide personalised outfit recommendations to each customer, enhancing the overall shopping experience. This project represents a major advancement in Lookiero's operations, leveraging data analytics and algorithmic predictions to deliver a customised and efficient personal shopping service.

A. Future work

Looking ahead, we believe that process innovations will greatly enhance Lookiero's existing business framework and open up new opportunities to increase revenue streams and attract more customers to their ecosystem. An innovative feature could be integrated into Lookiero's website or mobile application, inviting customers to interact with a diverse assortment of garments from the catalogue and assemble them to create their personal style narratives. This interactive engagement, a likely pastime for users during their downtime, is expected to provide a wealth of data to refine our models and drive cross-selling by incorporating a direct purchase option for these self-designed outfits. In addition, to address the challenges of the current customer referral system, which has not achieved its intended potential due to its cost, we are proposing to implement new, more robust tools with enhanced capabilities. By implementing these advanced tools, we would be able to alleviate the capacity and processing power issues and thereby realise a more effective and efficient referral system.

ACKNOWLEDGEMENTS

This research was made possible by the data provided by Lookiero and the shared expertise on retail analytics.

REFERENCES

[1] Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., ... & Sun, M. (2020). Graph neural networks: A review of methods and applications. AI open, 1, 57-81.

[2] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang y P. S. Yu, "A Comprehensive Survey on Graph Neural Networks", IEEE, 2020.

[3] Sperduti, A., & Starita, A. (1997). Supervised neural networks for the classification of structures. IEEE transactions on neural networks, 8(3), 714-735.

[4] Frasconi, P., Gori, M., & Sperduti, A. (1998). A general framework for adaptive processing of data structures. IEEE transactions on Neural Networks, 9(5), 768-786.

[5] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2008). The graph neural network model. IEEE transactions on neural networks, 20(1), 61-80.

[6] Micheli, A. (2009). Neural network for graphs: A contextual constructive approach. IEEE Transactions on Neural Networks, 20(3), 498-511

[7] Henaff, M., Bruna, J., & LeCun, Y. (2015). Deep convolutional networks on graph-structured data. arXiv preprint arXiv:1506.05163.

[8] Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: a comprehensive review. Computational Social Networks, 6(1), 1-23.

[9] Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. Advances in neural information processing systems, 30.

[10] Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. stat, 1050(20), 10-48550.

[11] Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2018). How powerful are graph neural networks?. arXiv preprint arXiv:1810.00826

[12] Barnard, J. (2023, December 22). Embedding. IBM. https://www.ibm.com/topics/embedding

[13] Liu, X., Sun, Y., Liu, Z., & Lin, D. (2020). Learning diverse fashion collocation by neural graph filtering. IEEE Transactions on Multimedia, 23, 2894-2901.

[14] Cui, Z., Li, Z., Wu, S., Zhang, X. Y., & Wang, L. (2019, May). Dressing as a whole: Outfit compatibility learning based on node-wise graph neural networks. In The world wide web conference (pp. 307-317).

[15] Polyvore dataset. URL https://github.com/xthan/polyvore-dataset.

[16] Gulati, S. (2024). Fashion Recommendation: Outfit Compatibility using GNN. arXiv preprint arXiv:2404.18040.

[17] Yang, X., Du, X., & Wang, M. (2020, April). Learning to match on graph for fashion compatibility modeling. In Proceedings of the AAAI Conference on artificial intelligence (Vol. 34, No. 01, pp. 287-294).

[18] Vasileva, M. I., Plummer, B. A., Dusad, K., Rajpal, S., Kumar, R., & Forsyth, D. (2018). Learning type-aware embeddings for fashion compatibility. In Proceedings of the European conference on computer vision (ECCV) (pp. 390-405).

[19] Chen, Y., Wu, Z., Shen, Z., & Jia, J. (2022, October). Learning from designers: Fashion compatibility analysis via dataset distillation. In 2022 IEEE International Conference on Image Processing (ICIP) (pp. 856-860). IEEE.

[20] Neo4j, https://neo4j.com/

[21] Fernandes, D., & Bernardino, J. (2018). Graph Databases Comparison: AllegroGraph, ArangoDB, InfiniteGraph, Neo4J, and OrientDB. Data, 10, 0006910203730380.

[22] Liu, L., & Qi, H. (2017, March). Learning effective binary descriptors via cross entropy. In 2017 IEEE winter conference on applications of computer vision (WACV) (pp. 1251-1258). IEEE.

[23] Xie, Y., Li, S., Yang, C., Wong, R. C. W., & Han, J. (2020, July). When do gnns work: Understanding and improving neighborhood aggregation. In IJCAI'20: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence,IJCAI 2020 (Vol. 2020, No. 1).

[24] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807-814).

[25] Clevert, D. A., Unterthiner, T., & Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). arXiv preprint arXiv:1511.07289.

[26] Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853

[27] Rossi, A., Tiezzi, M., Dimitri, G. M., Bianchini, M., Maggini, M., & Scarselli, F. (2018). Inductive–transductive learning with graph neural networks. In Artificial Neural Networks in Pattern Recognition: 8th IAPR TC3 Workshop, ANNPR 2018, Siena, Italy, September 19–21, 2018, Proceedings 8 (pp. 201-212). Springer International Publishing.

[28] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[29] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, 32.

[30] PyTorch, https://pytorch.org/tutorials/beginner/blitz/tensor_tutorial.html

[31] Morgan, N., & Bourlard, H. (1989). Generalization and parameter estimation in feedforward nets: Some experiments. Advances in neural information processing systems, 2.

[32] Prechelt, L. (2002). Early stopping-but when?. In Neural Networks: Tricks of the trade (pp. 55-69). Berlin, Heidelberg: Springer Berlin Heidelberg.

AI at Work: International study on AI at work as a vicious or virtuous loop for employees

Benedikt Hackl

Duale Hochschule Baden-Württememberg und TDHBW – Forschungszentrum Management Analytics.

Ravensburg and Munich, Gernany
hackl@dhbw-ravensburg.de

Joachim Hasebrook

TDHBW – Forschungszentrum Management Analytics

Munich, Germnay
hasebrook@management-analytics.de

***Abstract*— We need a more profound understanding of the impact of artificial Intelligence on professional tasks. Some studies assume that AI will replace a quarter of all work tasks and millions of jobs. These studies also predict large productivity gains of billions of US dollars per year. An international study led by TDHBW Research Center Management Analytics in collaboration with academic partners from eleven nations is investigating the impact of AI on work practices.**

Keywords—Augmented Intelligence, Assisted Intelligence, Work Scenarios, AI at Work

I. AI: HUMANS IN THE LOOP OR OUT OF THE LOOP

Global players like investment bank Goldman Sachs assumes that Artificial Intelligence (AI) will replace a quarter of all vocational tasks resulting in a replacement of 300 million jobs (fulltime equivalents) in Europe and the US alone. At the same time McKinsey Global Institute prognoses an increase in productivity worth 2.6 to 4.4 billion US-Dollar, per year. These scenarios focus mainly on the impact of autonomous AI fully replacing tasks like planning and controlling or jobs altogether. However, autonomous AI has not been and will not be the prevalent form of AI in use.

Already established job roles like Data Scientists and trending professions like Prompt Engineer indicate that human stay in the loop and will work with AI more closely and interactively. A recent study of the National Bureau of Economic Research (NBER) showed investment in AI lead to more educated workforces, with higher shares of workers with undergraduate and graduate degrees, and more specialization and significant increases in the share of workers at the junior level and decreases in shares of workers in middle-management and senior level. In our study, we want to focus on augmentation AI can provide to enhance human productivity and how these different forms of augmentation are currently used and will be used in the workplace.

In order to capture the different aspects of augmentation and their impact on work we suggest a task-AI-matrix derived from a thorough review of the current research literature. In our matrix, the levels of augmentation delivered by AI are:

- Rule-based Augmentation: Automation of repetitive, rule-based job tasks and business processes (Ng et al., 2021)
- Narrow Augmentation: assisted intelligence approach in which machines perform tasks but humans make the decisions (Hassani et al., 2000; Schmelzer, 2020; Jalal et al., 2018)
- Human-in-the-loop approach: intelligence augmentation designed to work together with humans supporting their performance and decisions (Lavenda, 2016; Karwowski, 2006; Gebauer et al., 2023)
- More-than-human approach: Extending the scope of abilities of an organization going beyond the abilities of humans or technical systems alone (Thompson & Graham, 2021; Ball et al., 2001; Venturini, 2022)
- Complete Automation (Autonomous AI): humans are completely replaced by machines including relevant aspects of their ability to think and feel (Raisch & Karwowski 2020; Raisch et al., 2020; Biocca, 1996; Bouschery at al., 2023).

These levels of augmentation can be applied to different areas of vocational tasks. These areas are leaning on stages or cycles of activities defined by Canonical Action Research (CAR; Davison et al., 2004).

- Description (Distinction): Better understanding of the subject matter e.g., automated data segmentation, document summaries and reviews (Hassani et al., 2000)
- Diagnosing: Improved error or problem analysis e.g., AI controlled sensors or digital twins (Gebauer et al., 2023)
- Action Planning: Improved and/or more adaptive planning e.g., generation of planning alternatives, critical path analyses (Citroen, 0211; Jalal et al., 2018; Thompson Graham, 2021)

Intervention (Action taking): Supporting and/or automating operational and managerial tasks (Arshad et al, 2022; Alshurideh et al., 2022)

- Evaluation: Better outcome and relationship analyses of operational and managerial tasks performed in the company and its ecosystem (Venturini, 2022)
- Reflection: The accuracy of the learning and the results of general findings, time to reflect on management tasks and/or implemented planning alternatives (Jarrahi, 2018; Thamm et al., 2021)

The background of the different scenarios is summarized in the table depicted in appendix 1.

II. THE STUDY THE RESEARCH CONSORTIUM

A. Types of AI

We want to better understand how different types of AI influence different professional tasks. For example, AI can be used to automate repetitive and rule-based jobs and processes. It can also be used to extend the scope of an organization's capabilities beyond just human capabilities.

Simply put, Artificial Intelligence is the simulation of human intelligence processed by machines, especially computer systems. In this survey, we take a broader view. We believe that "AI at work" can be a win-win situation for both organizations and employees when machines and humans can have a more efficient decision-making process less prone to errors, employees should experience less psychological overload, feel more confident in decisions, and make better decisions thanks to AI support (cf. figure 1).

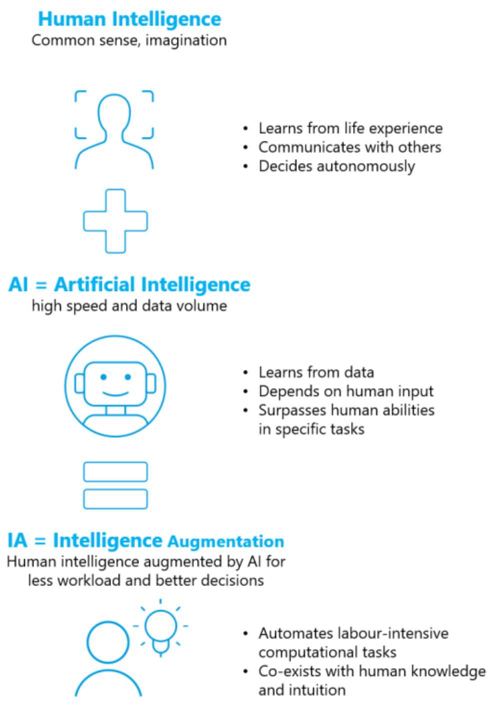


Figure 1: AI Augmentation brings together the strength of artificial and human intelligence.

There are several concepts of using solutions based on AI. Some use specially trained AI systems to automate tasks and certain processes. Others rely on a "human in the loop", that is to say a close interaction between humans and self-learning AI. The use of AI systems can be divided into four groups (cf. figure 2):

1. **Assisted intelligence** - support of routine tasks,
2. **Automated intelligence** - automation of repetitive rule-based processes,
3. **Augmented intelligence** - AI supports and enhances human decision-making,
4. **Autonomous Intelligence** - Learning independently of human input and deciding in ways that also affect humans.

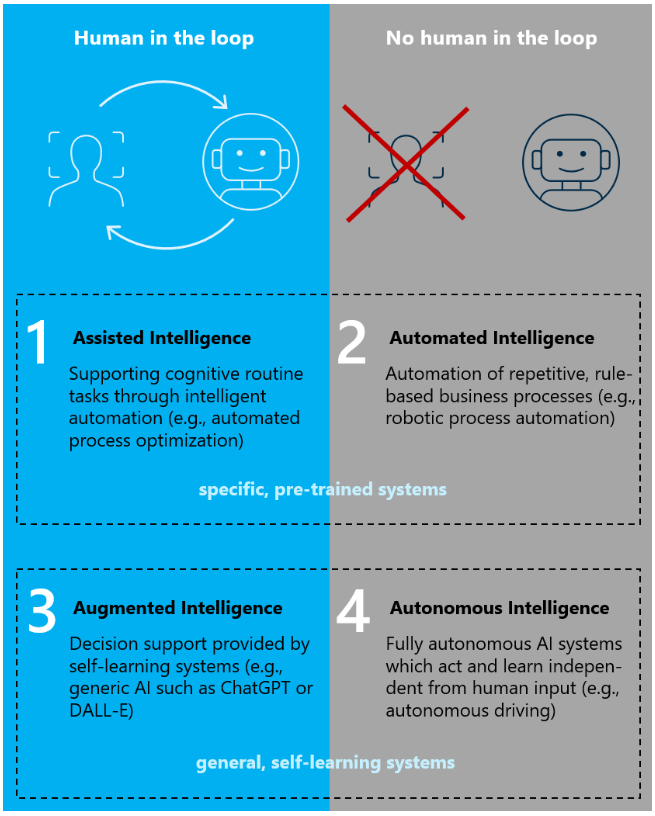


Figure 2: Types of AI tested in the study.

B. Work Scenarios

We ask questions about several scenarios how AI, AI-based analytics and AI-assisted decision making is applied at the workplace (cf. figure 3). The participants are asked to evaluate the questions based on their previous experience using AI. For each of the statements in survey, they rated where their companies applied – or selected the option: I don't know.

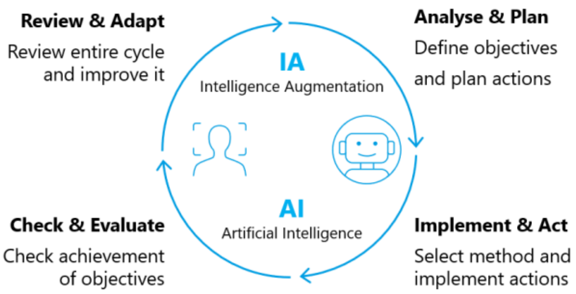


Figure 3: Work scenarios tested in the study about the application of AI.

- The first set of questions is related to the task area "Analyze and Plan": Using AI to set goals and plan actions answering the question: What should it look like?
- The second set of questions is related to the task area "Implement and Act": Using AI to select methods and implement actions answering the question: What and how do we do it?
- The third set of questions is related to the task area "Check and Evaluate": Using AI to check achievement of objectives answering the question: What has been achieved?
- The fourth and last set of questions is related to the task area "Review and Adapt": Using AI to review the entire cycle and improve it answering the question: How can we improve?

C. *Research Consortium*

In our study, we test scenarios in which different types of AI can be applied to different management tasks. We also ask the participants questions about their companies, their professional roles and about themselves. The study is supported and run by a number of international partners: MIT Center for Collective Intelligence (USA), Laboratoire en Intelligence des Données de Polytechnique Montréal (Canada), Bayes Business School (United Kingdom), Faculty of Management, Wrocław University of Science and Technology (Poland), European Research Center for Information Systems (ERCIS), University Muenster (Germany), PriceWaterhouseCoopers (pwc) Germany, Moroccan International Center for Artificial Intelligence, Mohammed VI Polytechnic University (Morocco), Council for Scientific and Industrial Research (CSIR), Pretoria (South Africa), College of Engineering, Abu Dhabi University (United Arab Emirates), Goa Business School, Goa University (India), Lebanon American University Adnan Kassar Business School, Lebanon American University (Lebanon)and Faculty of Management Hanoi University (Vietnam).

III. First, preliminary results

There are currently 302 questionnaires from 30 different countries. The most represented countries are: Germany (n=131 or 43%), India (n= 31 or 10%), Lebanon (n=24 or 8%), Switzerland (n=18 or 6%), United Arab Emirates. Emirates (n=16 or 5%).

A. *Is AI dangerous – or beneficial?*

Most of the participants think their job is safe and their work is supported by AI: 40% feel supported by AI and 18% feel safe in using it. But there are also 28% uncertain about AI's influence on their job prospects, and 14% feel their job might be threatened or endangered. Only a small minority of 6% think that AI is dangerous in general, but 20% feel uncertain. 28% agree that AI is just one tool and 37% agree that AI increases intelligence (summarized in figure 4).

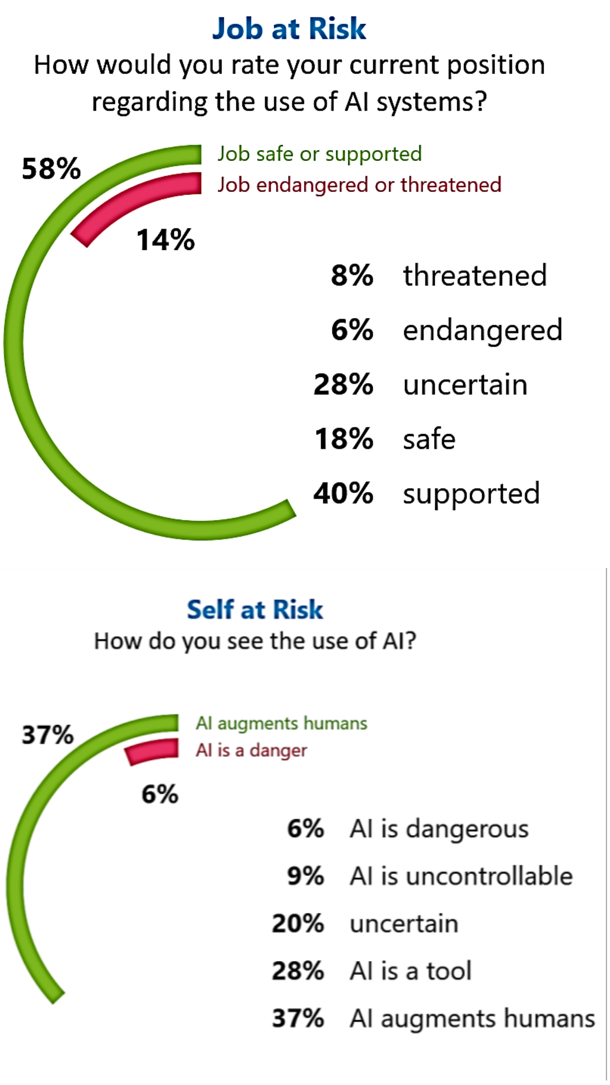


Figure 4: Job at risk and self at risk through AI.

B. *To what extent is AI implemented?*

We asked for the level of implementation of different types of AI in the four phases of the PDCA management cycle (Plan-Do-Check-Act). A large majority of the first 150 respondents already try out some form of AI: Between 86% (Analyse & Plan) and 72% (Implement & Act). But only a small minority has partly or fully implemented AI tools, yet. The level of usage is only slightly higher among younger colleagues (<30 years). But older colleagues (>50 years) are more skeptical about the possibility of replacing their jobs and wary of the way AI can be dangerous in general (see figure 5).

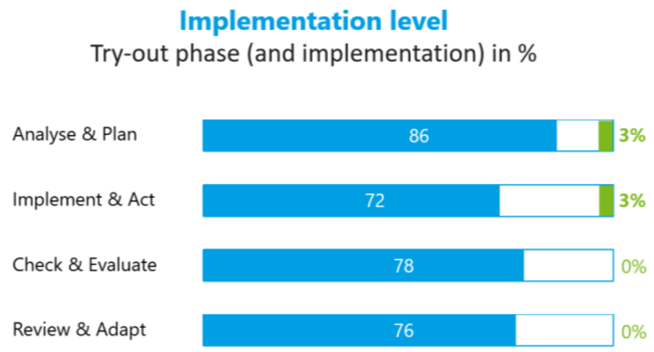


Figure 5: Implementation level of different types of AI.

C. *Does the organisation help to use AI safely and competently?*

We also asked if the use of AI was supported through AI-specific regulations and training. 26% found no policies, guidelines, or regulations and only 16% participated in AI-specific courses, although almost half (45%) had related external courses (see figure 6).

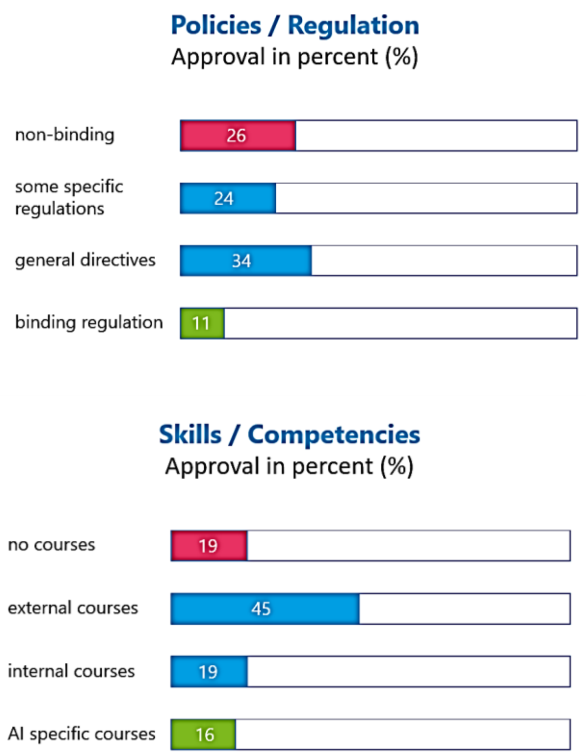


Figure 6: Policies, regulations, and skills concerning the application of AI.

These are just preliminary results and there is more to come – the more people respond to the survey about "AI at Work".

IV. Discussion and Outlook

We plan to continue our research a multinational multi-center, multi-industry study investigating the use of different levels of AI-based augmentation in different task areas. We also aim to investigate maturity levels along the value chain including operations, logistics/distribution, marketing & sales, and services as well as accounting & finance, human resources, and Org / IT, for instance training programs, data warehousing and data protection plans in place to enable augmentation. We also want to record residence, industry (e.g., simplified NACE code list), company/organization size (e.g., revenues, headcount, or FTE), and professional role of the respondent (worker / management, managerial level).

This results in a nested 2x2x4 design with the factors "human in the loop" (yes vs. no), system type (specific pre-trained vs. general self-learning) and the factor task area (1. analyze, 2. act, 3. control, 4. adapt). The 2x2 factors are leading to the AI system types (1. Automated, 2. Assisted, 3. Augmented, and 4. Autonomous intelligence). As co-variates we ask for the international codes describing country, industry, professional role as well as individual features such as age, gender, stay in the position, and a valuation of the risks or chances that AI means for their jobs and for society in general. Given a medium effect size (Eta) of 0.2 and a test power of 0.9 we need about 520 subjects in order to calculate this design.

The study is conducted as multi-lingual online-survey sent out by all participating partners to relevant companies and organization related to them. The partners will not and need not to uncover their contact lists. All data will be collected, stored, and processed fully anonymously. Only summarized and anonymous results will be published. All partners can use the data sets, for instance for their own country specific reports and articles. All partners of the study will publish the summarized results together in a technical paper as well as an article submitted to an academic, peer-reviewed journal. The project leads seek to find sponsorship and funding to support the technical costs of the study and built-up an on-going monitoring about 'AI at Work'.

V. Study Heads

Benedikt Hackl

Professor für Unternehmensführung und Personal an der DHBW in Ravensburg, Wiss. Leiter des Forschungszentrums Management Analytics. Schwerpunkte sind die Neupositionierung von Teamarbeit, Führungs- und Steuerungssystemen sowie Innovationsproduktivität.

Joachim Hasebrook

Professor für Personal- und Organisationsentwicklung an der Steinbeis Hochschule, Wiss. Leiter des Forschungszentrums Management Analytics. Schwerpunkte sind KI und HR, Teamarbeit und Lernen.

VI. References

Alshurideh, Muhammad Turki; Al Kurdi, Barween; Alzoubi, Haitham M.; Ghazal, Taher M.; Said, Raed A.; AlHamad, Ahmad Qasim; Hamadneh, Samer; Sahawneh, Nizar; Al-kassem, Amer Hani (2022). Fuzzy assisted human resource management for supply chain management issues. ANNALS OF OPERATIONS RESEARCH.

Arshad, Muhammad; Kun Yu, Chen; Qadir, Aneela; Ahmad, Waqar; Rafique, Muhammad (2022). The impact of big data analytics on organizational sustainability: The influencing factors of autonomous research and development and absorptive capacity. INTERNATIONAL JOURNAL OF ENGINEERING BUSINESS MANAGEMENT.

Ball, M.O., Chen, C.Y., Hoffman, R., et al., 2001. Collaborative decision making in air traffic management: current and future research directions. In: Bianco, L., Dell'Olmo, P., Odoni, A.R. (Eds.), New Concepts and Methods in Air Traffic Management. Springer Berlin Heidelberg, Berlin, Germany, p.17-30.

Biocca, Frank. "Intelligence augmentation: The vision inside virtual reality." In Advances in psychology, vol. 113, pp. 59-75. North-Holland, 1996.

Bouschery, Sebastian G., Vera Blazevic, and Frank T. Piller. 2023. "Augmenting Human Innovation Teams with Artificial Intelligence: Exploring Transformer-Based Language Models." Journal of Product Innovation Management 40(2): 139–153

Citroen, Charles (2011). The role of information in strategic decision-making. International Journal of Information Management, 31(6), 493-501.

Davison, Robert, Maris G. Martinsons, and Ned Kock. "Principles of canonical action research." Information systems journal 14, no. 1 (2004): 65-86.

Gebauer Niklas WA, Schütt, Kristof T., Stefaan SP Hessmann, , Jonas Lederer, and Michael Gastegger. "SchNetPack 2.0: A neural network toolbox for atomistic machine learning." The Journal of Chemical Physics 158, no. 14 (2023).

Hassani, Hossein, Emmanuel Sirimal Silva, Stephane Unger, Maedeh TajMazinani, and Stephen Mac Feely. 2020. "Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future?" AI 1, no. 2: 143-155. https://doi.org/10.3390/ai1020008

Jalal, T.D. Distinguishing between Narrow AI, General AI and Super AI 2018. Available online: https://medium.com/@tjajal/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22

Jarrahi (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577-586.

Karwowski, Waldemar. "The discipline of ergonomics and human factors." Handbook of human factors and ergonomics (2006): 1-31.

Lavenda, D. Artificial Intelligence vs. Intelligence Augmentation 2016. Available online: https://www.networkworld.com/article/3104909/artificial-intelligence-vs-intelligence-augmentation.html

Ng, K. K. H., Chen, C.-H., Lee, C. K. M., Jiao, J., & Yang, Z.-X. (2021). A systematic literature 16 review on intelligent automation: Aligning concepts from theory, practice, and future 17 perspectives. Advanced Engineering Informatics, 47, 101246.

Raisch Florian, Kurpiers, Christina, Biebl, Bianca Hernandez Julia Mejia. "Mode awareness and automated driving—What is it and how can it be measured?." Information 11, no. 5 (2020): 277.

Raisch, S., Krawkowski, S. (2020). Artificial Intelligence and management: The automation-augmentation paradox. Academy of Management Review. Preprint AMR-2018-0072.R4

Schmelzer, R. Artificial or Human Intelligence? Companies Faking AI. Forbes, 4 April 2020.

Shrestha et al. (2021). Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. Journal of Business Research.

Takayanagi et al. (2022). AI-Assisted Decision-Making and Risk Evaluation in Uncertain Environment Using Stochastic Inverse Reinforcement Learning: American Football as a Case Study. Mathematical Problems in Engineering.

Thamm et al. (2021). Concept for an augmented intelligence-based quality assurance of assembly tasks in global value networks. Procedia CIRP, 97, 423-428.

Thompson; T.L., Graham, B. (2021). Researching AI at Work. Alternative narratives for humans and AI systems as co-workers. In N. B. Dohn ezt al. (Eds.), Conceptualizing and innovating education and work with networked learning (171-187). Heidelberg: Springer.

Venturini, Francesco (2022). Intelligent technologies and productivity spillovers: Evidence from the Fourth Industrial Revolution. Journal of Economic Behavior & Organization, 194, 220-243.

