DS&AI

Co-funded by the Erasmus+ Programme of the European Union

Curriculum Development in Data Science and Artificial Intelligence

599600-EPP-1-2018-1-TH-EPPKA2-CBHE-JP

Identification of similar curricula and needs assessment in the area of Data Science and Artificial Intelligence

We teach today's students
With yesterday's knowledge
For a future we don't know.

Version 1: 15 September 2019

DS&AI

Co-funded by the Erasmus+ Programme of the European Union

PROJECT INFORMATION

Acronym	DS&AI
Project Title	Curriculum Development in Data Science and Artificial Intelligence
Contract Number	599600
Start Date	15 Nov 2018
Duration	36 months

DELIVERABLE INFORMATION

Deliverable Number	D1.4
Deliverable Title	Final report with recommendations
Submission Due Date	14 May 2019 (proposal) 15 September 2019 (revised)
Actual Submission Date	15 September 2019
WP Number and Title	WP1: Preparation
Lead Beneficiary	LEU
Dissemination Type	Report
Dissemination Level	International
Quality Reviewer 1	WU
Quality Reviewer 2	AUEB
First Quality Review Date	11 February 2020
Quality Review Pass Date	20 August 2020

DISCLAIMER

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect 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.

DS&AI

Co-funded by the Erasmus+ Programme of the European Union

Abstract

In this report we survey and identify existing master programmes in Europe and Asia in data science and artificial intelligence, as well as similar professional training programmes. Additionally, we do an assessment of the employment needs of ICT and other relevant industries in Sri Lanka, Thailand, and Indonesia for acquiring graduates from computer science related master programmes.

CONTENTS

1. Introduction

The average introduction to almost any book is somewhat of a bore

--- Boris Karloff

Across many industry sectors, including retail, agriculture, healthcare, banking, finance, and telecommunications, companies in South Asia are launching initiatives to improve their operations and offerings through data-driven techniques[1]. One of the main success factors in this digital transformation is having the right talent with critical skills in data science and artificial intelligence, with the ability to understand both data processing techniques and the business they are operating in.

1.1 Context

Thailand, Indonesia, and Sri Lanka are middle-income countries within the South Asia region. While data science and artificial intelligence based technologies are gaining traction and becoming strategic priorities these countries are still falling behind regional leaders in the technological market, such as Singapore and Malaysia, while the region as a whole is far behind China. Thailand, Indonesia, and Sri Lanka have well-established and growing higher education systems; increasingly focused on quality improvement. However, they fail in educating people who design, develop, deploy, and train data science and artificial intelligence-based technology, as evident from the acute shortage of data analysts in public and private companies[2,3].

These trends make clear that development of talent and capabilities is needed if data science and artificial intelligence are to reach their full potential throughout the region. The success of political and economic initiatives by governments in the south Asian region to transform the current export based economy into an innovation and knowledge based economy will critically depend on the readiness of the labour force for ICT in general and data-driven technologies in particular.

This document focuses on the identification of needs and gaps in master level education in data science and artificial intelligence through South Asia, using European experience on building such curricula in the subject areas. It will serve as guidance in developing the necessary knowledge, skills, and educational practices. This report is not intended purely as a description of the current data science and artificial intelligence graduate curricula in Asia and Western Europe. Rather, it aims to provide an account of a

[1] *The age of analytics: competing in a data driven world*, McKinsey Global Institute, 2016.

[2] *Artificial intelligence and South-East Asia's Future,* McKinsey Global Institute, 2017.

[3] Mathur, Aneja, Anwar, Shridhar, and Sanchez, *Future of Work in Sri Lanka: Shaping technology transitions for a brighter future*, TANDEM research report, International Labour Organization, pp. 1—98, 2019.

common core of what a Data Science and Artificial Intelligence master programme should provide, and how this core can be extended to satisfy the specific needs of countries of Asian partners. This report is compiled using the input from three different activities in the context of the Erasmus+ project on Curriculum Development in Data Science and Artificial Intelligence, registered under the reference number 599600-EPP-1-2018-1-TH-EPPKA2-CBHE-JP. These tree activities comprise:

1. the identification of data science and artificial intelligence curricula in Asia;
2. the identification of data science and artificial intelligence in Europe together with best teaching practices;
3. the identification of existing professional training courses in data science and artificial intelligence and the demand for internships in the Asian countries related to the project.

Data science is concerned with the extraction of useful domain knowledge from large, complex data sets. Drawing on the fields of machine learning and data mining, it provides critical support to decision makers in many professions, allowing them to make decisions based on statistically significant patterns in data. Applications include business intelligence, the detection of anomalies in patterns of behaviour, and for example the analysis of services in cloud computing.

Artificial intelligence is an even broader field concerning the theory and development of computational systems able to perform tasks inspired by human intelligence. We can distinguish three main paradigms within the field of Artificial Intelligence:

1. *symbolic AI* – consisting, among others, of inductive logic programming, robotic process automation and expert and fuzzy systems,
2. *statistical AI* – consisting of decision networks, probabilistic programming (including Bayesian program synthesis), computer vision (activities and image recognition as well as machine vision), natural language processing, and machine learning, and
3. *subsymbolic AI* – consisting of distributed artificial intelligence (including agent-based modelling, swarm intelligence and multi-agent systems), ambient computing, natural computing, multi-objectives optimization, affective computing, embodied intelligence, and autonomous systems.

Today the most widely applied discipline within AI is machine learning, a method of data analysis that automates analytical model building without being explicitly programmed. Machine learning comes in at least three “flavours”: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning refers to methods for learning a model from a set of training data containing both input and output results. Regression and classification are typical examples of supervised learning techniques. Unsupervised learning consists of building a model from a set data, via for example visualization or clustering. Reinforcement learning is about taking actions in an environment so as to maximize some notion of cumulative reward, as for example used in robot navigation and games (including chess, Go, and business games).

1.2 Methodology

The high-level methodology of the project contains of five steps, each described in a separate chapter of the report:

1. Describing of existing curricula on data science and artificial intelligence in Europe.
2. Identifying academic curricula on data science and artificial intelligence in the Asian partners' countries,
3. Investigating existing professional training programmes.
4. Assess industry labour needs.
5. Determining the gaps between existing programmes and requirements identified in previous steps.

The first step is about the activity concerning the **description of existing curricula on data science and artificial intelligence in Europe**. It is based on two questions:
Q1: What is the content of existing DS&AI curricula?
Q2: How is lectured in existing DS&AI curricula (good practices)?

The answer to the above questions has been obtained through a questionnaire sent to around 90 universities in Europe. We divided the universities in three large groups: Western Europe (Belgium, Germany, Netherlands, United Kingdom) Scandinavian region and Eastern countries (Czech Republic, Finland, Hungary, Norway, Poland, Portugal, Romania, Slovakia, Spain, and Sweden), and Balkan and Mediterranean countries (Bulgaria, Cyprus, France, Greece, Italy, Romania, Slovenia, Spain, and Turkey). Except for the Scandinavian region and Eastern countries, the response to the questionnaire was very low (35%, but none from Western Europe and the Balkans and only one from a Mediterranean country), and in some cases (4 out of 90) the universities we approached did not have any master programme related to data science or artificial intelligence. When we got no response, we collected all the information that was available ourselves via their website (although in few countries no information in English was available) and complemented the missing information by means of telephone interviews. In a few cases, we obtained detailed information via email. A copy of the questionnaire used can be found in Appendix A.

The second step is about **identifying academic curricula on data science and artificial intelligence in the Asian partners' countries**. Through desktop research and using a common template we have collected data from several master programmes on Data Science and Artificial Intelligence in the regions of Thailand, Sri Lanka, India, Malaysia, and Indonesia. Many of the programmes we collected data from are generic master programmes in computer science and contain only few courses related to data science or artificial intelligence. Only 13 out of 35 programmes we investigated are directly related to data science or artificial intelligence. Of these only two are full-fledged master programmes on artificial intelligence. We complemented the information about the programme with a discussion of the credit system in Asia.

The third step is like the second, but instead of looking for existing academic curricula it concentrates on **professional training programmes** on data science and artificial intelligence in the Asian partners' countries. Previous studies on professional training in general have shown that Indonesia, Malaysia,

Philippines, Thailand and Sri Lanka have a 'fairly developed' system of professional training characterized by low enrolment, high drop-outs, poor-quality of teachers, inequitable access for women and rural populations, limited private sector involvement, and very inadequate budgetary allocations. In other words, technical training in South Asia is neither supporting a high economic growth rate, nor expanding global markets, nor increasing employment opportunities.[4] Using a desktop search, we identified professional training courses specific to data science and artificial intelligence offered by either universities or private companies in south Asia. The goal is not to see if the situation is improved with respect to the above study, but to identify what are the topic offered by these professional training.

In the fourth step we identify the **labour needs from industry**. We proceed via a systematic process by gathering information through a questionnaire sent to several companies in Thailand, Indonesia, and Sri Lanka. We differentiate between the needs of the companies by their size in terms of employees, and we try to cover several sectors including both public and private companies. The questions concern both the needs of the companies for training courses for their own employees, as well as the needs they have in educating new students through internships. The requirements that company would like to have on the topic areas of the training courses and on the desired knowledge students should have before starting the internships reveals the difference (gap) between current knowledge/practices as described in the third step and current evidence based requirements revealed in this step. Most importantly we consider labour needs not only concerning knowledge but also those that can occur in skills (programming) or practice (tools).

The final step consists of **identifying the gaps that have to be addressed in the development of master's degree programmes in data science and artificial intelligence** in Thailand, Indonesia and Sri Lanka. With this goal, partner universities in each of these countries organized a focus group composed of 4 academics, 3 students and 3 representatives from ICT companies. This composition of the focus group guaranteed that all relevant stakeholders provided their input in understanding what is needed by a new education programme in data science and artificial intelligence.

[4] Ul-Haq, M., & Haq, K. (Eds.). (1998). *Human Development Report in South Asia 1998*. Delhi, India: Oxford University Press, page 100.

2. Data Science and Artificial Intelligence in Europe

Thanks to data science and artificial intelligence,
machines can now be programmed to the next thing right.
But only humans can do the next right thing.
--- Dov Seidman

2.1 Introduction

In agreement with the Framework for Qualifications of the European Higher Education Area Subject benchmark the objective of a master programme Data Science and Artificial Intelligence in Europe is to provide students with a suitable basis for a further career, both in research as well as in industry and society in general. As such, it provides the student with the specific knowledge and abilities, exemplified in the form of a master diploma that allows graduates access to a PhD programme in Artificial Intelligence and related disciplines. Also, it prepares graduates for a position in which they can earn their own subsistence, for example by working in knowledge-intensive companies.

A European Master programme in Data Science and Artificial Intelligence provides sufficient training in independent scientific reasoning, conduct, and communication to reach internationally accepted standards of academic skills at that level. Graduates can communicate original ideas in their own language (and often in English) to a public of specialists and non-specialists. Graduates are provided with the necessary knowledge and tools needed to formulate an informed opinion about the meaning, the ethical, and the social impact of Data Science and Artificial Intelligence in society and they are aware of their responsibility towards society. The above objectives are specified into different sets of learning outcomes or final qualifications. In Europe, the set of learning outcomes must comply with international standards presented in terms of the Dublin descriptors for the master's profile. These are general statements about the ordinary outcomes that are achieved by students after completing a master programme and obtaining a master degree.

The identification of European master programmes in Data Science and Artificial Intelligence is rather tricky because traditionally, the field of artificial intelligence is concerned with the study of cognitive processes that play a role in human perception, reasoning and action, and building intelligent systems. This implies that the field of artificial intelligence is closely related to other disciplines such as computer science, mathematics, psychology, linguistics and philosophy. Similarly, the field of data science belongs to both computer science and statistics. Consequently, there are several specialised study programmes divergent not only with respect to the field they are based on (computational sciences, cognitive science, and statistical science) but also in regard of the application areas (e.g., engineering, life science, and business science).

Most programs have a duration of two years (120 ECTS). There are exceptions, most notably in the United Kingdom, but also in other countries, where programmes have a nominal duration of one year, some worth a total of 90 ECTS, others only 60 ECTS. ECTS are credits representing the workload and defined learning outcomes ("what the individual knows understands and is able to do") of a given course or programme. They have been adopted by most of the countries in the European Higher Education Area. 60 ECTS are the equivalent of a full year of study or work. In a standard academic year credits are usually broken down into several smaller components and assigned to courses, internships, and thesis work. The way to assign credits to thesis projects and internships varies and appears programme dependent. In some cases, thesis and internship are combined together as a unique educational component.

2.2 Curricular aspects

The objective of the master programme is to provide students with a suitable basis for a further career, both in research as well as in industry and society in general. Graduates should be prepared for a number of different roles and careers at key positions in society. They should have

- a basic understanding of all the key areas of Data Science and Artificial Intelligence,
- an advanced understanding in some of the key areas of Data Science and Artificial Intelligence, and
- a specialist knowledge of at least one of the key areas in Data Science and Artificial Intelligence.

