

On the Creation of Classifiers to Support Assessment of E-Portfolios

Alexander Gantikow,^a Andreas Isking,^a
Paul Libbrecht,^a Wolfgang Müller^a

^a Media Education and Visualization
University of Education Weingarten
Weingarten, Germany
{lastname}@md-phw.de

Sandra Rebholz^{a,b}

^b Electrical Engineering, Media and Computer Science
Ostbayerische Technische Hochschule Amberg-Weiden
Amberg-Weiden, Germany
s.rebholz@oth-aw.de

Abstract—Text-classifiers are among the important services required of contemporary AI systems. Based on a trained classifier, one can perform relevant tasks such as highlighting relevant passages of a text, analyzing what is talked about or encouraging students to write texts which cover important concepts. This form of analysis is at the core of the AISOP project, AI-supported Observation of e-portfolios: Through the development of classifiers and the visualization of their results, the project aims at supporting e-Portfolio assessment in the university context. We report our investigations on the creation process of classifiers, on how their quality can be evaluated and enhanced, and on the domain specificities we have met.

Keywords—component; e-portfolios, learning, assessment, visualization, text classification, natural language processing.

I. INTRODUCTION

E-portfolios are digital works that learners create to express their knowledge and the process of acquiring this knowledge. As a composition, an e-portfolio lets learners express themselves with digital artifacts (texts, graphics, videos, ...) using a vocabulary that is both accessible to them and to their readers. As a digital composition, e-portfolios enjoy the ease of re-use to create a representation of their learning process and knowledge that is their own.

In a course, teachers can request students to create e-portfolios in order to testify their knowledge: They should represent what and how they have learned in the course [1]. The realization can be far beyond a mere repeat of the instructor's material, with the knowledge expressed in a re-appropriated manner and with selecting a focus topic of the student's own choice. The writing activity of e-portfolios is done "open-book" with references accessible. This ubiquity of knowledge imitates contemporary professional life, where keeping knowledge sources at hand and adapting them to the real world is a key exercise [2].

At the University of Education Weingarten and in many other universities, e-portfolios are used as means of documenting and assessing learning processes: Students have the task of creating e-portfolios to represent their knowledge as well as enrich their presentation by covering in depth a selected subject. This provides a representation of their knowledge which can be assessed. They do not simply repeat or formulate the general knowledge of the course, they also describe the process of acquiring this knowledge and of realizing chosen projects.

Teachers of the courses with e-portfolios assess the acquired competencies and knowledge of the students by

evaluating criteria such as the topical completeness, the subject-specific depth, the use of source material, the originality of the presented artifacts, the quality of the explanations, formal aspects, and the expressed reflection about the relationship to the subjects (e.g. the connections to their own environment). Upon reading the portfolios, teachers can assert how well these criteria are met and can classify what subjects were (more or less deeply) treated. However, this assessment is mostly done manually and therefore very time-consuming in classes with many participants. For instance, in a project module aimed at fostering digital competences of pre-service teachers, there are about 600 students per year that document their learning processes in e-portfolios. An assessment of one e-portfolio typically takes 1-2 hours.

The AISOP¹ project selects and applies text analysis techniques on the e-portfolio content in order to support teachers in assessing the learning outcomes displayed in the e-portfolios and to support the students in the process of creating their individual e-portfolio. Our approach is the following: based on the evaluation criteria for e-portfolios, the dimensions of analysis of the text content are derived and suitable techniques are chosen to deliver the required analysis outputs. As a first step, we apply the identification of the core concepts covered in a given e-portfolio and provide teachers with an interactive tool to easily locate the text passages in the e-portfolio where these concepts are explained. Because the concepts must be those of the course, we are building a custom text classifier based on the course-specific content. Our goal is to align and apply state-of-the-art text analysis techniques within the context of a practical setting in higher education.

