Bridging the skills gap in the Data Science and Internet of Things domains: A Vocational Education and Training Curriculum

Gkamas V.[1], Rigou M.[1], Paraskevas M.[2], Zarouchas T.[1], Perikos I.[1], Vassiliou V.[3], Gueorguiev .I[4], Varbanov P.[4], Sharkov G.[4], Todorova C.[4], Sotiropoulou A.[5]

[1] University of Patras, Department of Computer Engineering & Informatics, Rio, Achaia, Greece

[2] University of Peloponnese, Department of Electrical and Computer Engineering, Patras, Greece

[3] University of Cyprus, Department of Computer Science, Aglantzia, Nicosia, Cyprus

[4] European Software Institute - Center Eastern Europe, Sofia, Bulgaria

[5] Olympic Training & Consulting Ltd, Pyrgos, Greece

Abstract

Data Science and Internet of Things are among the key drivers regarding the skills and competences needed by IT professionals. Given the value that Data Science and Internet of Things have for the EU economy and society, the foreseen skills gap at these two domains, and the rapid and continuous evolution of these technologies, make the skills required by IT professionals increasingly sophisticated, and the need to be constantly updated imperative.

The SEnDIng project aims to address this skills gap, by providing IT professionals with knowledge, skills and competences that meet the current needs of Data Science and Internet of Things industries and are transferable and recognized among European countries. To achieve this goal, SEnDIng has designed and will deliver a learning outcomes-oriented multi-disciplinary VET curriculum. The curriculum covers three areas -Data Science, Internet of Things and Transversal Skills- and is divided in educational modules and learning units at three levels of proficiency: Introductory, Core and Advanced. The modularity of the curriculum permits each trainer to build a separate learning path that serves his/her specific training needs. The VET curriculum for each domain will be delivered in three phases: 100 hours online courses, 20 hours face to face training and 4 months work based learning.

Key Words: Data Science, Internet of Things, Vocational Education and Training, curriculum, online training, work based learning

Introduction

Data Science and Internet of Things (IoT) pose an unprecedented opportunity for all kinds of businesses, whether they are B2B, B2C or B2G. In addition, their value for the EU economy is huge; it is projected that the EU Data Economy will reach €739 billion by 2020 and will represent 4% of the overall European GDP (EC Digital Single Market, 2019), while IoT with a value of €120 billion will solely contribute to an increase of 7 points of European GDP by 2025, through productivity improvement and value distributed to end customers (ATKearney, 2016). These progressively build complex ecosystems of new business models and supply chains, from IoT vendors and system integrators, to the number of industries that exploit the huge volume of data coming from a myriad of devices to extract useful information using Data Science technologies.

However, one of the main barriers preventing the full exploitation of Data Science and IoT potential, is the skills gap observed at both domains. According to predictions, the demand for Data Scientists will increase up to 28% until 2020 (Columbus, 2017), while the 43% of the companies are reporting the lack of appropriate analytical skills as a key challenge (Press, 2015). At same time, the unfilled Data Science positions are estimated at 485,000. In addition, the need for IoT skills is huge, as 68% of businesses still struggle to hire IoT experts (Wright, 2018). These predictions, together with the rapid evolution of Data Science and IoT technologies and their application in many industries make the skills required by IT professionals increasingly sophisticated, and the need to be constantly updated imperative.

Although Higher Education Institutions (HEIs) and Vocational Education and Training (VET)

providers attempt to catch up with this new Data Science and IoT wave by providing relevant academic and vocational training programs, the complexity of the Data Science and IoT scientific domains and applications, the variety of the economic sectors exploiting IoT and data analytics, and the diversity of end-users and technical options available in both fields, result to consensual problems faced by all stakeholders in the value chain of education and training: the IT professionals in their career orientation, the organizations designing training programs at several educational levels and the businesses as recruiters of IT professionals. While in the Data Science field there are some developments in defining the required skills and competences (Manieri et.al., 2015; Mikroyannidis et.al., 2017), the respective developments at the IoT domain are very limited (World Economic Forum, 2016). In addition, although transversal or soft skills are among the key skills of the future employees (José and Serpa, 2018; Institute for the Future, 2011), the current Data Science and IoT training programs are only technical-oriented and do not commonly provide the learners with such skills.