Our starting point in this report is a set of core topics and skills, present in a master programme either as a dedicated course or as a substantial topic within one or more courses. We consider core topics, support knowledge and electives. Core topics are defined by master level courses on key areas of data science and artificial intelligence, while the elective courses give additional and specialistic knowledge on these areas. We consider also support modules, and academic skills. While core and elective topics are about courses grounded in data science and artificial intelligence, the support modules are about course or part of it that on a different related discipline that is needed to understand core and elective courses. The academic skills are specific neither to data science, to artificial intelligence nor to related disciplines.

Subject names and content are not standardized. For each programme, the more content related part of the above subjects is defined in terms of topics and learning objectives. The learning objectives of each topic (often divided into more courses) are related to the learning outcomes of the entire programme. The formal framework is provided by the Dublin descriptors, which give an overview of the academic competences of a bachelor (first cycle), master (second cycle) and PhD (third cycle) programme. In all three cycles, these descriptors include the following components: knowledge and understanding, applying knowledge, and understanding, making judgements, communication, and lifelong learning skills. For a two years master programme the Dublin descriptors are given below.

	Dublin Descriptors Master	**Master (2-years ; 120 ECTS credits ; second cycle ; EQF level 7**[3] Qualifications that signify completion of the master degree are awarded to students who:
D1	Knowledge and understanding	have demonstrated knowledge and understanding that is founded upon and extends and/or enhances that typically associated with the first cycle, and that provides a basis or opportunity for originality in developing and/or applying ideas, often within a research context
D2	Applying knowledge and understanding	can apply their knowledge and understanding, and problem solving abilities in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their field of study
D3	Making judgements	have the ability to integrate knowledge and handle complexity, and formulate judgements with incomplete or limited information, but that include reflecting on social and ethical responsibilities linked to the application of their knowledge and judgements
D4	Communication	can communicate their conclusions, and the knowledge and rationale underpinning these, to specialist and nonspecialist audiences clearly and unambiguously
D5	Lifelong learning skills	have the learning skills to allow them to continue to study in a manner that may be largely self-directed or autonomous

For example, the core topic Machine Learning may be described through a set of courses including one on Reinforcement Learning. This course could treat topics such as

- Core techniques and approaches in reinforcement learning
- The multi-armed bandits' problem
- Markov decision processes; dynamic programming, Monte Carlo methods
- Approximate solutions, including linear and non-linear function approximation
- Applications and case studies, including AlphaGo and Q learning for game.

The learning objectives depends on the topics but also on the lecturing and the assessment methods. For this course the learning objectives could be

- Understand the key features and components of reinforcement learning;
- Knowledge of theoretical foundations on basic and advanced reinforcement learning techniques;
- Understand the scientific state-of-the-art in the field of reinforcement learning;
- Define an application problem as a reinforcement learning problem;
- Understand how to apply leaned reinforcement learning techniques on a new problem.

[5] Here the abbreviation EQF refers to the European Qualifications Framework. More information can be found at https://ec.europa.eu/ploteus/en/content/descriptors-page

In this case, the first three objectives are related to the Dublin descriptor D1 while the last two are related to D2. The fourth objective is also related to D3.

2.2.1 Core topics

Core topics define the key areas of a curriculum in Data Science and Artificial Intelligence and are therefore present as a dedicated course or as a substantial topic within one or more mandatory courses in the programme. From the data we collected, it is not always obvious which of the 8 key core topics a course covers, and some of the conclusions below are merely our interpretation of the data due to lacking definition in terms of learning objectives and detailed topics of the courses forming a master programme. Without a detailed definition of each course, it is difficult to be certain where the dividing line lies between the areas.

There is a European (and possibly international) consensus that most of our proposed eight core topics define the key areas of data science and artificial intelligence and are therefore covered by mandatory courses in master programmes. In fact, the great majority of the master programmes we examined have courses related to

- Algorithmic Problem Solving (search, decision making, optimisation)
- Knowledge Representation, Extraction and Reasoning (data mining, deep learning)
- Machine Learning (supervised, unsupervised, reinforcement learning)

Many (but not all) master programmes cover other topics such as

- Intelligent Autonomous Agents and Multi-Agent Systems
- Interaction (Perception, Human-Computer Interaction, Communication)

and to a bit less extend computational linguistics. Only few universities have mandatory courses related to cognitive science or about giving the context of artificial intelligence (history, philosophy, and ethics). Important omissions in the above list of key areas are, for example, given by mandatory courses on traditional computer science and mathematical topics, such as databases, computer architectures, information theory, and security. More data science and artificial intelligence area omitted in the above list but covered by some mandatory courses are

- Graph Theory and Network Science (social network analysis, complex networks)
- Probabilistic models and Reasoning under Uncertainty (statistical learning, Bayesian inference, fuzzy logics).

2.2.2 Elective topics

The set of core topics does not need, by itself, to constitute a complete master programme. Every programme may include additional elective topics relating to artificial intelligence, data science,

computer science, cognitive psychology, but also courses related to application areas such as bioinformatics and business analytics.

Most of the European master programmes on data science and artificial intelligence have elective courses on

- Computational Intelligence
- Computer Vision
- Evolutionary Algorithms (Genetic Algorithms, Evolutionary Computing)
- Language and Speech Technology
- Neural Networks
- Text Mining and Information Retrieval
- Reinforcement Learning,
- Robotics

Other electives offered by few European master programmes cover topics on

- Cognition and Cognitive Modelling
- Computational and Cognitive Neuroscience
- Ethical, Legal and Social Aspects of AI
- Perception (Computational and Natural)
- Reasoning under Uncertainty
- Virtual Reality and Gaming
- Web and Artificial Intelligence

2.2.3 Support knowledge

Support knowledge define the basic body of know-how assumed to successfully understand core and elective topics. Only few programmes do not assume any prior knowledge (18% of the programmes in the Mediterranean and Baltic countries!). Virtually all other master programmes assume knowledge coming from Computer Science (Programming in Java, C++ or Python, Algorithms and Data Structures, Databases), and from Mathematics (Calculus, Discrete Mathematics, Linear Algebra, Probability Theory and Statistics). In general, knowledge of Logic (Propositional and Predicate Logic) is not required, perhaps a sign of the current interest of artificial intelligence in statistical and sub-symbolic techniques, methods, and tools. By far, most of the master programmes are taught in English, and, as a result, students from outside the European Union are often required to have an internationally recognized English certificate, like TOEFL or Cambridge.

Most of the master programmes provide some form of support knowledge as dedicated but not necessarily mandatory courses. This is especially true for courses on programming, database, algorithms and probability theory and statistics.

2.2.4 Academic skills

Apart from knowledge and understanding, every master programme supports, in different forms, the development of a set of general academic skills. Via specific courses or using an appropriate choice of work and assessment methods analytic skills, teamwork, modelling, written and oral communication are considered the most important academic skills, directly followed by empirical methods, argumentation, and presentation.

2.3 Didactic aspects

All the programmes we examined have their own didactic principles, sometimes defined at an institute level, or at the university level. Not many documents on didactic aspects were available on the general websites of the programmes, so the information available is somewhat dependent on those programmes that answered our questionnaire. Instead of presenting this partial information, below we give several examples of best practices concerning didactic aspects, not necessarily related to data science or artificial intelligence programmes.

2.3.1 Research-oriented education

The key concept of research-oriented education is to provide education embedded within a research environment. The research orientation is not only reflected in the thesis, but also in several courses, in which students write a scientific essay or paper. All courses are taught by full time faculty members, who are actively contributing to the international scientific endeavour. In this context, the staff is often required to have a University Teaching Qualification (UTQ) certificate as well as an International English certificate. The university often provides possibilities for Advanced University Teaching Qualification certificate.

Several courses and the thesis projects offer opportunities for students to contribute to current research of the staff. Students are stimulated to publish their research in international conferences and journals. An international environment is stimulated through exchange programmes and regular visits of external professors.

2.3.2 Problem-based education

In problem-based education (also known as Project-Centred Learning), the students learn to apply the knowledge they gained in the courses in realistic and challenging projects. In small groups students perform all aspects of medium to large scale projects. The projects can come from a variety of sources, which implies a good connectivity of the programme with companies, for example via internships. The method is typically divided in three parts: case studies, role-playing, and solutions. Case studies are presented to students in some form matching the real-world problem, roleplaying has students dividing their tasks as in a real-world situation, and solution involves the development of critical thinking skills and problem-solving abilities.

To build an effective curriculum, explicit attention has to be paid to integration, analysis and application of the courses. The goal is for students to become familiar with the culture of research and the practice as academic professionals.

2.3.3 Student activating education

The underlying didactic concepts of Student Activating Education are self-responsibility, self-management and goal directed education, cooperative learning including the pre-organized and self-organized form, and conceptual learning in a challenging learning environment. Activating methods organize the teaching process, so that the education objectives would be achieved mainly on the basis of students´ own learning. The emphasis is on thinking and problem solving. Methods include face to face teaching in small classes, e-learning, blended learning, and group work.

As for research-oriented education, the staff is often required to have a University Teaching Qualification (UTQ) certificate as well as an International English certificate. The use of a foreign language in teaching and of an international and multi-cultural class are seen as other activating methods. The university often provides possibilities for Advanced University Teaching Qualification (ATQ) trajectories in which lecturers further professionalize their didactic skills.

2.4 Conclusion

In this document we focused on the identification of the curricula of master level education in data science and artificial intelligence through Europe. While there is no common international framework describing a curriculum in data science and artificial intelligence most universities tend to agree on what are core courses (programming, knowledge representation and machine learning) and on the topic of the electives offered. Interestingly, many university cover also not technical aspects in data science and artificial intelligence, such as ethical, legal, philosophical, and social ones. Apart from commonalities in knowledge and understanding, all programmes we examined have their own didactic principles and good practices that we have categorized and shortly presented here.

3. DS and AI Master Curricula through South Asia

Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom

--- Clifford Stoll

3.1 Introduction

The use of systems based on data science and artificial intelligence techniques is rapidly growing in several sectors around the world, e.g., in the health care, transport and financial sectors. This revolution is accelerated by advances in data collection, storage and analysis, efficient algorithms, and an increasing computer processing power. Although the United States, China and Europe are the frontrunners in developing such artificial intelligence systems, these world-wide scientific advances also have a major impact on countries from Southeast Asia. As the fields of data science and artificial intelligence have the potential to contribute positively to the economic and social climate of those countries, there seems to be a great need for training professionals in this area.

The aim of this chapter is to identify existing master level curricula related to Data Science and Artificial Intelligence in Thailand, Indonesia, Sri Lanka and wider region. While in Europe modern higher education programmes have been developed on specific areas of computer science such as data science, data analytics, business analytics and artificial intelligence system, in many countries from Southeast Asia computer science education programmes are often developed in a generic way, e.g., based on ACM curricula recommendations. Most common master programmes are therefore named Computer Science, Information Technology, Information System, and Computer Engineering. These programmes often include several elective tracks, such as hardware and computer architecture, communication and networking, computer system or software. But there are no programs specific to data science and artificial intelligence. Often, students are only allowed to follow only a limited number of elective courses within their track, as such the postgraduate knowledge of these students cannot be compare with the postgraduate knowledge of European students who followed specific master programmes dedicated on data science or artificial intelligence.

3.2 Education and credit system in Thailand, Indonesia and Sri Lanka

In Asia, universities often offer two kinds of masters' degree: coursework-based and research-based. Research-based master programme does not require any course work and the admission is usually based on a specific bachelor's degree. Coursework-based master programmes requires to gather information by taking courses. In many universities, students must perform an additional independent study project or an independent research project, after their coursework. Coursework-based programmes are typically

shorter than the research-based programmes. The coursework-based programmes take 1-2 years while research-based programmes takes between 2-3 years. In Asia the nominal duration of a masters' degree is typically two years.

An academic year is usually divided into the first and second semester, they may rarely be an optional third summer session. Each semester length is 16-18 weeks and include an examination. Prospective student can apply for masters' degree twice a year. A number of minimum credits are between 36-45 in South-East Asia and nearly 60 in South Asia.

Each university in Asia is managed by its own polices and regulations, with the only constraint to be consistent with governmental strategies and policies. Quality assurance is usually supervised by government organization such as University Grant Assurance (UGC), Quality Assurance and Accreditation Council (QAAC), High Education Institute or Minister of Education. Some programmes have professional council to guarantee quality as required by the labour world (e.g., for the fields of engineering, medical or pharmacy). In Asia there is no Data Science and Artificial Intelligence council to guarantee quality of higher education.

Asia Credit Transfer System (ACTS)

A credit system in higher education is widely accepted as a way of describing educational programmes by attaching credits to its components. The credit value's role is to indicate the average amount of learning time required for a learner to achieve a credit value. Learning time is made up of lecture-based learning as well as independent learning, e.g., self-learning online, doing homework and research and work placement. There is not a unique standard for academic credit unit and accreditation system. Universities in the Association of South East Asian Nations, or ASEAN, member states utilize a credit transfer system. Traditionally, recognition of semesters abroad is been carried out on a case-by-case basis. When students study aboard, they must transfer credit into other credit system such as ECTS. Asian credit earning systems generally are defined as follows

- Lecturing or any academic activities that are equivalent to 1 hour per week throughout the whole semester or approximately 15 hours in a semester is measured as 1 credit
- An operation, experiment, or a laboratory activity that takes 3 hours per week throughout the whole semester is measured as 1 credit.