In this paper, we investigate: *What are the specific requirements arising from this educational context and how are text classifiers built and refined if they are to be used in a supportive manner for teachers and students?*

In order to address this question, we provide an overview of existing initiatives in the field of applying text analysis to support assessment. Then, section III depicts the criteria along which an e-portfolio is evaluated and in which conditions. Section IV describes the data selected to enter the training process while section V is a report on our process of annotating the training data. Section VI illustrates how a training result has been evaluated. Section VII presents lessons-learned. The paper concludes with a sketch of relevant future works.

¹ AISOP: AI Supported Observation of e-Portfolios, see aisop.de for more about the project.

II. EXISTING INITIATIVES TO SUPPORT ASSESSMENT

In this section we explore the landscape of existing projects that employ natural language processing (NLP) to support assessment that is applicable for e-portfolios.

The vision of a semi-automatic support to assess portfolios that this project also follows has been formulated in [3]. The formative assessment importance in the learning process has been established, e.g. in [4].

Text-analysis techniques have largely progressed, and techniques to extract structured information from texts based on machine learning have been investigated with promising results. For example, transformers [5] and language models building on transformers, like BERT [6] or Sentence-BERT [7] bring the ability to model each section of the e-portfolios in vector spaces and to extract comparisons and knowledge.

The design of visualizations and dashboards for the analysis of learning processes have been the object of a number of studies, among others [8]. Visualizations and dashboards aim at providing various perspectives on the e-portfolio content, both for teachers and for students. While teachers need to quickly grasp the overall content of the e-portfolio and identify relevant parts of the e-portfolio for closer inspection, students need feedback on the current state of their e-portfolio with regard to the underlying assessment criteria. Applying both, a suitable design and development process, as recommended in [8] and state-of-the-art visual design [9] and analytics concepts [10], is the prerequisite for developing dashboards that communicate useful insights and efficiently support both learning and assessment processes.

Text-analysis methods are among the classical applications of artificial intelligence. While they may deliver strong services, a challenge remains to establish an adequate level of trust that a user can put in them, since any machine-learning-based system is based on a training process that can be biased or missing under multiple aspects. Principles of trust as in [11] have to be considered.

However, the problem of trust is a long-known problem as demonstrated in [12] and may be related to how the system made available lets users expect being helped or advised: Instead of taking the role similar to a human coach, a system to analyze e-portfolios should take the role of an instrument [13] supporting a transparent synthesis and an active navigation. By the multiplicity and by the graphical nature of the analysis results, we aim to let users enjoy the benefits of the services without enduring considerable drawbacks. Such an application of artificial intelligence technique is far from automatic grading (which is a classical topic as shown in [14]).

III. ASSESSING AN E-PORTFOLIO

A. Criteria-Based Assessment

For assessing the students' competencies documented in individual e-portfolios, a list of evaluation criteria is used. The criteria are derived from the expected learning outcomes in the associated course and have been applied and improved during the manual assessment of the e-portfolios of past cohorts. For each criterion, students can reach a certain performance level which is defined by a

competency description. Typically, the evaluation criteria and the competency descriptions are represented by a rubric. Rubrics are scoring grids that are used as an assessment instrument. In order to ensure transparent assessment, the rubric is communicated to the students. We group the criteria into three categories: content-related criteria, formal criteria, and criteria related to the learning process.

Content-related criteria assess the completeness of the representation (coverage of all topics), the depth of content (information density of the text), the use of relevant literature and the genuine creation of artifacts. Formal criteria include aspects such as design (clarity and aesthetics), media usage (number and variety of included media artifacts), language (orthography, choice of words and linguistic style) and correct citing (consistency, completeness and correct allocation). Finally, criteria related to the learning process aim at assessing the curiosity, the openness, and the individual commitment that students exhibit when dealing with the learning content. It becomes visible if new (sub-)topics or examples have been added by the students or they wrote in-depth reflections.

Considering all these aspects in the assessment of e-portfolios, a holistic view of the students' learning outcomes and process is provided as reported in [3].