APPENDIX 1: MATRIX TYPES OF AI X WORK SCENARIOS

		Rule-based Augmentation	Narrow Augmentation	Human-in-the-loop approach	More-than-human approach	Complete Automation (Autonomous AI)
Description (Distinction)		- automate repetitive business processes, which plays a key role in mimicking routine manual tasks and workflow processes via the advancement of IT (see Ng et al. (2021), p. 4)	- "Assisted-intelligence Approach": machines might undertake the action, but humans are making the decisions. (see Hassani et al. (2020), p. 148) - improves what people and organizations are already doing (see Schmelzer (2020)) - Reactive machines: do not involve memory-based operations and were the first AI-based system machines, and thus have very limited capability (see Hassani et al. (2020), p. 145-146) - Limited-memory machines: are capable of learning from historical data to inform subsequent decisions (e.g. chatbots, virtual assistants, self-driving vehicles, etc.) (see Hassani et al. (2020), p. 145-146)	- Intelligence Augmentation is designed to work with people and focus on building systems that augment and support human cognition. (see Hassani et al. (2020), p. 147) - exploits IT for supplementing or support human cognition whilst leaving the human at the center of human-computer interaction. (see Lavenda (2016), online) - places humans at the core of the system and decision-making. (see Masih (2018), online) - is interpreted as a model that entails human interaction. (see Karwowski (2006))	- enables organizations and people to do things they could not otherwise do (see Schmelzer (2020)) - Theory of mind: can understand as entities with which they interact by discerning their needs, emotions, beliefs, and thought processes. (see Hassani et al. (2020), p. 145-146) - Understanding the larger shifts and the ethical implications demands sensibilities, theory, and methodologies to see human-technology together as the phenomena of interest (see Thompson et al. (2021), p. 184) - enable the researcher to examine and articulate how narratives, and therefore, knowledge about AI emerges and moves via complex social, technical and political constellations of actors, texts, and technologies as a form of assemblage (see Thompson et al. (2021), p. 185)	- human is completely automated away (see Raisch (2020)) - "ability to "think" and "feel" as humans": Self-aware AI -> involves AI systems that have evolved to the point where they are comparable to the human brain in that they have developed self-awareness. (see Hassani et al. (2020), p. 145-146) - believes in autonomous systems that can imitate or replace human cognitive functions. (see Lavenda (2016), online) - wants to produce an independent machine. (see Biocca (1996), p. 59-75) - Close linkage to particular groups (e.g. innovation team for new product development) (see Bouschery et al. (2023))
Diagnosing	*Objectives*	- Integration of different datasets to facilitate repetitive business processes - Use of logical rules to augment data/humans	- Design an AAI to improve the performance and capabilities of AI systems within a specific domain or task. - data curation	- Handle uncertainty and ambiguity: Humans can help handle ambiguous inputs, interpret context, resolve contradictions, or make subjective judgments that may be difficult for AI algorithms alone	- leveraging the capabilities of non-human entities to enhance performance, address complex tasks, and foster collaboration and resilience	- fully automated algorithm that replace human cognitive functions
	Privacy	- ensure that any data used in rule-based augmenation is collected, stored, and processed in compliance - consider the potential impact of the augmented	- these systems don't perform outside of the single task that they are designed to perform (see Jalal (2018), online)	- integrate machine learning-generated suggestions and human annotation decisions (see Niklas et al. (2023), p. 1)	- Understanding, acknowledging, and then pedagogically addressing these perceptions in order to clarify and educate workers and the publics (see Thompson et al. (2021), p. 184)	

		data on the privacy of individuals or groups				
Action Planning	*Respons-ibilities*	- Systems that analyze and visualize structured data	- Systems that are programmed to perform a single task (see Jalal (2018), online)	- Systems that that entails human interaction (see Karwowski (2006)). They blends exceptional human intelligence with the most reliable machine intelligence (see Hassani et al. (2020), p. 151)	- Systems that provide a way to conceptualize, attune to, and study the complex interactions that unfold between AI systems, workers, ways of working, workplaces, policies, and public narratives (see Thompson et al. (2021), p. 183)	- Systems that combine models and services to augment and extend human cognition
Intervention (Action taking)	*Prevention (bias)*	- Pattern analysis in data/datasets	- Curating data/datasets			
Evaluation & Reflection	*Level of in-volvement*	- machines provide explorations, alerts and other support for human decision makers		- Humans are directly engaged in training, tuning and testing data for a particular AI algorithm such as machine learning. (see Hassani et al. (2020), p. 151)	- More-than-human sensibilities align with methods such as controversy mapping (e.g. Venturini, 2022) or networked ethnography (Ball et al., 2001) (see Thompson et al. (2021), p. 184)	- autonomous decisions are made by machines

Fostering students' AI literacy: An approach to deeper learning in empirical research initiatives

Ulrich Bucher
Studienzentrum Dienstleistungsmanagement
DHBW Stuttgart
Stuttgart, Germany
ulrich.bucher@dhbw-stuttgart.de

Markus Schwarzer
Studienzentrum Dienstleistungsmanagement
DHBW Stuttgart
Stuttgart, Germany
markus.schwarzer@dhbw-stuttgart.de

Abstract— **In an era characterized by AI, AI literacy and its application to solve unstructured problems is becoming a key competence. The article describes a concept for promoting AI literacy through a deeper learning approach in which students solve unstructured problems from practice in empirical research projects using AI over five semesters and various courses. Based on the determination of the challenges of an empirical research project and the comparison with the support provided by AI, use cases for the application of AI are determined. With the AI-supported literature research, one use case is explained in more detail. The competence requirements resulting from this are then compared with the competence set of the AI Literacy Framework by Long and Magerko.**

Keywords — AI Literacy, Artificial Intelligence, ResearchBased Learning, Critical Thinking, Problem-Solving, Unstructured Problems

I. INTRODUCTION

AI literacy is becoming a central interdisciplinary competence in an environment that is increasingly being changed and shaped by AI [1]. Due to the rapid spread of AI, Ng even considers AI literacy to be a key competence that everyone should possess [2]. However, research into the design of competence programs to promote AI literacy is still in its infancy [3]. This article therefore aims to make a contribution to this field of research by developing a concept for promoting students' AI skills.

AI literacy is understood below, based on Long and Magerko, as a set of skills that enable a critical assessment and responsible and legally compliant use of AI, taking ethical aspects into account [4]. Based on this definition of AI, it is necessary to clarify which set of skills make up AI literacy and which skill levels are required. Just how difficult it is to answer these questions is made clear by the large number of frameworks that have been created for this purpose [2, 4, 5]. Long and Magerko, for example, identify 17 different competencies that emerge from five main questions: "What is AI", "What can AI do", "How does AI work", "How should AI be used" and "How can people perceive AI" [4].

However, if we look at the competencies derived from answering these questions, the framework appears incomplete and not very concrete. For example, Long and Magerko only link the question of how AI should be used to a single competence: ethics [4]. Ultimately, however, the question is only inadequately answered with reference to the ethical aspects. Rather, users must be able to develop suitable use cases, integrate them into their work processes in a meaningful way and they need user knowledge of the AI tools used. For example, how suitable prompts are created or which data the AI tools used are based on (e.g. training data used for a research assistant such as Elicit).

The application of such frameworks is made more difficult by the fact that they often do not adequately define the concept of competence. This is surprising insofar as AI literacy in Long and Magerko's definition is nothing more than a set of different competencies. Ultimately, however, it remains unclear what is meant by competencies in Long and Magerko's framework. The question arises as to whether these are understood as general or context-specific cognitive performance dispositions, whether the competencies refer to a broad range or specific classes of situations and requirements and whether they include a motivational orientation (or, in other words, which variant of the concept of competence [6] is used within the framework).

The dynamic development of AI and the associated changes (e.g. in work processes) also raise the question of whether current frameworks for AI literacy are even capable of identifying the relevant competencies. In the past, it has repeatedly been shown in volatile environments that it was neither possible nor useful to determine the relevant competencies [7].

Overall, the current AI literacy frameworks offer at best a rough orientation for the development of concepts to promote AI literacy. For this reason, and due to the urgent need to promote AI competence, concrete AI use cases were developed using a bottom-up approach. Their starting point is a concrete situation - the implementation of an empirical research project. Based on this, use cases for the application of AI are developed. To this end, the challenges of an empirical study are compared with the support services provided by AI. The skills required for these are then derived on the basis of the use cases.

Accordingly, a narrow understanding of the concept of competence is chosen in this article, which includes the situation, the tasks and the requirements. In this article, competencies are understood as context-specific cognitive performance dispositions that relate to the implementation of an empirical research project in the context of a university degree program and the associated challenges.

II. PROMOTION OF AI LITERACY THROUGH DEEPER LEARNING

The choice of deeper learning as a frame of reference for the development of a concept to promote AI literacy is motivated in particular by the fact that the use of AI is associated with a large number of dangers and risks [8]. For example if internal company information is entered into the AI, expert interviews are transcribed without question or copyrights are violated [1]. Responsible use of AI therefore requires a deep understanding of it. In addition to user knowledge, students also need an understanding of AI technologies, they must be able to critically reflect on the outputs of AI and they must be able to evaluate legal and ethical aspects of AI use. All of these skills require an in-depth examination of AI.

The core idea of deeper learning is to develop a deep understanding of a topic and to promote lasting learning [9]. Deeper learning aims to give students a conceptual understanding of the topic and should enable them to transfer what they have learned to new situations [10]. Strong anchors should be created from which the students will benefit for a long time. Deeper learning represents a counterpoint to superficial learning (often referred to as bulimic learning), in which knowledge is only memorized, never really understood and quickly forgotten again [10].

The choice of deeper learning as a frame of reference was also motivated by the fact that this approach promotes critical thinking [3]. The importance of critical thinking skills is emphasized in numerous contributions to AI Literacy. In the Long and Magerko framework, this is one of the competencies underlying AI literacy [4]. Users should be empowered to question the trustworthiness and intelligence of AI [4]. In our own educational research, we came to the conclusion that students predominantly do not have the necessary AI literacy skills in the area of critical thinking [11]. For example, students often accept the output of ChatGPT without reflection, they largely ignore inconsistencies between the task and the output of ChatGPT and they do not bother to think through the statements of the output logically enough [11].

It is therefore important to promote critical thinking skills among students. One of the ways in which deeper learning promotes critical thinking is that solving a problem is a process that requires methodical consideration and critical thinking in order to find a suitable approach to achieve the desired goals [12].

Deeper learning is also closely associated with the ability to solve problems [13]. Problem-solving skills are one of the six skills that result from deeper learning according to a panel of experts conducted by the Hewlett Foundation [13]. This reference is obvious, as in deeper learning students are regularly confronted with complex, unstructured problems from practice.

There is broad agreement in the scientific literature that standard tasks are increasingly being automated by AI [14]. As a result, the world of work is changing and, in particular, the tasks that remain for humans [15]. The latter include solving unstructured problems - a task that AI is currently unable to perform and is therefore increasingly moving to the forefront of human work [16].

III. CONCEPT

The aim of the concept is to provide students with the competencies and thus the knowledge, skills and abilities that enable them to solve the challenges associated with an empirical study with the help of AI.

A. Conceptual framework – Deeper Learning

Deeper learning was used to develop a conceptual framework. In order to enable deeper learning among students, the following framework conditions were defined in the conceptual design:

- Project-based learning: Students carry out an empirical research study in groups. In order to promote problem-solving skills, the study must address a problem from operational practice.
- Self-directed learning: Students have a great deal of freedom in choosing the problem, the research question, the design and the tools and methods used. Only the field of research is specified. This is made necessary by the fact that the students are supported by a supervising professor as a coach for the entire duration of the project. Otherwise, the content taught in the courses is based on the progress made by the groups. For example, methodological issues such as data cleansing or legal aspects such as copyright are addressed if they are to be applied within the scope of the project. This also follows the realization that in the perception of many students in a digital AI era, knowledge is only a prompt away, which reduces the willingness to learn in advance.
- Interdisciplinary learning: In order to enable a holistic understanding, the concept combines different subject areas and courses. The project is not only dealt with in courses on empirical research, but is also integrated into courses in law, subject-specific lectures and academic work. In addition, different lecturers are involved throughout the project, many of whom have very different professional backgrounds. In this way, students are confronted with a variety of perspectives, which helps them to develop their own points of view.
- Feedback and reflection: The groups receive regular feedback (e.g. on their progress and their approach) and are encouraged to reflect. The supervising lecturers do not dictate decisions to the students, but merely provide them with impulses. It is the students' own responsibility whether and how they take up these suggestions.
- Integration of the internet / AI: Students are encouraged to access resources via the internet or to use AI. One example of this is the use of the website statistikguru.de, which includes various instructions on how to carry out statistical data analyses with SPSS. By using this website, students are given guidance on how to proceed when carrying out a data analysis and how to interpret the results. At the same time, the integration of the internet / AI opens up the learning process and does not follow a fixed sequence, but allows new impulses to be integrated spontaneously.

B. Challenges that students face in empirical research projects

In this concept, AI literacy is promoted by teaching students how to master the challenges of an empirical study with the help of AI. It is therefore important to identify the challenges associated with an empirical study. To this end, the experiences from previous courses were used and compared with the literature.

Prior to the development and introduction of the concept to promote AI literacy, teaching-integrated research projects were carried out over several years in two courses in the third and fourth semesters. The design of the study was developed and the data collected in the third semester. The data analysis then took place in the fourth semester. As part of these teaching-integrated research projects, students had the option of pursuing their own research projects in small groups or participating in teaching-integrated research projects carried out in cooperation with a dual partner.

In line with empirical studies on the challenges of empirical work, the courses showed that students are confronted with numerous challenges. For example, in a study by Sitompul and Anditasari, around a third of students found the definition phase to be difficult [17]. Many students rush into developing a solution far too quickly, even though they have not even penetrated the problem yet [18]. Each phase of the research process has its own challenges and shortcomings in a single phase often cannot be compensated for. Cornford, for example, points out that the deficiencies in data analysis are often so severe that the data collected with considerable effort does not result in knowledge and is not used to build an argument [19]. In a study by Sitompul and Anditasari, two thirds of the students surveyed were of the opinion that interpreting, transforming and drawing conclusions from the data was difficult [17]. Even if the first four phases of an empirical study have been successfully completed, the high demands on scientific language still present hurdles for students [20].

scientific databases require knowledge of the keywords, which are not known at the beginning of a literature search.

The scientific literature was used to determine the support provided by AI in overcoming the challenges. Scientific research has dealt with the support provided by AI in numerous studies. A growing number of research papers use AI for empirical research, whereby the use of AI is not only extensively documented, but also comprehensively discussed and researched. In this way, extensive (empirical) knowledge on the use of AI in scientific empirical research has been acquired. Since the challenges of a student empirical research project largely coincide with the usual challenges of an empirical study, students can also benefit from this knowledge.

For example, Liang et al. addressed the question of whether large language models can provide helpful feedback on research work [22]. Their broad-based comparative study of human and machine feedback showed a high degree of overlap between the two types of feedback [22]. Although machine feedback cannot replace human feedback, it can complement it [22]. In particular, the large language models often found it difficult to criticize the research design of a

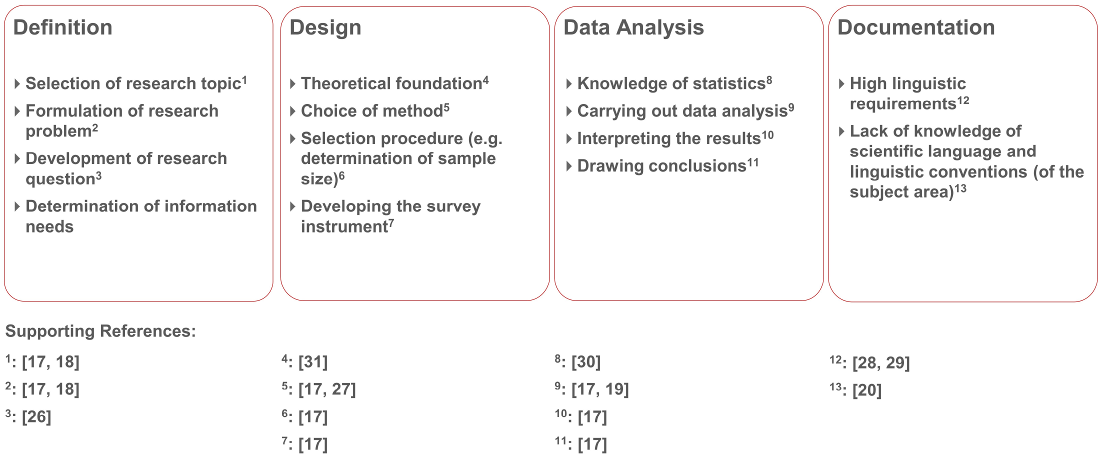


Fig. 1. Challenges that students face in the various phases of an empirical research project

Challenges that arise in the data collection phase were not taken into account in the concept development. The reason for this is that the support provided by the AI should primarily take place in the other phases, as data collection proved to be comparatively less challenging in the previous courses.

C. Support provided by the AI

The supervision of students during the empirical research projects made it clear that most of them were not yet aware that AI can provide significant support for many of the challenges of an empirical study. This corresponds with a study by Garrel, Mayer and Mühlfeld, which deals with the usage behavior of AI by students [21]. The use of students is largely limited to generative AI and AI translation tools [21]. All other AI tools were used by less than 1% of the students in the study [21].

Surveys in the supervised courses showed that, for example, research assistants such as Elicit or visualization tools such as ConnectedPapers were largely unknown to the students. Accordingly, the students found it difficult to identify seed papers, among other things, especially as the

study in detail [22]. A key advantage of machine feedback, on the other hand, is that it can be provided promptly if required, without placing demands on other people [22].

The results of scientific studies on the support services of AI show that it can positively influence both the efficiency of an empirical research project and the quality of its results [23]. For example, the use of speech-to-text AI tools can significantly reduce the time and effort required to transcribe interviews, freeing up resources for more demanding activities [24]. At the same time, however, this can also give rise to data protection or contractual problems [23]. Another example is the use of AI methods to recognize patterns in large data sets that would otherwise not have been detected.

D. Development of use cases and their integration into the curriculum

The comparison of the challenges mentioned under 3.2 with the support provided by AI led to the definition of use cases for the use of AI. The use cases aim to enable students to master the challenges of an empirical research project. The following figure shows the identified use cases and their location in the curriculum.

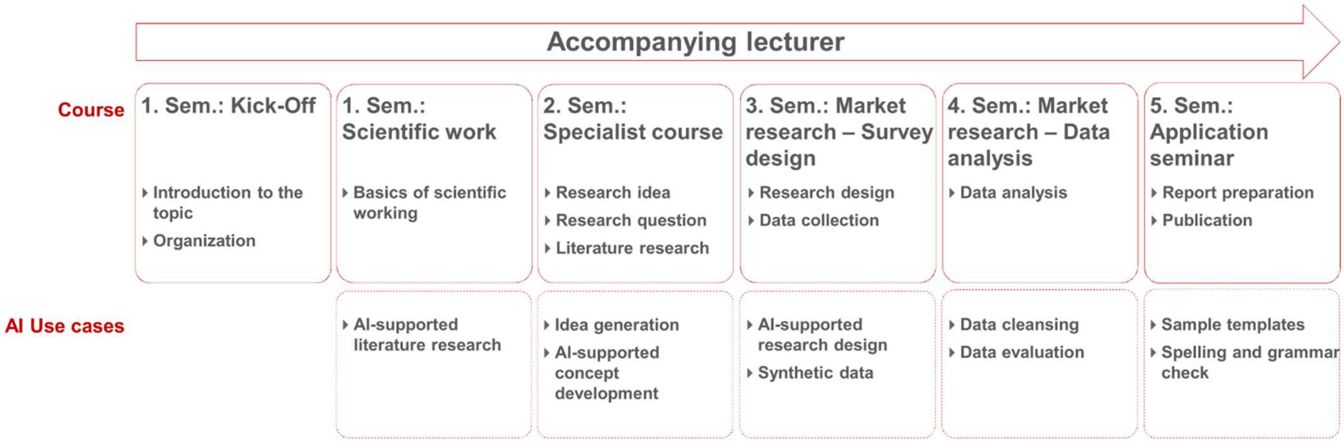


Fig. 2. Use cases and their location in the curriculum

The use cases describe the challenges addressed by the use of AI, the support provided by AI, the desired result of using AI, how AI is applied and the technical aspects of the AI tools used. Potential risks and ethical aspects are also discussed.

E. *Example use case*

One of the use cases, AI-supported literature research, is described below as an example.

Initial situation: The AI-supported literature research is the first application of AI in the empirical research project. At this stage, students usually have at best a rough idea of the goals they want to pursue with their research project. The AI-supported literature research aims to support the students in developing their research question. This is done by identifying the first publications in their field of research and examining their research objective, methodological approach and findings.

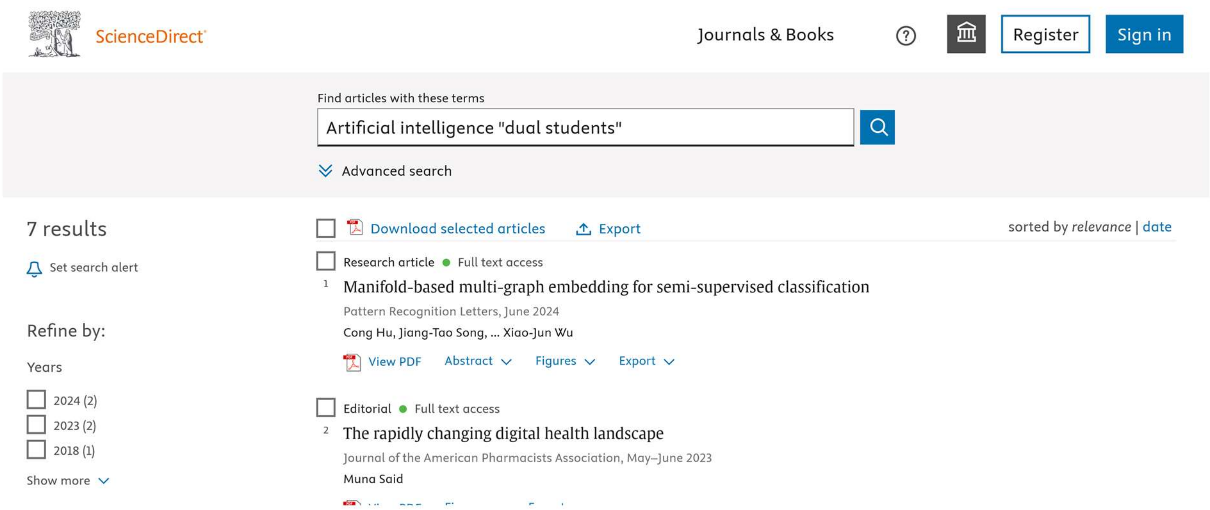
Procedure: The use case initially addresses the challenges that students usually encounter in this phase of the research project when using a scientific database. This offers the opportunity to discuss the functioning of scientific databases as well as the use of different search strategies (e.g. from general to specific, forward search, backward search). In this way, three central challenges are derived: The identification of a seed paper, the traceability of the literature search and the determination of the relevant scientific literature. We then show how these challenges can be solved through the use of AI. AI search engines such as Elicit and Consensus are presented in order to arrive at the first seed papers. Visualization tools such as ConnectedPapers, Litmaps and ResearchRabbit are then used to identify the collection of material. These also help to identify the relevant scientific literature. This is followed by a description of how to proceed with a systematic literature search. To illustrate this, keywords are extracted from the collection of material. A search string is generated from these using generative AI (whereby the generative AI is instructed to supplement the keywords). For each tool used, it is shown how the prompting was carried out in scientific studies, the functionality of the tools is explained (e.g. which training data they are based on), the strengths and weaknesses are discussed and the dangers and ethical aspects are addressed. It also shows how students integrate the use of AI tools into the literature research process.

Alternative scenarios: The AI-supported literature search is contrasted with the classic literature search as an alternative.

Prerequisites: Students must have at least a rough idea of the research objective they are pursuing with their empirical study. Previous knowledge of scientific work is favorable, but is not mandatory.

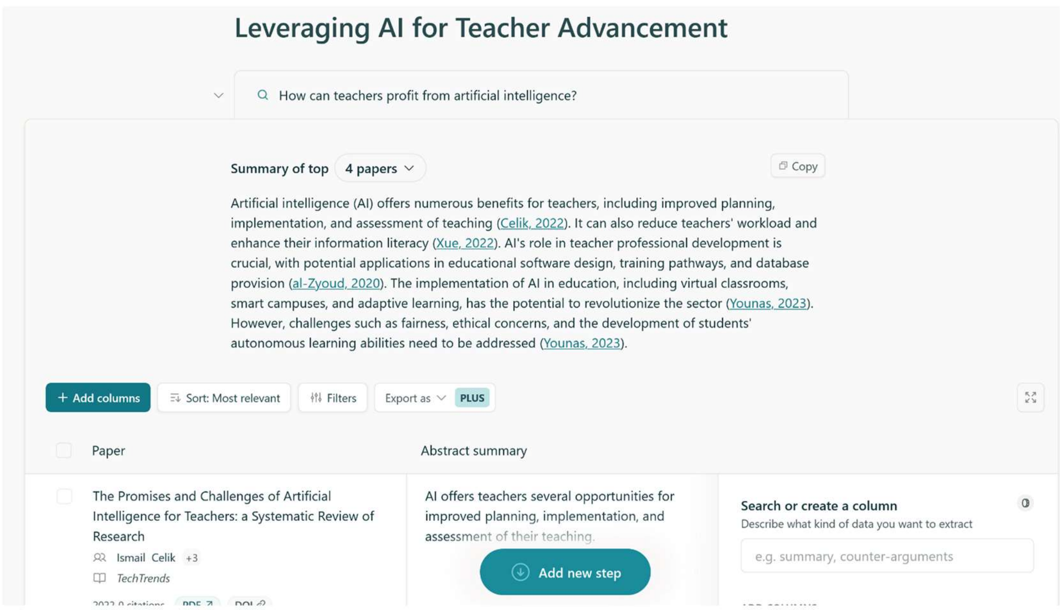
The following excerpts from the set of slides developed for the use case illustrate the process and the content covered.

Use case: AI-supported literature search



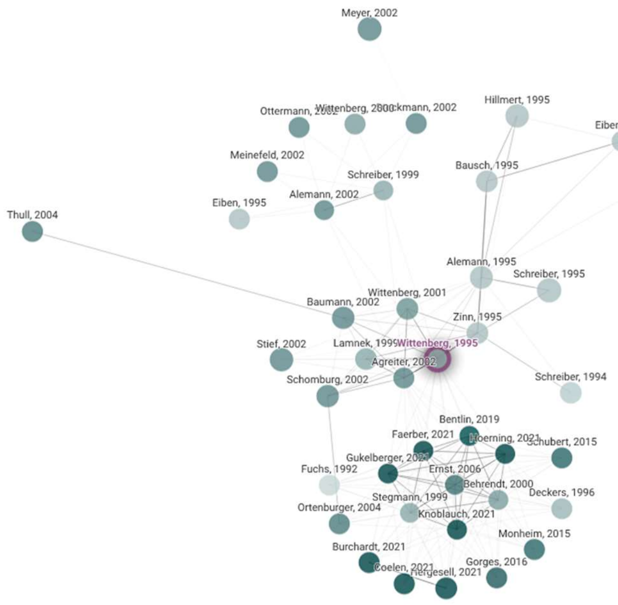
Source: Screenshot ScienceDirect (2023)

Finding a seed paper with Elicit



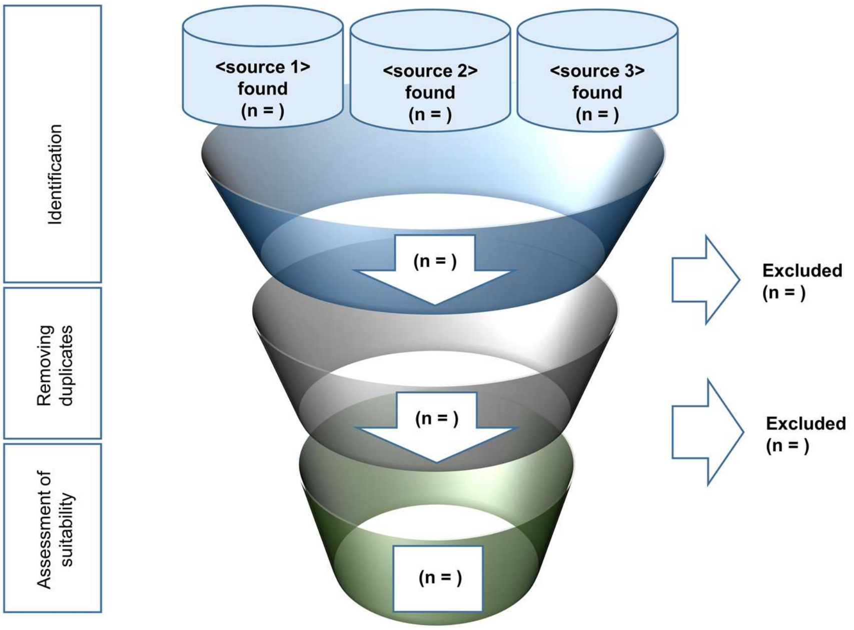
Source: Screenshot Elicit. In: Bucher, U./ Schwarzer, M./ Holzweißig, K. (2023): Künstliche Intelligenz für die wissenschaftliche Arbeit. Mit ChatGPT & Co. den KI-Turbo zünden, Stuttgart

Visualization tools: ConnectedPapers, Litmaps, ResearchRabbit & Co.



Source: Screenshot Connected Papers. In: Bucher, U./ Schwarzer, M./ Holzweißig, K. (2023): Künstliche Intelligenz für die wissenschaftliche Arbeit. Mit ChatGPT & Co. den KI-Turbo zünden, Stuttgart

Systematic literature search: flow chart



Source: based on Ziegler, A./ Antes, G./ König, I. (2011): Bevorzugte Report Items für systematische Übersichten und Meta-Analysen: Das PRISMA-Statement. In: Dtsch med Wochenschr 136 (08), e9-e15. DOI: 10.1055/s-0031-1272978.

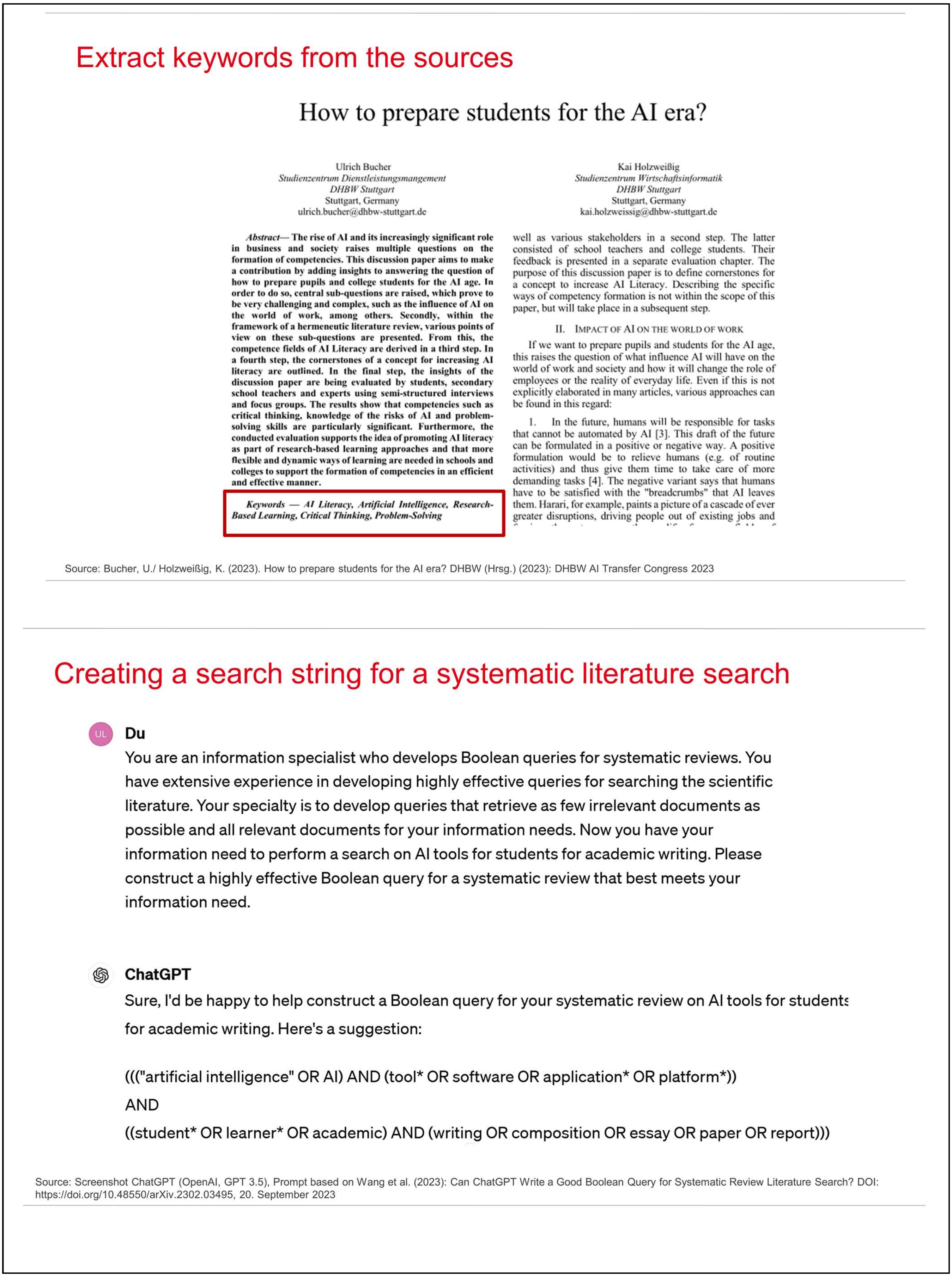


Fig. 3. Excerpts from the slide set on the use case of AI-supported literature research

By comparing traditional and AI-supported literature research, the skills that are added with the use of AI and therefore require special support were identified.

