

# Improving E-Learning Material Quality With the Aid of Deep Learning and Workflow Management

Melinda Pap, László Zsolt Nagy, Dániel Fekete

Eszterházy Károly University, Hungary

([pap.melinda,nagy.laszlo.zsolt,fekete.daniel](mailto:pap.melinda,nagy.laszlo.zsolt,fekete.daniel@uni-eszterhazy.hu))@uni-eszterhazy.hu

## Abstract

E-learning systems are available since decades, they became widely accessible and offered by most colleges and universities. It has been recognized in recent years that the future of e-learning lies in crowd-sourced learning, where multiple users can contribute to the global knowledge base. [6, 15] However, in such systems, the question of quality arises.

In the development of our e-learning system, one aim was to assure the high quality of the created contents. To achieve this goal, we have incorporated a workflow-engine, that is commonly used in business intelligence software, to manage business processes. We created a system that provides user interfaces for the definitions of state transitions and the corresponding user permissions. Thus, enabling versatile validation processes for the different types of materials, so that a newly created content can become an approved and verified learning object.

When we are talking about e-learning quality, we must not forget that this includes the aspects of accessibility. When we enable users to create new content in our system, we want to be sure that it is accessible. On the other hand, we do not wish to put the burden of image description entirely on the users. Therefore, we have incorporated a pre-trained “Show and tell” model [24] to generate natural sentences describing an image, to annotate our image database. This not only improves the accessibility of the system, but also makes it possible for our *Elasticsearch* search engine to provide relevant search results, even when it comes to images. Deep learning emerged recently as a universal tool for machine learning tasks. The used “Show and tell” model is one of such, a combination of a convolutional neural network (CNN) and a long short-term memory (LSTM) network.

In this paper we present the architecture of our e-learning system, highlighting the concepts towards quality assurance.

*Keywords:* E-learning, workflow, deep learning, crowd-sourced learning, image captioning

## 1. Introduction

E-learning systems has been around since decades. By today, they became the part of the everyday life of colleges and universities. It has started as the adaptation of the traditional distance learning model, such that online courses were organized in a standard way, following a specified curriculum in a predetermined pace. With the advances of web-technologies, new trends evolved in the field of e-learning as well. State of the art systems enable the user to conduct an individual learning procedure. Moreover, learners can communicate, collaborate and share content with each-other over the e-learning software.

In this study, we focus on a Learning Management System (LMS), more precisely a Learning Content Management System (LCMS). An LMS is used for delivering, tracking and managing training and education. An LCMS is more about the creation of contents and their sharing. These can be hosted in an LMS system as learning objects. It has been pointed out in the literature, that the future of such content creation and management and the development of learning objects lies in crowd-sourcing and machine learning. [18, 15]

Crowd-sourcing makes it possible to utilize the intelligence and wisdom of large groups of people in solving problems and creating and validating content. Despite all the advantages, it raises new challenges when it comes to quality control. This is due to the fact that the crowd can be composed of people with varying skills, intellects and objectives [5].

We have recently entered the era of Big Data, which means that way more data is produced than what can be processed by human operators. This is also true in the case of LCMS applications. Machine learning offers a wide range of tools that can help in this matter, especially when it comes to automatically annotating data. This not only aids the searchability of learning objects, but also increases their accessibility. For instance, by helping visually impaired people better understand the content of images. Inclusive education is also gaining attention nowadays and has been addressed in the literature [8, 21].

The goal of this paper is to propose an architectural design for an LCMS system that incorporates building blocks and methodologies responsible for the quality control and accessibility of learning objects.

## 2. Motivation

In the past years, we developed a national e-learning platform called *Ekernel* in Hungary for elementary schools. In this software, besides the basic LMS func-

functionalities, it is possible to create content in the form of online books, interactive exercises, uploaded media, etc. Furthermore, users can search for all types of contents and share them with each-other.

To enable search engines to find relevant information, we must focus our attention on the annotation of our learning objects, even when they are in the form of images, videos or soundtracks. This not only helps the search, but if done well, helps the accessibility of our educational portal. Since recent advances in the field of deep neural networks, the quality of automated image caption generation increased considerably.[2] In this paper we aim to show how such a tool can be infused in an e-learning architecture.