At this paper, we present a multi-disciplinary and learning outcomes-oriented VET curriculum for the training of IT professionals on transversal skills and Data Science and IoT technologies. The curriculum has been designed in the context of the EU project SEnDIng (http://sending-project.eu) which aims to address the skills gap of IT professionals at Data Science and IoT domains. To achieve this goal, SEnDIng has developed and will deliver to IT professionals a VET program using innovative teaching and training methodologies. Furthermore, a reference model has been designed for the vocational skills, e-competences and qualifications for Data Science and IoT experts, which is compliant with the European eCompetence Framework (eCF) and the European Skills, Competences, Qualifications and Occupations (ESCO) classification, ensuring transparency, comparability and transferability between European countries. The curriculum for each domain (Data Science and IoT) will be provided in the form of blended learning combining also work based learning: 100 hours of online self-paced courses, 20 hours of face to face training on transversal skills and 4 months of work based training. The trainees will be certified having successfully completed the three phases of training and a final exam.

The main contribution of SEnDIng VET curriculum is two-fold. In contrast with other training programs that provide only technical skills at the Data Science and/or IoT domains, the SEnDIng curriculum combines technical knowledge and skills with cross-domain transversal skills and competences. In addition, it includes a work based learning phase which commonly lacks from current vocational trainings at Data Science and IoT domains.

The SEnDIng VET curriculum

Curriculum is increasingly seen by stakeholders as a dynamic framework guiding the teaching and learning processes, and as a steering mechanism for quality. According to the European Centre for the Development of Vocational Training (2010), “a curriculum is a normative document (or a collection of documents) setting the framework for planning learning experiences”. Depending on the country, the type of education and training, and the institution, curricula may define, among others, learning outcomes, objectives, contents, place and duration of learning, teaching and assessment methods to a greater or to a lesser extent.

Methodology for curriculum development

The methodology applied for the development of SEnDIng VET curriculum is based on the constructive alignment approach. This approach maximises the conditions for quality learning by ensuring alignment throughout the process, from the forming of learning outcomes, to the choice of teaching methods to assessment. It assumes that when learning objectives, assessment methods, and teaching and learning activities are intentionally aligned, the outcomes of learning are improved substantially (Blumberg, 2009). The process of constructive alignment emphasizes that learners are central to the creation of meaning, and must be provided with opportunities to actively select, and cumulatively construct their own knowledge (Biggs, 1996). According to Biggs, “The fundamental principle of constructive alignment is that a good teaching system aligns teaching method and assessment to the learning activities stated in the objectives, so that all aspects of this system are in accord in supporting appropriate student learning.”

The elements involved in the process of constructively aligning the educational modules are three: (i) defining the learning outcomes, (ii) choosing the learning and teaching methods that can lead to

attainment of outcomes and (iii) assessing the learning outcomes of learners.

The process followed for developing SEnDIng curicullum consists of three stages (see Fig. 1):

- **Define curriculum goals and design learning outcomes**. We adopted the European Parliament Council (2008) recommendations on the establishment of the European Qualifications Framework for lifelong learning, to design the learning outcomes of the curriculum in terms of knowledge, skills and competences. According to these recommendations: (i) **knowledge** is defined as the outcome of the assimilation of information through learning; it is the body of facts, principles, theories and practices that is related to a field of work or study, (ii) **skill** is defined as the ability to apply knowledge and use know-how to complete tasks and solve problems and (iii) **competence** is defined as the proven ability to use knowledge, skills and personal, social and/or methodological abilities, in work or study situations and in professional and personal development.