Professional competence

- Know and understand how the AI tools used work
- Understanding the relevance of visualization tools (graphs) in the context of a literature search

Methodological competence

- Creating and interpreting graphs
- Creating a search string with the help of generative AI
- Designing the literature search process: from a seed paper to a systematic literature search
- Linking different AI tools (e.g. exporting Bibtex and importing it into Zotero)
- Combining traditional techniques and procedures with AI

Personal competence

- Critical thinking (in particular assessment of credibility and quality)

Competence to act

- Selecting AI tools: Which tool for which purpose?

F. *Comparison of competencies with the AI Literacy Framework by Long and Magerko*

If we compare the skills identified on the basis of the use case with the skills from the AI Literacy Framework by Long and Magerko, considerable differences become apparent. Numerous skills play no role or at best a subordinate role for the use cases. One example of this is knowledge of the steps of machine learning as well as the practices and challenges that each step entails [4: competence 9]. These two competencies play no role in the present concept. The same applies to recognizing AI [4: competence 1]. The concept also does not enable students to imagine future AI applications and their impact on the world [4: competence 7].

The following table compares the competence set of the AI Literacy Framework by Long and Magerko with the competence set derived from the use case for AI-supported literature research.

Competence set AI Literacy Framework (Long and Magerko) [LM20]	*Competence set AI Literacy Use Case literature research*
Recognizing AI: Distinguish between technological artifacts that use and do not use AI	X
Understanding Intelligence: Critically analyze and discuss features that make an entity "intelligent", including discussing differences between human, animal, and machine intelligence.	X
Interdisciplinarity: Recognize that there are many ways to think about and develop "intelligent" machines. Identify a variety of technologies that use AI, including technology spanning cognitive systems, robotics, and ML.	X
General vs. Narrow: Distinguish between general and narrow AI.	X
AI's Strengths & Weaknesses: Identify problem types that AI excels at and problems that are more challenging for AI. Use this information to determine when it is appropriate to use AI and when to leverage human skills.	This competence is also conveyed through the use case. Students are shown the range of functions of various AI tools and the strengths and weaknesses of the use of AI are worked out by comparing them with traditional literature research.
Imagine Future AI: Imagine possible future applications of AI and consider the effects of such applications on the world.	X
Representations: Understand what a knowledge representation is and describe some examples of knowledge representations.	X
Decision-Making: Recognize and describe examples of how computers reason and make decisions.	How AI tools make decisions is discussed extensively in the use case. For example, how Elicit determines the (top 4) results or how ConnectedPapers builds the graphs.
ML Steps: Understand the steps involved in machine learning and the practices and challenges that each step entails.	X
Human Role in AI: Recognize that humans play an important role in programming, choosing models, and fine-tuning AI systems.	This competence is defined differently in the application example than in Long and Magerko. In the use case, the human role refers to the user as a person. Here we discuss how AI should be used in terms of augmented intelligence.

Data Literacy: Understand basic data literacy concepts.	The data sources of the training data are discussed in the use case.
Learning from Data: Recognize that computers often learn from data (including one's own data).	X
Critically Interpreting Data: Understand that data cannot be taken at face-value and requires interpretation. Describe how the training examples provided in an initial dataset can affect the results of an algorithm.	Critical thinking is also a key competence in the application case. Students are encouraged to critically scrutinize the output of AI tools. This is encouraged by comparing different AI tools and contrasting them with the results of traditional literature research.
Action & Reaction: Understand that some AI systems have the ability to physically act on the world. This action can be directed by higher-level reasoning (e.g. walking along a planned path) or it can be reactive (e.g. jumping backwards to avoid a sensed obstacle).	X
Sensors: Understand what sensors are, recognize that computers perceive the world using sensors, and identify sensors on a variety of devices. Recognize that different sensors support different types of representation and reasoning about the world.	X
Ethics: Identify and describe different perspectives on the key ethical issues surrounding AI (i.e. privacy, employment, misinformation, the singularity, ethical decision making, diversity, bias, transparency, accountability).	Various ethical aspects are discussed in the context of the use case. These include the transparency of the use of AI. Students are made aware of the need to document the use of AI in the documentation of their research project.
Programmability: Understand that agents are programmable.	X

Note: The X in the figure above means that a competence does not play a role in the application.

Fig. 4. Comparison of the competencies of the AI Literacy Framework according to Long and Magerko with the competencies required for an AI-supported literature search

G. Integration of law lectures into the concept

Interdisciplinary learning is a central element of this concept for promoting AI literacy among students. The following section will therefore show how law lectures were integrated into the concept.

The starting point here is that the implementation of an empirical research project in the context of a problem from business practice is usually only a sub-project of a larger project. For example, empirical research is used to determine consumer preferences in order to design a target group-oriented marketing campaign. In this case, the empirical study would be part of an overarching marketing project.

Following on from the approach of project-based and interdisciplinary learning described above, projects should not only be dealt with in courses on empirical research, but should also be integrated into law lectures, subject-specific lectures and academic work. This offers new ways of imparting knowledge in the field of law, especially for non-law degree courses.

Up to now, the structure of legal lectures in business administration subjects has been similar to traditional law studies: BGB AT, law of obligations, property law, commercial and corporate law, employment law, special areas of law depending on the business administration specialization. And the examination form is often still the classic legal exam. This leads to the bulimic learning described above. The effect: exam passed, knowledge lost a short time later [25].

The advantages of project-related, interdisciplinary learning are illustrated here using an example. If we consider a marketing project, for example, the law lecture accompanying the project could be started directly with project-related legal aspects, for example from advertising or event law. This would have the advantage for students that they would have to combine knowledge from two different academic disciplines (business administration and law) in an interdisciplinary way and could apply it directly in practice.

This would be followed by copyright explanations in the context of scientific work. In this way, the project participants gain knowledge of copyright law, which is essential right now for a responsible approach to AI in general and for AI-supported literature research in particular. This is an essential condition for responsible support by AI described in section C. In this way, students gain an understanding of the importance of scientific source work right from the start of their studies, thus fostering reliable scientific work.

Lectures on data protection and employment law are suitable for the subsequent data collection and data evaluation. This is because there are data protection and employment law pitfalls, especially when entering data into an AI, which can be discussed and illustrated to the students using the project-related questions that arise. Finally, the publication of the project should be followed by an explanation of media law.

However, the revision of the curriculum must not stop at the content. As mentioned above, there is also a need for action when it comes to the legal forms of examination. For example, the portfolio examination offers advantages in the responsible use of AI. In this way, the deficits of students in dealing with AI tools identified under section C can be addressed. The preparation of presentations on individual legal problems initially offers the opportunity to use various

AI models. Students must also develop their own views and apply them to their project. The subsequent oral defense of the papers checks whether the students have actually understood, critically reflected on and questioned their explanations.

The procedure described here for a marketing project can be adapted to any other project. For example, IT law would be the introductory module for the legal aspects of a project in the context of business informatics. The other legal lectures could be held analogically to the marketing project described.

Result: Law applied in a project-based and interdisciplinary way, no longer just learned in relation to exams and thus better understood in practice.

IV. Conclusion

When developing a teaching concept to promote AI literacy among students, two different approaches can be chosen. One option is to use an AI literacy framework to develop the concept in a top-down process. Due to the weaknesses of the currently available AI competence frameworks, a bottom-up approach was chosen instead in this article. This begins with an identification of the challenges that students face when conducting an empirical research project. It then examines the extent to which AI provides support in solving these challenges. The comparison of the challenges with the support provided by AI leads to the definition of use cases. Experience gained from the use of AI in scientific studies and studies on the potential of AI were used to develop the use cases. Finally, the necessary skills were determined on the basis of the use cases.

If we compare the competencies determined using the bottom-up approach with those found in Long and Magerko's AI Literacy Framework, there are significant differences. The two sets of competencies differ in terms of their scope, the competencies they contain and their specificity.

From the authors' point of view, both approaches (top-down and bottom-up) are fundamentally justified. The advantage of the bottom-up approach is that it is very action-oriented and can identify the specific skills that are required for a particular purpose in a specific situation. In addition, the bottom-up approach provides orientation as to which competencies are highly relevant in practical application. This is because competence frameworks often consist of extensive lists of required competences that place high demands on students and are often difficult to fulfill. In line with deeper learning, it is important to prioritize the competencies in order to enable targeted and intensive support.

V. Outlook

The comparison of competencies required by AI competence frameworks with those arising from use cases should be taken up much more strongly by scientific research in the future. This will allow gaps to be identified in the AI competence frameworks and their practical suitability to be tested. In addition, the development of use cases can create a link between the general AI competence frameworks and practice. Ultimately, this will ensure that the concepts for promoting AI skills are both theoretically sound and practically relevant.

However, this requires that significantly more use cases for the use of AI are developed in the future. There are numerous advantages associated with the development of use cases. They are based on the concrete use of AI in specific contexts, which makes them highly relevant in practice. As a result, students build up targeted action skills, which also contributes to student motivation and commitment. In addition, the skills acquired can be used directly to solve the challenges that students face at university or in practice.

[1] U. Bucher and K. Holzweißig, “How to prepare students for the AI era?” Proceedings of the 2nd DHBW AI Transfer Congress, Sept. 29, 2023, Heilbronn, 2023, pp. 58-69.

[2] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu and M. S. Qiao, “Conceptualizing AI literacy: An exploratory review”, Computers and Education: Artificial Intelligence, 2, 2021.

[3] O. Almatrafi, A. Johri, and H. Lee, “A Systematic Review of AI Literacy Conceptualization, Constructs, and Implementation and Assessment Efforts (2019–2023)”, Computers and Education Open 6 (Juni 2024): 100173. https://doi.org/10.1016/j.caeo.2024.100173.

[4] D. Long and B. Magerko. “What Is AI Literacy? Competencies and Design Considerations”, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–16. Honolulu HI USA: ACM. https://doi.org/10.1145/3313831.3376727.

[5] S.C. Kong, Q. Wang, R. Huang, Y. Li and T. C. Hsu (Eds.), “Conference Proceedings (English Paper) of the 25th Global Chinese Conference on Computers in Education (GCCCE 2021)”, 2021.

[6] J. Hartig, and E. Klieme, “Kompetenz und Kompetenzdiagnostik”. Leistung und Leistungsdiagnostik, K. Schweizer (Ed.), 2006, pp. 127–43. Berlin/Heidelberg: Springer-Verlag. https://doi.org/10.1007/3-540-33020-8_9.

[7] R. Caldwell, “HR business partner competency models: re - contextualising effectiveness”, Human Resource Management Journal, 18(3), 2008, pp. 275-294.

[8] M. R. Theisen, “ChatGPT: Risiken, Gefahren und Chancen in Lehre und Forschung”, Wirtschaftswissenschaftliches Studium (52), 2023, pp. 17-23.

[9] E. Manalo, “Introduction: Establishing a case for sharing research-based instructional strategies”, Deeper learning, dialogic learning, and critical thinking, 2019, pp. 1-14. Routledge.

[10] J. McTighe and H. F. Silver, “Teaching for deeper learning: tools to engage students in meaning making“, Alexandria, VA: ASCD, 2020.

[11] Bucher, U., “ChatGPT – Are the students ready for the AI age?”, Proceedings of the 2nd DHBW AI Transfer Congress, Sept. 29, 2023, Heilbronn, 2023, pp. 2-9.

[12] M. M. Rahman, “21st century skill'problem solving': Defining the concept. Rahman, MM (2019). 21st Century Skill “Problem Solving”: Defining the Concept”, Asian Journal of Interdisciplinary Research, 2(1), 2019, pp. 64-74.

[13] B. C. Czerkawski, “Designing deeper learning experiences for online instruction”, Journal of Interactive Online Learning, 13(2), 2014.

[14] F. Bordot, “Artificial Intelligence, Robots and Unemployment: Evidence from OECD Countries”, Journal of Innovation Economics & Management, 37, 2022, pp. 117-138.

[15] Y. N. Harari, “21 Lektionen für das 21. Jahrhundert”, München: C.H. Beck, 2018.

[16] A. Gallego and Th. Kurer, “Automation, Digitalization, and Artificial Intelligence in the Workplace: Implications for Political Behavior”. Annual Review of Political Science 25, Nr. 1 (Mai 12., 2022), 2022, pp. 463-84.

[17] S. K. Sitompul and A. W. Anditasari, “Challenges in writing academic research: an axplotation of Master's students' experiences”, Getsempena English Education Journal, 9(2), 2022, pp. 136-148.

[18] A. Luse, B. Mennecke and A. Townsend, “Selecting a research topic: A framework for doctoral students“, International Journal of Doctoral Studies, 7(1), 2012, pp. 143-152.

[19] T. Cornford and S. Smithson, “Project research in information systems: a student's guide. Bloomsbury Publishing”, 2017.

[20] C. E. Snow and P. Uccelli, “The challenge of academic language. The Cambridge handbook of literacy”, 2009, p. 133.

[21] J. von Garrel, J. Mayer and M. Mühlfeld, “Künstliche Intelligenz im Studium Eine quantitative Befragung von Studierenden zur Nutzung von ChatGPT & Co.”. https://doi.org/10.48444/H_DOCS-PUB-395.

[22] W. Liang, Y. Zhang, H. Cao, B. Wang, D. Ding, X. Yang, K. Vodrahalli, S. He, D. Smith, Y. Yin, D. McFarland and J. Zou, “Can large language models provide useful feedback on research papers? A large-scale empirical analysis”, 2023. arXiv preprint arXiv:2310.01783.

[23] U. Bucher and K. Holzweißig, M. Schwarzer, “Künstliche Intelligenz und wissenschaftliches Arbeiten: ChatGPT & Co: Der Turbo für das wissenschaftliche Arbeiten”, München: Vahlen, Franz, 2024.

[24] S. Wollin-Giering, M. Hoffmann, J. Höfting and C. Ventzke, “Automatic transcription of qualitative interviews. Sociology of Science Discussion Papers”, 2023, URL: https://www.static.tu. berlin/fileadmin/www/10005401/Publikationen_sos/Wollin-Giering_et_al_2023_ Automatic_transcription. pdf.

[25] M. Schwarzer, “Neue Prüfungsformen an der DHBW: Die Portfolio-Prüfung im Wirtschaftsrecht”, in: J. Weber, B. Schiller and S. Fünfgeld, “Lehre, Forschung, Transfer & Management, Beiträge aus der Fakultät Wirtschaft der DHBW Stuttgart, Festschrift für Prof. Dr. Bernd Müllerschön”, 2020, pp. 251-255.

[26] J. Bachner, “The Common Mistakes Students Make When Crafting Research Questions”, APSA 2012 Teaching & Learning Conference Paper, 2012.

[27] T. Clear, “Framing a research study”, ACM Inroads, 5(2), pp. 22-23, 2014.

[28] M.A. Ezquerro, “Cuando se aproxima el lenguaje escrito al oral: competencias lingüísticas en el discurso académico del alumnado del grado en educación primaria”. Educ. Form. 6, Nr. 1 (Jan. 4., 2021): e3502. https://doi.org/10.25053/redufor.v6i1.3502.

[29] J. J. Jeyaraj, “It's A Jungle Out There: Challenges In Postgraduate Research Writing”, GEMA Online Journal of Language Studies, 18(1), 2018.

[30] C. F. J. Pols, “Students’ ability to analyse empirical data in practical work”. Journal of Physics: Conference Series, vol. 1287, No. 1, 2019, IOP Publishing.

[31] A. Rastri, Y. S. Rezeki, U. Salam, D. Riyanti and S. Surmiyati, “An analysis of students' problems in writing a research proposal”, Acitya: Journal of Teaching and Education, 5(1), 2023, pp. 57-71.

WindGISKI: Using AI to Propose Areas Suitable for Building New Wind Turbines

Daniel Gritzner*, Sandra Peters-Erjawetz‡, Carsten Fichter‡, and Jörn Ostermann*
*Institut für Informationsverarbeitung, Leibniz Universität Hannover, Germany
{gritzner,ostermann}@tnt.uni-hannover.de
‡Hochschule Bremerhaven, Germany
{speters,cfichter}@hs-bremerhaven.de

***Abstract*—To mitigate the effects of man-made climate change a switch to renewable energy sources is necessary. However, new wind turbines are not built as quickly as is necessary in Germany. There are many sources of delays or even conflicts, e.g., slow approval processes, lack of public acceptance, or wildlife conservation laws. In the project WindGISKI we therefore developed an AI model which is able to predict wind farm suitability scores for areas in Germany, as well as which features are most important for the model's prediction. By feeding this information into a geographic information system (GIS), this system can assist end users, such as region planners in local authorities or employees of wind energy companies, in their decision making and thus speed up the transition to renewable energy. In this paper we present a survey we conducted among industry experts, our AI model and a work-in-progress prototype of our vision of an AI-enhanced GIS. The expert survey was necessary to identify suitable samples for training our AI. It showed that urban structure and nature preservation are most relevant to wind energy projects, while social factors are barely relevant. Additionally, we designed a new metric for measuring our model's performance in the light of a very drastic class imbalance of samples which rendered existing metric unsuitable.**

***Index Terms*—artificial intelligence, deep learning, geographic information system, renewable energy, wind energy**

I. Introduction

The German government aims for the country to become neutral in terms of greenhouse gas emissions by 2045 in order to minimize the effects of man-made climate change. Producing energy from renewable sources, such as wind and solar energy, is an essential aspect in the strategy to achieve emission neutrality. However, the actual amount of new wind turbines being built every year lags behind the target metrics set by the government. Multiple reasons prevent a faster expansion of wind energy in Germany, e.g., slow approval of permissions for wind farm projects, high construction costs, a lack of public acceptance of wind turbines, or conflicts with nature protection laws or laws regulating disturbances caused by exposure to noise or shadows [1], [2]. As a consequence, too few potential areas for wind farms are proposed by local authorities and the actual wind energy projects are often delayed beyond their initial target date due to lawsuits. Even though the wind energy industry gained a lot of experience in the recent decades, the average time from starting the pre-planning of a new wind farm to finishing its construction increased from 5.5 years in 2015 to 8 years in 2022 [3].

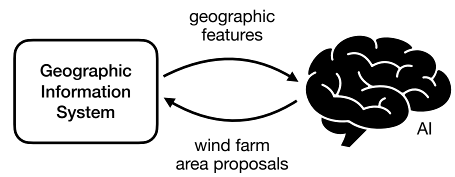


Fig. 1. The project WindGISKI is about using an AI to enhance a GIS to assist in finding new, suitable areas for wind farms.

The project WindGISKI[1] aims to use artificial intelligence (AI) to enrich a geographic information system (GIS) with additional information in order to assist people involved with wind energy projects, e.g., employees at local authorities or at wind energy companies, in their decision making process, as shown in Fig. 1. The goal is to identify low conflict areas in Germany on which new wind farms can be built quickly.

The interdisciplinary project team consists of research institutes from several fields, including engineering, computer science, social sciences, and life sciences, as well as companies working in the renewable energy business. Due to the broad range of interests affecting wind energy, a wide range of knowledge is necessary for the success of WindGISKI. The project covers many aspects: identifying features relevant to wind energy projects, collecting data suitable for AI model development, conducting interviews and surveys with industry experts for validation, implement noise propagation simulations, and more. In the end, the goal is not only to have an expert-validated AI model proposing areas suitable for wind farms, but to also have a separate reference booklet which instructs users of the AI on how to use it, as well as what other factors to consider when realizing a wind energy project that an AI cannot cover. As an example, involving the local population in a project such that it directly benefits from the nearby wind farm increases the acceptance and therefore how quickly the turbines can be built.

This paper specifically covers the AI model we developed for WindGISKI, our vision of how users may interact with the AI, and what information the AI can provide to assist

This research is funded by the Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV).

[1] https://www.windgiski.uni-hannover.de/

decision making. The remaining structure is as follows: in the next section we will discuss related work. Then we describe our method, i.e., the actual AI model and surrounding aspects relevant to its development. In the fourth section we will evaluate our model before finishing with a conclusion.

II. Related Work

The combination of an AI and a GIS has already been proposed in many cases. The closest work to ours is a very similar work using an AI-enhanced GIS to assist wind farm planners in Tuscany, Italy [4]. However, their spatial resolution is far coarser than ours at $10km$ compared to our $50m$. The next closest is a work predicting suitable wind turbine locations in the USA [5]. The authors are mainly concerned with studying the relationship of wind energy and wildlife conservation, though. Another work is concerned with predicting wind energy potential using AI and geospatial data [6]. Wind energy potential is an input feature for our model. Other works combining AI and GIS include efforts to improve decision making in other domains, e.g., city planning or disaster management [7]–[11]. A survey of state-of-the-art uses of geospatial data, including AI-based ones, can be found in [12]. A novel approach of combining AI with a GIS is using a large language model as a user interface to a GIS to help end user with creating visualizations or evaluations using existing data within a GIS [13].

III. Method

In this section we first discuss our dataset and the difficulties we faced with collecting suitable data for training an AI model. A survey conducted among industry experts, which we present in the second subsection, helped mitigate some of these difficulties. In the third subsection we explain the models we evaluated and in the last subsection we present our vision of how end users should be able to interact with our AI.

A. Dataset

In an initial phase of WindGISKI the project partners collected a list of features relevant to the success of a wind farm construction project. Each feature was classified as relevant to the AI model and/or relevant to the reference booklet, which was mentioned in the introduction, based on whether the feature is actually measurable and if so, measurable at a large scale, i.e., for the entirety of Germany. Features representing clear proposals without any room for decisions, such as involving the local population and having it directly benefit from the new wind turbines, were classified as only relevant to the reference booklet. Features from the categories meteorology, bodies of water, landscape preservation, nature preservation, wildlife conservation (birds and bats), forests, structure of urban development, traffic infrastructure, power grid infrastructure, topography, aviation, and military concerns were used by the AI model. Some features were discrete classifications, e.g., is an area a legally designated nature preservation area, while other features were continuous, e.g., distance to the closest residential building. A small subset of features was omitted entirely due to funding guidelines: WindGISKI must be non-political. Therefore features such as the voting behavior of the local population were omitted.

All of the data relevant to the AI was collected in a geo-referenced form, mostly as polygons. In order to be able train a model on this data we partitioned Germany in cells of size $50m \times 50m$ resulting in a tensor of shape $[H, W, C]$ with the height $H = 17359$, width $W = 12818$, and the feature dimension C. We stored each feature as a separate image in order to be easily able to choose which feature to load at each step of the AI training pipeline and which feature to omit to save processing time and memory. We converted all polygon coordinates to the EPSG:4839 coordinate system and rasterized all features into the data tensor. Data which was already rasterized was reprojected to the same coordinate system and resolution as our data tenstor.

We used a subset of $|C_e| = 33$ features to determine cells on which no wind turbine can be built for any reason, e.g., economic reasons (insufficient wind speed) or legal reasons (presence of residential buildings). For the actual AI model we used $|C_m| = 57$ features with some overlap between the features in C_e and C_m. Again, residential buildings are an example: they prevent wind turbines from being built in the same location (C_e) but the distance to the closest such building is also relevant to how suitable a cell is for wind turbines (C_m).

A naive approach for training a model would be a binary classification of cells based on whether they are part of an existing wind farm (positive class) or not (negative class). However, this approach is unsuitable since the premise of the project is that cells exists which are suitable for wind farms but no turbines have been built there yet. Therefore, we needed a way to identify samples which actually represent the negative class well. The cells omitted due to C_e are unsuitable as negative samples since the model would at best learn to look for overlapping features in C_e and C_m and therefore learn to identify what we already know. To mitigate this problem, we used the results of a survey we conducted among industry experts to create a rough scoring of cells to identify those cells which are likely negative samples but which are not excluded due to the presence of a feature in C_e.

We used wind farm cells as positive samples. However, we face a similar problem as we do with the negative samples. Due to advances in technology and changes in Germany's legal framework some existing wind turbines would no longer be built nowadays. We created a filtered subset of all existing wind turbines in which we excluded all wind turbines which were commissioned before January 1st, 2010, or whose total height (hub height plus rotor radius) is less than $150m$. The age filter accounts for the change in the legal framework. The height filter was used to exclude wind turbines which can no longer be built and run in an economically feasible way today.

Our dataset only contains coordinates of wind turbines but no wind farm identifiers as that concept is usually not represented in databases such as the Marktstammdatenregister. We therefore used the following approach based on heuristics provided by industry experts. We placed an ellipse around each wind turbine with its major axis aligned with the prevalent wind direction. For simplicity and based on expert knowledge, we assumed south-west winds as prevalent wind direction for all of Germany. The radius along the major axis was $5D$ and $3D$ along the minor axis with D being the rotor diameter. This ellipse represents the area in which no other wind turbine should be built due to negative interactions between nearby wind turbines. However, since wind farms are built compactly to maximize the number of turbines in a farm, these ellipses will overlap for close turbines within the same wind farm. Therefore, we use the overlapping of these ellipses as criterion to decide whether two turbines belong to the same wind farm.

Due to our assumptions wind farm assignments can be efficiently computed by rotating our coordinate system such that the first axis aligns with the prevalent wind direction and the second axis aligns with the minor axis. By scaling all coordinates by $\frac{1}{5}$ and $\frac{1}{3}$ along the respective axes all ellipses become circles of radius D in this transformed space. As a result, we only need to compare the Euclidean distance of any two turbines in this transformed space to the sum of their rotor diameters to decide whether their ellipses overlap or not.

After identifying wind farms, we computed the convex hull of each wind farm and intersected this hull with the union of all the wind turbine ellipses. The resulting area was used as the wind farm area. The intersection of the convex hull and the ellipses was necessary to get a better approximation of some wind farms, e.g., those shaped roughly like the letter L. We then rasterized the resulting wind farm areas into our data tensor. We performed this entire process once without either of the previously described age and height filters and once with both filters active. The difference in area between the two variants was marked as "ignored". This area contains a wind farm so we do not want to use it for negative samples but due to the aforementioned reasons we do not want to use it for positive samples either.

To compute distances to a feature, e.g., the distance to residential buildings or forests, we used an algorithm based on parabola intersections to compute a distance field for the respective feature [14], [15].

B. Expert Survey

To validate the results of WindGISKI's AI, interviews with and a survey among experts were conducted. The interviews ensured that no relevant features were missing on our list. We then conducted an online survey among industry experts. The link to the survey was sent to almost 900 persons, 66 of which responded. 45 of those 66 persons answered the survey fully, while the rest chose to drop out at some point. Due to the low number of full responses we consider our survey to be not be representative but still show trends. The respondents were employees or members of a diverse set of groups, e.g., wind energy companies, local authorities, nature conversation organizations, or law firms. According to the self-reported experience, about 50% of respondents have more than ten years of experience in wind energy and about 23% of respondents have between six and ten years of experience. Participants were asked about their goals in the wind energy industry, for whom they think our AI may be useful, how relevant each feature category is, and to rate each feature within each category. The last two questions, i.e., feature category relevance and feature rating, are particularly important for our AI model training.

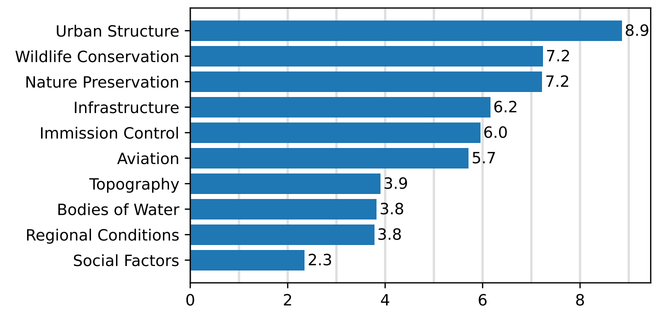


Fig. 2. Mean relevance of feature categories according to a survey we conducted among experts. We omitted the variance since we do not consider the survey representative.

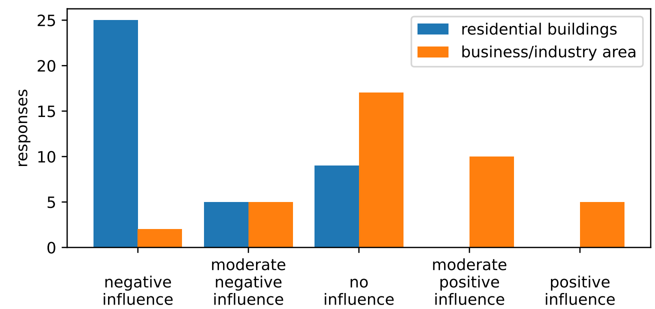


Fig. 3. Histogram of expert ratings of two features from the category "urban structure". The close proximity of residential buildings is considered to affect wind farm projects negatively while the close proximity to businesses or industry on average has no influence according to experts. There is a slight bias towards a positive influence, though.

For the feature relevance participants were asked to assign an importance or relevance with regard to their goals to each category on a ten step Likert scale from 1 (low relevance) to 10 (high relevance). The results are shown in Fig. 2. The most relevant categories are urban structure (where people live and work), wildlife conservation (e.g., birds and bats), and nature preservation. Social factors such as age distribution of the nearby population are of relatively low concern.

For the feature rating participants were asked to rate each feature on a five step Likert scale with regard to their goals:

1) Negative influence
2) Moderate negative influence
3) No influence
4) Moderate positive influence
5) Positive influence

Each feature was assigned to exactly one of the categories in Fig. 2 and all features were grouped by category in the survey. We also included all features we did not deem relevant to the AI in our dataset. A small excerpt of the results is shown in Fig. 3. The responses showed that some features have a negative influence, e.g., proximity to residential buildings (urban structure), while others affect wind energy projects positively, e.g., trust in local authorities (social factors). This even translated into trends for entire categories. Certain categories like urban structure or aviation largely contained features with a negative influence while social factors were deemed to mostly have a positive influence.

As mentioned in the previous subsection, choosing reliable negative samples for AI model training from our dataset was an issue. We therefore used the expert survey results to implement a rough scoring of each cell from our dataset to identify negative samples. In our scoring model, we assigned a factor ω_c to each category and to each feature ω_f. The score assigned to a cell is

$$\sum_{f \in F} \omega_{c(f)} \cdot \omega_f$$

with the subset F of features relevant to that cell and $c(f)$ being the category f is in. The distance up to which each feature is considered relevant, e.g., up to which distance a residential building is considered to be close and therefore relevant, was determined by an industry expert.

To determine the factors ω_c and ω_f we used the hyperparameter optimization tool SMAC [16]. First, we computed the relative amount of responses for each feature category c and each response $i \in \{0, 1, 2, \ldots, 9\}$ such that c_i is the relative amount of responses that rated the category c at $10 - i$. As an example, 63 out of 66 survey participants answered the question regarding the category $c =$ "urban structure". 41 out of those 63 responses rated the relevance of this category as 10, therefore $c_0 = \frac{41}{63} \approx 0.65$. We also determined relative amounts of responses f_i for each feature in the same way.