Although currently no global credit transfer system for the ASEAN region exists, certain systems are used to help streamline the process. In order to realize a student-centered higher education system in Asia, the Asian Cooperation Dialogue recently developed an Asia Credit Transfer System (ACTS) based on the successful three decades of European experience with the European Credit Transfer System (ECTS). ECTS uses 60 credits as standard for one year of study or 30 credits for each semester or 20 credits for each trimester. A typical master's degree would consist of 60 (one year) to 120 (two years) credits, and each credit is based on 28 hours of work for the student (based on the average working time of 1560 hours per year in Europe). ECTS is a very systematic and permeable credit transfer system. However, student workload for ECTS does not reflect an Asian workload (1560 hours per year in Europe vs. 1800-2100+ hours per year in Asia). Therefore the ACTS coefficient depends on the minimum number of credits in

each programme, and it is not standard even within the same country. In general, Asian master's programs have a duration of two years and they normally require 36-45 credits in total. Therefore, one credit in a master's degree is approximately 2.6 to 3.3 ECTS. For example

- 1-year study in Europe = 60 ECTS (standard)
- 1-year study at PSU in Thailand = 36 credits (the minimum required credits in total) / 2 years (the duration for master's degree) = 18 credits

Therefore, the coefficient to convert 18 credits to 60 credits is 3.33 ECTS per local credit. As such, a course at PSU consisting of three credits is equivalent to 10 ECTS.

3.3 Relevant master programmes in Thailand, Indonesia and Sri Lanka

This section focuses on the learning knowledge related to existing master programmes in data science and artificial intelligence in Thailand, Indonesia, Sri Lanka and wider region. The knowledge topics have been grouped by the course type which can be divided into three groups including core, elective and supportive course. The core courses are the mandatory courses in the curricula. The elective course are additional course that can be selected by students among several alternatives. The supportive courses are related to the basic knowledge which may be the pre-requirement for admission or a pre-master course or a provided course in the programme. Supportive course may construct under some condition such as academic English language or basic mathematic and statistic. There are five groups of core courses, i.e.:

- Business Intelligence (data modeling, data mining, data analytics and data visualization);
- Data science (programming, data analytics, data warehousing, management and cloud computing);
- Machine Learning ;
- Mathematics and Statistic (theory and methods, regression analytics, statistical analysis and statistical inference);
- Research area (research methodology, independent study, seminar and project).

The most popular topics among the core course are data management, data mining, data warehousing, machine learning and research methodology. There are various elective topics which are hard to categorize, because the non-uniform naming of courses. There is only one topic which is taught in all programmes: data and text analytics. Finally, all programmes make available three groups of supporting knowledge: programming, algorithmics, and mathematics and statistics.

3.3.1 Thailand

The fields of data science and artificial intelligence is currently a part of national strategic plan of Thai government, called Thailand 4.0. Thailand 4.0 is a 20-year strategy to accelerate the country's development from upper-middle-income country to high-income country status. The strategy is designed to promote and support innovations, creativities, research, higher technologies, and green technologies. Thailand 4.0 is correlated to global Industry 4.0 and focuses on 10 industrial areas, i.e., next-generation automotive, intelligent electronics, advanced agriculture and biotechnology, food

processing, tourism, digital, robotics, logistics, biofuels and biochemicals, and medical. The goal is to make possible a digital transformation of manufacturing by combining a wave of next-generation technologies, and thus heavily rely on the advancement and technologies from the field of data science and artificial intelligence. Therefore, the development study programs in these particular fields and the related areas are encouraged by the government.

Beside the formal approval of a new higher education programme on data science and artificial intelligence, the major challenge comes from the northern and north-eastern regions of Thailand, because of the limited number of experts who are eligible to teach and supervise research in master and doctoral programs. The solution is a dedicated development plan and training helping in recruiting more expert teaching staff

Next we summarize the curricula of existing master programmes in areas related to data science and artificial intelligence. These programme are offered by public universities in the northern and north-eastern regions of Thailand. All information was collected and analysed through desk research and interviews.

Master of Science Program in Business Analytics and Data Science

*By:*National Institute of Development Administration, Graduate School of Applied Statistics.

Duration: 2 years with 45 minimum credits.

Core topics:

- Mathematics of Computing (Discrete Math, Probability and Stats)
- Data Modeling and Management
- Data Mining and Information Retrieval
- Natural Language Processing
- Graph Theory and Network Science (social network analysis, complex networks)
- Applied Data Science and AI
- Human-Computer Interaction and Information Visualization

Elective topics:

- Language and Speech Technology (Voice Recognition)
- Machine Learning
- Data and Text Analytics
- Software Engineering (for DS&AI Systems)
- Applied Data Modeling and Management

Support knowledge: Background knowledge in programming, algorithms, mathematics, statistics and data communication (computer network) is required. For those without related background, preparatory courses are provided.

Master of Science Program in Data and Information Science

By: Rajamangala University of Technology Thanyaburi, Department of Mathematics & Computer Science.

Duration: 2 years with a minimum of 36 credits.

Core topics:

- Mathematics of Computing (Discrete Math, Probability and Stats)
- Data Modeling and Management

- Human-Computer Interaction (Visualization)

Elective topics:

- Context of Artificial Intelligence (History, Philosophy, Ethics)
- Data and Text Analytics
- Machine Learning
- Data Mining and Information Retrieval
- Internship

Support knowledge: Previous degree in computer engineering, software engineering, computer science, information technology, business computing, mathematics, statistics or related disciplines, or related working experience of at least 5 years is required.

Master of Science Program in Data Science

By: Srinakharinwirot University, Department of Computer Science.

Duration: 2 years with a minimum of 37 credits.

Core topics:

- Data Modeling and Management
- Machine Learning
- Cloud Computing

Elective topics:

- Applied Data Science and Artificial Intelligence
- Data and Text Analytics
- Image and Video Analytics
- Bioinformatics
- Security and Privacy
- Graph Theory and Network Science (social network analysis, complex networks)

Support knowledge: Background knowledge in programming, algorithms, mathematics and statistics is required. For those without related background, preparatory courses are provided.

Master of Science (Data Science)

By: Chiang Mai University (CMU).

Duration: 2 years with a minimum of 36 credits (24 coursework and 12 thesis, or 30 coursework and 6 independent study).

Core topics:

- Statistics for Data Science
- Data Management and Big Data
- Data Science Programming
- Everything Start with Data

Elective topics in study area: Analytics and e-Commerce

- Business Data Analytics
- Data Analytics for e-Commerce
- Data-driven Customer Relationship Management
- Data Methodology for Web Search Optimization
- Selected Topics in Business and e-Commerce 1 & 2

Elective topics in study area: Data Engineering and Architecture

- Advanced Embedded Systems
- Parallel Processing and Distributed Systems
- Information Security
- Database Management Systems
- Information Technology Infrastructure Library
- Software Project Management
- Data Governance
- Cloud Computing
- Selected Topics in Data Engineering and Architecture 1 & 2

Elective topics in study area: Data Analysis and Machine Learning

- Data Analytics and Machine Learning
- Artificial Intelligence
- Operational Research Techniques
- Statistical Decision Method
- Demographic Statistics
- Methods of Statistics
- Computer Packages for Advanced Statistical Analysis
- Forecasting Techniques
- Numerical Analysis
- Selected Topics in Data Analysis and Machine Learning
- Statistical Quality Engineering and Control

Support knowledge: Background knowledge in Science, Engineering, or other related fields with good Mathematics background or have work experience in the data science fields.

3.3.2 Indonesia and wider region

The need for higher education in the field of data science and artificial intelligence has triggered the development of higher education programmes in these areas. However, the government of Indonesia has some regulations related to the naming of a study programme. In Indonesia, before a new higher education study programme can be established, it should be first accredited by the National Accreditation Board of Indonesia. The name of the study programmes in the computing field must be as specified in the ACM Computing Curricula: Computer Science, Information Technology, Information Systems, and Computer Engineering.

Due to the above naming regulations, data science and artificial intelligence programmes are typically presented under existing computer science, information technology, and Information systems curricula. For this reason below we summarize only those curricula of existing master programme the areas of data science and artificial intelligence in the neighbourhood country Malaysia.

Master of Science Program in Data Science

By: Universiti Kebangsaan Malaysia.

Duration: 1 to 2 years with a minimum of 40 credits.

Core topics:

- Data science fundamentals

- Statistical Methods
- Structured data analytics
- Unstructured data analytics
- Information modeling and database
- Big data analytics and management
- Machine learning

Elective topics:

- Business Intelligence
- Executive Information Systems

Support knowledge: Background knowledge in programming, algorithms, database and data warehouse fundamentals.

Master of Data Sscience

By: University of Malaya.

Duration: 2 years with a minimum of 42 credits.

Core topics:

- Research Methodology
- Principles of Data Science
- Data Analytics
- Programming for Data Science
- Data Mining
- Machine Learning for Data Science
- Big Data Management

Elective topics:

- Parallel and Distributed Computing
- Big Data Applications & Analytics
- Network and Security
- Numerical Optimization

Support knowledge: Background knowledge in programming, algorithms, and methods in artificial intelligence.

Master in Data Science and Business Analytics

By: Asia Pacific University of Technology and Innovation (APU) – Malaysia.

Duration: 12 months over 3 semesters of 12 weeks each.

Core topics:

- Big Data Analytics & Technologies
- Behavioural Science, Social Media and Marketing Analysis
- Data Management
- Business Intelligence Systems
- Research Methodology
- Applied Machine Learning
- Data Analytical Programming
- Multivariate Methods for Data Analysis

- Capstone Project 1
- Advanced Business Analytics and Visualisation
- Capstone Project 2

Elective topics:

- Time Series Analysis and Forecasting
- Natural Language Processing
- Operational Research and Optimization
- Multilevel Data Analysis
- Strategies in Emerging Markets

Support knowledge: Background knowledge in statistics, programming, artificial intelligence methods, management.

3.3.3 Sri Lanka and wider region

The situation of higher order education in the fields of data science and artificial intelligence in Sri Lanka is very similar to the one in Thailand, with a government pushing towards a new digital revolution in industry but with a not well developed higher order education programmes specialised in data science and artificial intelligence to absorb the increasing request from the labour market in these areas. Below we summarize the curricula of the current master programmes in data science and artificial intelligence in Sri Lanka and wider region.

Master of Business Analytics

By: University of Colombo, School of Computing.

Duration: 2 years with a minimum of 22 credits in first two semesters.

Core topics:

- Business Statistics
- Organizational Data Management
- Fundamentals of Business Analytics and Data Science
- Data Programming
- Statistical Inference for Analytics
- Machine Learning and Pattern Recognition
- Data Warehousing and Mining
- Information Visualization
- Modelling and Simulation of Data
- Predictive Analytics
- Computational Social Sciences

Elective topics:

- Applied Optimization
- Open Source Intelligence
- Text Analytics
- Project on Business Analytics
- Independent Studies in Business Analytics
- Big Data Analytics
- Analytics for Process Improvement

- Intelligent Agents in Gaming

Support knowledge: English.

Master in Artificial Intelligence

By: University of Moratuwa.

Duration: 2 years with a minimum of 56 credits.

Core topics:

- Programming Essentials for Artificial Intelligence
- Essentials of Artificial Intelligence
- Mathematics for Artificial Intelligence
- Distributed Computing Concepts for AI
- Deductive Reasoning and Logic Programming
- Neuroscience & Neurocomputing
- Evolutionary Computing
- Artificial Cognitive Systems
- Fuzzy Reasoning
- Data Mining and Data Warehousing
- Software Agents and Swarm Intelligence

Elective topics:

- Cryptography and Security Mechanisms
- Semantic Web and Ontological Engineering
- Intelligent Solutions for Industry
- Natural Language Processing
- Embedded Robotics
- Inductive Logic Programming
- Kansei Systems

Master in Computer Science, specialization Data Science, Engineering, and Analytics

By: University of Moratuwa.

Duration: 2 years with a minimum of 60 credits.

Core topics:

- Data Mining
- Big Data Analytics Technologies
- Data Science
- Business Intelligence
- Machine Learning
- Statistical Analysis
- Statistical Inference
- Advanced Algorithms
- Neural Networks
- Research/Industry Projects

Master Business Analytics

By: Informatics Institute of Technology (IIT).

Duration: 2 years.
Core topics:

- Data Management
- Introduction to Big Data and Data Science
- Business Intelligence Tools and Applications
- Statistics for Business Analytics
- Business Modelling and Analytics
- Fundamentals of Data Warehousing
- Research Methods
- Web Mining
- MSc Project

Master Big Data Analytics
By: Informatics Institute of Technology (IIT).
Duration: 2 years.
Core topics

- Advanced Databases
- Big Data Programming
- Cloud Computing
- Data Analysis
- Data Mining
- Data Warehousing
- Research Methods
- Text Analytics
- MSc Project

Master of Technology, specialization in Artificial Intelligence
By: Indian Institute of Science – Bangalore
Core topics

- Digital Signal Processing
- Machine Learning
- Pattern Recognition
- Computer Vision
- Graphics
- Stochastic Systems
- Multimedia
- Real Time and Fault Tolerant Systems
- Sensor Networks
- E Commerce
- Speech Processing
- Real Time Signal Processing and Embedded Systems.