As a first step to support the measurement of content criteria such as completeness and coverage, a text classifier is built to identify the most important concepts covered in an e-portfolio. The identified concepts match the expected domain concepts that students should elaborate in their e-portfolio. This set of concepts represent the course-specific learning content and is specified by the course instructor. For visually representing the domain concepts and their relationships, a concept map (diagram that explains the general structure of a subject domain) can be used.

B. Enriching the learning through assessment

Having sketched the ingredients of an e-portfolio assessment, we now describe situations of how it is used in practice and identify possibilities in which the AISOP project can enhance it:

- In the writing/creation process, students often need to take distance from their e-portfolio so as to see it from another perspective [15]. Simple preview functions found in platforms such as Mahara [16] support this. AISOP aims at the enrichment of this preview: for example, by using visualizations, a student is informed which topic was covered from a concept map. As self-directed learners, they can continuously self-assess and iteratively adapt their work based on the feedback provided by the results of the text analysis (formative assessment).
- Students work and learn in groups. Receiving feedback from their peers, learning from each other and comparing to each other may enhance the learning process. Visualizations demonstrating one's own progress as compared to the learning group provides valuable information for organizing and planning the portfolio creation process.
- Teachers can take part in these formative processes, but also conduct summative assessment which aims to be complete and objective.

E-portfolio-based examinations may comprise both the e-portfolio as a collection of digital artifacts and an oral examination for personally presenting the e-portfolio and discussing its contents: This practice has proved to be successful at the University of Education Weingarten and other universities. AISOP plans to assist teachers in preparing for oral exams by allowing them to navigate e-portfolios by topic, by depth, by comparison and by nature of content to guide the exam in areas where there is a need to make sure that the knowledge is present and only decide spending time on certain areas.

Finally, another didactic practice is the use of e-portfolios as the sole source of assessment: For example, comparing students to ensure fairness of assessment and viewing e-portfolios through the lens of a decorated concept map can provide more effective and precise grading.

IV. A SYSTEM TO SUPPORT ASSESSMENT OF E-PORTFOLIOS

In the AISOP software architecture, each such process can be performed in a way that is integrated with the Mahara e-portfolios' authoring and sharing platform. The processes of this architecture are described in figure 1: It shows the flow of information for the service phase, where students and teachers use the system to obtain visualizations: this usage is depicted on the outer circle of the graphics with the interpretation and adjustment process allowing an iterative e-portfolio creation process. In the same way, teachers use the dashboard as an interactive support tool for reviewing and grading e-portfolios.

This figure also shows the development phase where teachers make learning materials available and connect other sources of information such as encyclopedias or textbooks. This set of materials enters the NLP extraction and training process so that classifiers can be built. This is depicted in the inner part of the graphics. It is important to note that the result of the text-classification is in the form of

a dashboard which is not for automatic grading but to give visual clues about the content of the e-portfolio.

The next sections describe this development phase, especially the extraction, annotation and training steps in more detail.

V. BUILDING THE TEXT CLASSIFIER

In order to identify the core concepts covered in an e-portfolio (see evaluation criteria in section III.A), we aim at building a custom text classifier. This classifier should identify concepts matching the course-specific topics (learning content) as described by the teacher. The topics and their relationships are represented in a concept map, thus, making the relevant topics available for further analysis and visualization. In the following, the process of building the text classifier is described step by step: starting from collecting suitable data sources, selecting relevant labels and annotating the dataset, to the training and refinement of the classifier and evaluating the classification results.

A. Data Collection for the Training

In a first step, we collected suitable text documents for annotation and the initial training of the model. For this purpose, text from both lecture slides and nine e-portfolios were extracted with the consent of the students and the teacher. In order to facilitate the text extraction process, we developed two software tools for copying coherent text blocks from the documents suggesting suitable labels for the blocks, as well as formatting and storing the output for further processing. Using these tools, large documents were broken down into chunks of less than 200 words. As a result of the data collection process, two data sets with a total of 1600 chunks were created from the slides and the e-portfolios.