As we opened towards the direction of crowd-sourced content creation, we soon realized that we need to set up a sophisticated protocol to publish the produced data. Our primary target audience is elementary school students, therefore it is of high importance to filter the content according to quality and relevance. We had observed the possibilities and found that we can take examples from the industry where quality control normally plays a crucial role. This led to the decision to adapt a Workflow Management System (WfMS) [7] to our needs. Such a WfMS enables the description of validation stages, such as professional and linguistic proofreading, plagiarism check etc. In the Ekernel system, we do not only have standalone learning objects, but we allow the creation of lecture notes and complete books. In these cases a validation hierarchy is required, when the check of subsequent chapters and lessons is the prerequisite of the final publication. In the proceeding sections, we aim to give an overview of how a WfMS can be integrated in an LCMS architecture in a customizable way.

## 2.1. Hypothesis

In designing the Ekernel system, we aim to give an easily understandable, accessible and intuitive user interface. We put high emphasis on the search capabilities of our system: the global search page is available from the main menu and in addition, we included additional search fields inside all pages of the interactive books.

To increase the user satisfaction with usability, all the created pages are responsive, thus all content is accessible from mobile devices. This is crucial, knowing that it is highly probable that the students will work on tablets or smartphones during class. Therefore, we expect that students will find the webpage easy to use and easy to understand. We expect that the teachers will be more confident in using this tool during course work, and will feel motivated to contribute to the global knowledge base with their own creations. We also suppose that both teachers and students will find the search functionalities easy to use and will be able to find relevant contents in our Ekernel system. We expect that the quantity and quality of learning objects will make the users more motivated to use our software.

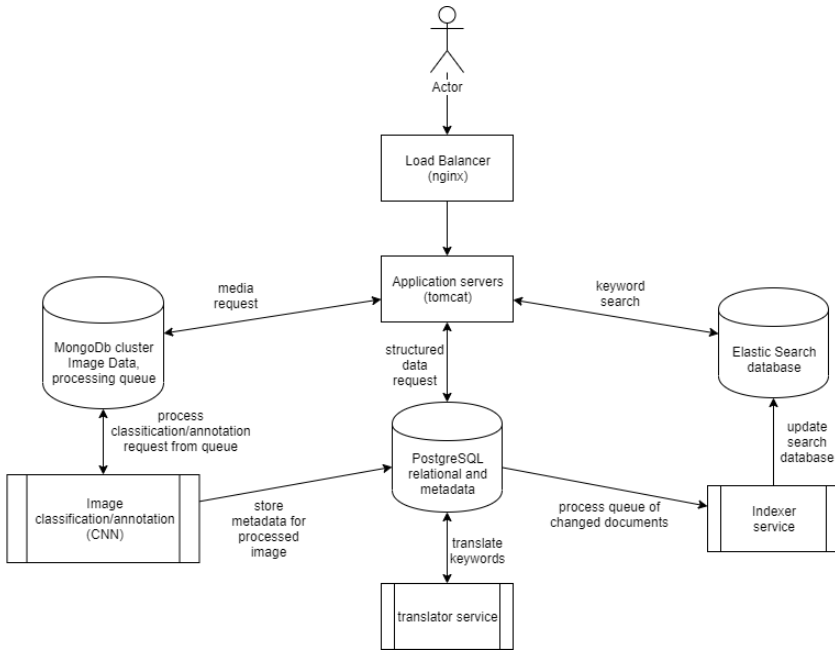


Figure 1: Ekernel architectural design

### 3. Architectural design

In Fig. 1. the core architecture of our Ekernel system is illustrated. In designing the system, we put emphasis on using the most recent technologies that came to existence by the innovations of Big Data and semantic web. Due to the versatility of data types used in an LCMS application, we decided to use a scalable noSQL database besides our relational database, that can support large scale and high-concurrency applications with heterogeneous data.[12, 4] The used MongoDB [3] noSQL database is used to store documents and images.

To follow the trends of semantic web, we have incorporated an Elasticsearch engine, that is a distributed, scalable, real-time search and analysis engine [10]. It comes with a built-in database to store the indexes. With its aid, the software is capable of indexing learning objects, not only the text-based ones, but also those that have corresponding meta-data given (e.g. keywords).

Text based data, and the metadata of all objects is stored in a PostgreSQL relational database.

#### 3.1. Image annotation

As shown in Fig. 1. we created a service over our MongoDB, for image annotation. This service is for two main purpose, one is two give keywords for an image, and

one is to describe it in a complete, natural language sentence. We have found that several publications have addressed the problem of image captioning before [27, 9, 2, 14, 25, 26, 1]. Due to the lack of computational power and training database with ground truth, we decide to use pre-trained models for our tasks.