Support knowledge: C and C++

3.4 Conclusions

Currently, data science and artificial intelligence are popular topics around the world. Many universities in Asia have specialised master's program to support the need for the labour market. However, there are several problems in current programmes. Firstly, some universities are regulated by governmental rules about naming, therefore, data science and artificial intelligence specialised programmes must be encapsulated under a general programme of computer science according to the ACM curricula recommendations. Secondly, there is no standard set for data science and artificial intelligence topics, therefore, current curricula have been designed by the present expertise within each university. Thirdly, most of the current master's programmes do not have internships as part of their curricula, creating an unnecessary distance between the labour and the academic world.

4. Professional Training Courses through South Asia

Practice is the hardest part of learning, and
training is the essence of transformation
--- Ann Voskamp

4.1 Introduction

Professional training courses focus on transferring specific knowledge and practical skills which allow individuals to engage in a specific occupational activity. Those courses are not only important in providing employment opportunities to individuals but also helps in enhancing the productivity of companies. In this chapter we provide an (absolutely not complete) review on challenges, outcomes and present situation in professional training courses on data science and artificial intelligence provided in some Asian countries (Thailand, Indonesia and Sri Lanka).

In recent years, artificial intelligence and data science technologies have been massively adopted particularly in manufacturing, financial services, and transportation and logistics industries. Many countries all over the world are tasked with the creation of a skilled labour force consisting of data scientists, business analysts or experts in machine learning. Universities fill the gap on the increasing demand by creating new specialised and attractive curricula on the area while business companies look for shorter professional training courses specialised to their own needs.

4.2 Thailand

In Thailand, the only university providing professional courses on data science is the Khon Kaen University (KKU). KKU offers a 3-day course on Big Data and Data Analytics, focused on the use of the programming language Python and on the tools Hadoop and Tableau. No formal certification is released.

Other professional courses are provided by private companies, with a duration between 2 and 3 days. The main topics covered are Data Modelling and Management; Machine Learning; Data engineering; Database; Management of information systems, Big Data and Data Analytics. The programming languages used are the expected ones (Python, R, Pytorch, Tensorflow), and as tool they all use the most common in the field (Hadoop Tableau, Spark, MongoDB, Amazon Web Services, Google Cloud Platform, Azure). The table below present a summary of the courses we analysed.

Company	Course name	Duration	Subject area	Certification
Intelligist Company Limited	Big data and Data Analytics	2 days	- Big Data and Data Analytics	-
IMC Institute	1. Machine Learning for Data Science 2. Advanced Analytics using Deep Learning 3. Big Data Architecture and Analytics Platforms	3 days 3 days 3 days	- Machine learning - Deep Learning - Big Data and Data Analytics	Yes
AIAT Academy	Intensive Training Python for Data Science	1 day	- Python	Yes
NSTDA Academy	Art of Data Science: ADS	3 days	- Big Data and Data Analytics	Yes
Techtalkthai training	Intro to deep learning	2 days	- Machine learning - Deep Learning	No
-	Smart speaker, Smart home, Chatbot or Smart vision	-	- Data modelling and management	-
ALTERYX, INC.	Alteryx Data + Analytics rEvolution Summit	-	- Data modelling and management	-

4.3 Indonesia

Syiah Kuala University delivers a course of 3 days on Big Data and Data Analytics using Java and Hadoop, giving a professional certificate at the end. Similarly, the Institut Teknologi Bandung provides a certificate for the successful attendance of a course on machine learning using Python as programming language. The table below summarize other professional courses offered by private companies in Indonesia.

Company	Course name	Duration	Subject area	Certification
Telkom Professional Certification Center (TelkomPCC)	Practical Data Science	3-5 days	- Data modelling and management - Data engineering - Database - Management of information systems;	Yes
Inixindo Jogja	Certified Data Science Specialist	2 days	- Data modelling and management - Data engineering - Management of information systems	Yes

Digital Talent Scholarship	topics on Artificial Intelligence	-	Machine learning	-

4.4 Sri Lanka

In Sri Lanka, no university is providing a professional course on data science or artificial intelligence. Some of the courses offered by the private sector are listed below. They all use basic of Python and/or R as standard programming language.

Company	Course name	Duration	Subject area	Certification
NIBM: National Institute of Business Management (NIBM)	• Advanced Certificate Course in Data Analytics • Data Analytics with R for Working Professionals • Analytics for Business Managers	4 months one day 2 days	- Data modelling and management - Machine learning - Natural language processing - Statistics	Yes
Institute of Applied Statistics, Sri Lanka	• Statistical Methods for Data Mining • Introduction to R • Statistical Methods for Big Data Analysis	4 days 2 days 4 days	-	-
NobleProg	• A Practical Introduction to Data Science Training • Data Science Programme Training	35 hours 245 hours	-	-
IASSL	Statistical course for data mining	4 days	-	-
University of Moratuwa	• Phyton Essentials • Big data & Analytics	80 hours 70 hours	-	-

4.5 Conclusion

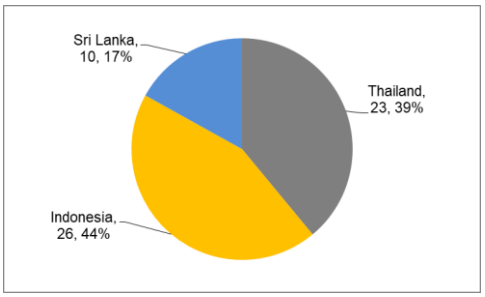
In Thailand, Indonesia, and Sri Lanka professional trainings data science or artificial intelligence are provided mostly by private companies, all focusing on data modelling and processing and machine learning.

5. Companies Needs Assessment through South Asia

Good companies will meet needs; great companies will create markets.
--- Philip Kotler

5.1 Introduction

Across South Asia, talent shortages in data science and artificial intelligence—caused by gaps and mismatches between labour market needs and the available supply of educated and trained professionals—constrain economic growth.[6] These shortages prevent companies from scaling up, meeting demand in new locations, and launching new products and services. The talent shortage will have a direct impact on their competitiveness and productivity. In this chapter we explore the current needs that companies have for data science and artificial intelligence experts.

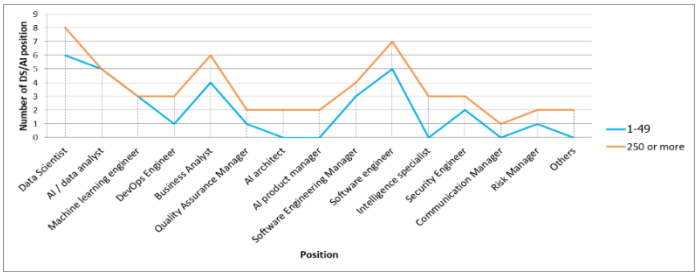


We sent a questionnaire to several companies with varying number of employees. We received a response from 59 companies from Thailand (23), Indonesia (26), and Sri Lanka (10). This survey has three main objectives: 1) To evaluate the companies' current status and future needs for personnel in data science and AI related positions; 2) To receive feedback from companies on how an internship programme could best operate during the MSc programme; 3) To receive feedback from companies on how a short duration professional course could be delivered so that it addresses the companies' most immediate training needs by upskilling its existing personnel in data science and artificial intelligence. Next, we present the results per country.

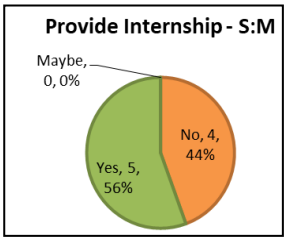
5.2 Thailand

In Thailand we received a response from 23 of the companies we contacted. They include 9 small size company with less than 49 employers, and 14 larger companies with more than 250 employers. Unfortunately, we received no response from mid-size companies with a number of employers between 50 and 249 employers. Some of the companies are from the private sector (fintech, banking, ICT, and automotive related companies) but few are also from the public sector (healthcare and government). The companies and their answers to our questionnaire can be found in Appendix A, here we report a summary.

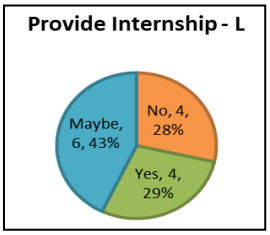
[6] Skinner, Saunders and Beresford (2004), Towards a shared understanding of skill shortages: differing perceptions of training and development needs.

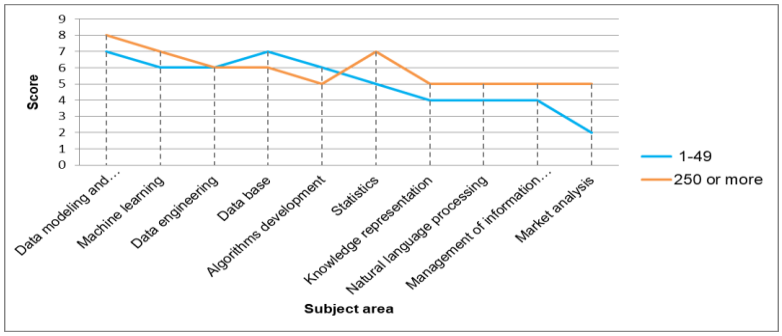


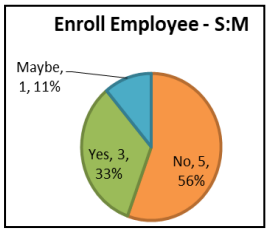
The diagram to the right summarizes the current available positions in data science and artificial intelligence. Overall, the highest demand for job positions in these fields are data scientist, software engineer, and business analyst, respectively. The demands of job positions follow a similar trend for small size as well as large size companies. For both the highest demand is for data scientist, AI/data analyst, software engineer, and business analyst.

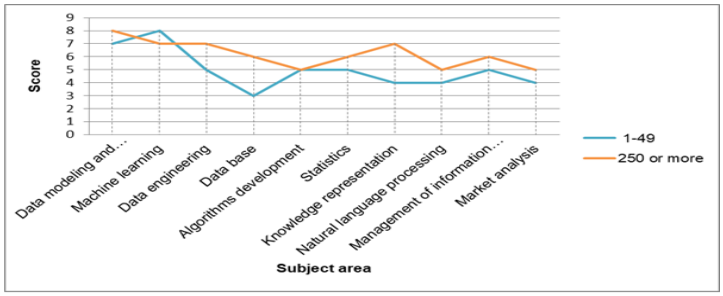


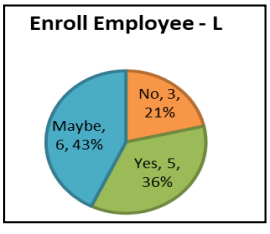
About half of small-size companies provide internship opportunities for 1-3 students with a duration of 3 months or less (diagram on the left). Large-size companies can offer more places to students for internships, but only 29% answered positively, although 43% answered that maybe it can be an option (diagram on the right). Before starting the internship, companies require from students to have a strong background in programming (Python, R, SQL, Anaconda, Tensorflow, Genetic Algorithm, and NoSQL), to be fast learners, to have basic skills to complete the end-to-end workflow (public cloud computing platform, SAP, Microsoft Office 365, Power BI, Microservice (Docker for Deployment project)), and to have knowledge on several subject areas as illustrated by the figure on the left. The priority lies on data modelling and management, databases, machine learning, data engineering, and algorithms development, and statistics.









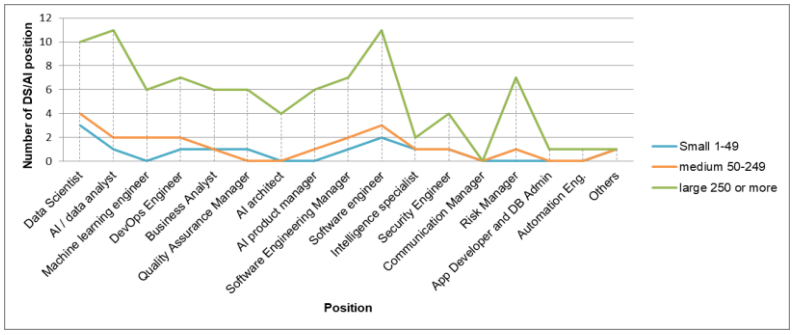


Only one third of the small size companies provide training courses, typically for 1-3 employees with durations varying from 1 week to 3 months (diagram on the left for small companies, and on the right for large ones). Regarding the training needs, most companies had demand for the professional

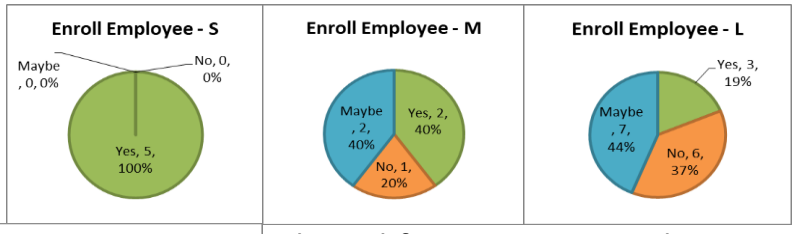
development courses for 1-6 employees with the duration of 1-2 weeks to 1 month. Companies think most of the courses are too focused on statistical theory and algorithm developments, and that training courses should also teach the data analyst to demonstrate their model. During training, the companies assume knowledge of programming languages like Python, R, Scala, PHP, noSQL, DB, DB Monet. The demanded subject areas during training courses are described in the figure above on the right. The most important subject areas are considered to be data modelling and management, machine learning, data engineering, and knowledge representation.