B. Providing Input Data to the Training

For the training of the custom text classifier, an appropriate input data set is required: A set of texts, each annotated with one or more classification labels.

Based on the course-specific concept map, the main topics of the learning units were identified and used as the set of labels for the annotation. For the course "Fundamental Concepts of Computer Science", for example, this resulted in a list of 19 "central" labels.

In our first experiment, we have used this limited set of labels, but know that involving larger sets will require us to work in a hierarchy: By extracting a hierarchical structure from the topic maps, we can annotate the training data set with more labels as recommended in [17] and thus obtain a refined model. This is inline with our approach to share the list of topics as graphical concept map and SKOS terminology [18].

The user interface for carrying out the training is a simple presentation of the text block that is subject to classification and a choice of the available labels; the result is an annotated set of texts. In the annotation tool Prodigy [17], this functionality is offered as the so-called *manual* recipe because each topic has to be chosen manually for each text of the set. The user interface is realized as an input form embedded in a very simple web page. It allows

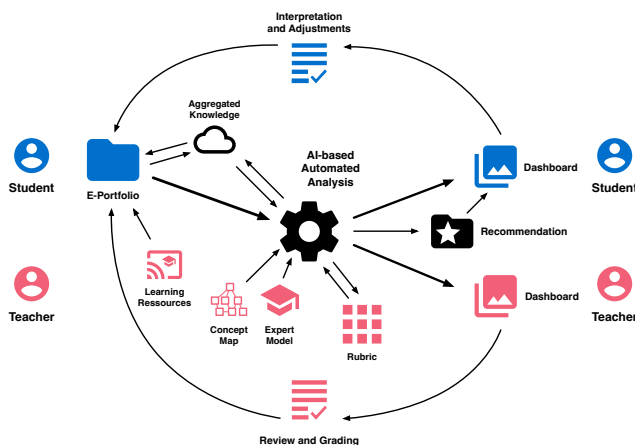


Figure 1: The processes: Left are the input to the automatic analysis system, right are the services it provides to teachers and students.

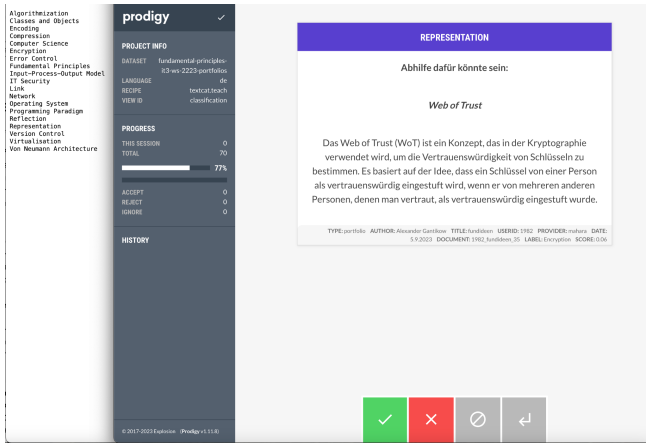


Figure 2: An annotation step in Prodigy’s teach recipe: After a first training, the tool asks to confirm a potential classification for which it has the least evidence.

even non-technical users to participate in the annotation process. The process also allows texts to be *ignored* and passed on to other annotators to work upon. After the initial annotation round, the process was concluded by two subsequent reviews by subject matter experts, each with different competencies.

As a next step, Prodigy offers several refinement workflows where a training pipeline is involved: The training pipeline applies NLP transformations and models as well as corpora of general interest such as BERT [6] on the underlying training data. It analyzes text inputs and classifies them with some degree of certainty. One of the refinement workflows is called *teach*. In this workflow, Prodigy selects and presents text examples that were classified with a very low degree of certainty. The annotator reviews the proposed classifications and corrects them, if necessary. Figure 2 illustrates this type of annotation activity.