We chose to use two pre-trained model, one for keyword generation and one for the whole sentence image captioning. For the former, we selected the InceptionV3 model [22]. This Convolutional Neural Network (CNN) was trained on the ILSVRC-2012-CLS image classification dataset. For the latter, we chose a “Show and tell” model, also called Neural Image Caption (NIC) provided by the authors of [24]. This method uses a CNN for image representations and a Long Short-Term Memory (LSTM) for generating the final image captions. The CNN is used as an encoder, then the output of its last hidden layer is used as the input of the LSTM decoder. The Flickr 8K/30K, and the MS COCO public image databases were used in the process of training.

We created a service over our MongoDB that processes the uploaded images and creates the captions for them. This data is stored in the PostgreSQL database, and is indexed by the Elasticsearch engine in the background.

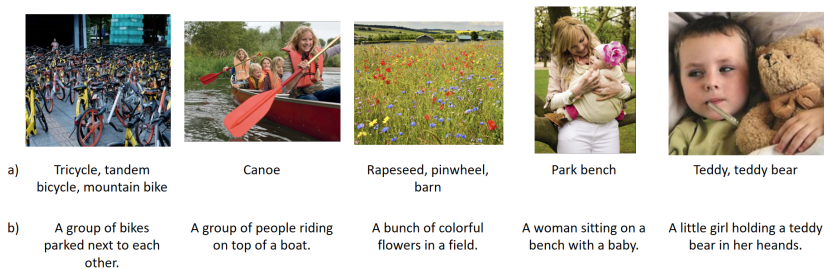


Figure 2: Results of the automated image annotations. Row a) shows the keywords returned by the InceptionV3. Row b) shows the results of the NIC image captioning.

Fig. 2. shows the results of image annotations. We can see that the whole sentence captions give sufficient description of an image. In many cases, we found that the captions are more generic and keywords give more insight to an image. For instance, the captioning returns “A group of people riding on top of a boat” and the keyword creating classification gives the exact type of boat that they are ridding: “canoe”.

### 3.2. WfMS

A workflow management system (WfMS) is a software system that helps the setup, execution and monitoring of processes and tasks. The goal of such a system is to increase productivity, agility, improve the information exchange within an organization and assure high quality.[7, 23, 19] Since we do not only operate an LCMS system, but we provide LMS functionalities, we did not limit our workflow design

to document processing. The creation and announcement of a course is organized via a workflow as well as the enrollment, evaluation and graduation of a student in our system.

During the design of a workflow, one has to take into consideration the followings: 1. execution order of activities, 2. permission-based availability of activities, and 3. permission-based availability of information. In our scenario, we found that we need a routing type system, that conducts the routing of the flow of information and documents.

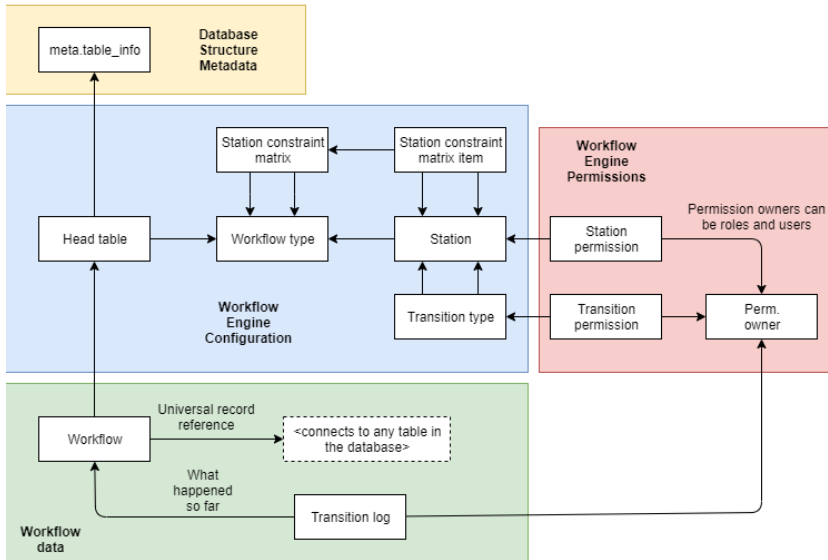


Figure 3: The database tables used for our flexible WfMS implementation in the Ekernel architecture

The database architecture of the proposed highly configurable WfMS is presented in Fig. 3. We start with the creation of a meta table (`meta.table info`) describing all the existing tables in our PostgreSQL database and assigning id-s to them. We store the types of possible workflows in the `workflow type` table. Users with admin permissions can define separate workflows for documents, images, book chapters but even for the admission of students to an online course in the system. For each workflow type, one can create stations and define transitions connecting them. Such stations can be: under editing, under professional proofreading, rejected etc. Both stations and transitions can be constrained, such that only users in a specified permission group can access the data when it is in a certain state or can perform a transition to a new state. All of the executed transitions are stored in the database, so one can trace back the history of any item in the workflow queue.