5.3 Indonesia

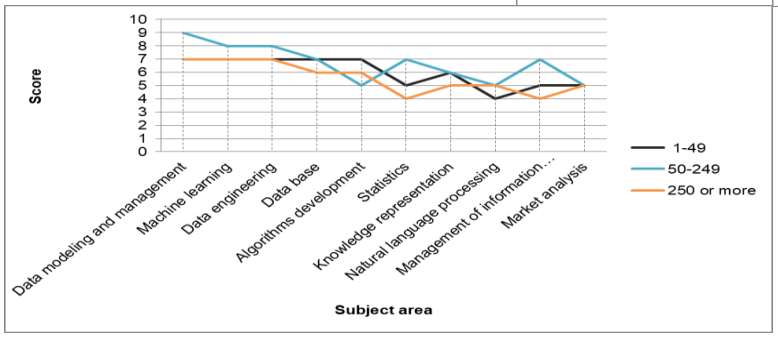
In Indonesia we received responses from 26 of the companies we contacted. They include 5 small size company with less than 50 employers, 5 mid-size companies with a number of employers between 50 and 250, and 16 larger companies with more than 250 employers. As for the case in Thailand, some of the companies are from the private sector (ICT, software and hardware vendors, aquaculture and agriculture related companies) but few are also from the public one (e.g. healthcare). The companies and their answers to our questionnaire can be found in Appendix D, here we report a summary.



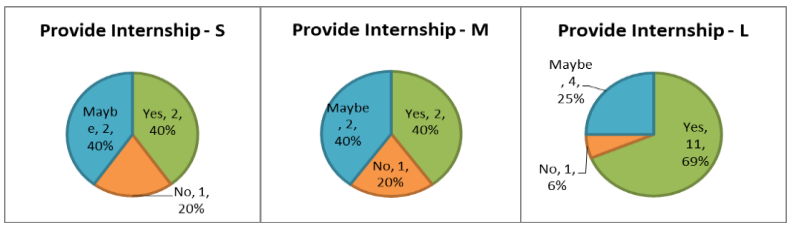
Overall, the highest demand for data science and artificial intelligent job positions are data scientist, software engineer, and artificial intelligence/data analyst, respectively. Considering only small-size companies (1-49 employees) and medium-size companies (50-250 employees), the highest demand for data science and artificial intelligent job positions is for data scientist and software engineer, respectively. Large-size companies (more than 250 employees), additionally, are also looking for artificial intelligence/data analysts.



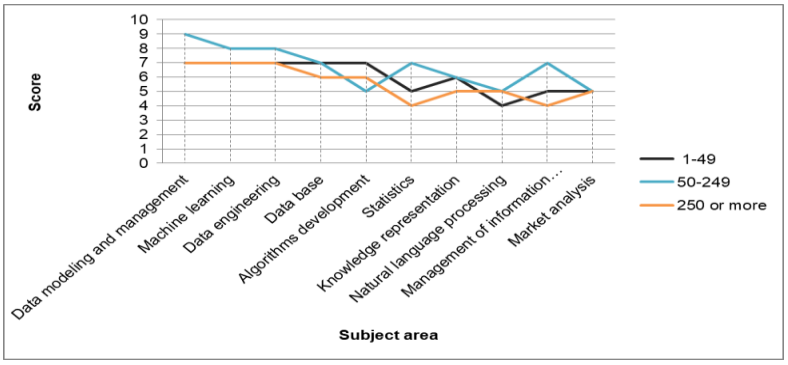
The demand for training courses for employee per size of the company is summarized in the diagrams on the right. Small-size companies had demand for training 1-3 employees in short training courses (1 week to 1 month). Most medium-size companies have needs to train different sizes of group of employees (1-3, 4-6, and 7-10 employees) for a period varying from 3 days to 2 weeks. Circa 37% of large-size companies do not have training



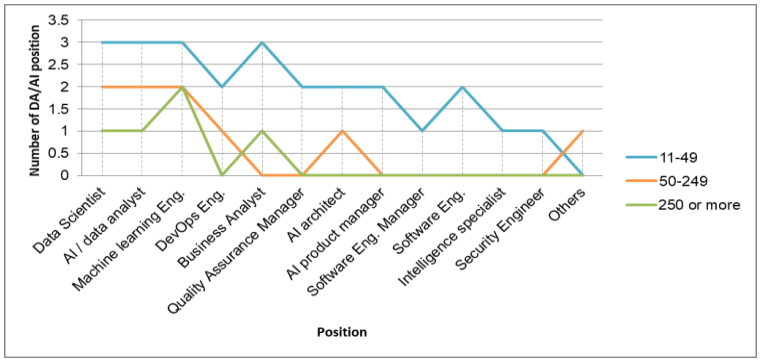
needs, while the remaining could positively consider enrolling about 1-3 employees for a training course of 1-week to 1-month. The most desirable topics among all the analysed companies corresponds to data modelling and management, machine learning and data engineering. The table on the left shows the average score given by the companies to topic areas for the training courses. Courses on statistics, natural language processing, and management of information systems are considered, on the average, least interesting for employees.



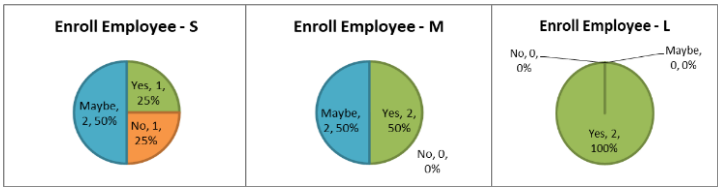
Most companies, independently from their size, might be able to provide 1 to 3 internship placements for a period of 3 months or less. Larger companies are however more in favour towards offering internships (94% consider that possibility). All companies would require students to have knowledge on data modelling, machine learning and data engineering. In addition, small size companies require knowledge on data bases and algorithms. Natural language processing is felt in general as not really necessary. Interestingly, also statistics is one of the knowledge areas that is not really required from students by large companies. The diagram on the right summarize the average score given by the companies on the required knowledge of a student for an internship. Knowledge is also required of programming languages like Python, Tensorflow, R, SQL, Java or any web-based programming Language, and of statistical and data visualization tools like Hadoop, D3.js, Microsoft Power BI, SAP, or similar.



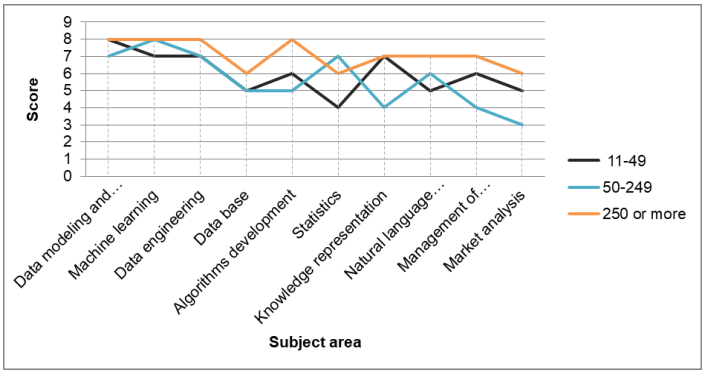
5.4 Sri Lanka



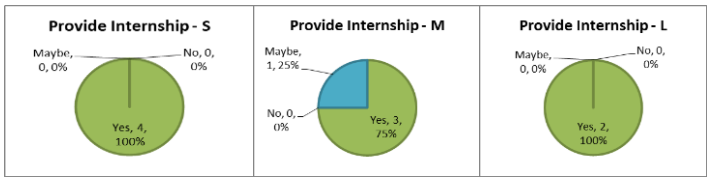
In Sri Lanka we received a response from 10 companies, 4 small-size, 4 mid-size, and 2 large ones. They are all private and all in positioned in an ICT elated area. Overall, the highest demands for data science and artificial intelligence job positions are machine learning engineer, data scientist, and artificial intelligence/data analyst, respectively. Small-size and large size companies additionally are in needs of business analysts. The companies and their answers to our questionnaire can be found in Appendix E.

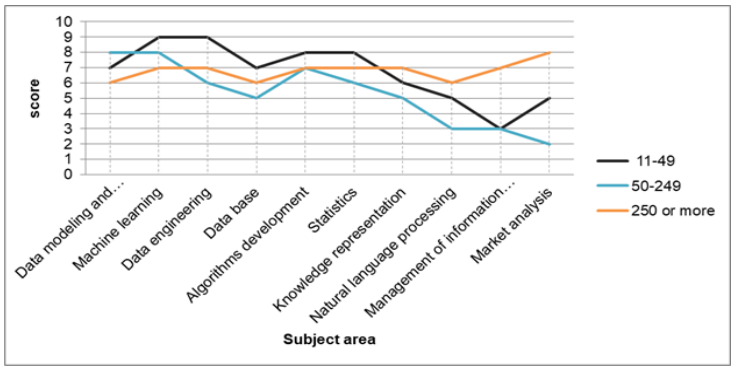


Most companies may consider enrolling some employees to training courses lasting between one to four weeks. All medium-size companies tend to support training courses for 1-3 employees per year with a duration ranging from a few days to a month. Large companies are interested in short training courses of at most 3 days. The most important subject areas for the courses are Data Modelling and Management, Machine Learning, and Data Engineering. Large companies find Algorithms also very important. The least important subject areas are Databases, Statistics and Market Analysis, probably because the employees already have knowledge on these subjects and do not need extra training. The table below summarize the average scores given by the companies (1 to 10) to different subject areas that could be considered for a professional training of their employees.



All small and large-size companies could offer 1-3 internship placements for students with the duration of 3-6 months. Mid-size companies could do the same but for more students: 1 to 10. Students would be required to have knowledge of programming language such as Python, TensorFlow, R, SQL, and to be able to use Matlab, Hadoop, and Git. There are large differences between what is considered important as knowledge of the students before an internship, depending on the size of the company. Large companies prefer their interns to have knowledge of Market Analysis, whereas small and medium size companies prefer students to have knowledge of Machine Learning, Algorithms, and consider Market Analysis as the least important subject. Other topics that are considered important are Data Modelling and Management and Data Engineering. The table on the left summarize the findings.





5.5 Conclusion

In Thailand, Indonesia, and Sri Lanka companies are willing to provide internships, especially to students with knowledge in data modelling and in machine learning. Strangely enough, natural language processing is not considered relevant, while applications using chatbot are very much in use by all companies.

6. Gaps Analysis in Master Curricula through South Asia

If you bend-over to analyse a gap too long, you'll probably fall into it.
--- Ryan Lilly

6.1 Introduction

Economic and social development increasingly depend on innovation. Universities have a potentially important role in driving innovation and development. They can do so both through their role in carrying out research and development and by training workers for the knowledge economy. In the previous chapter we have seen that companies in the South Asian region have a great demand of employees in the field of data science and artificial intelligence. The need for a master degree in these fields is vital for companies to grown larger and keep pace with the advancement in the rest of the world, more importantly in USA, China and Europe. Higher education is seen to have an ever more important role not only in human resource development but also in bringing students to companies through internships and in the regional workforce education via professional training.

Higher education in computer science across South Asia faces new challenges: maintaining and improving education quality, even in the face of serious financial constraint; improving the relevance of curriculum and instruction at a time of rapid change in labour market needs (as we explored in the previous chapter). Extraordinary effort needs to be made in diversifying curricula, and experimenting with new instructional delivery systems, as discussed in the European context in Chapter 2.

In view of the companies' needs analysis of Chapter 5, based on the current situation of the master education in data science and artificial intelligence as sketched in Chapter 3, and using the experience of similar European programmes as discussed in Chapter 2, in this chapter we present the result of a gaps analysis. We used a focus group discussion to identify needs and gaps that have to be addressed in the development of master's degree programme in data science and artificial intelligence in Thailand, Indonesia and Sri Lanka. In particular, the following issues and questions are investigated:

Issue 1: Identification of subject areas related to data science and AI that are most in need in the country (and wider region);
Issue 2: Identification of a set of skills that the curriculum should develop and promote;
Issue 3: Teaching and learning processes that are appropriate for the curriculum (project-based learning, professional certifications, trainings, practical and industrial projects, workshops, internships, research, theses, etc.);
Issue 4: Required resources, facilities, tools, as well as support that the universities should provide (what we lack or need)

Issue 5: Any concerns and opinions regarding internships and job opportunities that the curriculum should provide

The discussion comprises at least 15 participants from three groups, namely academics (7 or more), ICT industrial sector (4 or more), and students (4 or more), and hence allowing us to obtain significant inputs in developing a balanced Master's degree programme. The participants were asked to participate in a discussion and share their inputs and opinions regarding the development of data science and artificial intelligence curricula, including their opinion on job and skill demands, needs for new courses, pedagogy and instructional methods, internships, job opportunities, required tools, resources and facilities. Specifically, the five issues mentioned above were discussed in detail. Information about the focus group participants for each country can be found in Appendix F.