E	#	LOSS TOK2VEC	LOSS TEXTC...	CATS_SCORE	SPEED	SCORE
0	0	0.00	0.25	47.41	3543.74	0.47
3	1000	0.00	38.17	82.97	3827.73	0.83
14	2000	0.00	3.60	84.06	4206.53	0.84
39	3000	0.00	1.00	84.07	3730.80	0.84
65	4000	0.00	0.67	84.12	3955.11	0.84
91	5000	0.00	0.62	83.94	3894.79	0.84
117	6000	0.00	0.60	83.94	3967.37	0.84
143	7000	0.00	0.56	83.92	4127.00	0.84
169	8000	0.00	0.59	83.94	3827.35	0.84
195	9000	0.00	0.55	83.97	4239.65	0.84

Figure 3: cross-validation result.

Similarly, Prodigy allows several annotation data sets to be reviewed for conflicts so that consensus can be reached among experts. So far, our data set contains 1222 annotations and the training was done with a claim of “completeness” of 92%. During the process, we encountered many challenges which we report in section VII.

C. Evaluating the Training results

When the annotated data set is ready, a model can be trained using state-of-the-art tools. This model is the core of the classifier. The library spaCy [19] has been used to create a trained model and to evaluate it. In the following, we present the main evaluation methods: an inner one using cross-validation and an external one using actual tools offering a visualization.

A simple method to evaluate the quality of an annotated corpus is to use *cross-validation*: the training is done on a random part of the input data and the rest of the data is used to evaluate the quality of the training result. spaCy supports this process automatically and produces an output as in figure 3 which shows that there may be enough annotated data as the sub-sampling growth barely impacts the score.

D. Experimenting the classifier in a visualization

Visualizing the classification results obtained by the text classifier can now leverage its functionality and show its potential utility. In the AISOP project, we have developed a web service that is integrated in the Mahara e-portfolio authoring system, one of the classical tools in the e-portfolio community. The service allows users to transfer their

Figure 4: An example navigation process resulting of the classification.

portfolio to the AISOP analysis server, export the portfolio as HTML, enrich it by text classification information, and display the result as an interactive visualization. The visualization shows the concept map next to the HTML content of the e-portfolio. The map highlights the topics that are covered in the currently displayed part of the portfolio and allows for quick navigation through the portfolio by choosing a topic in the map. This enables the user to perceive the topic coverage and supports thematic navigation through the e-portfolio. An illustration of the first prototype is provided in figure 4.

VI. LESSONS LEARNED IN THE CREATION PROCESS

The first lesson-learned is that *a classification “per se” should not be thought of as the goal, but a classification for a given purpose*: Our experts steadily asked what the purpose of a classification can be so that they can resolve ambiguities, rest assured that complementary classifications can be done, or be informed on the effects of a bad classification. In the current case, it means missing or wrong colorings in the concept map visualization. While other systems may make a more operational process of the classification result (for example [14] where some works display a coverage numeric mark expressing a percentage of the topics found), the fact that a visualization is the application of the classifier limits the impact of classification errors.

The second lesson-learned is that *the topic structure must be considered as a whole*. It has been repeatedly observed that some texts could be classified in several possible topics, if each of them is considered in isolation, but are to be classified more strictly if the complete set of classification is considered. An example such text is about *media-types* (expressing MP3 streams as examples): While this text has a good overlap with the topic of *error-control* (and the teach recipe described in the annotated process suggested us this topic), the fact that *data representation* exists as a topic makes it a much more preferable topic and the suggestion can be rejected.

This lesson-learned has a non-trivial consequence: *Adding new topics may endanger all annotations* as each of them may be preferable to the existing annotations. Only persons having a global knowledge of all the content can decide whether a complete round of annotation is needed again.