We modelled ω_c has having an exponential decay in relevance, i.e.,

$$\omega_c = \sum_{i=0}^{9} c_i \cdot e^{-\lambda \cdot i}$$

with the hyperparameter λ being one of two hyperparameters optimized by SMAC. We modelled ω_f as

$$\omega_f = -\alpha \cdot f_1 - \beta \cdot f_2 + \beta \cdot f_4 + \alpha \cdot f_5$$

with $\alpha = 0.5 + \gamma$ and $\beta = 0.5 - \gamma$ where $\gamma \in (0, 0.5)$ is the second hyperparameters opitmized by SMAC. In this equation f_1 is the relative amount of "negative influence" responses, f_2 is the relative amount of "moderate negative influence", and so on (cf. the enumeration earlier in this subsection).

This modelling was chosen such that an emphasis is put on the higher ratings wrt. to ω_c and such that α and β are values between 0 and 1 with $\alpha > \beta$. Furthermore, ratings of negative influence (f_1 and f_2) lead to negative scores, ratings of no influence (f_3) were ignored, and ratings of positive influence (f_4 and f_5) lead to positive scores.

SMAC needs an optimization goal to be able to optimize hyperparameters. As a maximization goal, we choose the relative number of cells being assigned a score less than the mean score of the non-ignored wind farm cells. The hyperparameter values $\lambda \approx 0.359$ and $\gamma \approx 0.023$ maximize this goal with about 82.8% of cells being assigned a score less than the non-ignored wind park mean score. With these hyperparameter settings we computed a rough scoring of all cells in our dataset which we then used later to choose reliable negative samples for our AI model training.

C. AI Model

The goal of WindGISKI is to train an AI which proposes suitable areas for constructing wind farms. However, we do not simply want to replicate decisions by experts but rather to use the rough scoring from the previous subsection as a guide. The goal is for the AI to be able to discover new suitable areas which experts have not discovered yet so far. Therefore we implemented three different AI model variants with different levels of reliance on the survey-based scoring.

The first two variants are based on how the model is trained while being flexible wrt. the model architecture. The third variants requires a specific kind of model architecture in addition to a specific way of training:

1) Binary classification
2) Metric learning
3) Normalizing flows

Our first variant treats the problem as a binary classification. The model is trained to assign high logits (= scores) to positive samples taken from non-ignored wind farms and low, even negative, logits to negative samples chosen based on the rough scoring from the previous subsection. We train the model to minimize the cross-entropy

$$H() = - \sum_{cls \in \{neg, pos\}} \mathbb{1}_{y=cls} \cdot \log\left(p(cls|x)\right)$$

where x is a vector of length $|C_m|$ from our data tensor, i.e., a vector describing all the features of a single cell, and $y \in \{neg, pos\}$ defines whether the sample x is a positive or negative sample. $\log\left(p(cls|x)\right)$ is the model's prediction. The logits $\log\left(p(pos|x)\right)$ predicted by a model trained this way can be used as a score for each cell.

To choose which samples to use as negative samples we first determined the number of positive sample cells n in each federal state. Since federal state laws differ wrt. to wind energy we chose as many negative samples from each federal state as there positive samples in that state. We ordered all candidate cells, i.e., cells not excluded (cf. C_e in subsection III-A) which are not part of an existing wind farm, by their expert-based score in ascending order. We then randomly chose n negative samples from the bottom $3n$ samples of this ordered list, i.e., we randomly chose from a subset of candidates likely to be good negative samples.

Modern models in other domains such as natural language processing or computer vision often are based on a transformer architecture such as OpenAI's famous GPT-line of models [17], [18]. However, these models require a far larger amount of training samples than we can provide. We therefore used

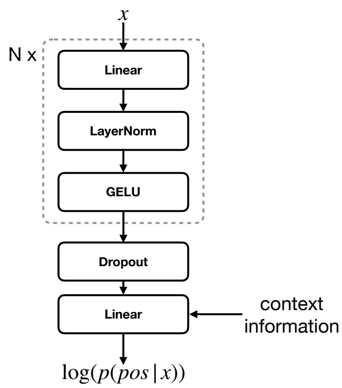


Fig. 4. The architecture of the MLP we evaluated. Details on N (how many times we repeated the Linear-LayerNormalization-GELU-Activation-block) and the hidden dimensions after each Linear layer can be found in section IV. The context information is an optional vector which is concatenated with the output vector of the Dropout layer to form the input of the last Linear layer.

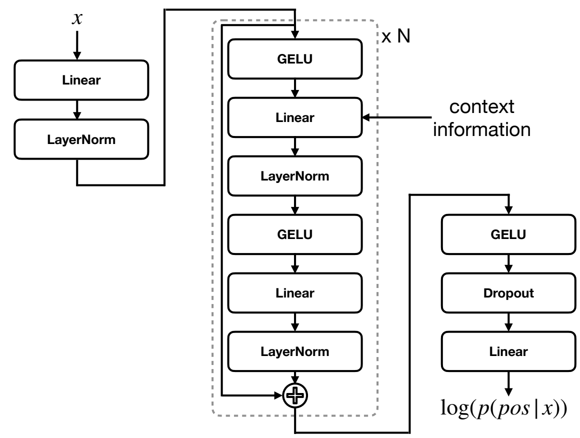


Fig. 5. Another MLP architecture we evaluated. Compared to Fig. 4, this architecture uses residual connections and a larger block that is repeated several times. Again, details on the configuration can be found in section IV. Also, the context information is optional and, if present, integrated in the same way as in the standard MLP, i.e., by concatenation.

an older but still performant multi-layer perception (MLP) [19] architecture as shown in Fig. 4. MLPs are even still used as components within transformers, e.g., the feed forward blocks in [17]. Inspired by the design of modern convolutional neural networks (CNNs) [20], [21], we used layer normalization instead of batch normalization and GELU instead of ReLU as activation function. Therefore, our modernized MLP consists of several blocks, each of which consists of a linear transformation layer followed by a layer normalization layer and a GELU layer. After the final such block, we apply dropout [22] before applying a final linear transformation of the model features into $\log\left(p(pos|x)\right)$ (it is not necessary to compute $\log\left(p(neg|x)\right)$ explicitly to compute the binary cross entropy). The input of our MLP is a vector x describing a single cell of our dataset. Optionally, we provide context information describing the surroundings if the cell represented by x. If we do so, we concatenate the context feature vector with the output of the dropout layer before applying the final linear transformation. Since this changes the number of weights of the last layer, in each experiment we decide whether to always use context information or to never use context information.

In order to be able to train deeper models with more overall layers, modern models use residual connections, i.e., the input of certain layers is added to the output of later layers [17], [23]. To take advantage of this, we also evaluated an MLP with residual connections as shown in Fig. 5. Keeping in line with [23] we place the residual connections such that their end, where two signals are added together, are placed just after a normalization and before an activation function application. For the residual connection to be well-defined, the number of features going into the repeated residual block and the number of features going out of the block have to be equal. In order to allow a different number of features, e.g., for applying non-linearities in a higher dimensional intermediate space, each repeated block contains two linear transformations. The first linear layer may map the features to a higher dimensional space, while the second layer then projects the features back into the same space that was used as an input to the block. The rest of the residual MLP design is the same as in our regular MLP (Fig. 4) with the exception of the context information, which, if available, is concatenated to the input of the first linear layer in every residual block.

For the optional context information we start with a small image centered around the cell x for which we want to make a prediction. Each "pixel" of this image is a cell, i.e., this image has $|C_m|$ channels. We use an image classification model as feature extractor, as is common for many computer vision tasks such as segmentation or object detection. We change the first convolutional layer of the classification model to accept $|C_m|$ input features and remove the final classification layers. We flatten the extracted feature map into a vector which is then used as the context information for the MLPs. Again, due to the lack of training data, we use an older, smaller, parameter-efficient model, namely Xception [24], as our feature extractor.

The second training approach we used for our models is based on metric learning. In metric learning a model is trained to directly assign scores to input samples, e.g., siamese networks learn to compute a similarity score for pairs of inputs. In our case, we compute scores for individual cells. We use the exact same network architectures as before, but we use a different loss function. The loss function

$$\mathcal{L}(x_1, x_2) = \mathrm{ReLU}(m(x_2) - m(x_1) + \Delta)$$

compares two samples x_1 and x_2. It penalizes the model m if it assigns a higher score $m(x_i)$ to x_2 than it does to x_1. The score $m(x_1)$ is supposed to be at least Δ larger than $m(x_2)$. Therefore, x_1 is supposed to be a better sample than x_2. We use two kinds of pairs of samples for our loss. First, we use inter-class pairs, i.e., x_1 is a positive sample and x_2 is a negative sample. In this case we use a large value for Δ. Secondly, we use intra-class pairs, i.e., both x_i are from the same class (positive or negative) but x_1 has a higher score than x_2 according to the expert-based rough scoring. We do use a

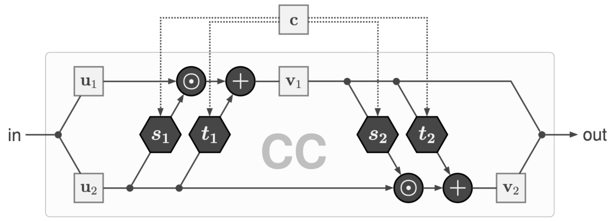


Fig. 6. Coupling blocks which can be used to implement a normalizing flow. Diagram taken from [25].

smaller value for Δ in this case but we do not use the actual difference of the rough scoring. We use the rough scoring as a ranking rather than a true absolute scoring.

The third and last training approach is based on normalizing flows [25]. Normalizing flows are invertible models, i.e., they map vectors of a specific length to vectors of the same length. However, the output vector is actually in the space of a known probability distribution, usually a multivariate standard normal distribution. This can be achieved by a sequence of coupling blocks as shown in Fig. 6. Each coupling block takes an input vector and splits it into two components u_1 and u_2. First, u_2, and optionally context information c, is used to compute the scaling and translation parameters of an affine transformation via the submodules s_1 and t_1. This affine transformation then transforms u_1 into v_1 which is used in a similar fashion to compute an affine transformation of u_2 into v_2. Then, v_1 and v_2 are concatenated into the output of the coupling block. This sequence of affine transformations is reversible.

Assuming that the output of the final coupling block is in the space of a multivariate standard normal distribution, the likelihood of the output vector can be computed. Minimizing the loss function

$$\mathcal{L}(x) = \mathbb{E}\left[\frac{\|m(x)\|_2^2}{2} - \log|J|\right]$$

with the input vector x representing a single cell of our dataset, the normalizing flow model m and the determinant $|J|$ of the Jacobian matrix $\frac{\delta m}{\delta x}$, results in maximizing the likelihood of all samples x shown to the model m. Normalizing flows learn the distribution of all the samples shown to it. The model will learn to assign high likelihoods to wind farm cells and low likelihoods to every cell that is dissimilar. We therefore do not need the rough scoring in this approach at all.

To increase the transparency of the black box AI models we use, we use integrated gradients to compute the importance of each feature. First, we compute the average features of all cells in our dataset as a baseline $\bar{x}$. We then, for each cell, linearly interpolate in multiple steps from the baseline to the actual cell features, compute the loss and backpropagate the gradients to the input vector x. The integrated gradient then is the sum of these input gradients over all interpolation steps. The relative absolute values of the individual features in the integrated gradient measure the relative importance of the features to each other. We further improved this measurement by using SmoothGrad. SmoothGrad applies the integrated gradient computation multiple times to each cell x but adds a small amount of random noise to x each time. The final result is the average of all integrated gradients for a given cell x.

D. Geographic Information System

In Fig. 7 we show a work-in-progress prototype of our vision on how to integrate our AI model's predictions into a GIS. The screenshot at the top shows an excerpt of Germany (northern tip of Germany including the island Sylt) and a user control panel (left-hand side). The excerpt shows Germany as a heatmap (from purple for low scores over blue and green to yellow for high scores) with excluded areas (cf. C_e in subsection III-A) shown in red. Gray is used for cells outside of Germany, including the sea. A zoomed out preview of all of Germany can be seen in the bottom-left of the screenshot.

From top to bottom, the control panel allows the user to select which AI model to use. In the screenshot a model trained solely on data from the federal state Schleswig-Holstein is selected in the drop-down menu. Next, the user is able to adjust the scoring ("Bewertung") used for the heatmap. Options are "absolute" (the score range is mapped to the interval $[0, 1]$ with 0 being rendered in purple and 1 being rendered in yellow), "relative" (each cell x is mapped to a value in $v \in [0, 1]$; v is the relative amount if cells in the training data which has an equal or lower score than the cell x), and "relative to wind farms" (same as "relative" but only the wind farm cells from the training data are used as a reference). The heatmap can even be turned off, rendering all heatmap pixels as black instead. The controls also allow the user to scale certain features (distance to residential buildings is shown in the screenshot) before the model assigns a score to each cell. This allows the user to simulate changes in laws, e.g., if a change in law required all wind turbines to be twice as far away from residential buildings, settings the corresponding feature's factor to 0.5 would simulate this change.

Next, the user is able to choose a target area ("Zielfläche") for further evaluation. The current target area is shown as a magenta square near the center-right of the screenshot. An actual end-user application would require the user to be able to specify areas freely as arbitrary polygons. The evaluation ("Auswerten") button opens a new window, a mock-up of which is shown in the bottom-left of Fig. 7. This window shows the distribution of scores in the target area, as well as which features were most important in the model's decision. A history of recently viewed target areas is also shown for an easy comparison of areas.

The optimization ("Optimieren") button leads to a series of dialogues in which the user can set parameters for areas to propose. Example parameters are the desired shapes of wind farms, the number of wind farms, constraints such as distances to existing wind farms, or the minimum area which should be proposed for new wind farms. An evolutionary algorithm is then used to find a configuration of areas satisfying all constraints and optimizing a desired goal, e.g., maximizing the average score of the proposed cells. As a last step, the user can inspect all areas proposed by the evolutionary algorithm

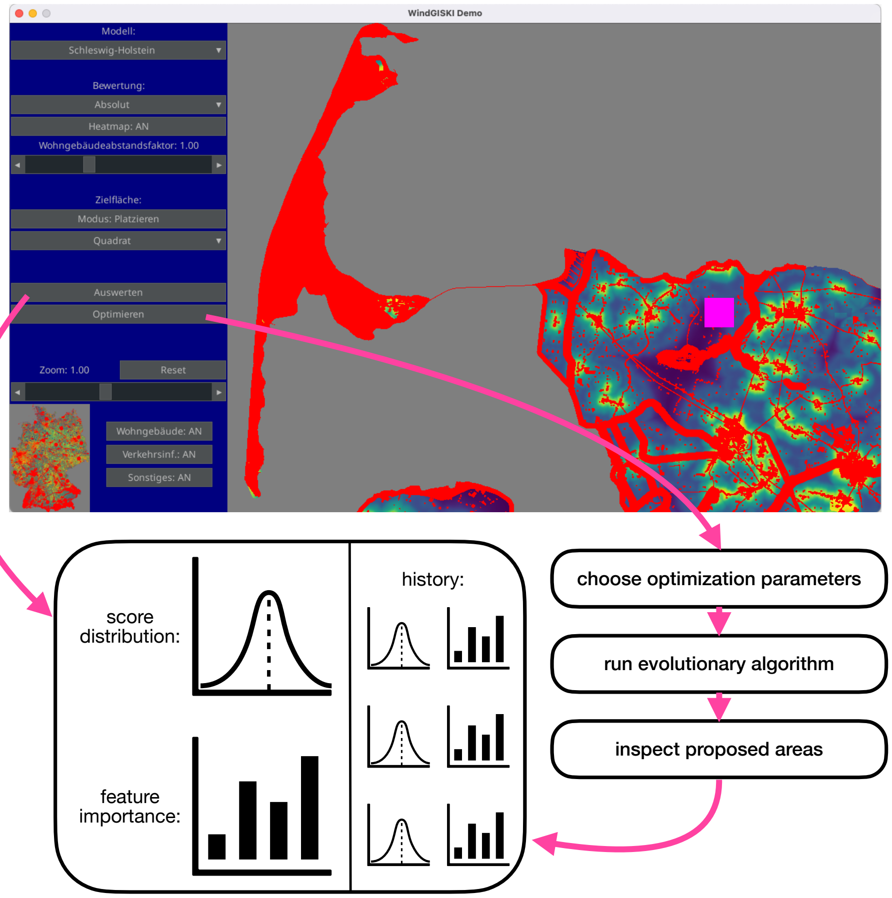


Fig. 7. A work-in-progress prototype demonstrating the integration of our AI model into a GIS. The heatmap uses mock data.

as they could do by choosing a target area directly and using the evaluation button.

Lastly, the last group of controls allows the user zoom the excerpt of Germany. The excerpt can be panned by holding a mouse button and moving the mouse or by clicking on the desired location in the small preview in the bottom-left. Next to the preview are controls allowing the user to disable the exclusion of cells due to certain features in C_e, i.e., fewer cells will be red. Again, this enables more freedom of choice for the user and allows adaptation to changes in laws.

IV. Evaluation

In this section we evaluate our model variants and further inspect the performance of the best variant in the different federal states of Germany. But first we introduce our evaluation metric, since existing metrics only provide a poor signal for optimization and/or comparison of models.

A. *Metrics*

In each experiment we applied the standard practice of partitioning our training data into an actual training subset and a validation subset (roughly 15% of samples). For the positive samples we assigned each wind farm to either the training subset or the validation subset so that all cells of a wind farm are either used for training or used for validation.

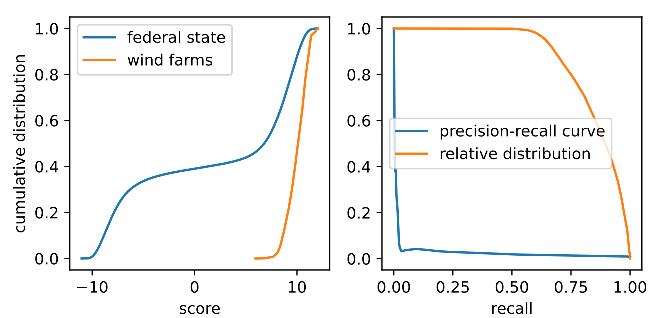


Fig. 8. Distribution of scores of model's prediction (left). The blue curve shows the distribution of all non-excluded cells of an entire federal state while the orange curve only shows the score distribution of the wind farm cells. The right shows two attempts at quantifying the quality of the model. The blue curve is a precision-recall curve used to compute the average precision while the orange curve is used to compute our evaluation metric.

We quickly noticed that regular metrics such as accuracy, precision, or recall did not provide useful information for comparing models and therefore for optimizing hyperparameters such as the number of layers. Almost all models quickly achieved 100% accuracy, even on the validation subset. We tried to shift to the average precision, which is commonly used in object detection. In Fig. 8 an example is shown. To compute the average precision, a precision-recall curve is created by systematically choosing a score threshold to classify cells into positive or negative. The average precision is the area under the precision-recall curve and is a value between 0 (bad) and 1 (good). However, since there are far more non-wind farm cells than wind farm cells in every federal state, the precision quickly drops to very low values. This is to be expected: even in the federal state Schleswig-Holstein, which has a relatively dense distribution of wind farms, there are almost 90 times as many non-excluded non-wind farm cells as wind farm cells. If we assume that just 1% if those cells are actually very suitable for new wind farms, the number of non-wind farm cells in the precision computation quickly outnumbers the total number of wind farm cells in the entire federal state. This can be observed in the example in Fig. 8. However, the actual distribution of scores in the example is actually desirable. The wind farms are assigned high scores while the federal state overall has some good, some mediocre, and some unsuited cells, just as we should reasonably expect.

Our metric uses a similar approach to the average precision to compare the relative location of two probability distributions while being independent of the absolute number of cells/samples in either distribution. We want most of the mass of our distribution of positive samples to be on the higher end of the overall distribution of all samples/cells, just as shown in the left subplot of Fig. 8. For the horizontal coordinate x of the orange curve in the right subplot, we use an approach similar to the recall. We choose score thresholds t such that $x\%$ of all cells in a federal state have a score of t or less. For the vertical coordinate, we compute the relative amount of wind farm cells which have a score equal to or higher than t. We can use the orange curve in the left subplot to do so. The meaning of the orange curve in the right subplot can be interpreted as follows: as you move along the horizontal axis, you go from the worst scores in the federal state to the best scores. The vertical axis then tells you the relative amount wind farms which are at least as good as this score. The curve for our metric is always monotonically decreasing. As the final evaluation metric we compute the area under this curve, just as is done for the average precision. We call our curve the relative distribution and therefore our evaluation metric the average relative distribution (ARD).

B. Model Variants

In a first experiment we compared the three different model training variants with the results shown in Figures 9 and 10. In these experiments we not only randomly picked a model training variant and model architecture, but we also randomly chose the learning rate as a first step in a hyperparameter optimization process. All MLP models were trained with $N = 4$ blocks with a decreasing number of output features of the linear layers $(256, 192, 128, 64)$. The residual MLP model used $N = 8$ blocks with the residual feature dimension set to 128, i.e., the first linear layer in each residual block expected an 128 dimensional input vector (before concatenation of the optional context information if present). The feature dimension between the two linear layers in each residual block was set to 256. The dropout probability was set to 0.1 for both model architectures. For the normalizing flow models, we used ten coupling blocks and all submodules s_i and t_i consisted of two linear layers with a leaky ReLU activation in between and no normalization. The number of input and output features of each submodule was defined by how the input vectors of each coupling block got split into u_1 and u_2. We used 256 features as an intermediate feature dimension between the two linear layers in each submodule. When using optional context information, we used Xception to generate feature maps. Xception eventually increases the feature dimension of the feature maps it computes to 728 and more. We decided to limit all convolutional layers to no more than 256 features to reduce the model size and account for the limited amount of available training data. The “images” created extended 32 cells in each direction (north, east, south, west) from the center cell x which was to be scored. Xception uses strided convolutionas in five places to downsample the input image, i.e., the resulting feature map had a height and width of 2 spatial units (downsampled by $2^5 = 32$ along each axis) and a feature dimension of 256 features. This was flattened into a 1024-dimensional contextual information vector. We trained all models for 25 epochs.

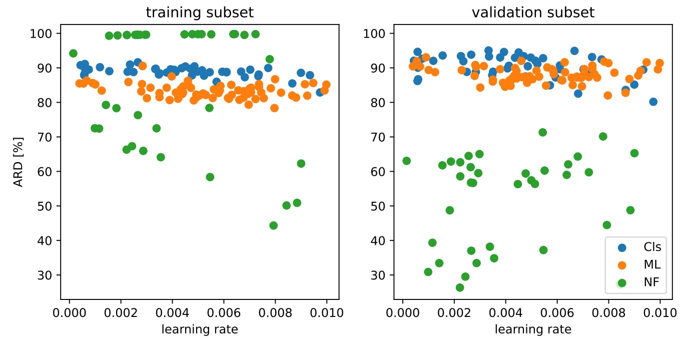


Fig. 9. The three different model training variants we evaluated: binary classification, metric learning and normalizing flows. We tested each variant with and without the optional context information. Binary classification and metric learning models were tested with our MLP architecture (Fig. 4) and our residual MLP architecture (Fig. 5). Each data point is the performance a model trained solely on the federal state Schleswig-Holstein using a random learning rate.

As can be seen in Fig. 9, normalizing flows are able to learn the training subset by heart but fail to generalize well to the validation subset. They also perform worse when adding context information. Both MLP-based model training variants performed well with a slight advantage to the binary classification-based approach. There is a slight increase in performance on the validation subset. For computing the validation performance we not only removed all cells belonging to training wind farms but also all cells within a $250m$ vicinity of those cells since we already noticed during development that models often tend to rate the immediate vicinity of existing wind farms very highly, i.e., they tend to suggest to simply

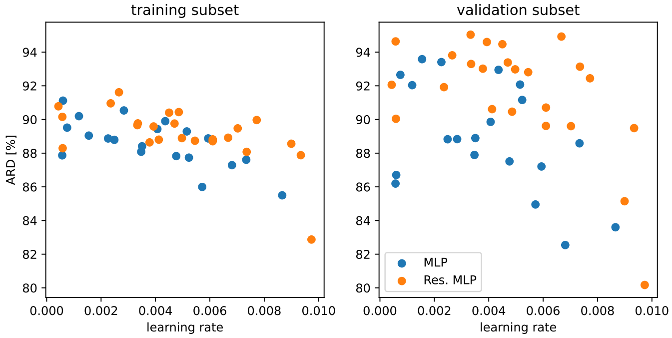


Fig. 10. Subset of the data shown in Fig. 9. Only binary classification data points are shown.

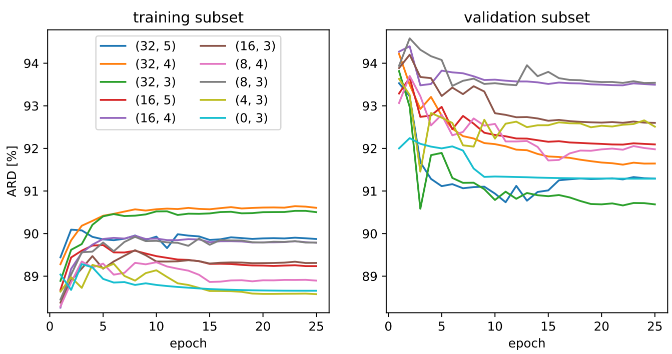


Fig. 11. Mean performance of the residual MLP model trained via binary classification. The configuration tuple specifies the extend of the context (first value) used in each cardinal direction and the number of downsampling steps used (second value). In the previous experiment (Figures 9 and 10) we only tested the configurations $(0, 3)$ and $(32, 5)$. The second value is irrelevant when using no context information (first value $= 0$).

increase existing wind farms instead of proposing new, well suited areas. Therefore, we remove the immediate vicinity of the training wind farms for validation purposes. As can be seen in Fig. 10, the residual MLP performed slightly better than the regular MLP. We therefore chose to focus on the residual MLP trained using binary classification as our best model variant from this point on.

We could not draw conclusions whether context information is actually helpful or not from the previous experiment. After deciding on the best model variant we ran another experiment in which we tested different context information configuration 15 times each. The mean performance across training epochs is shown in Fig. 11. When using less than the default five downsampling steps, we removed the later downsampling steps by setting the corresponding strides to 1 (from 2) while keeping the earlier downsampling steps. This is a common strategy also used in semantic segmentation models to increase spatial resolution (fewer downsampling steps) while keeping computational costs low (removing the late rather than early downsampling steps). While a large context of 32 helped with training performance, validation performance was actually best for the $(16, 4)$ and $(8, 3)$ configurations. Since there is no significant difference between those two configurations, we

TABLE I

MEAN PERFORMANCE AND STANDARD DEVIATION OF THE BEST MODEL (RESIDUAL MLP; BINARY CLASSIFICATION; OPTIMIZEZD HYPERPARAMETERS) TRAINED AND EVALUATED ON DIFFERENT FEDERAL STATES. WE OMITTED THE CITY STATES BERLIN, BREMEN AND HAMBURG BECAUSE ALMOST NO LARGE ENOUGH WIND TURBINES HAVE BEEN COMMISSIONED IN THEIR AREA IN RECENT YEARS.

federal state	training performance (ARD)	validation performance (ARD)
Baden-Württemberg	$93.4\% \pm 0.5\%$	$77.4\% \pm 1.4\%$
Bayern	$81.2\% \pm 2.8\%$	$79.0\% \pm 4.1\%$
Brandenburg	$88.7\% \pm 0.5\%$	$86.7\% \pm 1.1\%$
Hessen	$90.3\% \pm 0.6\%$	$75.1\% \pm 2.5\%$
Mecklenburg-Vorpommern	$87.5\% \pm 0.7\%$	$83.5\% \pm 1.3\%$
Niedersachsen	$84.2\% \pm 0.9\%$	$86.1\% \pm 0.9\%$
Nordrhein-Westfalen	$88.9\% \pm 0.5\%$	$84.3\% \pm 0.7\%$
Rheinland-Pfalz	$89.9\% \pm 0.7\%$	$90.2\% \pm 1.0\%$
Saarland	$93.1\% \pm 1.1\%$	$87.8\% \pm 1.5\%$
Sachsen	$95.9\% \pm 0.8\%$	$77.1\% \pm 6.1\%$
Sachsen-Anhalt	$91.5\% \pm 0.6\%$	$92.6\% \pm 1.0\%$
Schleswig-Holstein	$89.1\% \pm 0.5\%$	$95.4\% \pm 0.4\%$
Thüringen	$94.7\% \pm 0.6\%$	$86.5\% \pm 1.7\%$
all	$86.1\% \pm 0.5\%$	$83.4\% \pm 0.7\%$

chose $(8, 3)$ as our best configuration going forward. The downward trend of the validation performance across the epochs already indicates that we train the models for too long, an issue we fixed by further hyperparameter optimization.

C. Best Model

After deciding on the best model variant, we ran a random search to optimize the hyperparameters used for our model. We used the federal state Schleswig-Holstein for this hyperparameter optimization process. Our optimized hyperparemeters are as follows. We set the number of input and output features of both linear layers in the residual blocks to 480 and reduced the number of blocks N to 7. The dropout probability was decreased to 0.025 as well. Furthermore did we change the number of features used by Xception. The model starts with 32 features after the first convolution and increases this number roughly by a factor of 2 until reaching the final number of features of 2048. We changed this to 24 after the first convolution and an increase by a factor of 1.5 up to the final number of features of 411. We used a learning rate of 0.0019 with the optimizer AdamW [26], [27] and cosine annealing learning rate schedule [28]. We trained for 14 epochs with a mini-batch size of 1024, half of which were postive samples and the other half were negative samples.

Performance results of the optimized model can be found in Tab. I. The model converges to good solutions, even in federal states with very few wind farms such as Saarland or Sachsen. The performance on the training subset ranges from 81.2% in Bayern to 95.9% in Sachsen. The validation performance ranges from 75.1% in Hessen to 95.4% in Schleswig-Holstein and is therefore, as expected, slightly worse than the training performance. While there still is some room for improvement in some federal states, the performance is already good enough to make a prototype as described in subsection III-D very viable and useful.

V. Conclusion

In this paper we presented WindGISKI, a project which aims to use AI to enhance a geographic information system to assist users in identifying areas suitable for the construction of new wind farms. We collected more than 60 geographical features for use in a large dataset covering Germany. We then conducted a survey among experts and used the results to identify a small subset of samples to use as positive and as negative samples for training a deep neural network model with residual connections. This model is able to assign suitability scores to every $50m \times 50m$ square in Germany. A work-in-progress prototype will make the AI's prediction accessible to end users to assist them in choosing suitable areas and help them understand why an area is considered suitable or not.

In future work we want to validate our AI model based on expert knowledge. The survey we conducted unfortunately had too few participants. We therefore are considering using statistical measurements which do not rely on absolute values such as the rank correlation for validation. A high rank correlation, i.e., the AI model and the experts rank feature importance and/or cell scores similarly, would point to the AI model actually reflecting expert knowledge. Another, more involved, approach could be to have experts manually score certain areas and compare their results to the AI's predictions. Or the experts could choose suitable areas from a larger region and, once the evolutionary algorithm part of our prototype is done, we could compare their choice to our prototypes optimization routine.

References

[1] Fachagentur Wind, "Hemmnisse beim ausbau der windenergie an land - ergebnisse einer branchenbefragung." https://www.fachagentur-windener-gie.de/fileadmin/files/Veroeffentlichungen/Genehmigung/FA_Wind_Ergebnisse_Branchenumfrage_06-2022.pdf, 2022. Accessed May 17th, 2024.