6.2 Focus group discussion Thailand

This section presents the findings from the focus group discussion in Thailand from 3 different perspectives: ICT industrial sector, student and academics, respectively.

Issue 1: Identification of subject areas related to data science and AI that are most in need in Thailand (and wider region)

ICT Industrial Sector Perspective:

- Jobs with currently strong need and most preferable by the industry sectors are Data Engineers. They should have the abilities to perform ETL (Extract Transform Load), Data warehouse manipulation (Big data repository), and data preparation skills which are found lack at present. Currently, many companies employ data scientists working as a role of data engineers. Data Scientists are normally expected to use the latest data technologies so as to solve business problems for strategic advantages.
- The curriculum should provide both fundamental and advance core courses composing of SQL, NoSQL, Data Ingestion, DDL, Data Warehouse Logic, Programming, SQL for Hadoop big data platform. ICT Industrial deal with bank sector proposed Probability and Statistics, Programming in Python, Linear Algebra. Machine Learning, Neural Network, Deep Learning, Genetic Algorithm, Text Mining and IR, Data Structures and Algorithms, NLP (Natural Language Processing) and Text mining are also essential to learn.
- One thing that should be noted is that students should understand how to utilize when applying Machine Learning algorithm in their works by using ready-made libraries or off-the-shelf component. These useful tools can lead to further development as a product such as Python, Hadoop, Matlab, and Weka are needed to learn in order to improve capabilities of the products. Students should enhance Business Analysis or Market Analysis which will certainly be used in real work. In practice, working together between the Business user and

Data scientist will have a greater chance of success rather than trying to get Data scientist to understand Business on their own.

- Start-up companies usually require all tasks in one employee with wide and various knowledge, such as business, legal, marketing, and the ability to produce products. Machine Learning, Deep Learning and Genetic Algorithm
- Data governance also should be taught in this program due to concerning the capability that enables an organization to ensure that high data quality exists

Student Perspective:

- Basic subjects such as Data modelling and management (Data structure, Data Pre-processing) and more advance courses related to Big Data, Data mining, Mechanical learning and Robotics are expected to be taught according to student needs. The special topics should be brought from industry-based projects on different domains to help student better understand the core concepts of Data Science and Big Data.
- Students should be given the opportunity to choose the subjects that interest them. Moreover, teaching courses should not specificied in some particular domestic job markets so that students can find jobs worldwide with higher income.

Academics Perspective:

- Teaching and learning should lay down the foundation of core knowledge together with the ability of using the various tools and software i.e. Python that will be implemented to develop the product in practical. Moreover, collaboration with industrial section can lead to students' better learning by applying different real case studies.
- The example of important fundamental courses recommended by lecturers are programming development such as Algorithm, Logic, Data structure, Programming and Data management such as Database / SQL. - SQL, NoSQL. In addition, Discrete mathematics, Linear algebra, Probability, Statistics, Operation research are referred by ICT industrial that should be included into core courses such as Introduction to Data Science, Machine Learning, Data Mining and so on.
- The target groups of learner may be various depending on learner's objectives and business companies involved in that region i.e., some learners with their own data/information will have a desire to learn how to analyse; they focus on business analysis and require the algorithm at the basic level. Others may focus on in-depth algorithms for the maximum efficiency product customization and development and these in-depth learning should be also leaded to research capabilities as well.

Issue 2: Identification of a set of skills that the curriculum should develop and promote

ICT Industrial Sector Perspective:

- The priorities of skills recommended from the industrial sector are communication, analytical skills, empirical skills, presentation and Argumentation, respectively. Moreover, Data visualization (Infographics) should be taught to effectively present all analyzed results to customers or relevant users. In addition, Teamwork, Problem solving skills and Attitude are important as well.

Student Perspective:

- According to students' points of view, the communication skills are very important. systematic thinking, effective communication and problem solving should be taught through this curriculum.

Academics Perspective:

- The important skills that student should gain through this curriculum are an in-depth understanding, Analytical skills, Empirical methods, Modelling, Teamwork, Written and oral communication, and Argumentation and presentation.

Issue 3: Teaching and learning processes that are appropriate for the curriculum (project-based learning, professional certifications, trainings, practical and industrial projects, workshops, internships, research, theses, etc.)

ICT Industrial Sector Perspective:

- According to industrial sectors' points of view, students should gain the practical skills through working on real world problems in an industry collaborative environment through projects. Project base learning should be taught due to it is a dynamic approach solving real-world problems to gain knowledge and skills. Through this learning experience, students will be able to investigate and respond to an engaging and complex question, problem, or challenge.

Student Perspective:

- Students prefer continuous project assignments to typical study. In addition, assignments should suit each student's ability. Workshops and invited speakers should be organized in order to provide students usable skills, and recognized qualifications.
- Industrial internship is also vital in order to boost student's practical skills while doing the projects with industrial partners.
- A variety of tools applied in industry nowadays should be provided during the teaching and learning process.

Academics Perspective:

- The teaching and learning process should offer two different tracks: Industrial project/Internship track and Research track.
- Students should experience to practice by using practical problems from industries. In case of big data study, there should be enough dataset for doing experiments. Therefore, it is very

important for the university to have collaborations with various industries leading to better teaching and learning.

- However, the business confidentiality is a sensitive issue. This leads to the cooperative research agreement between the university and industrial sectors instead of requesting data set /information used for teaching and learning.

Issue 4: Required resources, facilities, tools, as well as support that the universities should provide (what we lack or need)

ICT Industrial Sector Perspective:

- Instead of paying for expensive software in learning, students may suggest to use some free software or education version from various vendors such as AWS educate, Google cloud for education, Hadoop, Python, Spark cluster, Cloudera, Tensorflow, etc.
- Seeking strength in partnerships and expanding research capabilities, the collaboration ecosystems should be well planned.
- The start-up labs and product show case should be set to meet the entrepreneurial requirements and must be equipped with laptops, database, ecosystem (AI big data), licensed software and necessary things.

Student Perspective:

- Inviting speakers from industrial sectors to give a talk can provide students more relevant ideas.
- The laboratory equipment should be prepared and ready to use for the most effective learning such as hardware resources for high computation, Machine learning, Neural Network and Parallel Processing.
- The tools used in teaching and learning should be a tool that is in line with those commonly used in the industry.

Academics Perspective:

- To practice data processing in big data experiments, the high-performance laboratory equipment should be prepared, and the tools should be suitable for learning various algorithms both basic and advance level. Academics proposed that ICT Industrial should cooperated and support with university to provide up-to-date tools.

Issue 5: Any concerns and opinions regarding internships and job opportunities that the curriculum should provide

ICT Industrial Sector Perspective:

- Inviting experts from companies is also a good idea to inspire students with valuable practical knowledge focused on graduate ready skills for future careers.

- Students should be encouraged to join various community groups in order to update new knowledge and modern tools. Moreover, students should have a chance on some practical works related to the special topic that they are interested in.
- In order to apply for a job, join an internship among other competitors, a student portfolio is also important because that reflects all accomplishments, skills, experiences, and attributes. Students should learn how to professional create.

Student Perspective:

- Students expect to have the Industrial Training/Internship opportunity practiced by private companies/organizations so that they can understand organizational environment and perform well towards organization objectives.

Academics Perspective:

- Industrial Internship is good way to give opportunity with real case study for student's practice, which job opportunities in data science and artificial intelligent are needed to solve to explore insight data and can be applied how to get business value following the ICT's requirement. However, the academics concern about the taking course time that should be combined to research work.

Conclusions

The teaching and learning of the data science and artificial intelligence curriculum must be able to meet both the needs of the industry and the interest of students. Teaching and learning should lay down the foundation of core knowledge that is sufficient for learning data science and artificial intelligence, including (1) Mathematics and statistics groups such as Discrete mathematics, Linear algebra, Probability and Statistics, Operation research (2) Programming development groups such as Algorithm, Logic, Data structure, Programming and (3) Data management groups such as Database and SQL.

The need for knowledge of Business analytics and Machine Learning depends on whether basic or in-depth learning depending on the student group characteristics and the objectives of the student's study. Therefore, the developed curriculum should have different tracks to be chosen between (1) Applying data analytics techniques to solve problems, analysing data for the industry and (2) Researching to develop deep knowledge in DS / AI science, which leads to enhance the higher performance for the particular industry.

The laboratory equipment should be prepared and ready to use for the most effective learning. The tools used in teaching and learning should be a tool that is in line with those commonly used in the industry. Students should have the skills to use those tools that can lead to the product development, not just a tool that running or testing as experiments. The course must be built up the collaboration with the industry in applying the real case study for teaching or for students to experiment. In particular, students should practice as internship in a company, or do some industrial projects.

Not only hard skills that students can gain through education, training programs and internship training, but also soft skills for data science and artificial intelligence students are very important. The soft skills

that are suggested to be developed including (1) Communication skills, (2) analytical skills/ empirical skill, and (3) Presentation skills, both in terms of how to present and how to use visual tools, respectively. Students should have Data visualization / Infographic skills in order to be able to effectively present the results of data analysis to the user or executive member.

6.3 Focus group discussion Indonesia

This section presents the findings from the Indonesian focus group discussion from 3 different perspectives: ICT industrial sector, student and academics, respectively.

Issue 1: Identification of subject areas related to data science and AI that are most in need in Indonesia (and wider region)		
Industrial Sector Perspective:	Students Perspective:	Academics Perspective:
• Data scientist and AI-Engineer are highly needed. • Candidates should understand Probability theory and Statistics and have ability in Programming (preferable Pyhton). • As AI engineer, the candidates should have abilities and understanding in AI-related knowledge, such as Machine Learning, Neural Networks, Information Retrieval, Natural Language Processing, Computer Vision, Language and Speech Technology, Computational Intelligence • As Data Scientist, candidates should have abilities and understanding in Text mining and Information Retrieval, Machine Learning, Data Modelling and Management. • Start-ups tends to request full-stack capability from the candidates, therefore if the programme can provide wide spectrum of data science and	• Related to curriculum, students tend to follow what was prepared by the University or data science and artificial intelligence programme. • Based on questionnaire, as long as the programme will lead them for becoming data science and artificial intelligence related master, they will follow everything which is necessary and necessity that have been prepared by University or study programme.	• To join the master programme in data science and artificial intelligence, candidates must already be familiar with programming. • For USU&Unsyiah: Candidates need to have knowledge of basic subjects such as: Probability Theory and Statistics, Linear Algebra, Discrete Mathematics. DS related subjects such as Management Information Systems, Knowledge Representation, Data Structures and Algorithms, Data Modelling and Management. AI related subjects such as Machine Learning, NLP, Business and Market Analysis, Computer Vision, Ethical, Legal and Social Aspects of AI, and some selective subjects such as; Language and Speech Technology, Robotics, and Virtual Reality and Gaming. • For ITB: knowledge in Data structures and algorithms is

artificial intelligence knowledge, they do not have any objection at all.		pre-requisite for the candidates in order to join the master programme in data science and artificial intelligence. After joining the programme, students learn several subjects which are related to data science and artificial intelligence, such as discrete mathematics, text mining and Information retrieval, probability and statistics, machine learning, NLP, business analytics, computational intelligence, computer vision, knowledge representation, big data, and AI engineering which related to deployment model for large scale environment.

Issue 2: Identification of a set of skills that the curriculum should develop and promote

Industrial Sector Perspective:	Students Perspective:	Academics Perspective:
• Industrial partners also agree to make the curriculum to be 80% about technical skills and 20% about soft skills. • As per priority, they put analytical skills as the first skills that should be developed by the candidate or be developed by curriculum, followed by Empirical methods, Modelling, Teamwork, Written and Oral Communication, Argumentation and Presentation. • If possible, after graduation, they also able to have capability to teach and share knowledge with other people.	• The curriculum should provide every aspect of skill that is needed or required by the industry. • Students tend to believe in the curriculum that will be provided by the University.	• Academics tend to agree to make the curriculum to be 80% about technical skills and 20% about soft skills. • Those skills include: Analytical skills, Modelling, Empirical Methods, Teamwork, Written and Oral Communication, Argumentation and Presentation, Reimplementation of Existing works, and continuous learning in term of comprehension of scientific articles. •

Issue 3: Teaching and learning processes that are appropriate for the curriculum (project-based learning, professional certifications, trainings, practical and industrial projects, workshops, internships, research, theses, etc.)

Industrial Sector Perspective:	Students Perspective:	Academics Perspective:
• Short course/training from the university is one of the alternatives to upgrade the skill, but they tend to see the content of that short course rather than the certificate that will be received. • Internship is one of the ways to recruit the talents, therefore they are very welcome if the internship is a compulsory in this study programme. • Industrial partners also accept students who want to work as part-timer. • Exploration of tools in regards with data science and artificial intelligence can be beneficial for industrial partners.	• The curriculum should provide every aspect of skill that is needed or required by the industry. • Students tend to believe in the curriculum that will be provided by the University. • The university should organize classes which invites industrial partners. They are necessary to get insight as well as to motivate the students.	• Academic staff tend to easily accept any methodologies to be conducted during the programme. • For USU & Unsyiah: they suggested project-based learning for final project or theses writing. An internship is necessary for the students to be aware with the industrial problem, and this internship should be done within 2 – 3 months during the semester break, and it will be great if offer of internship for the students are coming from Industrial partners. • For ITB: they suggested project-based learning, practical-and-industrial projects, and internship as the teaching methodologies for developing the practical and awareness of the real application. As research and theses, as well as joining short course from the university can be used as an alternative to strengthen the understanding in the subjects. ITB also agree for the internship to be done within 2 – 3 months during the semester break, to expose the students with the real-world application.