The third lesson-learned is that *establishing the process depends on many factors, several of which can only be observed while actually practicing it*. For example, experiencing the annotation process has brought us the importance of UI aspects such as keyboard shortcuts. As another example, some of the quantitative indicators, e.g. the completeness in the *teach* process or the precision in cross-validation seem to provide useful comparison stabs. However, these do not necessarily become effective in the application of the classifier in the visualization. They only give a hint and hope.

Finally, the fourth lesson learned is that *the result of the annotation process is specific to the educational contexts it was built for*. Different course contents will likely give

different annotations and very probably different topic lists. For this reason, while we intend to share the annotated texts so that others can apply them in their training pipeline, we are sure that only IT-education at the beginner level of university will be a possible field of application.

VII. CONCLUSION

In this paper, we presented an approach and first results from a research initiative focused on the automated AI-based analysis of e-portfolios, targeted to provide teachers with tools and dashboards to reduce work related to the assessment of such digital learning artifacts.

We presented general steps and processes as well as indicators used to perform an extensive analysis of e-portfolios and a corresponding system architecture.

To our knowledge, the taken approach has been little investigated so far, with [20] being among the rare studies that go beyond fully automatic assessment using NLP. In particular, the design of visualizations and dashboards fed by AI-based analytics techniques to analyze e-portfolios have not been explored. The concrete nature of e-portfolios opens possibilities to better understand, depict and communicate patterns in documented students’ learning processes. This not only eases assessment, but will also likely enrich students’ vision, leaving them tangible traces of their learning. These traces form the carrier of a dialogue where the feedback provided by visualizations help students affirm their knowledge in a visible way.

This paper has documented the elaboration process to create AI artifacts that form an essential part of the dashboard’s visualizations: The text classifiers which enable each paragraph of a portfolio to be categorized make it possible to visually represent the topics and navigate based on them.

The project’s developments are done in an open source approach: They are made for open-source environments, include open-source web-services², and will include open-content corpora such as the results of the annotations. Standardized formats such as SKOS (for the hierarchical concept-maps), JSONL for the annotations, or *data-attributes* for the classification result are used so that other classification tools can be leveraged.

An amount of open questions remain which we intend to tackle in and after the project:

The source content has, so far, been restricted to content produced within the university: slides of the courses and pre-existing e-portfolios. They shall be released under an open-content license. Datasets of annotated Wikipedia pages will be released under the same license as the pages. However, do we have a possibility to share the annotation results after having extracted and annotated books protected by copyright? This is probably not shareable as it contains full sentences of complete chapters. It is not even clear if the model deduced from such annotation is independent in license. The importance of this problem can be measured by the amount of textbooks used in lectures in the university environments.

² The software is available from <https://gitlab.com/aisop/>.

The assessment situations are not yet completely explored. We expect to discover and be informed of new methods of assessment provided by new visualizations. To this end, we are investigating the use of rapid prototyping tools based on the Tableau software [21]. It will bring us the possibility of combining classification and more traditional analysis results (e.g. amounts of words, links or codes, semantic density or the number of references). We expect, moreover, to be able to productively apply comparisons of e-portfolios of several students or of a student across several periods.

Finally, other NLP or other AI-based routines are likely to provide us with new information on the texts. We expect that if privacy barriers are lowered, the use of cloud-based large models can provide richer information; for example optical character recognition services may recognize pieces of codes in pictures and objects in photographs. Similarly, large language models may provide more connections to concepts or a more detailed analysis.

ACKNOWLEDGMENT

This research was partially funded by the grant 16DHBKI015 (AISOP) of the German Federal Ministry of Research and Education. This paper is realized in memoriam of Thierry Declerck who contributed to the project before suddenly passing away in the summer 2023.