[2] Bund-Länder-Kooperationsausschusses, "Bericht des bund-länder-kooperationsausschusses zum stand des ausbaus der erneuerbaren energien sowie zu flächen, planungen und genehmigungen für die windenergienutzung an land, an die bundesregierung." https://www.bmwk.de/Redaktion/DE/Downloads/E/EEG-Kooperationsausschuss/2021/, 2021. Accessed May 17th, 2024.

[3] Fachagentur Wind, "Typische verfahrenslaufzeiten für genehmigung und realisierung, 30. wind-energietage 2022." https://www.fachagentur-windenergie.de/fileadmin/files/Veranstaltungen/2022-11-10_Kompetenztag_Windenergietage/FA_Wind_Kompetenztag_Verfahrenslaufzeiten_Quentin_10-11-2022.pdf, 2022. Accessed May 17th, 2024.

[4] R. Mari, L. Bottai, C. Busillo, F. Calastrini, B. Gozzini, and G. Gualtieri, "A gis-based interactive web decision support system for planning wind farms in tuscany (italy)," *Renewable Energy*, vol. 36, no. 2, pp. 754–763, 2011.

[5] M. A. Boggie, M. J. Butler, S. E. Sesnie, B. A. Millsap, D. R. Stewart, G. M. Harris, and J. C. Broska, "Forecasting suitable areas for wind turbine occurrence to proactively improve wildlife conservation," *Journal for Nature Conservation*, vol. 74, p. 126442, 2023.

[6] S. Grassi, F. Veronesi, R. Schenkel, C. Peier, J. Neukom, S. Volkwein, M. Raubal, and L. Hurni, "Mapping of the global wind energy potential using open source gis data," in *2nd Int. Conf. Energy Environ. bringing together Eng. Econ*, no. June, p. 6, 2015.

[7] P. Emami and A. Marzban, "The synergy of artificial intelligence (ai) and geographic information systems (gis) for enhanced disaster management: Opportunities and challenges," *Disaster Medicine and Public Health Preparedness*, pp. 1–3, 2023.

[8] S. R. Samaei and M. Ghahferokhi, "Ai-enhanced gis solutions for sustainable coastal management: Navigating erosion prediction and infrastructure resilience," in *2th International Conference on Creative achievements of architecture, urban planning, civil engineering and environment in the sustainable development of the Middle East*, 2023.

[9] M. Hamano, S. Shiozawa, S. Yamamoto, N. Suzuki, Y. Kitaki, and O. Watanabe, "Development of a method for detecting the planting and ridge areas in paddy fields using ai, gis, and precise dem," *Precision Agriculture*, vol. 24, no. 5, pp. 1862–1888, 2023.

[10] H. Ouchra, A. Belangour, and A. Erraissi, "An overview of geospatial artificial intelligence technologies for city planning and development," in *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1–7, IEEE, 2023.

[11] S. Taşan, "Estimation of groundwater quality using an integration of water quality index, artificial intelligence methods and gis: Case study, central mediterranean region of turkey," *Applied Water Science*, vol. 13, no. 1, p. 15, 2023.

[12] R. Bill, J. Blankenbach, M. Breunig, J.-H. Haunert, C. Heipke, S. Herle, H.-G. Maas, H. Mayer, L. Meng, F. Rottensteiner, *et al.*, "Geospatial information research: state of the art, case studies and future perspectives," *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 90, no. 4, pp. 349–389, 2022.

[13] Z. Li and H. Ning, "Autonomous gis: the next-generation ai-powered gis," *International Journal of Digital Earth*, vol. 16, no. 2, pp. 4668–4686, 2023.

[14] R. Fabbri, L. D. F. Costa, J. C. Torelli, and O. M. Bruno, "2d euclidean distance transform algorithms: A comparative survey," *ACM Computing Surveys (CSUR)*, vol. 40, no. 1, pp. 1–44, 2008.

[15] T. Saito and J.-I. Toriwaki, "New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications," *Pattern recognition*, vol. 27, no. 11, pp. 1551–1565, 1994.

[16] M. Lindauer, K. Eggensperger, M. Feurer, A. Biedenkapp, D. Deng, C. Benjamins, T. Ruhkopf, R. Sass, and F. Hutter, "Smac3: A versatile bayesian optimization package for hyperparameter optimization," *Journal of Machine Learning Research*, vol. 23, no. 54, pp. 1–9, 2022.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[18] O. et al., "Gpt-4 technical report," 2024.

[19] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[20] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11976–11986, 2022.

[21] M.-H. Guo, C.-Z. Lu, Q. Hou, Z. Liu, M.-M. Cheng, and S.-M. Hu, "Segnext: Rethinking convolutional attention design for semantic segmentation," *Advances in Neural Information Processing Systems*, vol. 35, pp. 1140–1156, 2022.

[22] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[24] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.

[25] L. Ardizzone, C. Lüth, J. Kruse, C. Rother, and U. Köthe, "Guided image generation with conditional invertible neural networks," *arXiv preprint arXiv:1907.02392*, 2019.

[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[27] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[28] I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," *arXiv preprint arXiv:1608.03983*, 2016.

AI Systems Engineering and Dataspaces – Two Sides of the same Coin

Thomas Usländer
Information Management and Production Control (ILT)
Fraunhofer IOSB
Karlsruhe, Germany
https://orcid.org/0000-0001-6864-6631

Philipp Hertweck
Information Management and Production Control (ILT)
Fraunhofer IOSB
Karlsruhe, Germany
https://orcid.org/0000-0002-1051-8052

***Abstract*— AI projects in industrial production lead very often, in particular in small and medium sized enterprises, only to prototypes or to demonstrators, although in most cases they propose promising solution concepts. The emerging discipline of AI Systems Engineering addresses this problem via a systematic engineering approach that gives guidance where to start and how to proceed. This article describes how the concepts of dataspaces and digital twin systems may support the engineering of AI-based solutions. It focuses on the particular challenge of providing usable data, e.g., as training datasets for machine learning, whether within the shop-floor, a company or via domain-specific dataspaces encompassing multiple companies. Furthermore, it highlights how digital twin systems relate to the data provisioning and the operational phase in AI Systems Engineering.**

Keywords—AI Systems Engineering, dataspace, data usage control, digital twin systems

I. Introduction

There are many success stories about the use of AI methods in industrial production, especially machine learning (ML). However, these developments do often not get beyond a prototype status. There are difficulties in the transition to operational use with all its boundary conditions such as controllability, robustness and maintainability of the technical systems. In addition, the question arises as to whether statements and recommendations from AI-based solutions can really be relied upon from an engineering perspective, performance indicators, corresponding quality standards and compliance to regulations (e.g., European AI Act).

For AI-based inspection systems, a study by Maddox AI illustrates the current situation [1]. While 70 percent of the participants in a survey consider AI-based inspection systems to be ready for series production, only 17 percent are currently using them. In addition to the high costs (56%), which questions the economic use, the reasons lie firstly in the lack of AI expertise in the companies (51%) and secondly in the high effort required for the preparation and provisioning of the data.

Lack of AI expertise can be understood in different ways in the technical environment. On the one hand, a certain level of competence in dealing with machine learning methods and their mathematical and statistical foundations is, of course, necessary. On the other hand, such an AI expertise alone is not enough. The innovation and optimization potential of ML methods in industrial production can only be tapped if it is systematically embedded in the engineering of the overall information technology (IT) system, too. This requires a dedicated methodology, which we call AI Systems Engineering [2]. By this term we mean the "systematic development and operation of AI-based solutions as part of systems that perform complex tasks." As a methodology to be applied, AI Systems Engineering is recommended when applications have a high criticality, a high organizational complexity or a strong linkage to the physical world.

The need to consider system engineering aspects has recently been stressed by a Gartner study on AI-maturity of enterprises, even in the light of the impressive evolution of generative AI offers [22]. Only 9% of the organizations may be considered “AI mature”. One of the four foundational capabilities that make these organizations different is to “focus on AI engineering, designing a systematic way of building and deploying AI projects into production.”

IT trends have a major impact on AI Systems Engineering. This paper tackles the following research question: What does it need on the conceptual level in order to improve the acceptance and systematic engineering of AI systems? More concretely, the paper discusses the hypothesis that it basically needs two major constituting concepts that are linked as “two sides of the same coin”: firstly, a system engineering method tailored to AI systems (called AI Systems Engineering) and, secondly, a trusted environment (organized in dataspaces) to acquire and manage the data that is required for the AI methods [3]. These concepts are especially but not exclusively described in the application domain of industrial production, such that they may benefit from standards and technologies of the Industrie 4.0 and the Industrial Internet of Things (IIoT).

In section II the paper provides an overview about the state of the art in AI Systems Engineering, data provisioning in ML environments, dataspaces, related initiatives and digital twins. Section III illustrates the conceptual system models behind AI Systems Engineering, data spaces and digital twins. Section V describes how the described concepts of AI system engineering, dataspaces and digital twin systems fit together, before the joint applicability in an industrial example is described in section V. Section VI provides a conclusion and an outlook, when applying the concepts to projects of the AI Alliance Baden-Württemberg.

II. STATE OF THE ART

A. AI Systems Engineering

In the literature, AI Systems Engineering is closely linked to the research topics defined by Jan Bosch et al. in their research agenda for AI engineering, especially the topics dedicated to domain-specific AI engineering [4]. These topics lack, however, a dedicated view of the engineering discipline, e.g., the mechanical, electrical or chemical engineering disciplines.

In order to define AI Systems Engineering more precisely and to advance it as an emerging engineering discipline on its own, the Competence Center for AI Systems Engineering CC-KING has been formed [2]. CC-KING develops the scientific foundations and methods for AI Systems Engineering, develops software tools to support and apply them, and demonstrates the results using practical use cases, primarily from the domains of industrial production and mobility.

The challenges of AI system engineering are summarized in [2], classified into technical development challenges, e.g. difficulties to predict or quantify the performance and quality of an AI-based approach, and organizational development challenges, e.g. increased risk management as a consequence of these uncertainties. Therefore, one of the first and important demands for AI Systems Engineering is the specification of a systematic engineering approach. CC-KING proposes the process model PAISE® - Process Model for AI Systems Engineering, which combines approaches from computer science and data-driven modeling with those of classical engineering disciplines [5]. PAISE® suggests the phases in which engineering should proceed from problem and goal description to installation in order to systematically integrate AI processes into overall systems and which phases should be processed sequentially (as in a classical waterfall model) and which one in an agile manner (see below).

B. Data provisioning in ML environments

In order to be able to use ML methods for a use case, extensive training data sets are required (cf. dataset box in Figure 1). While in unsupervised learning they can be used, e.g., for automated classification of typical system states (then semantically annotated by the plant operator to start-up phases, normal operation phases, fault operation phases, etc.), in supervised learning they may be used for quality prediction of products. However, obtaining data has become the key bottleneck in many ML applications [6]. Lack of data typically means, there is either a need for more data, or there is abundant data but unlabeled or weakly labelled. Hence, a large part of the effort for engineering the overall system flows into the PAISE® phase of data provisioning with the sub-aspects of data acquisition for experiments and target system as well as data evaluation according to defined target metrics.

Due to the increased networking of plants and machines in Industrie 4.0, the questions of data origin, data ownership and data usage rights play a significant role here. Ultimately, it is a question of which data space the data comes from and which rules (policies) apply in the respective data space.

C. Dataspaces

Here, a dataspace is not understood as the physical location of the data storage, e.g., in the company ("on premise") or in a cloud. Rather, a dataspace refers to a data integration concept or, somewhat more concretely, a data-based concept for collaboration between the participants in the dataspace. By exchanging and sharing data, common goals may and shall be achieved [7]. Of course, a dataspace also requires a suitable distributed software infrastructure that supports and implements the requirements and assurances of trust, interoperability, and data sovereignty [8]. In general, the exchange of data with partners that do not know each other cannot be considered being trustworthy. If a user knows the recipient of the data, e.g., it is a customer, supplier or partner, trust in the data exchange can be established through bilateral organizational and technical agreements (e.g., encryption and exchange of passwords via separate communication means). However, this is cost-intensive and not scalable. The industrial dataspaces currently being created (see below in section D), therefore aim to enable trustworthiness "by design" and "by operation" through technical, trust-building measures in the dataspace infrastructure. This means that the participants gain trust in the sharing and exchange of data via the dataspace infrastructure, even without the participants having to know each other. This clearly demonstrates the applicability of the classic definition of trust of Luhmann [23] to technical infrastructures as "an attitude that permits risk-taking decisions for the purpose of reducing the complexity that would otherwise be entailed by the necessary control mechanisms".

The two major aspects of data sovereignty are [9][10]:

- Data usage control: How and to what extent can the infrastructure assure that, once data is accessed, it is only used by the data consumer according to the purpose intended and permitted by the data provider, i.e., according to his data usage policy?
- Data provenance tracking: How can the infrastructure support the consumer's ability to identify the origin of the data, and to use it according to the associated legal and contractual conditions?

D. Dataspace initiatives

Gaia-X is a European initiative to build a federated and secure data and service infrastructure [11]. It aims to provide an ecosystem with basic definitions, architectures, and technologies as well as common rules and policies to establish interoperable dataspaces. In addition to the common framework, provided by the Gaia-X association, each dataspace (considered as federation of participants) extends the policies, architecture, and data model to its own (domain specific) needs. This allows the realization of the DS hierarchy, presented in the previous section as well as the interoperability over the different tiers.

Each entity inside Gaia-X is described by a machine readable and -interpretable self-description (SD). This enables discoverability, as well as interoperability of services and datasets both inside dataspaces and inside the whole Gaia-X ecosystem. Cryptographic signatures, used in SDs ensure trust and integrity in this distributed environment. Federation services are responsible to realize the key functionality of data spaces: decentralized identity and access management, discoverability and search functionality for services and data sets as well as sovereign data exchange.

Conceptually relying upon the Gaia-X principles, the Catena-X automotive network [12] set-up an open collaborative data ecosystem dedicated to the automotive industry. Consequently, other dataspace projects for other industrial sectors such as mechanical engineering, aerospace, healthcare or process industry are being developed as part of the overarching Manufacturing-X funding program of the German Ministry for Economic Affairs and Climate Action (BMWK), in the global context brought forward within the International Manufacturing-X Council (IM-X) [24].

E. Digital Twins

Despite of multiple survey publications on Digital Twin definitions and uses (e.g., [13], [3]), there is, up to now, no common definition and conceptual approach of Digital Twins that is jointly accepted and applied across different domains and industries. We refer to the original conceptual definition coined by Michael Grieves [14], who specifies a digital twin to be a „digital entity of a physical system that exists on its own. This entity includes all the information of the physical system, which is connected to the physical system throughout the product lifecycle." Furthermore, he proposed that the elements of the conceptual Digital Twin approach comprise "a real space, a virtual space, the link for data flow from real space to virtual space, the link for information flow from virtual space to real space and virtual sub-spaces."

In order to enable an efficient implementation of applications in the virtual space, it is beneficial to have an interoperability approach between digital twins. Hence, a standardized approach is recommended. In industrial production there is the proposal of the Asset Administration Shell (AAS) concept of the German initiative Industrie 4.0 that provides a technology-neutral meta-model for this purpose, internationally being standardized in the IEC 63278 series [15].

III. SYSTEM MODELS

A. System Model of AI Systems Engineering

The understanding of the role of AI methods in overall IT systems is essential for AI Systems Engineering, e.g., in industrial production it is important to know how the IT components are related to the production facilities, be it control systems, condition monitoring systems or maintenance systems. From an AI system engineering perspective, AI methods are only built into parts of these technical systems, often in independent sub-systems that also need to be treated separately in terms of engineering (see Figure 1, [2]). According to PAISE®, these are also typically developed in an agile manner and iteratively integrated and versioned via checkpoints with other subsystems.

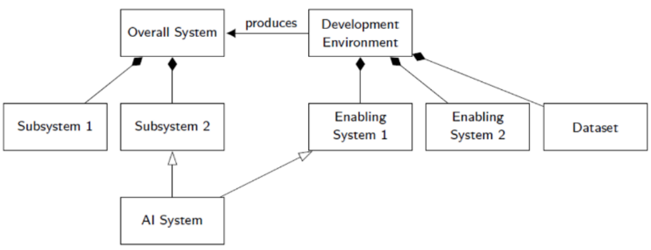


Figure 1: System model of AI Systems Engineering [2]

In addition, the AI subsystems also play an essential role as "enabling systems" in the development environment, together with the (training) datasets that are indispensable for ML methods, especially in case of supervised learning. In the classical sense of systems engineering, AI Systems Engineering includes the design of the overall system, since it is ultimately this system that must fulfill the decisive performance and quality criteria, such as reliability, IT security, and controllability. This is also compatible with statement of Simon Ramo who specifies systems engineering as "the design of the whole as distinguished from the design of the parts" [16]. It is the overall system that, in critical environments, must be approved and certified by a testing authority, not just the AI subsystem itself.

B. Dataspace model

In order to deal with data spaces in industrial production, one first needs a model of industrial production facilities. Here we consider a very simplified model with three levels: Hardware and software components (e.g., gripper, controller, PLC program) assembled to form a machine (e.g., robot), and production plants consisting of logically or physically related machines (e.g., a press shop consisting of welding machines and robots). For simplicity, not considered here are factories, recursions (plant is part of another plant), the role of humans as assets in the factory, or logistical aspects for the transport of material and goods.

In all of these three levels, data is generated to be shared and exchanged, so that there are analogously three levels of data spaces (see Figure 2):

- DS1: A level 1 dataspace is created by the collaboration of components (C) within a machine (M) and the IT components of the machine itself. Objectives: Monitoring, control and optimization of the machine, configuration of components, preventive maintenance services, etc.
- DS2: A level 2 dataspace is created by the collaboration of machines in a production plant (PP) and the IT components of the production plant (e.g., a production control system or a manufacturing execution system (MES)). Objectives: Monitoring, control and optimization of production, services for product quality prediction, energy optimization, etc.
- DS3: A level 3 dataspace is created by the collaboration of production plants across factories and company boundaries. Objectives: Supply chain management, traceability, sustainability calculations (e.g. product carbon footprint), resilience enhancement and flexibility services.

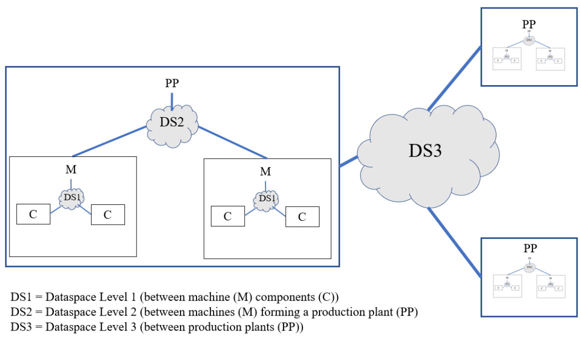


Figure 2: Dataspaces in industrial production plants

All services for achieving these objectives may be realized classically model-driven or also data-driven on the basis of AI methods. Typically, requirements on trust, interoperability, and data sovereignty are stricter in DS(x+1) than in DS(x) because the participants of the respective data spaces change more dynamically and come from different companies and trust spaces. Thus, the requirements on the infrastructure of the dataspaces also increase to compensate for this loss of trust and interoperability and data sovereignty.

In DS1, all software and hardware components and their interaction including data exchange will be intensively tested as part of the machine integration test, also called qualification test [17], especially if the components are delivered by different manufacturers. The qualification test may be carried out at the component or at the machine manufacturer site. Sometimes, it may even undergo a standard or proprietary certification process. The regulatory framework, such as the Machinery Directive [18], leaves little freedom for maneuver here. Changes during operation undermine extensive tests for safety reasons alone and are rather rare. This may, however, change in future driven by trends towards modular production when machine components are getting smarter and may authenticate themselves when (automatically) being installed into a machine during operation.

At the DS2 level, we expect significantly more changes, which will also affect the use of data for AI Systems Engineering. The following questions arise in DS2:

- Which data of which machines are available at all?
- Are they an official part of the machine products and therefore future-proof?
- Is the meaning of the data (semantics) known and clearly specified?
- Are there agreements with the manufacturers and suppliers of the machines and the components installed in them on the use of the data?
- How can the data be technically accessed, i.e. by which interfaces and service operations (e.g. IEC 62541 OPC UA) can they be read and managed?
- Are these interfaces version-safe and assured or can they change in the course of version changes of the machine operating systems?
- If sensor data is involved: What is the quality of the data in terms of accuracy and sampling frequency with the physical world?

While DS2 still typically takes place and is regulated within a manufacturing company, DS3 is cross-company and very dynamic by its very nature. An example of a DS3 is the Catena-X automotive network referred to in section II. DS3 typically includes the participants of the supplier network of a company, activated in logical supply chains and materialized in logistical transport chains. Due to uncertainties, problems and risks in these supply chains, the participants may change quite often. Hence, the qualitative and functional requirements upon the dataspace infrastructure in order to establish the necessary trust level are highest here. For AI Systems Engineering, this means that the provision of training data and operational data from a DS3 environment shall be largely automated, e.g.:

- Trust shall be automatically established through the DS3 infrastructure services.
- Data interoperability shall be established through semantic annotation into standardized knowledge models or through semantic mappings.
- Data sovereignty of data providers shall be done by evaluating appropriate data usage control policies and enforcing the associated policy rules.

There is an increasing need to configure machines within production plants according to the detailed characteristics of the incoming material and sub-components purchased from suppliers. In order to avoid cost-intensive incoming material tests and sophisticated and expensive measurements, there is a trend to get the data needed directly from the supplier as a data add-on to the material and/or component (see section V).

C. Digital Twin System Model

The framework how to integrate digital twins in an overarching digital twin system is described in the Digital Twin System Reference Model (DTS-RM) [3] (Figure 3).

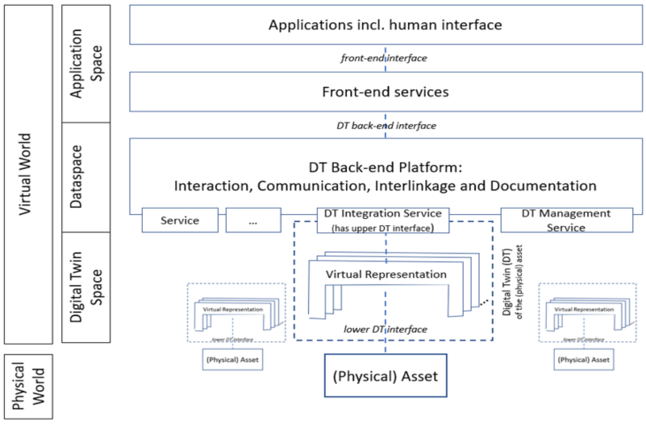


Figure 3: Digital twin system - reference model (DTS-RM)

The DTS-RM divides the virtual world into

- a Digital Twin Space, providing the Digital Twin instances including the virtual representations (e.g., structured according to the Industrie 4.0 AAS meta-model) of the physical assets, accessible and managed via the lower DT interface (e.g., by means of IEC 62541 OPC-UA),
- a Dataspace, providing the DT Back-end Platform comprising, among others, the DT Integration Service (e.g., by means of the AAS interfaces), management services for the Digital Twin instances and their interactions, and
- an Application Space, providing the front-end services as use-case specific views and the applications including the human interface.

The DTS-RM provides a linkage between the concept of digital twins and dataspaces. In the next section, it is illustrated how AI Systems Engineering fits into this DTS-RM.

IV. Relationship to AI Systems Engineering

We claim that AI Systems Engineering and dataspaces are two sides of the same coin in the sense that one side cannot exist without the other. The relationship is described according to the following three aspects:

1. Data provisioning
2. Usage of ML results
3. IT system design

A. Data provisioning

The data provisioning phase is explicitly described in the Process Model for AI Systems Engineering (PAISE®). Due to the increasing flexibilization of industrial production, the importance of dataspaces is increasing at all levels DS1-DS3, too. To cope with this, the PAISE® phase of data provisioning shall be extended and/or profiled for dataspaces, i.e., the data sovereignty demands need to be considered in the data provisioning phase. What is the provenance of the data, and, may it be used it for this purpose, e.g., as training dataset? Complex computations to increase sustainability and resilience in production ecosystems require AI-based applications, for which the data provisioning through the dataspaces form the appropriate basis. Data spaces and AI Systems Engineering thus form a multifaceted potential in their interplay and are mutually dependent in their further evolution.

B. Usage of ML results

Based upon data sets and context information ML components deliver results that need to be further processed in a processing pipeline in an overall system. For example, a ML-based preventive maintenance service requires sensor data from a machine (DS1) or plant (DS2) plus context information (e.g., air quality data from the facility management from DS2) and delivers a maintenance recommendation to be delivered to a maintenance service provider by means of a DS3. Both the sensor data as well as the recommendation may be considered to be properties of a digital twin instance about the machine and the plant. The overall data processing pipeline is illustrated in Figure 4.

C. IT system design

The DTS-RM presented above may be used for the IT system architect as a blueprint for the overall system architectural design. There are two three architectural design decisions, in order to obey demands for flexibility and interoperability:

1. To organize the dataset used for the training as well as for the operational phase of an ML component as well as the result information according to the concepts of a digital twin standard, e.g., the Industrie 4.0 AAS.

2. To organize the data processing pipeline as a service in an DT Backend platform (see figure 4).

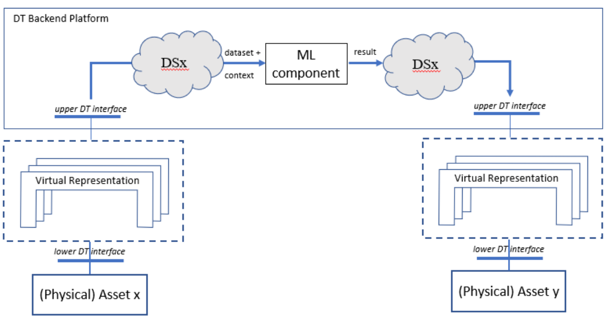


Figure 4: Processing pipeline based upon digital twins

3. To organize the data provisioning and the result delivery by means of dataspaces with well-defined policies, such that organizational changes do not harm the system architecture.

Furthermore, AI-based software, like any other IT systems are not static, but subjects to change. The increasing flexibilization of industrial production requires constant change and adaption. For AI based solutions this does not only concern the software engineering discipline, but also the development and retraining of AI components. The usage of MLOps (Machine Learning Operations) methods and tools assist through this continuous change [19]. Those have to be extended, to benefit from emerging dataspaces and the associated availability of new data sets.

V. INDUSTRIAL EXAMPLE

The following example is presented to highlight the interplay between AI Systems Engineering, dataspaces and digital twin systems in a industrial use cases at the Schwabmünchen Pre-Materials Plant of ams OSRAM [20].

In this plant, more than 3000 different products are produced, e.g. tungsten and molybdenum wires or filaments for lighting applications used, among others, in the aerospace and automotive industry. This plant supplies all OSRAM plants with these pre-materials. Due to changes in the lighting technology from fluorescent lamps to LEDs more flexibility is required. On the one hand, there is a fundamental decrease in the production volume, and, on the other hand, an increase in quality requirements, e.g. less tolerances. In order to cope with these changes, ams OSRAM follows a rigid digitalization strategy, e.g. to extend the process control, to increase the process capabilities and to enable full traceability.

Additional sensors have been developed and installed to collect all relevant process data as a pre-requisite to apply data analytics tools and use ML methods. Referring to the dataspace levels introduced in section III.B, this kind of sensor data refers to DS1 and DS2, depending on the plant structure.

However, the ams OSRAM plant managers recognized that the digitalization of their plant alone is not sufficient. Instead, connectivity, also with between all machines, plants and their suppliers, is key for further development. For example, there is a need for single piece product backtracking for tungsten products. Here, the whole value stream starting from the tungsten power production of a supplier, over the rod production up to the heavy and fine wire production has to be considered, including the production of the necessary moldings. In order to optimally configure the production line within ams OSRAM, it is necessary to get quality data from the suppliers, e.g., tungsten powder producer(s).

In the fixed supply chain relationship with a high level of trust, the delivery of this data may be fixed in bilateral contract(s). With this decision, the data quantity to be mastered at ams OSRAM has increased by 400% whereby the product resolution is 10 times higher than before, which is the basis for quality optimization. Furthermore, ams OSRAM also gives information to their customers, e.g. the OSRAM Plant Automotive for LED production, and vice versa, gets data back from these plants. As the data sources to be considered for the ML models gets more diverse, and the effort to manage bilateral (data sharing) contracts with different trust levels does not scale very well, this tendency also increases the organizational complexity of the AI-based solutions at the ams OSRAM plant. A possible counter-measure may be the establishment of a a dataspace of level 3 (DS3) for this industry branch.

Furthermore, ams OSRAM runs a pilot project to set-up and manage a digital twin for the production of the OSRAM H16 halogen lamps. This includes two other plants (located at Bruntal and Herbrechtingen). By combining the expertise of production control specialists, IT professionals and data engineers, all the data that is necessary to feed the digital twins is sent to and managed in MS Azure cloud environment, Based upon this DT backend platform (see section III.C and figure 3), business intelligence tools (of the DTS-RM application space) are used by process engineers and data scientists to implement the use cases, respectively, e.g. comparison of input and output quality time series by means of dynamic time warping.

VI. Conclusion and Outlook

The paper claims that the emerging discipline of AI systems management may strongly benefit from the concepts of dataspaces and digital twin systems in order to reduce the organizational complexity as one of its application dimensions in industrial production. The industrial example at the ams OSRAM plant illustrates the potential when bringing these concepts together in real-world deployments.

Looking at the research agenda for domain-specific AI engineering presented in [4] it is obvious that the at least the research topics "federated collection and storage of data" as well as "federated ML models" may also benefit from such a synergetic consideration.

In order to enable such a synergetic evolution of these concepts, it is proposed to organize and combine them within a reference model for AI Systems Engineering, as also proposed by the edition 2 of the German Standardization Roadmap on AI in its standardization need 05-01 [21].

Following the logic that AI applications and AI systems need both data and an engineering methodology, the AI Alliance Baden-Württemberg (https://ki-allianz.de/) has recently started two sub-projects dedicated to these two aspects: 1) data platform, and 2) AI challenge project.

The "data platform" project aims at the design and implementation of an operational, customizable platform for companies, preferably SMEs, public and scientific institutions with the following objectives according to a private-public partnership model:

1. generation, management and sharing of data and AI models,
2. in compliance with ethical and legal compatibility,
3. providing added-value through services, cross-sector linking and standardized metadata (e.g., provenance, quality, documentation), and
4. (experimental) access to computing resources for the execution of AI models.

The "data platform" project is domain-independent, i.e., in addition to industrial production the domains of healthcare, mobility and smart city/smart regions are considered, too. Although the conceptual and technological foundations are similar across the domains, domain-specific standards and initiatives have to be considered, e.g., the use of Industrie 4.0 standards for the domain of industrial production, or the use of FHIR (Fast Healthcare Interoperability Resources) specifications (https://hl7.org/fhir/) for the domain of healthcare.

The "AI challenge" project uses the methodology of AI Systems Engineering in a series of workshops comprising regional stakeholders from industry, research and the public domain. The idea is to analyze and jointly discuss thematic challenges according to region-specific topics, e.g., smart cities/smart regions, resource efficiency, circular manufacturing or healthcare resilience. Following the PAISE® process model, a project roadmap is derived and systematically mapped to the needs for data. If this data is owned by different organizational entities, the data access and usage policies of the different data spaces shall be applied.