  One of the main challenges we faced with the definition of Data Science and IoT learning outcomes is the broad spectrum of disciplines that can be included in the curriculum of these technology domains and their accordance with the specific needs of the respective work places that they are applied. For these reasons, the learning outcomes of the curriculum were designed in two phases. At the first phase, we developed the macro-level design of the learning outcomes based on a research of related programs and a survey among companies that currently run Data Science or IoT projects, or are interested to run such projects in the near future. Then at the second phase, we developed the micro-level design of the learning outcomes that is the design of the learning outcomes of each separate educational module of the curriculum.
- **Developing teaching methods and forms of assessment**. We developed the teaching methods and forms of assessment covering the three phases of SEnDIng VET program: online, face to face and work based learning. The assessment methods are in accordance with the learning outcomes of each module, aiming to foster a deep approach to learning.
- **Reviewing and refining the curriculum**. The review cycle is defined as a systematic approach to evaluating, reviewing and revising VET curricula within a specific timeframe which aims to identify gaps and weaknesses with a view to increase their effectiveness and continuously improve VET trainees learning experiences (UNESCO-IBE 2013). Normally it involves several phases including: research and selection; revision and development; implementation; as well as evaluation and monitoring. Reviewing is not a retrospective process, but it is an integral part of the module development, providing information before, during and after the process.

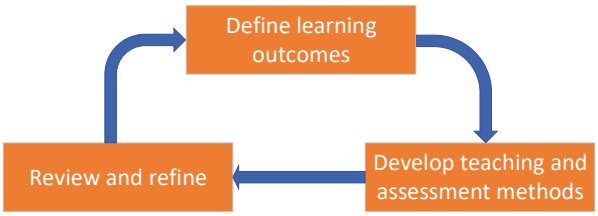

Figure 1: Curriculum development process

Curriculum characteristics and structure

The key characteristics of SEnDIng curriculum are:

- **Multi-disciplinarily**. The modules developed cover both technical knowledge and skills at Data Science and IoT domains as well as transversal skills and competences. The transversal skills aim to build upon academic and experiential learning and to prepare the IT professionals for engaging within the business environment in a creative way, communicating effectively with the internal and external environment of a business and acting in a collaborative way.
- **Modular**. For each domain, the curriculum is separated in educational modules. The modules have been designed by selecting and ordering the types of learning activities that trainees undertake to achieve the learning outcomes of each module.

  Each module is further divided in training units at three levels of proficiency:
  - **Introductory** (I): The educational module is introduced and its most important aspects

are given.
    - **Core** (C): All core aspects, principles and methods of the module are covered in sufficient detail as necessary to apply the knowledge and skills on the job. The learner becomes able to discuss matters with other stakeholders and acquire more knowledge when necessary.
    - **Advanced** (A): Advanced aspects of the module are covered in sufficient detail as necessary to apply the knowledge and skills on the job.

  This permits the learners to create their own learning paths according to their needs. As the learning outcomes aimed by the transversal skills training are horizontal, the educational modules of the transversal skills are not split into the three training units (Introductory, Core and Advanced).
- **Learning outcomes-oriented**. The curriculum give emphasis on what an individual should know, understand and/or be able to do at the end of a learning process. Such curricula consist an effective way to avoid potential mismatches between academia and industry, and furthermore to promote active learning and inclusive training.

Considering the aforementioned curriculum definition, the structure of a training unit includes the following elements:

- **Objectives**. The objectives of the training unit indicate its general direction or orientation in terms of its content.
- **Learning outcomes**. The learning outcomes of the training unit in terms of knowledge, skills and competences.
- **Content**. The content of the training unit.
- **Learning methodologies**. The learning methodologies applied for the delivery of the training unit.
- **Assessment methodologies**. The methodologies applied for the assessment of the learning outcomes.
- **Duration**. The duration of the training unit. The indicative duration of each training unit is as follows: (a) Introduction: 1 to 2 hours, Core: 3 to 10 hours and Advanced: 5 to 10 hours.
- **Pre-requisite**. Pre-requisite knowledge and skills to be able to attend the training unit.