In our system, we enable the users to visualize a specific the workflow as directed graph and configure them via an extensible user interface. Fig. 4 show parts of

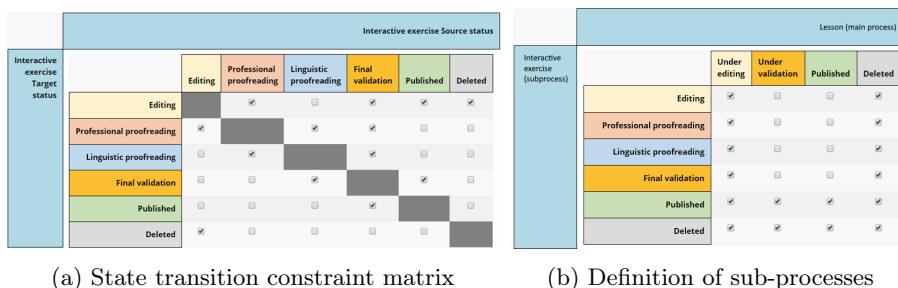


Figure 4: User interfaces for defining the states and transitions in the Ekernel system. This particular workflow is for the interactive exercises.

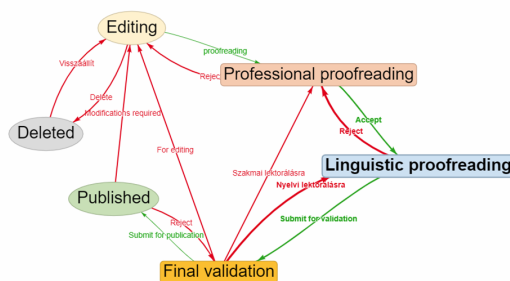


Figure 5: Graph visualization interface of state transitions

the user interfaces. One can define the possible states, then configure the possible transitions via a matrix representation (see Fig. 4a) and set the corresponding privileges. It is possible to create dependent workflows, when the state of one item defines the possible states and transitions of a related item. For instance, if we have a lesson that contains an interactive exercise, we can specify that the lesson can only move to a published state, when the contained exercises are published as well. This particular setup is shown in Fig. 4b.

## 4. Methodology

The Ekernel system was first tested from January to June, 2019 by 310 Hungarian public schools. During this test, 446 teachers and 3706 fifth grade students used our Ekernel system for a period of 5 months. At the end of the test run, we offered the participants to fill out a voluntary questionnaire that addressed their satisfaction with our created LCMS.

The basis of the construction of our questionnaire was the theoretical framework of user experience. This research framework distinguishes between perceived ergonomic quality, perceived hedonic quality and perceived attractiveness of a prod-

uct. [13] The ergonomic quality focuses on the goal and task oriented aspects of product design. High ergonomic quality enables the user to reach his or her goals with efficiency and effectiveness. The focus of hedonic quality is on the non-task oriented quality aspects of a software product, for example the originality of the design or the beauty of the user interface.

In the context of the questionnaire it is necessary to consider both pragmatic and hedonic aspects if we want to measure how satisfied users are with a our system. It should allow the users to express feelings, impressions, and attitudes that arise when using our product.[17, 20]

The questionnaire consisted of 18 questions for both students and teachers for our user experience assessment. Out of these, 12-12 used a 5 level, and 5-5 used 3-level Likert scale [16], where the scale extended from the “not at all” to the “very much”. To be precise, since we were working with students aged between 10-11, we decided to use Smiley-faced Likert scale [11] for the 12 questions with five levels (illustrated in Fig. 6 ). For the additional 5 questions we used textual answer representation. The questionnaire was filled out online, in our Ekernel system, simultaneously during classwork, supervised by the teachers.



Figure 6: Five-level Smiley-faced Likert scale used in our survey. The image title was set to corresponding text - in the scale: “not at all” to “very much” - to be shown on mouse hover.