These two projects of the AI Alliance Baden-Württemberg practically demonstrate the need to combine a methodology with data spaces when systematically engineering AI systems. Hence, two sides of the same coin.

Acknowledgment

This work was partly supported by the Competence Center Karlsruhe for AI Systems Engineering (CC-KING, https://www.ai-engineering.eu) and the AI Alliance Baden-Württemberg, sponsored by the Ministry of Economic Affairs, Labour and Tourism Baden- Württemberg.

References

[1] Maddox AI: KI-Diskrepanz - Warum nutzen nicht mehr Firmen KI-basierte Inspektionssysteme? Internet: https://invision-news.de/ki-deep-learning/ki-diskrepanz/ (last access 17.05.24)

[2] J. Pfrommer, T. Usländer, and J. Beyerer, "KI-Engineering – AI Systems Engineering: Systematic development of AI as part of systems that master complex tasks," at - Automatisierungstechnik, vol. 70, no. 9, 2022, pp. 756-766. https://doi.org/10.1515/auto-2022-0076 (last access 17.05.24)

[3] T. Usländer, M. Baumann, S. Boschert, R. Rosen, O. Sauer, L. Stojanovic, and J.C. Wehrstedt, "Symbiotic evolution of digital twin systems and dataspaces," Automation 2022, *3*, 378-399. https://doi.org/10.3390/automation3030020 (last access 17.05.24)

[4] J. Bosch, I. Crnkovic, and H.H. Olsson, "Engineering AI systems: a research agenda," ArXiv, 2020, abs/2001.07522.

[5] C. Hasterok and J. Stompe. "PAISE® – process model for AI Systems Engineering , " at - Automatisierungstechnik, vol. 70, no. 9, 2022, pp. 777-786. https://doi.org/10.1515/auto-2022-0020 (last access 17.05.24)

[6] C. Shani, J. Zarecki, and D. Shahaf. "The lean data scientist: Recent advances toward overcoming the data bottleneck," Communications of the ACM 66, 2 (February 2023), 92–102. https://doi.org/10.1145/3551635 (last access 17.05.24)

[7] B. Otto. "The evolution of data spaces," in: B. Otto, M. ten Hompel, and S. Wrobel (eds.), "Designing data spaces", Springer, Cham. https://doi.org/10.1007/978-3-030-93975-5_1 (last access 17.05.24)

[8] M. Huber, S. Wessel, G. Brost, N. Menz, N, „Building trust in data spaces," In: B. Otto, M. ten Hompel, and S. Wrobel (eds.), "Designing data spaces", Springer, Cham., https://doi.org/10.1007/978-3-030-93975-5_9 (last access 17.05.24)

[9] T. Usländer, A. Teuscher, „Industrial data spaces," in: B. Otto, M. ten Hompel, and S. Wrobel (eds.), "Designing data spaces", Springer, Cham. https://doi.org/10.1007/978-3-030-93975-5_19 (last access 17.05.24)

[10] C. Jung and J. Dörr, "Data usage control," in: B. Otto, M. ten Hompel, and S. Wrobel (eds.), "Designing data spaces", Springer, Cham. https://doi.org/10.1007/978-3-030-93975-5_8. (last access 17.05.24)

[11] Gaia-X European Association for Data and Cloud AISBL, "About Gaia-X," https://gaia-x.eu/what-is-gaia-x/about-gaia-x/ (last access 17.05.24)

[12] Catena-X automotive network, https://catena-x.net/en/library (last access 17.05.24)

[13] B.R. Barricelli, E. Casiraghi and D. Fogli, "A survey on digital twin: Definitions, characteristics, applications, and design implications,"

IEEE Access 2019, *7*, doi:10.1109/ACCESS.2019.2953499, pp. 167653–167671.

[14] M. Grieves, “Origins of the digital twin concept,” 2016, doi:10.13140/RG.2.2.26367.61609.

[15] IEC 63278-1:2022-07, “Asset administration shell for industrial applications - part 1: asset administration shell structure”.

[16] R.C. Booton and S. Ramo, “The development of systems engineering,“ in: IEEE Transactions on Aerospace and Electronic Systems, vol. AES-20, no. 4, pp. 306-310, July 1984, doi: 10.1109/TAES.1984.4502055.

[17] A. Birolini, “Qualification Tests for Components and Assemblies,” in: Reliability Engineering. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-03792-8_3 (last access 17.05.24)

[18] Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC . https://eur-lex.europa.eu/eli/dir/2006/42/oj

[19] D. Kreuzberger, N. Kühl, and S. Hirschl, "Machine learning operations (mlops): overview, definition, and architecture," *IEEE Access* (2023).

[20] I. Hild, “Smart Data – How to use AI to get the most out of your data,” 6th Annual Smart Manufacturing Excellence Summit, 22-24 March 2023, Munich, unpublished.

[21] DIN/DKE, “German Standardization Roadmap on Artificial Intelligence 2nd edition,” 2023. To be obtained at https://www.din.de/en/innovation-and-research/artificial-intelligence/ai-roadmap (last access 17.05.24)

[22] Gartner study (2024). https://www.gartner.com/en/newsroom/press-releases/2024-05-07-gartner-survey-finds-generative-ai-is-now-the-most-frequently-deployed-ai-solution-in-organizations (last access 17.05.24)

[23] Luhmann, N. (1973): Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität (2. Ausg.). Ferdinand Enke Verlag Stuttgart.

[24] https://www.plattform-i40.de/IP/Redaktion/EN/Standardartikel/manufacturing-x-international-mx.html (last access 25.06.24)

Educators' Reaction to an Introductory AI Training Session at a Large Vocational Institution in The Small Island State of Malta

Gerard Said Pullicino
Institute for Information and Communication Technology MCAST
Paola, Malta
gerard.said@mcast.edu.mt

***Abstract*— With the growing integration of Artificial Intelligence (AI) in education a pressing need to react to the realities of students using such tools in various contexts, training was seen as a pressing need. An initiative was carried out to train lecturers in each of the 8 institutes in Malta's largest vocational college on the basics of AI tool use and an indication on the effectiveness or lack thereof of the AI detection tool in current general use, when faced with AI generated text. The intention of the author has been to explain and highlight the changes in AI to train educators in its use, but this process brought about an interesting discussion which highlighted some facts about the thoughts and attitudes of educators in the vocational context and how their understanding of AI tools was different to the options available to students using it. The transcripts of the sessions were coded with contributions from the vocational lecturers indicating valuable insights into the current perception of educators on AI and how best to take such perceptions in the right direction and assist them to embrace AI tools to improve their pedagogical practice.**

Keywords—Generative AI, ChatGPT, Chatbots, Education, Teacher reactions

I. Introduction

The introduction of Large Language Models in the classroom was initially prompted by the launch of ChatGPT (Introducing ChatGPT, 2022) in November 2022. When the large language model was launched it heralded the possibility of a fundamental change to society and specifically to education. Large Language Models can very conveniently offer the answer to many technical problems based on the user crafting a specific prompt, with clear implications on student teaching, learning and assessment.

Although there were large language models before ChatGPT none of them had reached the vocational classroom. Students were not using such tools yet since they were not generally available. The availability of ChatGPT prompted a sudden change in perspective with students suddenly becoming aware of the fact that Large Language Models existed and were an extremely convenient tool to generate technically correct answers to tasks and assignments that hitherto took a significant amount of time and research to get done properly.

The proportion of students understanding the potential of such models to generate answers to assignment questions led to significant concerns expressed by the academic body at MCAST. A significant proportion of lecturers requested the use of an AI detection tool to detect AI use. The reason for the series of courses that were delivered was in fact to explain to lecturers that AI detectors had a non-zero false positive rate (Pan et al., 2024) and that it was difficult for the author to explain the reason for a high AI percentage. AI detectors are necessarily a 'closed box' and therefore the explanation of such percentages could not be adequately explained by the author's own experiences.

As the IT and Development Operations administrator at MCAST the author has been exposed to the impact of several very large disruptions in the educational context. The initial COVID crisis sensitized the author to the massive changes that needed to be done very quickly for education to continue in the new remote learning age. This made the author very sensitive to the impact of AI and enhanced the perception of the importance of training educational staff in the use of AI as a whole.

This research contains an analysis of the autoethnographic insights into the way the researcher's perceptions evolved while delivering training in AI to educators at his workplace. The actual process of presenting AI and discussing the issues raised by the activation or deactivation of AI detectors for assessments changed the initial perceptions of the researcher. These insights were then compared to a recently released research paper (Mollick & Mollick, 2024) that suggests AI activities linked to the Effective Teaching (Coe et al., 2020) toolkit, which provides pragmatic solutions to some of the issues raised by the educators.

A. Research Objectives

The research aims to determine the approaches and attitudes of educators to AI and how those attitudes contributed to a change in approach by an experienced AI research who is also an educator, and how it assisted in the process of clarifying the training required by educators to adapt to the AI enabled present.

1. To assess the initial impact of AI on educators and students in a large vocational institution through the experiences of training these educators.
2. To identify potential ways that solutions and training can be provided to tertiary lecturers dealing with the possibility of students plagiarising or misusing AI.

Through the above objectives a preliminary model explaining the attitudes and approaches of an experienced operator in the context of e-learning and how they change when faced with many inputs from different educators can be proposed, which could then be analysed and researched further based on more sources of data.

This paper adopts a Grounded Theory approach to research, which is inherently scalable, and invites further research in an area that is sure to attract various different players. Grounded Theory's iterative constant comparative process and inherent scalability, with its assumption that 'all is data', can ironically be compared to the training process of an AI model. Given the success that AI models have had in developing realistic stochastic models of the world, one can hypothesize that Grounded Theory is a form of subjective 'AI' applied to the world.

II. Literature Review

The researcher's innate attitude is coming from a varied background which includes a first degree in Psychology, followed by another undergraduate qualification in Software Engineering, and a Master's in E-Learning. This background serves to provide the context in which the researcher in this case comes from a constructivist philosophical viewpoint (Charmaz 2014), due the focus on empathy that comes out of his first degree. There is also a certain positivist bias due to the software engineering background of the researcher, that tends to indicate that a balance needs to be sought in the approach that underpin the analysis.

In order to put the researcher's background in the right context, a clear analysis of the literature in the context of AI in education is an important part of this paper, as it serves to contextualize the preliminary biases and insights that drive the attitudes and approaches of the researcher. The basis of this research is an insight into the concepts explained in the AI training sessions along with a constant comparison of these concepts to the concepts outlined in the secondary data.

The research follows Glaser's (Glaser, 2015) perspective that 'all is data'. In this case, and for the purposes of this Grounded Theory paper, both secondary and primary data are in fact data to be coded and compared to determine the latent patterns that indicate that information is relevant and can be used to study this complex and dynamically changing phenomenon.

It is in fact due to the relative paucity of data with respect to educators and how they engage with AI that such a relatively limited method of data collection is used in this paper, as this paper and autoethnographic analysis can serve as a starting point in order for the Grounded Theory methodology of constant comparison to be applied to more varied experiences in the context of vocational teaching and education, especially with the rich and varied data source that is MCAST, as a vocational institution.

In line with a constructivist GT approach (Charmaz 2014), this literature review aims to enrich data gathering and inductively provide insights into emerging themes on AI assisted vocational learning this entails, as mentioned previously, an iterative constant comparison of secondary data in the form of academic literature with the field of study, which leads towards theoretical sensitivity about the field being studied, followed by the qualitative generation of primary data through an experimental analysis, and a comparative retrospective on the output of the primary data on the secondary sources.

Researchers performing studies evaluating AI in Education (AIED) tools have been investigating their ability to improve the quality of learning through customization and personalization features as they target individual learning styles, as posited by Chen et al. (2020). The nature of these systems is congruent with the argument about the need to focus on the individual and the social constructivist approach as described by Khaled et al. (2014). A more recent paper by Ethan and Lilach Mollick (Mollick & Mollick, 2024) has shown some fascinating examples of building interactive educational material in the context of AI.

The research approach outlined in this paper includes coding the modern research paper by Prof. Mollick, with all the suggestions on the use of AI that come in through the conclusions outlined in that paper, and comparing those codes to the codes that were generated from an analysis of the transcripts of 6 interactive online sessions that were carried out in February 2024 to all the Institutes at MCAST, one of the largest vocational colleges in Malta.

The challenge of pedagogy has always been the identification of the appropriate teaching methodologies that are relevant to the context of the students that we are teaching. Taking the Great Teaching Toolkit approach in consideration (Coe et al., 2020), it is clear that to perform good pedagogy, even in the age of AI, there are certain principles that are important to keep in mind for teachers. The Great Teaching Toolkit outlines several important aspects required for good teaching, explaining that there are four important principles that are required for students to have an enriching educational experience. These four principles are as follows:

1. Understanding the content
2. Creating a supportive environment
3. Maximizing opportunity to learn
4. Activating hard thinking

In this case, given the relative difficulty in determining what the outcome of correct teaching using AI activities are (our goalposts, so to speak), the data to consider can be categorized under the following main sets:

a) Teacher perceptions of AI's effect on their students
b) Student perceptions of AI's effect on their studies
c) Theoretical sensitivity to cutting edge pedagogies that are aware of AI's effects

This paper considers the first set by coding transcriptions of the reactions of Teachers to a seminar regarding AI's effect on students and teaching, and the third set by coding and examining papers that implement pedagogies based on AI. The second set can be catered for using a GT approach by qualitatively interviewing students, but this has not been done so far in the context of this research. It is envisaged that the second set will be covered as part of the scope of STAR as a research project, but at a later stage, and the results can be considered as part of a subsequent publication that links the points made in this paper with points made in a paper that focuses exclusively on qualitative interviewing of the experiences of Students using AI in the context of their studies, and specifically in their use of the AI model being developed in this study.

III. Methods

The methodology outlined in this paper can be termed Computer Enhanced Grounded Theory, in that the methodology itself for the study leveraged a number of AI tools in order to be able to provide for correctly tagged and organized information in MaxQDA 2024 (Verbi, 2021).

MCAST (2021) is one of the largest vocational institutions in Malta, with a current active population of nearly 10,000 registered full time students. The student body is organized by 'Institute', with each institute focusing on definite subject areas that are related to each other. There are 8 institutes and centres, 6 of which are located in the main campus at Paola, with the Institute of Creative Arts located in Mosta and the Gozo Institute located on Malta's sister island. Each of these Institutes has a complement of educators, all of whom are grappling with the concept of AI and how to apply it to education. The requirement of the CPD sessions was related to a requirement by the MCAST QA department to create a viable AI policy in order to ensure that students using AI have a fair framework governing the use of such tools in the context of their assignment submissions.

In order to provide viable training regarding AI, the researcher was asked to prepare a number of hour-long sessions explaining how AI works and the implications of the use of 'AI Detectors' like Turnitin on assessment. Several issues with respect to false positives had been flagged in the institution, and given that the researcher is also involved in an operational context, such issues were directly in the researcher's purview. It was deemed prudent at the time to ensure that all AI detectors were turned off until more information could be sought at that point in time.

The sessions were carried out in the week between the 19th and 25th February 2024. All the sessions were recorded and transcribed using OpenAI Whisper (Radford et al., 2022), which yielded a good quality transcription given that all the sessions were delivered in English. The code used did not discriminate between speakers, however there were a number of sessions where a lot of discussion took place, and this could be identified in the subsequent step of the analysis.

All the transcribed texts were imported into MAXQDA, with the text being re-read and coded by the researcher using the MAXQDA AI assist function. The iterative process was based on highlighting text in a line-by-line fashion, utilizing the 'AI Assist' function to generate candidate codes, and always selecting a single viable code to highlight the specific passage that was deemed the most representative. It was noted that this intervening step made the AI assist function a very useful critical companion for the initial step of generating 'in-vivo' codes, refining the researcher's thinking and abduction in choosing the correct code, since the function often provided the correct classification of the text that would have happened on the second iteration when generating codes from in-vivo codes.

Given Glaser's (B. G. Glaser & Strauss, 2009) principle that all is data, the paper by Ethan and Lillach Mollick (Mollick & Mollick, 2024), highlighting some preliminary strategies for using AI in education, was also included as a document to be transcribed in the same MaxQDA model, with the paper being included last to determine whether the series of lectures (which were conducted prior to the paper's publication) reflected the themes and ideas that were included in the paper. When all the codes were generated, they were re-analysed by the researcher in a subsequent pass, with the different codes being collected into categories as required. The collected codes were then analysed to generate a set of summary grids highlighting the passages that supported the development of the codes, through the extensive use of the 'smart coding view' window in MaxQDA, which allows for the user to only see the specific coded segments related to a specific code, which in turn supports the argument being built in a stepwise fashion through the tool.

The creative coding tool was also used to generate a preliminary model, even though it is clear that there are insufficient cases in this case to generate any lasting conclusions given that more data and more comparison is required in order to construct a preliminary model. There are however some interesting insights brought about by the application of the Grounded Theory constant comparison model to this information, which can lead to the generation of a preliminary model of the effect of AI in the classroom, at least on teaching and assessment.

IV. Results

When comparing results, a number of analysis tools are available in MaxQDA to create a functional model of the codes that were generated in the tool. An iterative process was adopted, with the researcher iteratively 'pruning' the codes that were initially generated to create an initial conceptual model. Code hierarchies, though initially adopted, were discarded at this point as it was clear that there was no single parent and that coded segments were relevant to a number of different issues that were highlighted by the lecturers in the focus groups.

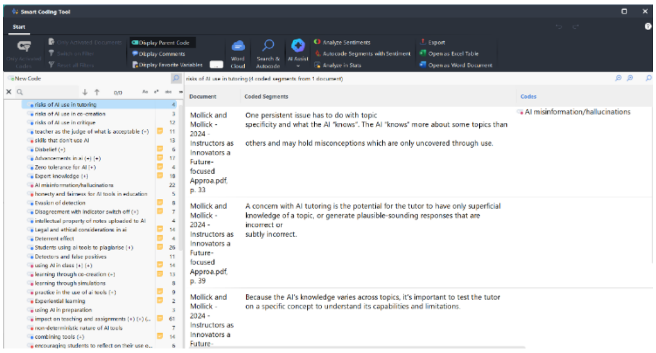
The document analysis was split into two 'sets', to distinguish the secondary data, in this case the Mollick paper, against the analysis of the sessions carried out with MCAST lecturers. Several 2 cases models were generated to compare the different sessions to each other and ultimately to compare all these comparisons to the included paper. The idea here was to see the code overlap between the different sessions and whether the insights from the different cases (with each session being taken as a case) was indicative of a specific pattern that could indicate a preliminary model / analysis. The code frequency table is reproduced below, highlighting the relative code frequencies once all overlapping codes were iteratively combined based on the conceptual meanings in the code segments. An analysis of the codes in MaxQDA's smart coding tool yielded the following list of codes and their relative frequencies:

Table 1: Code Frequencies

Code	Frequency
Impact on teaching and assignments	61
Students using ai tools to plagiarise	26
Tackling the situation	26
supervision of AI tool use	24
AI misinformation/hallucinations	22
Tricks for exposing GPT use	21
process instead of product	19
Expert knowledge	18
Advancements in ai	17
Legal and ethical considerations in ai	14
using AI in class	14
combining tools	14
skills that don't use AI	13
learning through co-creation	13
risks of AI use in critique	12
teacher as the judge of what is acceptable	11
Detectors and false positives	11
using tools in combination	11
learning through critique	10
practice in the use of ai tools	9
Evasion of detection	8
learning through simulations	8
Freely available tools vs paid	8
Disagreement with indicator switch off	7
non-deterministic nature of AI tools	7
students better understanding of tech	7
prompting for reflection	7
learning through mentoring, coaching and tutoring	7
Disbelief	6
encouraging students to reflect on their use of AI tools	6
integration agent	6
honesty and fairness for AI tools in education	5
Practical applications of ai	5
risks of AI use in tutoring	4
Zero tolerance for AI	4
intellectual property of notes uploaded to AI	4
Deterrent effect	4
Value judgment	4
User experience	4
AI tutoring	4
risks of AI use in co-creation	3
using AI in preparation	3
Utilization potential	3
critiquing an AI scenario	3
Experiential learning	2
Customization by individual instructors	2
ai as a student	2

The codes were re-checked though the use of the MaxQDA smart coding tool, which lists the codes on the left and generates a list of the summary codes on the right as per the following screenshot

Figure 1: Screenshot of the smart coding tool



The smart coding tool has the advantage of showing all the codes on the left while only showing the highlighted segments on the right. This allowed the researcher to re-check the coded segments in each of the documents and to determine whether the highlighted data incidents corresponded correctly to the thematic codes that are listed on the left. Some codes were linked to multiple thematic areas because there was a significant overlap between different concepts coming from different areas. As can be noted from the screenshot, the relevant codes are listed in a column in the smart coding tool.

Once the codes were re-checked, it was time to generate a visual representation of the data to facilitate visualization of patterns in the data in order for such visualizations to be presented. In this case, the MaxMAPS feature in MaxQDA was used to generate multiple 2-cases models, in order to compare the codes that were elicited in the multiple cases.

MaxMaps generates a knowledge graph linking the case to the specific code, while checking the overlaps between two cases. In this case, a number of maxmaps were generated, a selection of which are reproduced in the results below.

The following figures highlight a comparison of the sessions carried out based on the MaxQDA two-cases model feature, that identifies the overlaps between coded segments in multiple cases. In this case, the multiple cases are the different sessions carried out in the different institutes. One of the two cases models is included for brevity, however for the purposes of the study, all the combinations of two-cases models were generated, with slight differences indicating the priorities of the different content areas of the different institutes being highlighted when generating such two cases models.

Figure 2: Two cases model ICT (computing institute) & IBMC (business institute)

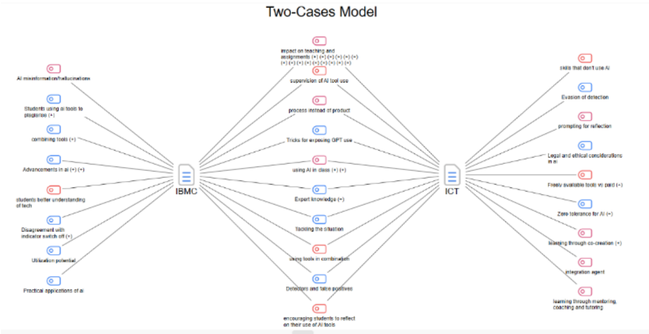
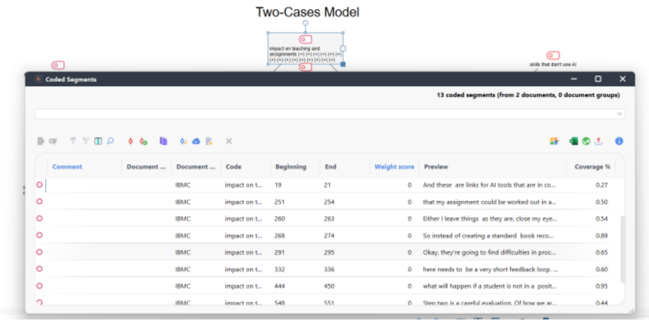


Figure 3: Inset showing the link from the code to the segments in the transcript

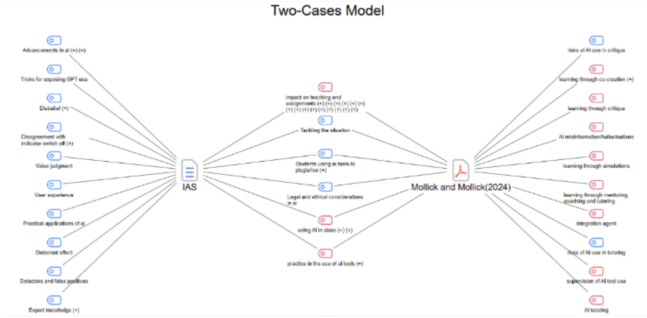


As can be seen in the two cases model, in cases where the institutes showed a largely technical background, the common codes that were highlighted were related to understanding how to supervise AI use, discussing false positives and the use of AI detection, and methods of guaranteeing the fact that students have access to expert knowledge. The codes in the middle of the two cases model are elements that are shared between the two cases, however the codes on the edges of the diagram are codes that are only prevalent in one of the cases. It is interesting that in the case of ICT, the idea of zero tolerance for AI, and the idea of legal and ethical considerations, as well as a focus on skills that don't use AI were deemed to be more important than was the case during the IBMC session, indicating different priorities in the cohorts of lecturers or perhaps a different set of contributions between the different lecturers.

This by no means indicates that in the case of these institutes, there is more of a focus on one item or the other, but it does indicate the direction the discussion took in each of the Institutes and can serve as an interesting pointer for subsequent research into the area and the differences in attitudes between staff in the different institutes.

Taking the comparison in more of a theoretical dimension, the two cases model comparing the case with the largest number of data incidents (IAS) with the Mollick paper, which contains a number of practical suggestions on how to include AI in teaching, indicates the following two cases model.

Figure 4: 2 Cases model, IAS & Instructors as Innovators (Mollick & Mollick, 2024)



In this case, there was a significant overlap between the concerns at IAS regarding how to use the AI tools, but the codes on the right hand side, indicating the paper's suggestions with respect to building AI enabled exercises, indicate that such suggestions have yet to feature in the lecturer perceptions at IAS, who are still trying to get to grip with the new AI enabled reality, while trying to stick to the idea that AI text can effectively be detected, seemingly hoping that the current models of instruction can continue to be followed.

V. Discussion and Conclusion

This preliminary paper and data analysis is but an initial step in an institution like MCAST's approach in dealing with the changes in educational context that are coming due to the profound impact AI is having on our society. The indications from this preliminary qualitative coding exercise is that there are a lot of concerns from our lecturing body, and such concerns were elicited at an early stage by raising awareness of the tools available and the methods at hand that leverage generally available AI tools. The results above indicate that there is an overall sense of disbelief and a wish that AI detectors continue to function, in order for lecturers to have a technical solution that protects them from students submitting artificially generated assignments.

There also seems to be a certain lack of awareness as to how AI tools can facilitate the process of course and teaching development, even though there is a general acknowledgement that AI tools have a profound effect in the classroom. The story told through these codes is the fact that there are a lot of lecturers who require guidance, either through policy or through practical guidance, to deal with tools that they are in general unfamiliar with. The insight that lecturers are experts who can guide students is also very clear, and indicates that AI is no replacement for a skilled lecturer, as can be seen from the risks associated with unsupervised use of AI, as well as the issues related to an AI 'hallucinating' answers. Unsupervised students run the risk of being given incorrect information, and it is ultimately down to the lecturers to clearly show students that the information they may have accessed from a 'trusted' source could be incorrect.

AI is now generally available, and it must be said that there are clear indications that several students are comfortably using AI tools in order to facilitate their studies. This means that more work needs to be done in this area in order to train educators effectively to understand how to deal with the existence of AI while continuing to do their job . Modern educators have to learn to adapt to uncharted territory, where students are able to generate their own learning materials to build information that is useful to them. This takes the agency away from the educator and puts it in the hands of the student, but it also raises important points to consider on the role of the educator and the importance of ensuring that pedagogical expertise is used to filter out the possibility that the AI gets things wrong.

Further studies are important at this point in time. This means that although this initial analysis serves as a first step, it is important to also include the student perspective and therefore introduce specific student-centred qualitative interviews. Such interviews are scheduled to be carried out through the second phase of the STAR project, where students will be requested to use a locally built large language model to perform a task related to the research methods but will then be interviewed qualitatively on their use of AI tools in general. Given a viable set of interviewees, it is hoped that a preliminary model for AI use in education can be put together from the various data sources available. Grounded Theory's flexibility in integrating multiple sources of data can serve us in good stead to build a viable hypothesis based on a set of data sources that are very different to each other.

VI. Acknowledgments

This paper was financed by Project STAR. Project STAR is financed by the Malta Council for Science & Technology, for and on behalf of the Foundation for Science and Technology, through the FUSION: R&I Research Excellence Programme

VII. References

[1] Coe, R., Rauch, C. J., Kime, S., & Singleton, D. (2020). Great teaching toolkit: Evidence review. http://bibliotecadigital.mineduc.cl//handle/20.500.12365/17347

[2] Glaser, B. G., & Strauss, A. L. (2009). The Discovery of Grounded Theory: Strategies for Qualitative Research. Transaction Publishers.

[3] Glaser, P. B. G. (2015). Choosing Classic Grounded Theory: A Grounded Theory Reader of Expert Advice by PhD, Hon. PhD Barney G. Glaser.

[4] Introducing ChatGPT. (2022, November 22). https://openai.com/blog/chatgpt

[5] MCAST – The Malta College of Arts, Science & Technology. (2021, March 23). MCAST. https://mcast.edu.mt/

[6] Mollick, E. R., & Mollick, L. (2024). Instructors as Innovators: A Future-focused Approach to New AI Learning Opportunities, With Prompts (SSRN Scholarly Paper 4802463). https://doi.org/10.2139/ssrn.4802463

[7] Pan, W. H., Chok, M. J., Wong, J. L. S., Shin, Y. X., Poon, Y. S., Yang, Z., Chong, C. Y., Lo, D., & Lim, M. K. (2024). Assessing AI Detectors in Identifying AI-Generated Code: Implications for Education (arXiv:2401.03676). arXiv. http://arxiv.org/abs/2401.03676

[8] Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust Speech Recognition via Large-Scale Weak Supervision (arXiv:2212.04356). arXiv. https://doi.org/10.48550/arXiv.2212.04356

Contribution of AI in improving the user experience of augmented reality applications

1st Frank Samson
Aurelia laboratory
Baden-Württemberg Cooperative State University (DHBW)
Heidenheim, Germany
frank.samson@dhbw-heidenheim.de

2nd Deogratias Shidende
Aurelia laboratory
Baden-Württemberg Cooperative State University (DHBW)
Heidenheim, Germany
deogratias.shidende@dhbw-heidenheim.de

3rd Sabine Moebs
Aurelia laboratory
Baden-Württemberg Cooperative State University (DHBW)
Heidenheim, Germany
sabine.moebs@dhbw-heidenheim.de

***Abstract*—This paper studies the use of Artificial Intelligence (AI) techniques in Augmented Reality (AR)-based applications to enhance usability and user experience. Despite their fundamental differences, combining AR and AI has proven beneficial in overcoming common usability challenges faced in AR environments. The AR part makes the real world more digital, blur these worlds and AI endows intelligence to make this interaction better. This study investigates the AI contribution to AR, explaining that some of these contributions could also be useful addressing user experience challenges due to hardware constraints and inconsistency in interface designs as well as technical issues such as motion sickness and lack of tracking energy accuracy. Utilizing a literature review of articles published in recent years in databases such as IEEE Explore, Google Scholar, and ACM Digital Library, we investigate how AI has been applied to these problems. The review highlights AI has a potential impact on object recognition, scene understanding diagnostics of user engagement in AR setups. This new pipeline not only generalizes impedance from the system level but also improves the user-experience making it a hopeful strategy for AR applications in future.**

Keywords— artificial intelligence techniques, augmented reality, user experience

I. INTRODUCTION

Artificial intelligence (AI) is the technology designed to enable machines to think and perform tasks that were traditionally carried out by humans [1]. Artificial Intelligence (AI) enhances Augmented Reality (AR) systems by enabling advanced functionalities such as object detection and location binding. This relationship allows AR technologies to interact more intelligently with the real world, improving both performance and user experience [2]. It is important to note that these are two different technologies. AR's definition revolves around its derivation from mixed reality technology as it combines the real word and the virtual world by super imposing digital contents on top on the real-world environment view in which the user is located [3].