Curriculum Educational Modules

The Data Science educational modules developed are the following (see Fig. 2):

- **Introduction to Data Science (DS-EM1).** This module is an introduction to Data Science: different roles involved in Data Science projects, common Data Science applications and methodologies.
- **Python for Data Science (DS-EM2).** This module aims to provide to the learners knowledge and skills for big data analysis, data visualization and machine learning using Python.
- **Statistics for Data Science (DS-EM3).** This module aims to provide to the learners knowledge and skills for inferential statistical analysis, visualization and machine learning, using the R language and environment for statistical computing and graphics.
- **Storing and Retrieving data (DS-EM4).** This module aims to provide to the learners knowledge and skills to work with Hadoop ecosystem and its applications at storing and processing large volumes of data distributed across commodity servers.
- **Applied Machine Learning (DS-EM5).** This module aims to provide to the learners knowledge and skills to work with machine learning technologies.
- **Data Visualization (DS-EM6).** This module aims to provide to the learners knowledge and skills to build effective and engaging data visualizations.

The IoT educational modules developed are the following (see Fig. 2):

- **Introduction to IoT (IoT-EM1).** This module is an introduction to the IoT: the basic principles of IoT applications, the different roles involved in IoT projects and their duties.
- **IoT Devices (IoT-EM2).** This module introduces the “Things” in the IoT, the different IoT devices (sensors, actuators, peripherals), their electronics, as well as, the different

microcontrollers and how they can interact with the IoT environment.

- **IoT Communication Technologies (IoT-EM3).** This module introduces the plethora of communication protocols and standards that are used for signalling and data exchange in IoT systems in a comprehensive and visual way.
- **Architectural Design and Applications in IoT (IoT-EM4).** This module presents the key aspects of an IoT system architecture (IoT edge devices, gateways) with emphasis on server-side infrastructure solutions (cloud computing service models, deployment models and public cloud providers).
- **IoT Security and Privacy (IoT-EM5).** This module presents common security risks raised in an IoT system and possible measures to create a more secure environment. It aims to create a sense of awareness to the learner of the possible security breaches and how to avoid them by adapting appropriate security measures.
- **IoT Business Value (IoT-EM6).** This module aims to make the learners aware of IoT business value, by presenting the advantages of exploiting the IoT into the business world providing real use cases.

The transversal skills educational modules developed are the following (see Fig. 2):

- **Effective communication and presentation (TS-EM1)**. This module aims to provide to the learners knowledge and skills for effective communication and presentation.
- **Change management (TS-EM2)**. This module aims to provide to the learners knowledge and skills for change management.
- **Team working (TS-EM3)**. This module aims to provide to the learners knowledge and skills for effective team working.
- **Goal setting (TS-EM4)**. This module aims to provide to the learners knowledge and skills for effective goal setting.
- **Creative thinking (TS-EM5)**. This module aims to provide to the learners knowledge and skills for creative thinking.