REFERENCES

1. A. Totter and C. Wyss, "Opportunities and challenges of e-portfolios in teacher education. Lessons learnt," *Research on Education and Media*, vol. 11, no. 1, pp. 69–75, Jun. 2019, doi: [10.2478/rem-2019-0010](https://doi.org/10.2478/rem-2019-0010).
2. S. Ravet. e-portfolio for a learning society. eLearning Conference, Brussels, 2005.
3. W. Müller, S. Rebholz, and P. Libbrecht, "Automatic Inspection of E-Portfolios for Improving Formative and Summative Assessment," in *Emerging Technologies for Education*, vol. 10108, T.-T. Wu, R. Gennari, Y.-M. Huang, H. Xie, and Y. Cao, Eds., in Lecture Notes in Computer Science, vol. 10108, Cham: Springer International Publishing, 2017, pp. 480–489. doi: [10.1007/978-3-319-52836-6_51](https://doi.org/10.1007/978-3-319-52836-6_51).
4. D. Wiliam, "The role of formative assessment in effective learning environments," *Educational research and innovation*, pp. 135–159, 2010.
5. R. Pappagari, P. Želasko, J. Villalba, Y. Carmiel, and N. Dehak, "Hierarchical Transformers for Long Document Classification." arXiv, Oct. 23, 2019. Accessed: Oct. 21, 2023.
6. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423).
7. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Conference on Empirical Methods in Natural Language Processing*, 2019.
8. K. Williamson and R. Kizilcec, "A Review of Learning Analytics Dashboard Research in Higher Education: Implications for Justice, Equity, Diversity, and Inclusion," in *LAK22: 12th International Learning Analytics and Knowledge Conference*, in LAK22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 260–270. doi: [10.1145/3506860.3506900](https://doi.org/10.1145/3506860.3506900).
9. G. Nieto, K. Kitto, S. Buckingham Shum, and R. Martinez-Maldonado, "Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling," Mar. 2022, pp. 219–229. doi: [10.1145/3506860.3506895](https://doi.org/10.1145/3506860.3506895).
10. C. Tominski and H. Schumann, *Interactive Visual Data Analysis*. in AK Peters Visualization Series. CRC Press, 2020. doi: [10.1201/9781315152707](https://doi.org/10.1201/9781315152707).
11. S. Vincent-Lancrin and R. van der Vlies, "Trustworthy artificial intelligence (AI) in education," no. 218, 2020, doi: <https://doi.org/https://doi.org/10.1787/a6c90fa9-en>.
12. J. Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation*. USA: W. H. Freeman & Co., 1976.
13. P. Rabardel and G. Bourmaud, "From computer to instrument system: a developmental perspective," *Interacting with Computers*, vol. 15, no. 5, pp. 665–691, Oct. 2003, doi: [10.1016/S0953-5438\(03\)00058-4](https://doi.org/10.1016/S0953-5438(03)00058-4).
14. O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," *Int J Educ Technol High Educ*, vol. 16, no. 1, p. 39, Dec. 2019, doi: [10.1186/s41239-019-0171-0](https://doi.org/10.1186/s41239-019-0171-0).
15. M. Sharples, *How We Write: Writing as Creative Design*. Routledge, 1999.
16. Catalyst inc and community, Mahara ePortfolio System, webpage consulted 2023-10-21.
17. Explosion AI GmbH, Prodigy, Radically efficient machine teaching. An annotation tool powered by active learning. Web page <https://prodi.gy/> consulted 2023-10-20.
18. A. Miles, S. Bechhofer. SKOS Simple Knowledge Organization System. W3C standard, <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>.
19. Explosion AI GmbH and community, spaCy, Industrial-strength natural language processing. Web-page <https://spacy.io/> consulted 2023-10-20.
20. A. Smithies, I. Braidman, A. Berlanga, D. Haley, and F. Wild, "Using language technologies to support individual formative feedback," in *9th European Conference on eLearning 2010, ECEL 2010|Eur. Conf. eLearn., ECEL*, United Kingdom: Academic Conferences and Publishing International Ltd, Nov. 2010, pp. 570–577.
21. Salesforce, Inc., Tableau Desktop, Insights at the speed of business. Web page <https://www.tableau.com/products/desktop> consulted 2023-10-20.