Issue 4: Required resources, facilities, tools, as well as support that the universities should provide (what we lack or need)		
Industrial Sector Perspective:	Students Perspective:	Academics Perspective:
• Industrial partners hope that universities will provide a specific room which is dedicated for discussion of data science and artificial intelligence, room that caters students' discussion, especially if it is related to technical issue.	• Facilities such as high-end workstations, dedicated discussion rooms and laboratories for data science and artificial intelligence, suitable internet connection for downloading big datasets.	• Universities in Indonesia still lacking in certain ways, such as suitable workstations for students for big data analysis, cloud computing resources, private servers for big data, internet connection that is more reliable, research lab that accommodates data science and artificial intelligence programme specifically, and also licenses in regards with software and tools.

Issue 5: Any concerns and opinions regarding internships and job opportunities that the curriculum should provide		
Industrial Sector Perspective:	Students Perspective:	Academics Perspective:
• Industrial partners tend to prefer longer period for internship (around 3-6 months). Although it is very rare, but students with longer period of internship do not mean to be performed better in meeting their expectation. In regards with this time period, industry will accept any period that is designed by the University (minimum 2 months).	• For internship, students only follow what the curriculum, if it is 3 months, then they will happily do it. • Job opportunities are not an issue in our Universities, but more access to giant multi-national companies will be better. • Students prefer to become entrepreneur, applying their knowledge after maximum 5 years working for the company.	• Inviting external experts to classes to discuss future career possibilities is considered as very good idea to inspire students.

Conclusions

Based on Industrial needs in Indonesia, Data scientists and AI-engineers are still needed for multiple years ahead by IT-related jobs/companies. Therefore, from the industry point of view, data science and artificial intelligence programmes are beneficial for the industry in Indonesia. Among Universities, we agree to put capabilities in programming as compulsory for the students before joining the data science

and artificial intelligence programme, however discrepancies happen in certain level of knowledge as pre-requisite by students before joining the programme because of each university policy. Detail of the course mapping can be found in Table 1.

Between academics and Industry, we both agree for the curriculum to be 80% for technical skills and 20% for soft skills that can be included within every subject as class-project or tasks. To produce graduates who are more ready for the job, academics and industry also agree to put industrial internship inside the curriculum. Even though Industrial partners prefer longer period for the internship (6 months) but industrial partner agree to follow the period of internship as designed by University (3 months). Besides the internship, Industrial partner also open for the students to participate as part-timer in their office. The industry also encourages students to explore more in regards with the data science and artificial intelligence tools, so the students should follow the current trends in data science and artificial intelligence. Short course in specific domain of data science and artificial intelligence is interesting for industrial partners, as one way to upgrade the skill of the employee. The industrial partners tend to look for the contents of the short course regardless the certification.

The laboratory equipment, suitable internet connection should be prepared and ready to use for the most effective learning. The tools used in teaching and learning should be a tool that is in line with those commonly used in the industry. Students should have the skills to use those tools that can lead to the product development, not just a tool that running or testing as experiments. The course must be built up the collaboration with the industry in applying the real case study for teaching or for students to experiment.

Not only hard skills that students can gain through education, training programmes and internship training but also soft skills for data science and artificial intelligence students is very important. The soft skills that are suggested to be developed including (1) Communication skills, (2) analytical skills/ empirical skills and (3) Presentation skills, both in using visual tools and in preparing content, respectively. Students should have Data visualization / Info graphic skills in order to be able to effectively present the results of data analysis to the user or executive member. (4) Teaching skills, for them to be able to teach their junior/colleagues.

6.4 Focus group discussion Sri Lanka

This section presents the findings from the focus group discussion in Sri Lanka from 3 different perspectives: ICT industrial sector, student and academics, respectively.

Issue 1: Identification of subject areas related to data science and AI that are most in need in Sri Lanka

ICT Industrial Sector Perspective:

- Industry partners strongly believe in the necessity of a well-defined data strategy, which enables the organizations to structure the data centric activities and treat data as an asset rather a commodity. Without such long-term view, organizations find very difficult to focus activities on

a common goal. Hence, the identification and implementation of suitable data strategy must be included in the curriculum, as per their opinion.

- Emphasis was given on the domain knowledge and the ability to apply data science knowledge within the domain, as it is important for students when they have to work in the industry projects.
- Blockchain was another subject among industry concerns. This topic was described as the next biggest thing after the internet and it has already created lot of excitement within the context.
- All industry participants agreed upon the value of studying data security, as it is important to protect data. Industries were dealing with lot of sensitive information of their customers.
- Laws and ethics in data science was identified as essential to restrict how you handle data. This is very important in the industry, and the employees who wish to join in the fields of data science should have awareness in them. Ex: GBDR (General Data Protection Regulation), For instance, nowadays the customers have a right to forget their data at a given company upon their request.
- Since the technology is continuously changing, the students need to have the knowledge in a wide range of tools and the ability to adapt to new tools quickly. For that reason, the students should essentially have the exposure to different tools like, Hadoop, Spark, Cloudera along with the practice in handling large data sets.
- Real time data handling is widely used in the industry and for that, the tools like Kafka is also useful for complex event analysis and processing.
- Data Governance – This refers to the overall management and caretaking of the data covering its usability, integrity (i.e. making sure the data is of a good quality, that you know where it has come from and that you have the right to use it as you need) and security. As discussed, this can be inserted in the curriculum as a part of data laws and ethics.
- Data Science Methodology – This defines the essential steps that must be followed towards a comprehensive study. The methodology steps can be; Problem to Approach – Business understanding, Data preparation – Cleaning & transforming data, Data understanding – Can the available data answer the question at hand?, Analytic approach – Statistically/machine learning driven?, Modelling – In which way we can use the data to answer the given question, Evaluation – Does the model answer the question, or does it need adjustments?, Deployment – Making the solution to be used by the stakeholders, Feedback – Feedback from the users and refine the model as needed.
- Lack of Mathematics knowledge in the employees make difficulties in understanding the domain problems related to machine learning. Mathematical optimization, linear algebra and multivariate calculus are also important for a master programme in this scenario.
- Subjects related to statistics and mathematics such as Statistical Data Analysis, Statistical modelling and statistical learning theory, Mathematical Optimization were proposed.
- In addition to theory, the applications of Deep Learning, Machine Learning and Data Science were mentioned as necessary for the course.

Student Perspective:

- Students pointed out the importance of learning Advanced Database Management Technologies such as NoSQL databases (e.g., HBase, Cassandra) for storing and processing big data in distributed environments.
- Containerization has risen in recent years as a ready to use solution to work around platform differences across a variety of applications from single user development environments to highly

distributed production environments (e.g. Google Cloud, Docker). Students were interested in studying trendy topic like this.

- More practical knowledge is important for doing a research successfully. May be students are not studying what they want for their research implementation. For that, more practical sessions on a variety of tools in DS&AI is advantageous for students.
- Vector calculation was also suggested as important for data analysis.
- Students also proposed Block chain programing, as it would be a latest topic in DS&AI in near future.
- Students further suggested financial data analytics with AI as a part of regression analysis which would be useful in analysing financial data.
- In addition, topic such as Expert systems in AI, Deep learning for AI and Mahout for large scale Machine Learning were also proposed as essential for in depth knowledge in DS&AI.

Academics Perspective:

- Not having enough data for students to practice data science concepts and to learn the techniques is a problem. Companies might also be reluctant to share data due to their sensitivity. Hence, we have to encourage students to use open data sets that are now largely available.
- Advanced network technologies, High Performance Computing and Distributed Computing were identified as important subjects for students to have the knowledge used when working with large data sets. At least, the basics of them were proposed to include in the curriculum.
- To understand advanced statistical tools, enough Mathematics background is essential. Nonlinear programming and graph theory were identified as significant for neural networks.
- Matrix factorization was also proposed as important for big data analysis.
- Basic financial knowledge and economics with foundation in financial knowledge was identified as useful to interpret values and matrixes.
- Business management to give emphasis on the importance of business and understand the business perspective in terms of DS&AI was suggested for the curriculum.
- Scientific Report Writing should be developed for students to be able to effectively describe the data appropriately.
- Operating hardware components that are related to programing like GPU was found useful for DS&AI.
- Big data processing and large-scale machine learning and Deep Learning were suggested as core topics that should be included in the course. Under the course, following topics such as; Distributed databases (e.g. HBase), Data Transfer (flume, sqoop), Large Scale machine learning (Mahout, MLib), Real time data analytics (kafka, storm), Security – (zookeeper) and Administration – (Ambari) were proposed.
- Introduction to Artificial Intelligence should be included to provide foundation knowledge in AI. Advanced AI with topics in the areas of knowledge representation and general artificial intelligence was also suggested.

Issue 2: Identification of a set of skills that the curriculum should develop and promote

ICT Industrial Sector Perspective:

- One essential skill which was pointed out here was Presentation skills. In a corporate setup, how you tell your story according to the gained insight from data to a non-data savvy business leader is critical. No matter how important your insight is, if the final communication fails everything starts falling around it. So, building your story with right visualization in a clear concise way is a critical skill to develop.
- Documenting data science results and implementations in a standard way was proposed as very important for better understanding and to avoid unnecessary changes.
- Having good analytical skills to understand the business problem, being innovative, meeting deadlines and being productive, project management and effective communication skills in scientific communication were also suggested as required skills for the industry.

Student Perspective:

- M.Sc candidate should be able to do effective presentations and communication. In addition, team working, scientifc writing skills and and the self-education skills are important.

Academics Perspective:

- Ability to use a version control system
- Ability to adapt into changing technology

Issue 3: Teaching and learning processes that are appropriate for the curriculum (project-based learning, professional certifications, trainings, practical and industrial projects, workshops, internships, research, theses, etc.)

ICT Industrial Sector Perspective:

- Hackathon related data science
- Github open source projects for every student.
- Project based learning always gives an opportunity to practice and replicate end-to-end (problem approach -> deployment + feedback) data science methodology.
- Professional certification should be awarded and ideally it could be in a digital form so participants can add them into their on-line profile (saves paper too).
- Suggest having cross department collaboration to reinforce researches where Data Science students can provide the data specific insights to the research.

Student Perspective:

- All the suggested teaching and learning processes given here can be used.
- If the students can also get the professional courses, that would be very helpful in their studies as well as in getting to industrial opportunities.

Academics Perspective:

- A compulsory module for research should be offered in the course for doing extensive individual studies on a researchable topic. The industry companies were invited to share research areas with students and supervise them accordingly. An evaluation criterion should be set to publish at least one research paper.
- Offering students a compulsory mini project to get the exposure for different technological environments common in the industry, along with assignments shared by the industry.

- In addition to classroom teaching, project based learning should be used.
- Sort term training courses should be available for students as well.
- Problem based learning

Issue 4: Required resources, facilities, tools, as well as support that the universities should provide (what we lack or need)

ICT Industrial Sector Perspective:

- If we can initiate a discussion with Google, Amazon, Microsoft, IBM, and other large companies to partner up to utilize their cloud infrastructure to the participating students at a concession rate it will ease up lot of infrastructural burden to be provide to students. We never be able to catch these tech giants on these aspects. This would add a great value to the course content itself.
- Use machine learning as a service. Ex: AWS (Amazon Web Services) – SageMaker

Student Perspective:

- Learning Management System to make different aspects of the course online. (lecture material, evaluation, etc)
- It would be better if there are more poweful machines for DS&AI projet implementations.

Academics Perspective:

- Facilities will be required to get high performance computing power if necessary.
- Data repositories and clean data are much needed.
- The university has a Data Science Unit (DSU) that is implementing a data sharing platform, where the industries can share their preferred data sets, and ultimately this would be beneficial for researchers. It consists of a computer cluster using 3 powerful machines with the following specifications[7]:

Model	Processor(s)	Memory	Hard drive capacity	GPU	GPU Memory
Dell Poweredge R740xd	2 x Intel Xeon Gold 6130 2.1G, 16C/32T	128 GB	3 TB	Nvidia P40	24 GB
Dell XPS 8920	3.6 GHz Intel Core i7-7700 Quad-Core.	32 GB	2 TB	Nvidia GeForce GTX 960	8 GB
Dell XPS 8910	3.4 GHz Intel Core i7-6700 Quad-Core	16 GB	2 TB	Nvidia GeForce GTX 660	2 GB

We have set up a Hortonworks cluster for Big Data analytics. Moreover, we tested the applicability of Hadoop tools such as Pig, Hive, HBase, Sqoop, Flume, and Mahout, and prepared documentation explaining students/staff how to use those tools on our cluster. In order to

[7] Numbers mentioned in the table may be slightly different.

distribute computing resources between multiple users to run their programs concurrently, we are using SLURM as the scheduler and resource manager. We have also prepared documentation for how to run programs using SLURM. All the documents are available in our Data Science Unit website.