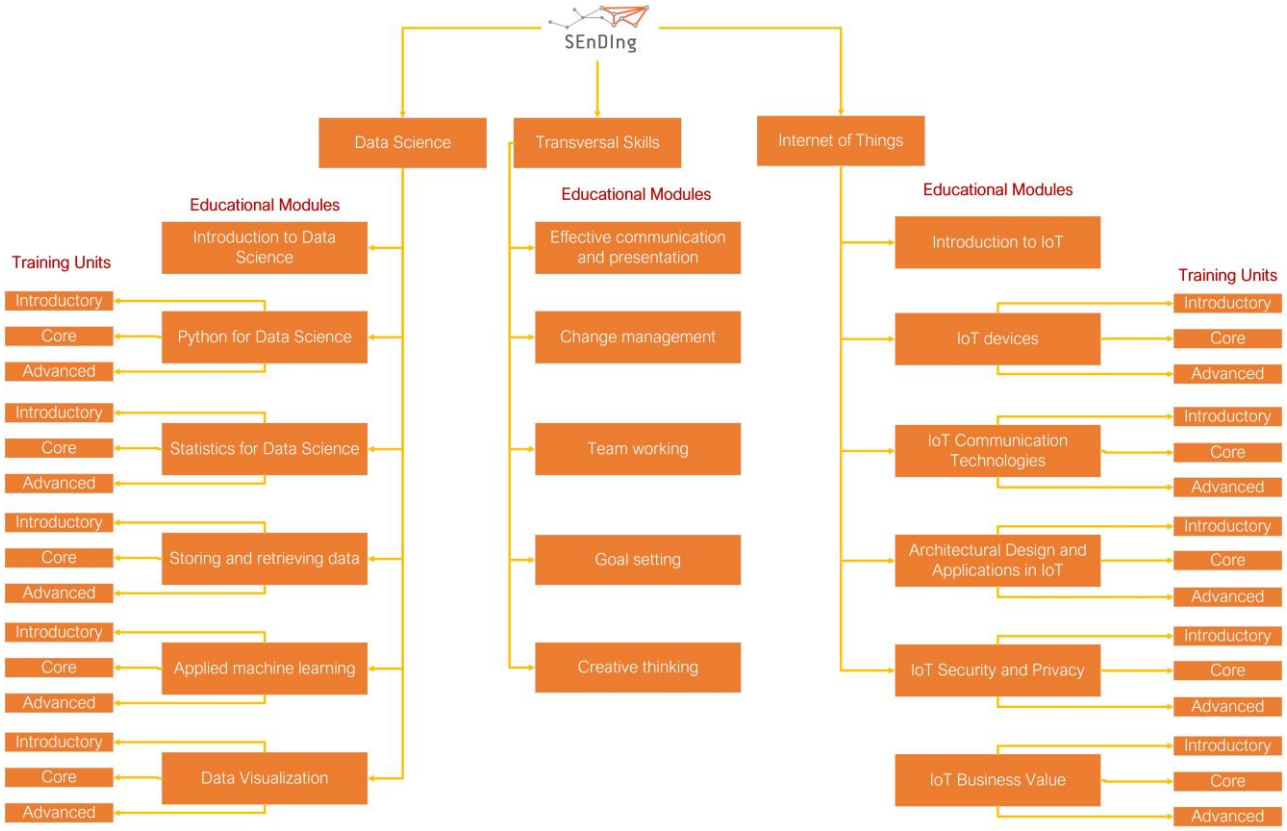

Figure 2: Curriculum educational modules

For a complete description of the educational modules, please refer to SEnDIng project (SEnDIng, 2017).

Curriculum European Qualifications Framework (EQF) level

Considering the qualification levels that the SEnDIng curriculum addresses, we defined that EQF level 5 is the most appropriate as the main beneficiaries of the vocational trainings will be IT professionals with working experience, or graduates from Higher Education Institutions. The learning outcomes relevant to qualifications at EQF level 5 are defined as follows (European Commission, 2008): (a) **knowledge**: comprehensive, specialised, factual and theoretical knowledge within a field of work or study and an awareness of the boundaries of that knowledge, (b) **skill**: a comprehensive range of cognitive and practical skills required to develop creative solutions to abstract problems and (c) **competence**: exercise management and supervision in contexts of work or study activities where there is unpredictable change; review and develop performance of self and others.

Conclusion

This paper presents a modular and learning outcomes-oriented VET curriculum for the training of IT professionals at transversal skills, Data Science and IoT. The curriculum will be delivered in autumn 2019 into three phases by combining online self-paced courses with face to face training and work based learning. The advantages of SEnDIng curriculum rely on (a) its modularity which permits the learners to build their own learning paths, (b) the combination of technical knowledge and skills with transversal cross-domain skills and competences and (c) the work based learning phase enabling the learners and especially those who lack professional experience to apply the knowledge and skills gained through the online and face to face training at real working environments. The SEnDIng curriculum can be easily exploited by training providers at each level of education (VET, higher education, adult education) in order to extend the offering of their training programs.