The questions in the case of teachers can be organized in 3 main categories:

1. **satisfaction and confidence** (e.g. “*How confident did you feel during teaching?*”)
2. **perceived student satisfaction** (e.g. “*Did the students understand the material better with the interactive books than with the paper ones?*” ,
3. **platform usability** (e.g. “*Did you find what you were searching for easily? Could you navigate where you wanted with ease?*”)

The students had another set of questions that we grouped to the following 3 categories:

1. **understanding** (e.g. “*Did you understand the material better with the interactive book than with the paper books?*”)
2. **motivation and mood**, (e.g. “*Did you find the materials interesting and exciting?*”)
3. **platform usability** (e.g. “*Did you find what you were searching for easily?*”)



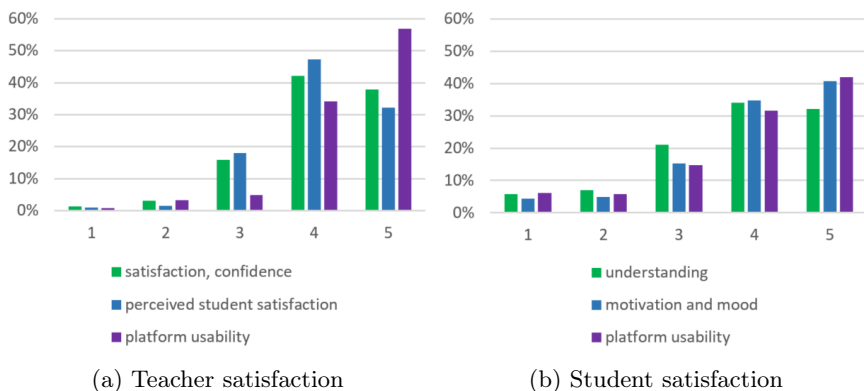


Figure 7: The results of the 5 level Likert scale questionnaire, where the level 1 stands for the “not at all” and level 5. is for “very much”.

To be able to filter out those who gave random answers, we included opposing questions as well, one in the 5-level part and its pair in the textual answer part.

## 5. Empirical findings

Until the time of our final survey, our editorial community of 250 users has written and validated 83 books, containing 4 568 lessons. Furthermore, 13 490 interactive exercises were created and 147 612 images were uploaded to our servers.

Overall 41 teachers and 258 students participated in our anonymous survey. The answers of 8 participant were filtered out due to inconsistency. The results of the survey are presented in Fig. 7. It shows that 66% of the students claimed that they had a better understanding of the topics with the use of the online, validated materials, than with printed books. Furthermore, 74% of the students and 91% of the teachers answered to the platform usability related questions positively.

As a hypothesis, we stated that teachers will be motivated to participate in content creation. To the question “*How much would you like to contribute to the global knowledge base with your own content?*” 59% of them responded positively.

In the textual questions part, we asked the participants whether they find this software better than the digital tools they have used so far. In this case, the possible answers were: “no”, “the same” and “yes”. 56% of the teachers and 61% of the students answered “yes”.

## 6. Future work

In the future, we plan to extend our editorial user-base. We would like to allow all the registered users to submit their creations (books, lesson plans, interactive exercises, etc.) for a reviewing process.

The image caption also leaves some things to be desired, e.g. higher classification accuracy has been reported in the literature recently. Fig. 8. shows to examples of complex scenes, where the annotations have low accuracy.

The currently used image annotating deep learning models can only operate on natural images, however our image database contains a great number of illustrations as well, thus, we continue our search for networks that can handle such data. Furthermore, we plan to use our image database to fine-tune a classification model. We would also like to continue our progress towards accessibility and create machine learning supported video sub-titling services.



<b>Keywords:</b>	ear, spike, capitulum, corn	slot, one-armed bandit
<b>Caption:</b>	A young boy sitting on top of a teddy bear.	A group of people sitting at a table in a library

Figure 8: Image annotations with low accuracy.

## 7. Acknowledgements

The research was supported by the grant EFOP-3.6.1-16-2016-00001 (“Complex improvement of research capacities and services at Eszterhazy Karoly University”).

## References

- [1] ASADI, A. I., AND SAFABAKHSH, R. A deep decoder structure based on wordembedding regression for an encoder-decoder based model for image captioning. *ArXiv abs/1906.12188* (2019).
- [2] BAI, S., AND AN, S. A survey on automatic image caption generation. *Neurocomputing* 311 (2018), 291–304.
- [3] BANKER, K. *MongoDB in action*. Manning Publications Co., 2011.
- [4] CATTELL, R. Scalable sql and nosql data stores. *SIGMOD Rec.* 39, 4 (May 2011), 12–27.
- [5] DANIEL, F., KUCHERBAEV, P., CAPPIELLO, C., BENATALLAH, B., AND ALLAHBAKHSH, M. Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions. *ACM Computing Surveys (CSUR)* 51, 1 (2018), 1–40.
- [6] DOWNES, S. E-learning 2.0. *Elearn magazine 2005*, 10 (2005), 1.