Despite their differences, AR and AI can be integrated to enhance AR applications by improving user experience through smarter interactions. Additionally, User interfaces significantly impact user experience, as noted by [4] . The interaction of users with AR applications is largely dependent on the design of the application and hardware where it runs from [5]. However, the AR application interface design variations can lead to inconsistencies that affect user's experience. For example, some applications may overload the user with text, while others might focus predominantly on the aesthetics of holograms, providing minimal contextual information [6]. Such discrepancies can confuse users, complicate the learning process of the AR system, causing bad user experience, and ultimately decrease the application's usage.

According to ISO 9241-210:2019, 3.15, UX is defined as "*user's perceptions and responses that result from the use and/or anticipated use of a system, product or service*". Usability primarily concentrates on the learnability and ease of use of a product, whereas user experience (UX) encompasses the entire process of product creation, from conception to interaction with users. Problems in user experience often stem from design flaws, which can lead to poor usability and negatively impact the overall UX. The creation process of AR experience includes core functions such include Tracking, Rendering, and Visualization. Tracking involves real-time user mapping within a scene, tracking the user's position to establish a reference point for viewing digital content relative to the user's location. Rendering is the process of aligning virtual content with the real world from the user's viewpoint. The final function, Visualization, is the real-time generation of virtual content to be overlaid on the real world. These functions depend on how the application is designed as well as the hardware it runs from as previously mentioned. Historically, AR systems utilized computer vision techniques known as Simultaneous Localization and Mapping (SLAM) to analyse visual features across camera frames, mapping and tracking environments effectively [7]. SLAM performs optimally in static environments but faces challenges in dynamic settings, particularly with moving objects and in unstructured or uncertain environments [8]. Different UX challenges on AR applications have been reported by scholars, for instance [9] conducted UX evaluation on mobile AR and identified several technical challenges affecting user experience including screen size limitations, tracking issues and battery drain from power consumption during rendering. [10] reported UX challenges in AR applications, such as motion sickness, field of view limitations, cognitive load, and physical constraints, need to be addressed. [11] mentioned attention, sensation, perception, and action as daily drivers of UX on consumer-based applications, however these key drivers are still challenge in AR application. The integration of AR with AI offers an opportunity to leverage AI techniques to enhance the user experience, mitigating these issues effectively [3]. It is true that AR applications faces user experience challenges, and they can be impacted by the recent advancement of AI techniques. This paper wants to explore these techniques contributions for improving the user experience of AR applications.

II. METHODOS

A literature review was performed that summarizes how Artificial Intelligence (AI) AI techniques are incorporated into the usability and user experience of Augmented Reality (AR) applications. The main research question was: How to utilize AI for improving the user experience in AR applications? To address this, a systematic search was performed, using specific keywords focusing on 'Augmented Reality', 'Artificial Intelligence,' 'user experience', 'usability' and 'integration'. Search was limited to English-language articles of publications between 2020 and 2024, chosen for capturing the current trends which was done through major databases Google Scholar, IEEE Xplore, and ACM Digital Library. The search included two strings: a first one consisting of 'Augmented Reality' AND 'Artificial Intelligence', and a second one with 'Augmented Reality' AND 'User Experience' OR 'Usability'.

The study found 1,769 articles and they were all imported into RAYYAN for screening of duplicate. This process located and eliminated 106 duplicates, leaving 1,663 papers. These were then sorted by titles and abstract which led to 1459 papers being filtered out, leaving 204 papers to go through reviewing the full texts. A further 189 papers were excluded after reviewing the full texts of the remaining 15 publications. Ultimately, only 15 Publications were identified which met the eligibility criteria for inclusion in this review. The search process is illustrated in Figure 1.

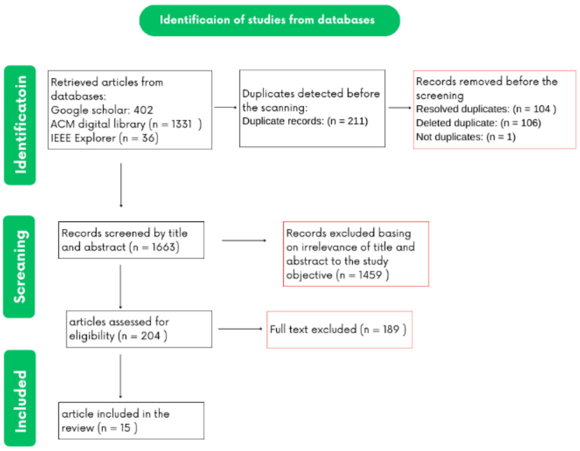


Figure 1: PRISMA flow diagram.

III. RESULTS AND DISCUSSION

A. *User experience challenges in augmented reality applications*

AR devices play are essential resources to enable usage of AR applications [12]. They are designed to capture real-world images, project virtual images, process sensor information, and display content in augmented reality in real-time. Various hardware components, including peripheral tracking and orientation devices, waveguide and Diffraction Optical Elements (DOE) glass wafers and the hand gesture recognition system for natural interaction and user experiences [13], [14]. However, the adoption levels of AR hardware remain far lower than other types of software systems as AR hardware [15] . AR devices are expensive which results in only a small subset of users having such devices hence AR applications are rarely used. In contrast to this, the so-lower popularity of a user, the higher difficulty to achieve it with a comprehensive evaluation of the user experience, because slightly users can provide feedback on this technology. The traditional marker-based AR-based applications have been consistently more common than marker-less applications as due to the GPS reliant characteristics of the wireless experiences' marker-less experiences such as the applications described in this paper are reported for stability and performance challenges, when outdoor [16]. As a result, marker-less AR applications are highly likely to be frustrating or disappointing for users, when they see virtual objects drifting or getting lost. Bad user experiences contribute to more unfavourable beliefs of AR technology overall and result in lower penetration rates in both the consumer and enterprise market [17] , [18]. Gestures and hand movements are common interactions for AR systems. However, detecting these gestures correctly and interpreting them are difficult [16] , especially in a mobile setting, where lighting conditions change quickly and where clutter in the background could confuse the gestures. A study by [19] found that users had trouble with specific interactions performed, such as propagation, scrolling, and resizing. Some virtual objects require complex manipulation such as rotation, scaling, and translation. While the feeling of control is important for delightful user experiences since users feel like they are directly "manipulating" the digital content on the display [20] but challenges with occlusion and gorilla arm problems mean that these may not always be reliable solutions [20] Limited content in AR applications can also reduce the richness of user interactions. The complexity of Creating AR content ranges between high and medium complexity and, creating high-quality content is time-consuming and technically complex [21] [22] Content like 3D models and animations with interactive elements requires specific skills and tools. This is mostly due to rendering delays, as an AR application is executed the user sees movement differently than his own movement and this difference between actual and expected motion using AR applications referred to as Motion sickness [23]. This leads to user will disorientation or dizziness and nausea, which work and fall in the similar [24] . Furthermore, too much information has been reported to induce slow rendering times, overwhelming or cluttered delivery thus potentially under or misinforming [25]. Nowhere do these UX implications appear to be as problematic as in areas with more immediate aspirations like on-demand education and training [26].

B. *Artificial intelligence contribution in addressing the UX challenges in augmented reality applications*

AI is enhancing AR experiences through predictive analysis and personalized content delivery bringing about a concept of AI-powered AR that involves leveraging artificial intelligence (AI) techniques to enhance various aspects of AR, including object recognition, scene understanding, and interaction [27]. By integrating AI capabilities into AR systems, developers can create more intelligent and adaptive experiences that better cater to user needs and preferences. [26] Advanced algorithms and computer vision techniques have markedly improved the precision and reliability of marker-based and marker-less AR. Sophisticated image processing. Machine learning, and deep learning algorithms are enabling more robust object recognition and tracking, even under challenging conditions, significantly enhancing applications in domains like healthcare and manufacturing. deep learning techniques have started to revolutionize the field of AR by introducing higher levels of automation and efficiency in object recognition and scene understanding. Convolutional Neural Networks (CNN) have proven particularly effective in image classification tasks, making it easier for AR systems to identify and superimpose digital information over real-world objects with greater accuracy. Generative Adversarial Networks (GAN) are facilitating more realistic virtual object generation, contributing to a more immersive AR experience [28]. Reinforcement learning algorithms have also been adapted to optimize tracking algorithms in AR, making them more resilient to environmental noise and variable lighting conditions. [29] had an Object Detection Module with OpenCV intending to develop an efficient object detection system capable of swiftly identifying and tracking objects within the augmented reality (AR) environment. Also integration of cutting-edge deep learning models like Faster R-CNN and YOLO to enable real-time object recognition and localization, ultimately contributing to immersive AR experiences

The concept of AI-powered AR suggests that AI can enhance AR-based applications by utilizing smart user feedback, predictive analytics, and classification techniques. It involves (1) online learning (training) of the AI model based on systematic data collection, (2) generating predictions for similar users and use cases, and (3) focusing on content generation for the AR creation process using modern technologies such as NLP to enhance scene understanding [30]. This could add detailed descriptions to scenes and improve user experience through techniques like text/speech to image conversion, image-to-image diffusion, and photogrammetry. These techniques could enable the creation of AR services with a high level of generalization[31].

Furthermore, the integration of AI in content generation processes has significantly enhanced customer relationships for numerous businesses. This is achieved through text-based generative AI tools, such as NLP-powered chatbots [32]. AI content generation tools are categorized based on the type of content they produce, ranging from tools specialized in generating images and enhancing photo quality [33], to those focused on generating text, which rely on NLP algorithms [34]. The use of AI generated contes in AR display has provided easy implementation of subtitle captions for videos and audio using NLP powered speech recognition tools such as IBM, Google, Vosk toolkit and VisualGPT [35] [36] [37]. By creating digital twins of real-world objects, this approach not only enhances the practicality of augmented reality content but also streamlines the content creation process by offering reusable digital objects. Furthermore, the implementation of AI-driven AR digital assistants facilitates seamless assistance for people across diverse linguistic backgrounds, making the technology universally accessible [38]. [39]. [40] Making the AR experience more intelligent to reduce redundant operations is one solution to enhance the user experience. semantic segmentation (Attention U-Net deep neural network) was used to assist automatic information placement in AR using a case study within precision agriculture as an example. The precise location of the crop area in the user view is determined by semantic segmentation, which helps to place information in the AR environment automatically. [27] thereby prompting the combination of deep learning-based object detection and instance segmentation with wearable AR technology to improve the performance of complex tasks. This challenge was addressed in this work using convolutional neural networks in the detection and segmentation of objects in actual environments. Experimental results showed satisfactory segmentation and accurate detection.

IV. CONCLUSION

AI with AR is another phenomenal breakthrough in digital interaction technology. In this study, we explain the Compound augmentation effects on improving the user experience majorly for the AR applications by using AI.

AI can prove to be a wonderful answer to some of the common usability issues pestering other AR applications that have uses in areas such as health, manufacturing, and education.

The results of this study highlight that AI aids largely in enhancing aspects of AR in object recognition, scene understanding and user interaction. Better and more interactive AR experiences with highly refined and high-dimensional image detection and content generation using advance AI algorithms. Furthermore, the ability of AI to perform predictive analysis, while it personalized how content is delivered, can also be beneficial in terms of streamlining user experience and ensuring that AR apps adapt to the singular needs of individual users.

Our findings also demonstrate the need for a design framework for AR and AI applications to mitigate the constraints imposed by AR hardware and software interfaces. All these enhancements of development are vital for alleviating some of the cognitive load for the user, keeping people sicker for lesser, but also for consistent AR experiences.

The AR and AI combination not only elevates the technical performance of the AR application but also significantly increases user satisfaction and user engagement. Additional research should also investigate new AI implementations that can improve AR technology, making it more available globally. This continued melding of AR and AI with this latest iteration, has the possibility of transforming how we interact with both worlds, combining into a new era of next level user experience engagement.

ACKNOWLEDGMENTS

This research work is funded by Baden-Württemberg Cooperative State University (DHBW) through the Education Competence Network (EdCoN). Specifically, special thanks go to DHBW Heidenheim, which is hosting the AuReLiA - Augmented Reality Learning and Accessibility research laboratory, and the partner university Nelson Mandela African Institution of Science and Technology (NM-AIST) Arusha.

REFERENCES

[1] C. Zhang and Y. Lu, “Study on artificial intelligence: The state of the art and future prospects,” *J Ind Inf Integr*, vol. 23, p. 100224, Sep. 2021, doi: 10.1016/J.JII.2021.100224.

[2] Z. Makhataeva, T. Akhmetov, and H. A. Varol, “Augmented Reality-Based Human Memory Enhancement Using Artificial Intelligence,” *Authorea Preprints*, Oct. 2023, doi: 10.36227/TECHRXIV.23586063.V1.

[3] J. S. Devagiri, S. Paheding, Q. Niyaz, X. Yang, and S. Smith, “Augmented Reality and Artificial Intelligence in industry: Trends, tools, and future challenges,” *Expert Syst Appl*, vol. 207, p. 118002, Nov. 2022, doi: 10.1016/J.ESWA.2022.118002.

[4] J. L. Derby and B. S. Chaparro, "The Challenges of Evaluating the Usability of Augmented Reality (AR)," *Proceedings of the Human Factors and Ergonomics Society*, vol. 65, no. 1, pp. 994–998, 2021, doi: 10.1177/1071181321651315.

[5] N. Patkar, L. Merino, and O. Nierstrasz, "Towards requirements engineering with immersive augmented reality," *ACM International Conference Proceeding Series*, pp. 55–60, Mar. 2020, doi: 10.1145/3397537.3398472.

[6] R. Vertucci, S. D'Onofrio, S. Ricciardi, and M. De Nino, "History of Augmented Reality," *Springer Handbooks*, pp. 35–50, 2023, doi: 10.1007/978-3-030-67822-7_2/FIGURES/11.

[7] J. Liu, Y. Xie, S. Gu, and X. Chen, "A SLAM-Based Mobile Augmented Reality Tracking Registration Algorithm," *https://doi.org/10.1142/S0218001420540051*, vol. 34, no. 1, Jun. 2019, doi: 10.1142/S0218001420540051.

[8] M. S. Bahraini, A. B. Rad, and M. Bozorg, "SLAM in Dynamic Environments: A Deep Learning Approach for Moving Object Tracking Using ML-RANSAC Algorithm," *Sensors 2019, Vol. 19, Page 3699*, vol. 19, no. 17, p. 3699, Aug. 2019, doi: 10.3390/S19173699.

[9] A. Dirin and T. H. Laine, "User Experience in Mobile Augmented Reality: Emotions, Challenges, Opportunities and Best Practices," *Computers 2018, Vol. 7, Page 33*, vol. 7, no. 2, p. 33, May 2018, doi: 10.3390/COMPUTERS7020033.

[10] B. Hari Chandana, N. Shaik, and P. Chitralingappa, "Exploring the Frontiers of User Experience Design: VR, AR, and the Future of Interaction," *2023 International Conference on Computer Science and Emerging Technologies, CSET 2023*, 2023, doi: 10.1109/CSET58993.2023.10346724.

[11] S. Bialkova, "From attention to action: Key drivers to augment VR experience for everyday consumer applications," *Proceedings - 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2022*, pp. 247–252, 2022, doi: 10.1109/VRW55335.2022.00058.

[12] E. Battegazzorre, D. Calandra, F. Strada, A. Bottino, and F. Lamberti, "Evaluating the Suitability of Several AR Devices and Tools for Industrial Applications," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12243 LNCS, pp. 248–267, 2020, doi: 10.1007/978-3-030-58468-9_19.

[13] A. H. Behzadan and V. R. Kamat, "Reusable Modular Software Interfaces for Outdoor Augmented Reality Applications in Engineering," *Congress on Computing in Civil Engineering, Proceedings*, pp. 825–837, 2007, doi: 10.1061/40937(261)98.

[14] E. Battegazzorre, D. Calandra, F. Strada, A. Bottino, and F. Lamberti, "Evaluating the Suitability of Several AR Devices and Tools for Industrial Applications," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12243 LNCS, pp. 248–267, 2020, doi: 10.1007/978-3-030-58468-9_19.

[15] R. K. Khamaisi, E. Prati, M. Peruzzini, R. Raffaeli, and M. Pellicciari, "UX in AR-Supported Industrial Human–Robot Collaborative Tasks: A Systematic Review," *Applied Sciences 2021, Vol. 11, Page 10448*, vol. 11, no. 21, p. 10448, Nov. 2021, doi: 10.3390/APP112110448.

[16] E. L. C. Law and M. Heintz, "Augmented reality applications for K-12 education: A systematic review from the usability and user experience perspective," *Int J Child Comput Interact*, vol. 30, p. 100321, Dec. 2021, doi: 10.1016/J.IJCCI.2021.100321.

[17] T. Lukava, D. Z. Morgado Ramirez, and G. Barbareschi, "Two sides of the same coin: accessibility practices and neurodivergent users' experience of extended reality," *J Enabling Technol*, vol. 16, no. 2, pp. 75–90, Aug. 2022, doi: 10.1108/JET-03-2022-0025/FULL/XML.

[18] Y. F. Wu and E. Y. Kim, "Users' Perceptions of Technological Features in Augmented Reality (AR) and Virtual Reality (VR) in Fashion Retailing: A Qualitative Content Analysis," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/3080280.

[19] P. Sinlapanuntakul, J. Korentsides, and B. S. Chaparro, "Exploring the user experience (UX) of a multi-window augmented reality environment," *Front Virtual Real*, vol. 4, p. 1194019, Aug. 2023, doi: 10.3389/FRVIR.2023.1194019/BIBTEX.

[20] C. Wu, M. Huang, W. Sun, and R. Yang, "Feeling of Control in Augmented Reality and Virtual Reality for Virtual Object Manipulation," *2023 9th International Conference on Virtual Reality, ICVR 2023*, pp. 193–199, 2023, doi: 10.1109/ICVR57957.2023.10169460.

[21] V. Davidavičienė, J. Raudeliūnienė, and R. Viršilaitė, "Evaluation of user experience in augmented reality mobile applications," *Journal of Business Economics and Management*, vol. 22, no. 2, pp. 467–481, Feb. 2021, doi: 10.3846/JBEM.2020.13999.

[22] R. Whitehand, G. Albanis, N. Zioulis, W. Bailer, D. Zarpalas, and P. Daras, "Towards user generated AR experiences: Enable consumers to generate their own AR experiences for planning indoor spaces," *IMX 2021 - Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*, pp. 299–304, Jun. 2021, doi: 10.1145/3452918.3465495.

[23] H. Jeon, S. Park, Y. Lee, H. K. Kim, M. Hussain, and J. Park, "Factors Affecting Motion Sickness in an Augmented Reality Environment".

[24] Y. C. Chen, H. M. Sun, and Y. H. Shih, "The effect of motion sickness on presence and user experience for head-mounted virtual reality," *Int J Hum Factors Ergon*, vol. 9, no. 2, pp. 111–127, 2022, doi: 10.1504/IJHFE.2022.122390.

[25] J. Cao, K. Y. Lam, L. H. Lee, X. Liu, P. Hui, and X. Su, "Mobile Augmented Reality: User Interfaces, Frameworks, and Intelligence," *ACM Comput Surv*, vol. 55, no. 9, p. 189, Jan. 2023, doi: 10.1145/3557999.

[26] C. E. Mendoza-Ramírez, J. C. Tudon-Martinez, L. C. Félix-Herrán, J. de J. Lozoya-Santos, and A. Vargas-Martínez, "Augmented Reality: Survey," *Applied Sciences 2023, Vol. 13, Page 10491*, vol. 13, no. 18, p. 10491, Sep. 2023, doi: 10.3390/APP131810491.

[27] Nurhadi, D. Stiawan, M. Y. Idris, and Saparudin, "Object Tracking in Augmented Reality: Enhancement Using Convolutional Neural Networks," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 10, no. 4, pp. 858–866, Nov. 2022, doi: 10.52549/IJEEI.V10I4.4104.

[28] Y. Lu, D. Chen, E. Olaniyi, and Y. Huang, "Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review," *Comput Electron Agric*, vol. 200, p. 107208, Sep. 2022, doi: 10.1016/J.COMPAG.2022.107208.

[29] B. Tiple *et al.*, "AI based Augmented Reality Assistant," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 13s, pp. 505 – 516, Jan. 2024, Accessed: May 20, 2024. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/4617

[30] Y. Sulema, A. Pester, B. Laforge, and F. Andres, "Augmented Reality User's Experience: AI-Based Data Collection, Processing and Analysis," *Springer Series on Cultural Computing*, pp. 31–46, 2023, doi: 10.1007/978-3-031-27166-3_2.

[31] K. Č. Pucihar, V. Geroimenko, and M. Kljun, "FUSE: Towards AI-Based Future Services for Generating Augmented Reality Experiences," *Springer Series on Cultural Computing*, pp. 285–306, 2023, doi: 10.1007/978-3-031-27166-3_16/FIGURES/10.

[32] T. Lalwani, S. Bhalotia, A. Pal, V. Rathod, and S. Bisen, "Implementation of a Chatbot System using AI and NLP," *SSRN Electronic Journal*, May 2018, doi: 10.2139/SSRN.3531782.

[33] Z. Wu, M. Fan, R. Tang, D. Ji, and M. Shidujaman, "The Art of Artificial Intelligent Generated Content for Mobile Photography," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14014 LNCS, pp. 438–453, 2023, doi: 10.1007/978-3-031-35572-1_29.

[34] H. Dang, K. Benharrak, F. Lehmann, and D. Buschek, "Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries," *UIST 2022 - Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, Oct. 2022, doi: 10.1145/3526113.3545672.

[35] J. Li, "Augmented Reality Visual-Captions: Enhancing Captioning Experience for Real-Time Conversations," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14037 LNCS, pp. 380–396, 2023, doi: 10.1007/978-3-031-34609-5_28/FIGURES/2.

[36] A. Kumar *et al.*, "A Survey on IBM Watson and Its Services," *J Phys Conf Ser*, vol. 2273, no. 1, p. 012022, May 2022, doi: 10.1088/1742-6596/2273/1/012022.

[37] J. Chen, H. Guo, K. Yi, B. Li, and M. Elhoseiny, "VisualGPT: Data-Efficient Adaptation of Pretrained Language Models for Image Captioning." pp. 18030–18040, 2022. Accessed: Apr. 02, 2024. [Online]. Available: https://github.com/Vision-CAIR/

[38] H. Lin, S. Wan, W. Gan, J. Chen, and H. C. Chao, "Metaverse in Education: Vision, Opportunities, and Challenges," *Proceedings - 2022 IEEE International Conference on Big Data, Big Data 2022*, pp. 2857–2866, 2022, doi: 10.1109/BIGDATA55660.2022.10021004.

[39] D. Liu, C. Long, H. Zhang, H. Yu, X. Dong, and C. Xiao, "ARShadowGAN: Shadow Generative Adversarial Network for Augmented Reality in Single Light Scenes".

[40] Y. Liang, Z. Xu, S. Rasti, S. Dev, and A. G. Campbell, "On the Use of a Semantic Segmentation Micro-Service in AR Devices for UI Placement," *2022 IEEE Games, Entertainment, Media Conference, GEM 2022*, 2022, doi: 10.1109/GEM56474.2022.10017522.

AICOMP - An Investigation into Future Skills for a World Shaped Through AI

Ulf-Daniel Ehlers
SCHOOL OF ECONOMY
Baden-Wurttemberg Cooperative State University
ULF-DANIEL.EHLERS@DHBW-KARLSRUHE.DE

Emily Rauch
SCHOOL OF ECONOMY
Baden-Wurttemberg Cooperative State University
EMILY.RAUCH@DHBW-KARLSRUHE.DE

Abstract: The AIComp competency model, developed by the NextEducation Research Group under the initiatives "AI-Campus" and "AI-Campus Hub Baden-Württemberg," is rooted in an empirical study involving over 1,600 professionals from Baden-Württemberg, Germany. This study surveyed the skill requirements emerging from the increasing integration of artificial intelligence into the workplace and everyday life. AIComp, an acronym for Artificial Intelligence Competences, represents a comprehensive framework defining 12 future-relevant AI competence fields crucial for a diverse audience. The framework, based on a behavioral approach, deliberately excludes competences specific to individual professions. Notably, AIComp is the first future skill model for AI derived entirely from a mixed-method study. It is organized into three broad competence areas encompassing 12 competence fields, with a total of 36 items.

Keywords: future skills, artificial intelligence, AI competences, competence model, quantitative study

I. Introduction

The rapid advancement in artificial intelligence (AI) has brought the discourse on "future skills" to the forefront, permeating all areas of society, education, and the workforce. This technological transformation presents both challenges and opportunities, impacting individuals' private and professional lives. Embracing this change is not only an opportunity but also a necessity for maintaining social participation.

Numerous studies and initiatives underscore the profound impact of AI on our work and daily lives (see Ehlers et al. 2024a), prompting critical questions about the skills required in a world where machine intelligence increasingly performs human tasks. We must define competence in this new context, exploring its relationship with constantly evolving and contextually accurate knowledge. Additionally, it is essential to reassess the determinants of "successful behavior," "professional success," and "personal fulfillment" in light of AI's influence. We must also consider how AI can contribute to these areas.

The multifaceted role of AI in our environments, from acting as creative partners and learning tutors to providing individualized feedback, highlights that the efficacy of AI-supported systems largely depends on human interaction and utilization. This necessitates the development of specific competences to fully leverage AI's potential. These competences include the ability to provide precise instructions to AI systems, adapt to the rapidly evolving AI landscape, and foster critical thinking skills.

II. Objectives of the AIComp study

The AIComp study meticulously addresses the essential questions regarding the competencies required to navigate social and professional transformations, and to foster personal development in a creative and productive manner. These competencies, referred to as "future skills" (see Ehlers et al., 2024a), are understood as comprehensive dispositions and readiness to act, rooted in knowledge, experience, values, and attitudes. These future skills enable individuals to tackle the transformation tasks they encounter, emphasizing the need for a thorough examination of the skills necessary in an AI-influenced living and working environment.

The AIComp "Future Skills Framework" not only identifies critical future skills for interacting with AI but also delves into the realm of shaping the future with AI. This involves designing action contexts in both private and professional spheres that promote AI integration, developing necessary skills, and actively guiding changes through and with AI. The research envisions a future where human intelligence and empathy gain new significance through interactions with AI systems.

The study's objective was to gain a comprehensive understanding of the current skills landscape, identifying development needs related to AI. These needs encompass personal development, specialization in AI subjects, and skills for shaping the social environment, such as within one's organization. The competency model is based on data from over 1,600 individuals, who provided subjective assessments of their competencies, future relevance, and experiences. These assessments formed the foundation for developing a competency model that outlines the skills required for the effective use and understanding of AI technologies.

Entitled "AIComp - Future Skills for a Living Environment Characterized by AI," this study makes a significant contribution to the ongoing debate about the skills needed in a rapidly changing world. It provides insights into diverse AI applications and offers a robust foundation for targeted educational measures spanning early childhood education, school education, vocational training, higher education, and continuing professional development. The study highlights the evolving role of education in the AI era, supporting the development of essential competencies and emphasizing the importance of understanding the human-AI interaction, which is crucial for overcoming challenges and seizing opportunities in an AI-driven future.

This understanding leads to the clear conclusion that the effectiveness of AI in the living world is intrinsically linked to people's AI skills. "AIComp - Future Skills for an AI-influenced World" offers valuable insights into the necessary future skills and provides a solid basis for designing educational measures to prepare individuals for a world increasingly influenced by AI. The study pursued several key objectives: obtaining comprehensive information on the use of and attitudes towards AI, collecting data on the skills required to operate successfully in an AI-dominated world, and developing a competency model based on respondents' subjective assessments of relevance, experience, and confidence in dealing with AI. These objectives aim to provide a broad and meaningful picture of the current skills landscape and the development needs in the field of AI, particularly among employees in Baden-Württemberg.

III. The Concept of Future Skills as Competences for the Future

Since around 2015, a new development has emerged: in addition to traditional subject curricula, frameworks for "future competences" or "future skills" have been introduced. These frameworks emphasize the skills needed to thrive in a constantly evolving world (Ehlers, 2020, 2022). The term "future skills" has various definitions, but for the purpose of this study, we adopt Ehlers' (2020) definition: competences that enable individuals to act successfully in highly dynamic and complex problem situations. This focus on future skills is evident among both university graduates (Ehlers, 2020; Huber, 2016, 2019; Schlaeger & Tenorth, 2020; Wild et al., 2018) and in vocational education and training (Ehlers, 2022), in German-speaking countries and internationally (Ehlers, 2022).

Competences are often oversimplified as merely the ability to act. However, they actually constitute a complex set of dispositions that include not only skills at the respective level of knowledge and expertise, but also the individual's subjective willingness to act. This willingness is influenced by a combination of knowledge, motivation, will, attitudes, and values (Ehlers, 2020). Determining whether someone "has" a competence requires evaluating their performance, or the successful execution of an action (see Fig. 1).

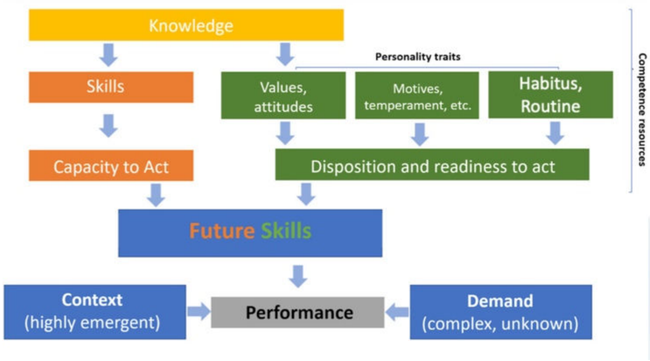


Fig. 1: Future skills as a competence for action (Ehlers 2020)

The triad of knowledge, skills, and attitudes (KSA) (see Binkley, 2012) represents the fundamental building blocks upon which individuals' capacity and willingness to act are based. Future skills, or future competencies, are those competencies specifically geared towards particular fields or contexts of action, enabling individuals to act in a self-organized and successful manner within highly dynamic environments. These competencies draw on cognitive, motivational, volitional, and social resources and are deeply embedded within the learner's value system.

The AIComp study employs the following definition: Future skills are competencies that enable individuals to solve complex problems autonomously and to act successfully in highly emergent contexts. These skills are rooted in cognitive, motivational, volitional, and social resources, are value-based, and can be acquired through a learning process (Ehlers, 2020).

From this perspective, an individual's competence is not merely a combination of discrete knowledge elements and individual skills and abilities. Rather, it constitutes a holistic disposition for action, deeply rooted in the relevant subject matter. In the formula "Knowledge, Skills, Attitudes" (KSA), attitudes are not simply to be viewed as an individual subjective factor; instead, they form the foundation of every competence development process (Mulder & Winterton, 2017; Binkley, 2012).