Acknowledgements

This research was conducted with the support of the Erasmus+ programme of the European Union (contract no 591848-EPP-1-2017-1-EL-EPPKA2-SSA). The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

The authors would like to thank the SEnDIng project partners for their valuable comments for the design of the curriculum.

References

ATKearney, (2016). The Internet of Things: A New Path to European Prosperity. *ATKearney*. Retrieved from: https://www.atkearney.com.

Biggs, J. (1996). Enhancing teaching through constructive alignment. Proceedings of Higher Education, 32, 347- 364.

Blumberg, P. (2009). Maximizing learning through course alignment and experience with different types of knowledge. *Proceedings of Innovative Higher Education,* 34, 93-103.

Columbus, L. (2017, May 13). IBM Predicts Demand For Data Scientists Will Soar 28% By 2020. *Forbes*. Retrieved from https://www.forbes.com/sites/louiscolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/#e75fad07e3bd.

European Centre for the Development of Vocational Training. (2010). *Learning outcomes approaches in VET curricula: A comparative analysis of nine European countries*. Retrieved from https://www.cedefop.europa.eu/files/5506_en.pdf.

European Commission. (2008). *The European Qualifications Framework for Lifelong Learning (EQF)*. Retrieved from http://ecompetences.eu/wp-content/uploads/2013/11/EQF_broch_2008_en.pdf.

European Commission Digital Single Market (2019). *Building a European data economy*. https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy.

European Parliament Council. (2008). *Recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning*. Retrieved from https://eur-

lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2008:111:0001:0007:EN:PDF

Institute for the Future (2011). *Future Work Skills*. Retrieved from http://www.iftf.org/futureworkskills.

José, Sá M., Serpa. S, (2018). Transversal Competences: Their Importance and Learning Processes by Higher Education Students. *Proceedings of Educational Science,* 8, 126.

Manieri, A., Brewer, S., Riestra, R., Demchenko, Y., Hemmje, M., Wiktorski, T., Ferrari, T, Frey J. (2015). Data Science Professional uncovered: How the EDISON Project will contribute to a widely accepted profile for Data Scientists. *Proceedings of the IEEE 7th International Conference on Cloud Computing Technology and Science*.

Mikroyannidis, A., Domingue, J., Phethean, C., Beeston, G., Simperl, E. (2017). Designing and Delivering a Curriculum for Data Science Education across Europe. *Proceedings of the 20th International Conference on Interactive Collaborative Learning*.

Press, G. (2015, April 30). The Supply And Demand Of Data Scientists: What The Surveys Say. *Forbes*. Retrieved from https://www.forbes.com/sites/gilpress/2015/04/30/the-supply-and-demand-of-data-scientists-what-the-surveys-say/#777725316b90.

SEnDIng Erasmus+ project. (2017). *D2.3: Vocational curricula/educational modules for Data Science and Internet of Things VET program*. Retrieved from http://sending-project.eu/attachments/article/71/SEnDINg_DLV2.3-1st_version.pdf.

UNESCO, International Bureau of Education. (2013). *Glossary of curriculum terminology*. Retrieved from http://www.ibe.unesco.org/fileadmin/user_upload/Publications/IBE_GlossaryCurriculumTerminology2013_eng.pdf

World Economic Forum. (2016). *The Future of Jobs*. Retrieved from: http://reports.weforum.org/future-of-jobs-2016.

Wright, B. (2018, March 9). How a talent shortage is hampering IoT development. *IDG Connect*. Retrieved from: https://www.idgconnect.com/idgconnect/analysis-review/1003535/talent-shortage-hampering-iot-development