- [7] DU, W., AND ELMAGARMID, A. Workflow management: State of the art vs. state of the products. *HP LABORATORIES TECHNICAL REPORT HPL* (1997).
- [8] FERRETTI, S., MIRRI, S., MURATORI, L. A., ROCCETTI, M., AND SALOMONI, P. E-learning 2.0: you are we-lcome! In *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)* (2008), ACM, pp. 116–125.
- [9] FU, K., JIN, J., CUI, R., SHA, F., AND ZHANG, C. Aligning where to see and what to tell: Image captioning with region-based attention and scene-specific contexts. *IEEE transactions on pattern analysis and machine intelligence* 39, 12 (2016), 2321–2334.
- [10] GORMLEY, C., AND TONG, Z. *Elasticsearch: the definitive guide: a distributed real-time search and analytics engine.* " O'Reilly Media, Inc.", 2015.
- [11] HALL, L., HUME, C., AND TAZZYMAN, S. Five degrees of happiness: Effective smiley face likert scales for evaluating with children. In *Proceedings of the The 15th International Conference on Interaction Design and Children* (2016), pp. 311–321.
- [12] HAN, J., HAIHONG, E., LE, G., AND DU, J. Survey on nosql database. In *2011 6th international conference on pervasive computing and applications* (2011), IEEE, pp. 363–366.
- [13] HASSENZAHL, M. The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction* 13, 4 (2001), 481–499.
- [14] HOSSAIN, M. Z., SOHEL, F., SHIRATUDDIN, M. F., AND LAGA, H. A comprehensive survey of deep learning for image captioning. *ACM Comput. Surv.* 51, 6 (Feb. 2019).
- [15] HUSSAIN, F. E-learning 3.0= e-learning 2.0+ web 3.0?. *International Association for Development of the Information Society* (2012).
- [16] JAMIESON, S., ET AL. Likert scales: how to (ab) use them. *Medical education* 38, 12 (2004), 1217–1218.
- [17] LAUGWITZ, B., HELD, T., AND SCHREPP, M. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group* (2008), Springer, pp. 63–76.
- [18] MORRIS, R. D. Web 3.0: Implications for online learning. *TechTrends* 55, 1 (2011), 42–46.
- [19] PESIC, M. *Constraint-based workflow management systems : shifting control to users.* PhD thesis, Department of Industrial Engineering and Innovation Sciences, 2008. Proefschrift.
- [20] RAUSCHENBERGER, M., SCHREPP, M., PÉREZ COTA, M., OLSCHNER, S., AND THOMASCHIEWSKI, J. Efficient measurement of the user experience of interactive products. how to use the user experience questionnaire (ueq). example: Spanish language version.
- [21] ROGERS-SHAW, C., CARR-CHELLMAN, D. J., AND CHOI, J. Universal design for learning: Guidelines for accessible online instruction. *Adult Learning* 29, 1 (2018), 20–31.
- [22] SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J., AND WOJNA, Z. Rethinking the inception architecture for computer vision. *CoRR abs/1512.00567* (2015).
- [23] VAN DER AALST, W., VAN HEE, K. M., AND VAN HEE, K. *Workflow management: models, methods, and systems.* MIT press, 2004.

- [24] VINYALS, O., TOSHEV, A., BENGIO, S., AND ERHAN, D. Show and tell: Lessons learned from the 2015 mscoco image captioning challenge. *IEEE transactions on pattern analysis and machine intelligence* 39, 4 (2016), 652–663.
- [25] WU, Q., SHEN, C., WANG, P., DICK, A., AND VAN DEN HENGEL, A. Image captioning and visual question answering based on attributes and external knowledge. *IEEE transactions on pattern analysis and machine intelligence* 40, 6 (2017), 1367–1381.
- [26] XU, K., BA, J., KIROS, R., CHO, K., COURVILLE, A., SALAKHUDINOV, R., ZEMEL, R., AND BENGIO, Y. Show, attend and tell: Neural image caption generation with visual attention. In *International conference on machine learning* (2015), pp. 2048–2057.
- [27] YOU, Q., JIN, H., WANG, Z., FANG, C., AND LUO, J. Image captioning with semantic attention. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 4651–4659.