Understanding future skills requires recognizing that they are dynamic and context-specific. These skills allow individuals to adapt and thrive in constantly changing environments. The underlying cognitive resources include critical thinking and problem-solving abilities. Motivational resources involve intrinsic and extrinsic motivations that drive engagement and persistence. Volitional resources pertain to the willpower and self-regulation required to initiate and sustain actions. Social resources encompass interpersonal skills, collaboration, and communication abilities essential for functioning in social and professional contexts.

Moreover, the development of future skills is a continuous process, influenced by various learning experiences and contexts. This development is not static; it evolves as individuals engage with new challenges and environments, constantly reshaping their cognitive, motivational, volitional, and social frameworks. Thus, the cultivation of future skills is integral to preparing individuals for the uncertainties and complexities of modern and future work environments.

In conclusion, future skills are multifaceted competencies that integrate knowledge, skills, and attitudes, forming a robust framework for effective action in dynamic and emergent contexts. These competencies are essential for enabling individuals to navigate and succeed in a world characterized by rapid change and complexity.

IV. Methodology

The process was conducted with rigorous methodological precision, employing an iterative approach to ensure thoroughness and accuracy. An extensive analysis of the current state of research resulted in the compilation of 167 competence formulations from existing competence models, which were subsequently iteratively clustered into 13 distinct competence fields. This was achieved using the Rapid Evidence Assessment (REA) method, an evidence synthesis technique analogous to a systematic review (SR) but distinguished by its streamlined approach.

The REA, designed for expedited results, makes strategic concessions in terms of breadth, depth, and completeness of the search to facilitate quicker outcomes. This method balances efficiency with scientific rigor, allowing for a comprehensive yet timely synthesis of evidence. For example:

- Search Strategy: The review was conducted using a limited number of databases, excluding unpublished research to streamline the process.
- Inclusion Criteria: Only specific research designs, such as meta-analyses or controlled studies, were included to maintain a high standard of evidence.
- Data Extraction: The extraction was limited to key data points, such as the year of publication, population studied, sector, study design, sample size, moderators/mediators, main results, and effect sizes.
- Critical Evaluation: The quality assessment focused on methodological appropriateness and quality, ensuring that the included studies met rigorous standards despite the expedited timeline.

While the REA's constraints make it more susceptible to bias compared to a comprehensive SR, it offers a pragmatic alternative when time and resources are limited. An SR typically requires a team of researchers working over several months, or even longer, to exhaustively identify all relevant published and unpublished studies. In contrast, an REA can be effectively conducted by experienced researchers within a few weeks. Given the frequent constraints on time and financial resources within organizations, the REA is often the preferred methodology for the critical appraisal of scientific literature.

In the qualitative preliminary study, the following 12 steps were meticulously executed, as detailed in Table 1:

1. Define the scope and objectives of the study.
2. Develop a comprehensive search strategy.
3. Identify relevant databases and sources.
4. Conduct the search and compile the initial pool of studies.
5. Apply inclusion and exclusion criteria to filter studies.
6. Extract key data from the included studies.
7. Perform a quality assessment of the studies.
8. Cluster the competence formulations into thematic fields.
9. Iteratively refine the clusters through expert consultation.
10. Synthesize the findings into coherent competence fields.
11. Validate the competence fields with external experts.
12. Document the methodology and results in a detailed report.

These steps ensured a robust and systematic approach to the analysis, facilitating the identification and clustering of competencies into well-defined fields. This methodical process underscores the scientific rigor and systematic nature of the study, contributing to the advancement of knowledge in the area of future competences.

Table 1: 12 steps of the REA

No.	Name & Description	AIComp Procedure
1	**Background:** Sets the context for the study.	Identify approaches for describing and inventorying "competences for AI" from the research literature.
2	**Research Question:** Specifies the objectives of the study.	Objective: to obtain a comprehensive picture of the range of approaches to "competences" in the field of AI, to develop an inventory list of competence items, and to organize these in the form of a competence structure model.
3	**Inclusion Criteria:** Filters which evidence should be included (e.g., date, type, focus area, etc.).	Criteria for the relevance of studies and approaches: time period (2019 to May 2023), language (German, English), keywords (artificial intelligence and AI, and combinations such as AI and competence, AI and competence framework, AI and skills, AI and abilities, AI and learning, AI and education, AI and training, AI and learning objectives, as well as English equivalents, particularly AI competences, competencies, literacies).
4	**Search Strategy:** Identification of database searches, publications.	Multi-dimensional search process: Internet search engines, meta-databases, journals, conference reports. Iterative, snowball-like reference search including grey literature.
5	**Study Selection:** Abstract review; full-text reading for those that meet the inclusion criteria.	The search focused on elaborated competence frameworks with lists of specific competence items. Additionally, publications explicitly dealing with "AI literacy" or "AI competence" but not containing lists or frameworks were classified as relevant.
6	**Data Extraction:** Extraction of all relevant data and results from the evidence base.	Extraction of the "competence items" from the developed AI competence models and creation of a list with 157 competence items.

7	**Critical Evaluation:** Application of quality metrics; critical interpretations.	Qualitative content-analytical reduction of the item pool. Review of the scope and quality using a test-based allocation of 34 competence items to the four competence dimensions (knowledge, skills, attitudes, and values).
8	**Results:** Results of the evidence assessment. Tensions in the evidence base are highlighted.	Further reduction of the item base from 157 to 83 items by checking and filtering the items for their relationship to the competence dimensions of knowledge, skills, and attitudes. Elimination of items that are purely descriptive or overarching in nature.
9	**Synthesis:** Summarizing the evidence and constructs; main findings on research questions.	Deductive development of competence fields with the data: Structuring of the item pool of 83 competence items based on 17 profile fields of the NextEducation model. Assignment of items to competence fields resulted in 13 competence fields. Further refinement of field designations based on item content while retaining the structure of the NextEducation model. Result: 13 competence fields with a total of 83 competence items.
10	**Conclusions:** Concise statements that convey the main findings.	Development of comprehensive competence field designations and necessary internal structuring for some competence fields. Result: Formulated description and systematization of AI-related competencies.
11	**Limitations:** Description of the limitations of the FGD method for this study.	Limitations in the construct clarity of the competence field formulations were checked via qualitative interviews. Focus and coverage of the competence items were also examined.
12	**Practical Implications:** Recommendations for action.	Development of a qualitatively based action-theoretical competence model, providing a basis for further quantitative research steps.

As a consequence of the Rapid Evidence Assessment (REA) research process, an initial 13-field model was derived from the current state of research. This model was subsequently refined through oral consultations with experts and stakeholders, which led to its validation and reduction to 12 fields of competence. This phase employed qualitative methods, including semi-structured interviews and group discussions, to gain a comprehensive understanding of the relevant competences.

The subsequent empirical quantitative study was designed based on the preceding analysis. A comprehensive questionnaire was developed, comprising 44 main questions covering eight areas of AI use and attitudes, as well as three items per area of expertise. These items included: (1) self-perception/sovereignty, (2) assessment of future importance, and (3) individual experience with competence (see Fig. 2). Additionally, the questionnaire gathered socio-demographic data and information on opinions and usage experiences of the respondents. The quantitative phase of the AIComp study was conducted as an online survey, facilitated through an online questionnaire distributed by multipliers and supported by a social media campaign. The field phase spanned from May to July 2023, involving the recruitment of approximately 31,900 individuals, resulting in 6,653 participants, of whom 1,644 completed the questionnaire in full (Ehlers et al., 2024a).

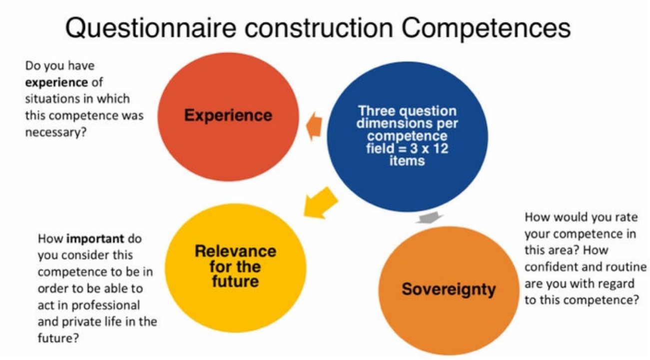


Fig. 2: Structure of the three questions for each competence field

The data analysis was carried out in the following steps:

1. Univariate and Bivariate Analysis: The initial phase of data analysis involved univariate and bivariate evaluations of all questions in relation to variables such as age, gender, position, and organizational affiliation.
2. Construction of an AI Activity Index: An AI activity index was constructed based on responses to questions regarding the frequency (question 2 of the questionnaire) and type (question 3 of the questionnaire) of AI use.
3. Attitudes Towards AI: Analysis of the main statements on attitudes towards AI was conducted.
4. Explorative Evaluation: Development of a statement system for the main statements and trends, which was then refined and validated discursively by the research team.
5. Principal Component Analysis: For the construction of an updated competency model, we analyzed 59,184 individual subjective assessments of AI-related competencies. Principal component analysis was utilized to construct and compare different variants of the competency model with 4 to 12 factors (i.e., clusters indicating separable competency areas). A 12-factor solution was ultimately selected based on statistical indicators (screeplot) and content quality, confirming the explanatory power of the future skills approach and resulting in new allocations and attributions in some areas (see the separate report on the development of the AIComp skills model).

The data was used to construct an AI activity index (KIX), representing an innovative approach to assessing the use of and engagement with AI. The KIX proved to be a good predictor of respondents' assessments of competence. The index is based on the combination of two central variables: intensity of use and type of use of AI. Intensity of use is measured by frequency, with response options ranging from "a few times already" to "several times a week or more." The type of use distinguishes between passive, active, and creative use of AI technologies. By scoring both variables, six levels of AI engagement are created, ranging from non-use to high intensity and formative use. These six levels were then summarized into three KIX levels of activity (see Fig. 3).

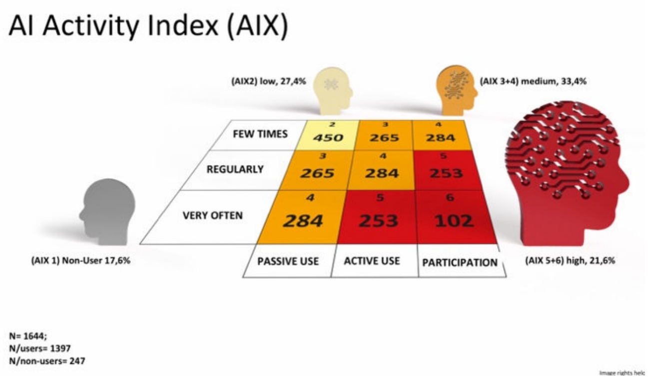


Figure 3: Construction of the AI Activity Index (KIX): Matrix and categorization into levels

The application of the KIX enables a differentiated analysis of AI use across various subsamples and subgroups. The distribution of the total sample across these groups is as follows:

- High AI Activity Index: This group comprises 355 respondents (21.6% of the total sample, N=1644), indicating intensive use and high expertise in AI technologies and applications.
- Medium AI Activity Index: This group includes 549 respondents (33.4%), indicating moderate familiarity and integration of AI in their activities.
- Low AI Activity Index: This group consists of 450 respondents (27.4%), suggesting limited or incomplete integration of AI in their processes.
- Non-users: This group includes 247 respondents (15.0%), indicating no use of AI, which could be due to various factors such as lack of resources, skills, or necessity in their work context.

In summary, the distribution shows that a significant proportion of respondents use AI to varying degrees, with a slight preponderance in the medium AI activity index category. However, a notable portion of respondents do not use AI, highlighting potential areas for the development of AI expertise and usage.

V. Results of the Study

The data analysis involved several stages, including univariate and bivariate analyses, and the construction of an AI activity index. This paper focuses on the principal component analysis (PCA) used to construct the competence model. The comprehensive data analysis can be accessed at www.ai-comp.org or in the publications by Ehlers et al. (2024a, 2024b).

The PCA condensed the 59,184 individual responses from the 1,644 participants into high-variance dimensions that are uncorrelated with each other. These dimensions reflect the relationships among individual competence items, which were subsequently grouped into competence fields and structured into broader competence areas.

Two modes of factor analysis were conducted: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). This process allowed the formation of item groups based on respondents' answers, identifying clusters of related items. After evaluating several alternatives, the research team selected a 12-factor solution. This decision was based on the statistical model fit, internal coherence of the factors, and the ability to explain more than 70% of the variance through the 12-factor solution (Ehlers et al., 2024b). These item groups formed the basis for the competence fields in the AIComp model (see Fig. 3).

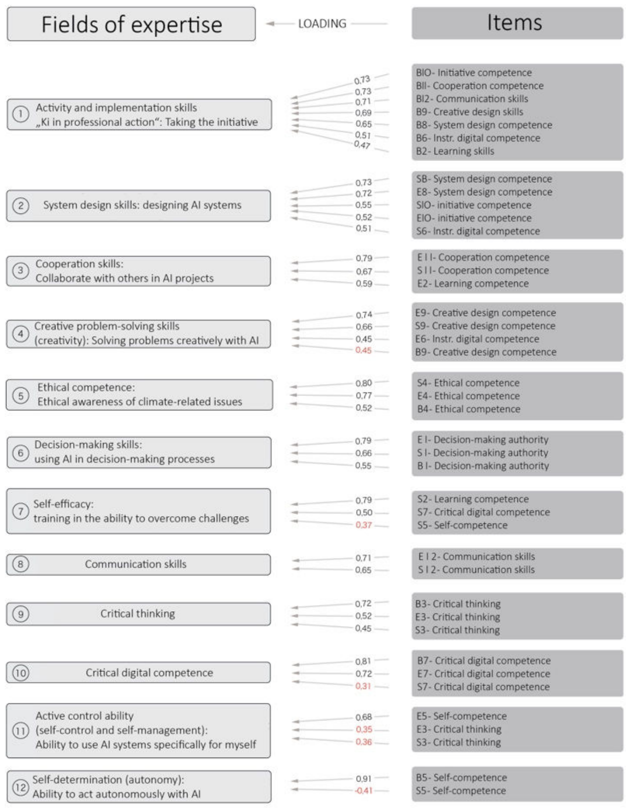


Fig. 3: AIComp component structure (Ehlers et al. 2024b)

The AIComp model of future skills is conceptualized as a "competence structure model." This model is composed of three main components (fig. 4):

1. Twelve Future Skills Fields: These fields represent distinct areas of competency required for success in an AI-influenced world.
2. Three Broader Areas of Competences: These broader areas categorize the twelve fields into overarching domains of expertise.
3. Specific Competence Items: Each competence field includes detailed competence items, which are the specific skills, knowledge, and attitudes necessary for effective action.

Fig. 4: Competence Structure Model AIComp, with 3 Areas (Inner Layer), 12 Fields (Middle Layer), and 36 Items (Outer Layer)

Definition and Elaboration of Competence Fields: A "competence field" is defined as a thematic disposition for action, encapsulating the ability and willingness to act in specific contexts. Each competence field includes detailed descriptions of the relevant knowledge, skills, and attitudes, which are further elaborated in the full research report. This comprehensive documentation includes illustrative examples of competence descriptions, drawn from relevant literature (Ehlers et al., 2024b).

This structured approach ensures that each competence field is thoroughly defined and contextualized, providing a robust framework for understanding and developing future skills in an AI-driven environment. The AIComp model not only categorizes these competencies but also offers a detailed roadmap for their application and development, ensuring that individuals are well-equipped to navigate the complexities of an AI-integrated world.

The competence structure model addresses the question of which competencies are essential for individuals to be, become, and remain capable of acting successfully in a working and living environment permeated by AI. It determines subjective assessments of successful professional and private behaviour in a world that is increasingly influenced by AI.

Top Level: Three Areas of Expertise
According to the model future skills can relate to three aspects: (1) either to individual developmental aspects of the acting subject (e.g. the ability to self-reflect on something experienced in the past, or ethical competence), (2) to dealing with an object, a subject or a task (e.g. design thinking skills), or (3) to the social environment or organisation in which the individual is acting (e.g. cooperation or communication skills) (for elaboration see Ehlers et al. 2024b).
The future skills we have identified as relevant for future skills in an AI-influenced world can therefore be categorised into one of these three dimensions (fig.5):

- Dimension 1: Developing personal capabilities for AI-related domains of action. These are skills that enable individuals to act confidently in an AI-influenced world and to use AI concepts and tools responsibly and reflectively for their own purposes.
- Dimension 2: Working and designing with and for AI. This is about skills to (further) develop work tasks and organisational processes.
- Dimension 3: Shaping one's own social environment with and for AI. This is about competences to use AI appropriately in one's own private or professional social environment and to creatively design new AI-related fields of activity in cooperation with others.

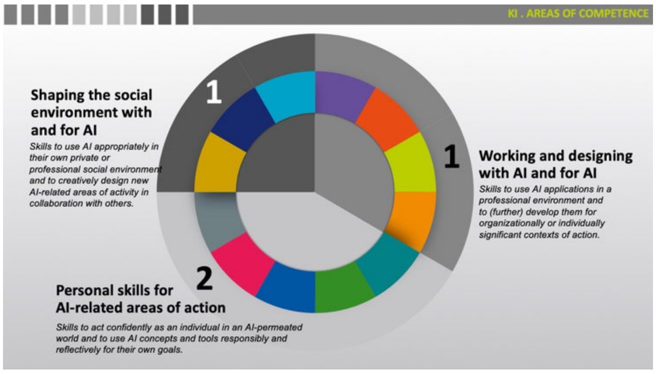


Fig. 5: Top level of the competence structure model (Ehlers et al. 2024b)

Medium Level: 12 areas of expertise
A total of 12 competence fields are assigned to these competence areas. A competence field is defined as a thematic disposition for action, with the corresponding ability and willingness to act being described in the form of a definition for each (fig. 6).

Fig. 6: Twelve areas of expertise in the AIComp model (Ehlers et al. 2024b)

Table 2 describes the 12 competence fields and the corresponding definitions.

Table 2: Description of the Future Skills AIComp (Ehlers et al. 2024b)

1	Activity and implementation competence "AI in professional behaviour":	The disposition to act proactively in the field of artificial intelligence (AI) and to integrate innovations in this field into one's own work context is the disposition of activity and implementation competence for AI in professional activities. This competence is of a fundamental nature and encompasses the knowledge, skills and attitude required to

	Taking the initiative	orientate oneself with regard to the growing importance of AI in professional and private contexts. It also requires an approach to the topic that is open and critical, as well as the internalisation of the importance of continuous personal initiative for further training in this area.
2	System design expertise: designing AI systems	System design competence is the ability to plan and integrate both conceptual and technological AI systems in a professional context and to implement them in specialised fields of activity. System design competence comprises the knowledge, skills and inner attitude to analyse existing systems (both technical systems and organisations) and to assess the potential and limits of the use of AI systems for these, to actively help shape and implement them.
3	Creative problem-solving skills (creativity): Solving problems creatively with AI	Creative problem-solving competence is the ability to use AI systems for creative problem-solving, idea generation and vision development. It comprises knowledge, skills and inner attitudes that make it possible to solve complex problems by combining technical and human systems.
4	Critical digital competence: being able to assess the benefits and challenges of technical applications	Critical digital competence is the ability to understand, analyse and critically evaluate the inherent logic of AI systems with regard to their use of data and their impact on organisations and society. It includes the necessary knowledge, skills and inner attitudes to be able to critically and differentiatedly assess AI systems in relation to a given context of values and application. (Example: being able to assess and analyse the influence of AI technologies on the handling of data).
5	Decision-making skills: using AI in decision-making processes	Decision-making competence is the disposition to use AI applications and systems to weigh up alternative choices and make decisions. The competence includes the necessary knowledge, skills and inner attitudes to consciously take responsibility for the decisions supported and/or made by AI.
6	Self-efficacy: Conviction that you can overcome AI-related challenges with your own abilities	Self-efficacy as a competence is the disposition to master the challenges associated with AI that arise in one's own context of action through one's own actions with conviction, courage and confidence. It encompasses the necessary knowledge, skills and inner attitudes to overcome AI-related challenges and to utilise AI systems in an appropriate manner for one's own questions and tasks.
7	Critical thinking: questioning how AI influences actions and decisions	Critical thinking as a competence is the disposition to reflect on the underlying ways of thinking, value systems and behaviours in AI-influenced spheres of activity and to be able to evaluate how they influence actions and decisions. It encompasses the necessary knowledge, skills and inner attitudes to analytically and critically assess AI-related circumstances, systems and applications as well as their effects.
8	Active control ability (self-control and self-management): Using AI systems specifically for me	Active steering ability is the disposition to personalise AI applications, systems and associated processes for one's own personal and professional development and to be able to shape them confidently and largely independently of external influences. This includes the knowledge, skills and inner attitudes for independent motivation and planning, cognitive load management and a high level of personal responsibility.
9	Self-determination (autonomy): Self-determined action with AI	Self-determination as a competence is the disposition to deal autonomously and confidently with AI applications without allowing oneself to be patronised. It requires the knowledge, skills and inner attitudes to develop a critical awareness of one's own personal boundaries and to act in a self-determined manner in relation to suggestion and decision-making processes with and through AI applications.
10	Ethical competence: Ethical awareness of AI-related issues	Ethical competence is the ability to recognise, articulate and critically reflect on ethically relevant issues and questions in connection with AI technologies and related processes. It encompasses the knowledge, skills and inner attitudes to deal intensively with the ethical implications of the use of AI applications and systems and includes an awareness of responsible behaviour in relation to AI.
11	Cooperation skills: working together with others in AI projects	Cooperation competence is the disposition to work in cross-departmental/interdisciplinary development partnerships and cooperations on AI transformation projects and new projects in relation to AI, also across organisational or cultural boundaries. It encompasses the necessary knowledge, skills and inner attitudes as well as the willingness to learn and develop further in this regard.
12	Communication skills: formulate and discuss specific topics on AI	Communication competence is the ability to communicate with others on AI-related topics in different contexts in a way that is appropriate to the situation, including views that differ from one's own. It encompasses the knowledge, skills and inner attitudes required to empathise with and communicate other perspectives on AI and related issues.

Level 3: 36 competence items

Each competence field in the AIComp model comprises specific competence items that describe the necessary skills, knowledge and attitudes in detail. These items are carefully selected indicators that depict various aspects of the respective competence. Through the factor analysis each item has a certain weight or "load" on the respective factors. Factors with a higher load have a greater definitory power than factors with a lower load. Figure 3 shows loading strength for each item on the factors.

VI. Critical Reflection on the Study and Conclusion

The rapid development of artificial intelligence (AI) necessitates a re-evaluation of essential future competencies, commonly referred to as "future skills." These skills are crucial for navigating educational and professional challenges while leveraging the opportunities presented by AI. Active participation in this transformation is imperative for individuals to secure their place in society and fulfill their personal and professional potential.

The AIComp study identifies and describes the competencies required for an AI-influenced environment. By integrating qualitative and quantitative methodologies, the study constructed a competency model delineating the abilities necessary to understand and utilize AI technologies. This model provides a foundational framework for educational initiatives designed to prepare individuals for an increasingly AI-driven world.

The AIComp study highlights the dependency of AI efficacy on AI-related skills. It offers valuable insights into future skills and establishes a robust basis for designing educational measures to prepare people for a future dominated by AI. However, several open questions and areas for improvement necessitate further analysis. One critical consideration is the long-term validity of the identified future skills, given the rapid pace of technological advancement. Longitudinal studies would be beneficial in determining the durability and adaptability of the skills model over time.

A notable limitation of the competence study is its context dependency. It remains to be seen how transferable the results are to other regions or cultures. Further research is needed to explore whether and how required competencies differ across global contexts. The practical implementation of the study's findings in education and training also warrants attention. There is a lack of concrete recommendations for translating these findings into practice.

The AIComp study's integration of qualitative and quantitative methods is a strength, but it also increases complexity. A more detailed description of the methodological integration and specific steps taken to ensure coherence between methods would substantiate the study's scientific claims. While principal component analysis is a well-established method, it involves certain risks related to the interpretation of components. Transparent presentation of decision-making processes and sensitivity analyses would further enhance the robustness of the results.

The definition and operationalization of competencies in the AIComp study are clear and well-founded. However, questions remain about their ability to capture the actual complexity and dynamics of competencies in real-world contexts. Continuous review and adaptation of the items and competence fields are essential to ensure their ongoing relevance and accuracy.

In conclusion, the AIComp study provides a methodologically robust foundation for identifying and analyzing future skills in an AI-influenced world. Nonetheless, unresolved issues and areas for enhancement should be addressed in future research to further validate and apply the results effectively.

VII. Bibliography

1. Amelang, M., & Bartussek, D. (1981). Differential psychology and personality research. Stuttgart, Berlin, Cologne, Mainz: Kohlhammer.
2. André, E. & Bauer W. et al. (eds.). (2021). *Competence Development for Artificial Intelligence - Changes, Needs and Options for Action.* Whitepaper from the Learning Systems Platform, Munich. DOI: https://doi.org/10.48669/pls_2021-2.
3. Baumgartner, P. (2014). Taxonomy of teaching methods: A plea for didactic diversity. Münster: Waxmann.
4. Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First Century Skills. In P. Griffin & E. Care (Eds.), Assessment and Teaching of 21st Century Skills. Methods and approach. Dordrecht: Springer, 17-66.
5. Bortz, J. (1999). Statistics for social scientists. Berlin, Heidelberg: Springer.
6. Brosius, F., & Brosius, G. (1995). SPSS Base System and Professional Statistics. Bonn: Thomson.
7. Comrey, A.L., & Lee, H.B. (2013). A first course in factor analysis. 2nd Ed. New York, NY: Psychology Press.
8. Dunteman, G.H. (1989). Principal Components Analysis. Quantitative applications in the Social Sciences. Newbury Park, CA: SAGE Publications.
9. Ehlers, U.-D. (2020). Future Skills, Future of Higher Education - Future Higher Education, https://doi.org/10.1007/978-3-658-29297-3_5
10. Ehlers, U.-D. (2023). Future Skills Framework. Competences for the AI world. Issue 305 ManagerSeminars
11. Ehlers, U.-D., Lindner, M., Rauch, E. (2024a). AICOMP - Future Skills for a living and working environment characterised by AI. Research report 1. Karlsruhe
12. Ehlers, U.-D., Lindner, M., Rauch, E. (2024b). AICOMP - Future Skills for a living and working environment characterised by AI. Research report 1. Karlsruhe
13. Ehlers, U.-D., Lindner, M., Sommer, S. & Rauch, E. (2023). AICOMP - Future Skills in a World Increasingly Shaped By AI Ubiquity Proceedings, 3(1): 230-239. URL: https://ubiquityproceedings.com/articles/10.5334/uproc.91
14. Erpenbeck, J., & Sauter, W. (2015). E-learning and Blended Learning: Self-directed Learning Processes for Knowledge Building and Qualification. Heidelberg: Springer.
15. Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. Psychological Bulletin, 103(2), 265-275.
16. Hartmann, P., & Watling, M. (2022). Applying AI: Culture, Change, Communication. Whitepaper of the "Applied AI" initiative. Munich: Applied AI.
17. Huber, L. (2016). "Studium Generale" or "key qualifications"? An attempt at orientation in the field of higher education. In U. Konnertz & S. Mühleisen (Eds.), Education and key qualifications. On the role of key qualifications at universities. Frankfurt a.M., New York: Peter Lang, 101-122.
18. Huber, L. (2019). "Education through science" as a quality of study. The higher education system (67), 154-159.

19. Jackson, J.E. (1991). A User's Guide to Principal Components. New York, NY: John Wiley & Sons.
20. Jolliffe, I.T. (1986). Principal Component Analysis and Factor Analysis. In: Principal Component Analysis. Springer Series in Statistics. New York, NY: Springer.
21. Kerres, M. (2014). Mediendidaktik: Konzeption und Entwicklung mediengestützter Lernangebote. Munich: Oldenbourg Verlag.
22. Laupichler, M., Aster, A. & Raupach, T. (2023). *Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy*. In Computers and Education: Artificial Intelligence 4 (2023), 100126. DOI: https://doi.org/10.1016/j.caeai.2023.100126.
23. Long, D. & Magerko, B. (2020). *What is AI Literacy? Competencies and Design Considerations.* In Conference on Human Factors in Computing Systems - Proceedings. ACM, Honolulu, Hawaii, pp. 1-16. DOI: https://doi.org/10.1145/3313831.3376727.
24. McKenney, S., & Reeves, T.C. (2012). Conducting Educational Design Research. London: Routledge.
25. Mulder, M., & Winterton, J. (2017). Introduction. In M. Mulder (Ed.), Competence-Based Vocational and Professional Education. Bridging the Worlds of Work and Education. Berlin, Cham: Springer, 1-43.
26. Reinmann, G. (2018). Experimental research in educational science. In G. Reinmann & M. Mandl (Eds.), Textbook for Learning and Teaching with Technologies. Munich: Oldenbourg.
27. Schlaeger, J. & Tenorth, H.-E. (2020). Education through science: The benefits of research-based learning. Berliner Wissenschafts-Verlag.
28. Vuorikari, R., Kluzer, S. & Punie, Y. (2022). *DigComp 2.2. the Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes*. Publications Office of the European Union, Luxembourg. https://publications.jrc.ec.europa.eu/repository/handle/JRC128415.
29. Wild, S., Deuer, E. & Pohlenz, P. (2018). Understanding the academic success of full-time teachers in the field of business studies at the Baden-Württemberg Cooperative State University (DHBW) - An attempt at typification. Journal for Evaluation - ZfEv, 17(2), 269-288.

Author Index

Program Committee

Slim Abdennadher	German International University
Mandy Badermann	University of Tübingen
Sara Brockmans	DHBW Karlsruhe
Irantzu Calvo Santamaria	MU Enpresagintza
Bernhard Drabant	DHBW Mannheim
Alexander Dück	DHBW Mannheim
Annemarie Friedrich	Universität Augsburg
Eva Gengler	Friedrich-Alexander-Universität Erlangen-Nürnberg, FemAI - Center for Feminist Artificial Intelligence
Sabine Gillner	Axon Ivy
Frank Gottfried	SAP
Beate Grawemeyer	London Knowledge Lab, Birkbeck College, University of London
Joerg Homberger	HFT
Michael Karl	Deutsches Zentrum für Luft- und Raumfahrt e.V. / Institut für KI Sicherheit
Thomas Kessel	DHBW Stuttgart
Monika Kochanowski	DHBW Stuttgart
Falko Kötter	DHBW Stuttgart
Bozena Lamek-Creutz	DHBW Mannheim
Stojan Maleschlijski	SAP
Johanna Muckenhuber	FH Joanneum (University of Applied Science) Graz
Prof. Dr. Carsten Mueller	Baden-Wuerttemberg Cooperative State University Mosbach
Dennis Pfisterer	DHBW Mannheim
Annette Plankensteiner	DHBW Stuttgart
Dirk Reichardt	DHBW Stuttgart
Sven Schmeier	DFKI
Jennifer Schoch	DHBW Karlsruhe
Stephan Schulz	DHBW Stuttgart
Marcus Strand	DHBW Karlsruhe
Nathan Sudermann-Merx	DHBW Mannheim
Philippe Thomas	Institut für Informatik, Humboldt-Universität zu Berlin
Andreas Weber	DHBW Karlsruhe
Katja Wengler	DHBW Karlsruhe
Philipp Wicke	LMU, Ludwig Maximillians Universität Munich