The cluster supports effective utilization of resources for deep learning projects using python frameworks such as Pytorch, Keras, Tensorflow, and Theano. Moreover, we provide additional support for the users who remotely run R programs using Rstudio Server that is running on our cluster. Recently, we conducted a deep learning lab session for the M.Sc. in Computer Science 2018 students, where around 50 students remotely accessed our cluster and ran their programs.

We assume our resources would be adequate to start an MSc in Data Science and AI course. In the future, we may purchase additional machines and integrate them with our cluster. We hope we may not need to rely on a cloud platform. If need increases, we may consider using cloud infrastructure along with our cluster. We keep updating our cluster to incorporate modern tools that are necessary for the MSc in Data Science and AI.

Issue 5: Any concerns and opinions regarding internships and job opportunities that the curriculum should provide

ICT Industrial Sector Perspective:

- If the MSc student is already working in another company, providing projects for them with real data would be challenging.
- From the very beginning we must develop a strong communication strategy on the program so, stakeholders (including the industry) see the benefits right away. Academic and business sense needs to be properly intertwined in the overall message.
- Deep Learning use cases like face recognition can be done without data sharing issues.

Academics Perspective:

- Trainig in the idustry should be a compulsory module and its duration has to be a minimum of six months.
- Implement a proper NDA and MOU with the industry partners.

The focus group discussed an additional issue that is worth reporting:

Issue 6: Identification of short-term professional training courses in data science and artificial intelligence

ICT Industrial Sector Perspective:

- When a company sends someone for a short-term course, they expect quick results in return. A course on Exploratory Data Analysis was suggested to provide competencies in identifying data patterns, visualization and interpretation. Basic mathematics and statistics knowledge would be pre-requisites for interpretation of results. Delivering of the course must need interactive sessions, not just listening mode lectures. Minimum required duration was decided as seven days which should facilitate:

- Language: R or Python (Python is most widely used in the industry, due to distributed environments)
- Data preparation/ Data cleaning: Pandas/ dplyr (Domain knowledge is also important)
- Visualization: ggplot2, Seaborn, Matplotlib, Power BI, Tableau
- Data analytics: SAS, SPSS

- Courses that can support to gain the abilities in working on industry grade cloud infrastructure (IBM Watson, Microsoft Azure, etc.).
- Courses that facilitate the basic practice of using tools such as Github
- It would be ideal if the courses can cater to three types of audiences, such as, employees with software developing knowledge, employees with statistics and mathematics backgrounds and employees from management and physical sciences streams.
- A proper evaluation mechanism should also be designed to particular course, which would be credible to the company. If the exams are automated and available online, that would be ideal for professionals.
- Courses on Electronic Data Interchange would be useful as in most cases we assume that data is in the big data system, but actually we need to get a lot of effort to transfer data to there.
- Short courses for Python basics and standards, as well as for, TensorFlow, caffe, paddle, Keras, Pytorch and deep learning frameworks were mentioned.
- Object oriented programing basics courses for statistics and mathematics students.

Student Perspective:

- Providing detailed knowledge in a set of Hadoop ecosystem that are widely used in local industries (e.g., Apache Kafka for real time data processing).
- Course on Advanced Machine Learning in Python was also proposed, with the tools such as scikit-learn, Pytorch and Vowpal Wabbit. Content of the course was proposed as, Traditional machine learning algorithms, artificial neural networks, deep neural networks, convolutional neural networks, long-short term memory networks, word embedding, ELMo, Transformer networks and BERT.
- Courses in Big data and Hadoop Ecosystem was suggested along with the knowledge on the topics such as; Distributed databases (e.g. HBase), Large Scale machine learning (Mahout, MLib), Data Transfer (flume, sqoop), Real time data analytics (kafka, storm), Security (zookeeper) and Administration (Ambari)

Academics Perspective:

- If a particular course can provide elective and supportive subjects as well, then the participants will be able customize the course according to their competence level.
- A short course on how to use Python in a developer environment of a real data science project.

Appendix A Questionnaire for European Curricula

Q1: **Core topics** define the key areas of a curricula in Data Science and Artificial Intelligence. Which of the following core topics and skills are present in your master programme either as a dedicated course or as a substantial topic within one or more *mandatory* courses?

- ☐ Algorithmic Problem Solving (Search, Decision Making, Optimisation)
- ☐ Cognitive Science
- ☐ Computational Linguistics
- ☐ Context of Artificial Intelligence (History, Philosophy, Ethics)
- ☐ Intelligent Autonomous Agents and Multi-Agent Systems
- ☐ Interaction (Perception, Human-Computer Interaction, Communication)
- ☐ Knowledge Representation, Extraction and Reasoning (Data mining, deep learning)
- ☐ Machine Learning (supervised, unsupervised, reinforcement learning)
- ☐ Other, namely: ….

Q2: The set of core topics does not need, by itself, to constitute a complete master programme. Every programme may include additional **elective topics** relating to artificial intelligence, data science, computer science, cognitive psychology, but also to application areas such as bioinformatics and business analytics. Which of the following topics and skills are present in your master programme either as a dedicated course or as a substantial topic within one or more *not necessarily mandatory* courses?

- ☐ Cognition and Cognitive Modelling
- ☐ Computational and Cognitive Neuroscience
- ☐ Computational Intelligence
- ☐ Computer Vision
- ☐ Ethical, Legal and Social Aspects of AI
- ☐ Evolutionary Algorithms (Genetic Algorithms, Evolutionary Computing)
- ☐ Language and Speech Technology
- ☐ Neural Networks
- ☐ Perception (Computational and Natural)
- ☐ Reasoning under Uncertainty
- ☐ Reinforcement Learning
- ☐ Robotics
- ☐ Text Mining and Information Retrieval

- ☐ Virtual Reality and Gaming
- ☐ Web and Artificial Intelligence
- ☐ Others, namely: ….

Q3: **Support knowledge** defines the basic body of know-how needed to successfully understand the programme courses. Which of the following support knowledge are students assumed to have *when starting* your master programme?

- ☐ Programming in Python
- ☐ Programming in Java
- ☐ Programming in C++
- ☐ Programming in other languages
- ☐ Databases
- ☐ Algorithmic
- ☐ Data Structures
- ☐ Digital Systems
- ☐ Propositional and Predicate Logic
- ☐ Calculus
- ☐ Discrete Mathematics
- ☐ Linear Algebra
- ☐ Probability Theory and Statistics
- ☐ Other, namely ….

Q4: Which of the following support knowledge is provided by your master programme as a dedicated but *not necessarily mandatory* courses?

- ☐ Programming in Python
- ☐ Programming in Java
- ☐ Programming in C++
- ☐ Programming in other languages
- ☐ Databases
- ☐ Algorithmic
- ☐ Data Structures
- ☐ Digital Systems
- ☐ Propositional and Predicate Logic
- ☐ Calculus

- ☐ Discrete Mathematics
- ☐ Linear Algebra
- ☐ Probability Theory and Statistics
- ☐ Other, namely ….

Q5: Apart from knowledge and understanding, every master programme support in different form the development of a set of general academic skills. Which of the following academic skills is topic in specific course or addressed by an appropriate choice of work and assessment methods throughout the programme?

- ☐ Analytic skills
- ☐ Empirical methods
- ☐ Modelling
- ☐ Teamwork
- ☐ Written and oral communication
- ☐ Argumentation and presentation
- ☐ Others, namely ….

Q6: How many *European Credits* is your master programme?

- ☐ 60 EC
- ☐ 90 EC
- ☐ 120 EC
- ☐ Other, namely ….

Q7: How many European Credits are dedicated to a *mandatory internship*?

- ☐ 0 EC (we do not have a mandatory internship)
- ☐ 1-10 EC
- ☐ 10-15 EC
- ☐ 15-20 EC
- ☐ More than 20 EC

Q8: How many European Credits are dedicated to the *master thesis*?

- ☐ Less than 10 EC
- ☐ 10-20 EC
- ☐ 20-30 EC

- ☐ 30-40 EC
- ☐ More than 40 EC, namely ….

Q9: Which is the *language of instruction* your master programme?

- ☐ English
- ☐ French
- ☐ German
- ☐ Spanish
- ☐ Other, namely ….

Q10: How does your programme ensure the continuing *professionalization of the lecturers* in a manner consistent with your didactic methods?

- ☐ Requiring a University Teaching Qualification (UTQ) certificate for all lectures
- ☐ Providing possibilities for Advanced University Teaching Qualification certificate
- ☐ Requiring an International English certificate
- ☐ Other way, namely ….

Q11: How does your programme organize the *teaching and learning process*?

- ☐ Via summative evaluation procedures (e.g. via oral or written student evaluations at the end of course)
- ☐ Via formative evaluation procedures (e.g via post-course (statistical) analysis of student works, assignments, examinations)
- ☐ Via diagnostic evaluation procedures (e.g. via tests prior to instruction to find students' strengths, weaknesses, knowledge, and skills)
- ☐ Other ways, namely ….

Appendix B European universities we considered

Belgium

Master Artificial Intelligence, KU Leuven

Bulgaria

Master Data Science, Varna Free University "Chernorizets Hrabar"

Cyprus

Master Intelligent Systems, University of Cyprus

Master Business Intelligence and Data Analytics, The Cyprus Int. Institute of Management

Master Data Analytics, UCLan Cyprus

Czech Republic

Technical University of Liberec

Czech Technical University in Prague

Egypt

Robotics, Control and Smart Systems, The American University in Cairo

Finland

University of Oulu

France

Master Artificial intelligence, Sorbonne University

Master Artificial Intelligence, Centrale Supelec

Master Artificial Intelligence & Advanced Visual Computing, École Polytechnique

Master Data Sciences and Business Analytics, ESSEC Business School / Centrale Supelec

Master in Big Data Analytics for Business, IESEQ School of Management

Master Data Analytics and Artificial Intelligence, EDHEC Business School

Master Data Science and Artificial Intelligence, Université Cote d'Azur, Nice

Master Machine Learning and Data Mining, Université Jean Monnet

Germany

Master Robotics, Cognition, Artificial intelligence, Technical University Munich

Master Intelligent Adaptive Systems, Universiteit van Hamburg

Master Automation and Robotics, Universiteit Dortmund

Master Intelligent Systems, Bielefeld University

Greece

Master Business Analytics, Athens University of Economics and Business

Master Business Analytics and Decision Sciences, International Faculty CITY College, The University of Sheffield

Hungary

Budapest University of Technology and Economics

Italy

Master Data Science and Business Analytics, Università Commerciale Luigi Bocconi

Master Data Analytics for Business and Economics, Catholic University of the Sacred Heart

Master Business Analytics and Big Data Professionals, MIP Milan Polytechnic University

Master Digital Technology Management, University of Bologna

Norway

University of Stavanger

NTNU: Norwegian University of Science and Technology

Poland

Warsaw University of Technology

University of Zielona Góra

Wroclaw University of Science and Technology

Gdańsk University of Technology

Opole University of Technology

Portugal

ISCTE – Instituto Universitário de Lisboa

University of Beira Interior

University of Aveiro

Universidade NOVA de Lisboa

University of Lisbon

University of Lisbon

Universidade do Minho

Universidade do Minho

Romania

1 Decembrie 1918 Univeristy of Alba Iulia

Dunarea de Jos of Galati University

Politehnica University Timisoara

Master Artificial Intelligence and Distributed Computing, West University of Timişoara

Master Artificial Intelligence, Politehnica University of Bucharest

Slovakia

Technical University of Kosice

University of Žilina

Slovenia

Master Business Informatics, University of Ljubljana

Master Social Informatics, University of Ljubljana

Spain

University of Deusto

Catholic University of Ávila

Master Science in Management (spec. Business Analytics), Pompeu Fabra University

Master Big Data Management and Analytics, Polytechnic University of Catalonia

Master Artificial Intelligence, Polytechnic University of Catalonia

Master Business Analytics, Ramon Llull University

Master Business Intelligence and Analytics, SolidQ

Master's Degree in Artificial Intelligence, Universitat Rovira i Virgili

Master Computer Security Engineering and Artificial Intelligence, Universitat Rovira i Virgili

Master Artificial Intelligence, University of Barcelona

Master Business Analytics and Big Data, IE University

Sweden

Karlstad University

Chalmers University of Technology

Turkey

Data Science, İstanbul Şehir University

Business Analytics, İstanbul Şehir University

Business Analytics, Sabanci University

Big Data Analytics and Management, Bahcesehir University

The Netherlands

Master Artificial Intelligence, Utrecht University

Master Artificial Intelligence, Universiteit van Amsterdam

Master Artificial Intelligence, Radboud University Nijmegen

Master Artificial Intelligence, Rijksuniversiteit Groningen

Master Cognitive science and Artificial Intelligence, Tilburg University

Master Artificial Intelligence, Vrije Universiteit Amsterdam (VU)

Master Artificial Intelligence, Universiteit van Maastricht

Master Computer Science: Data Science, Leiden University

Master Computer Science: Advanced Data Analytics, Leiden University

United Kingdom

Master Artificial Intelligence, University of Edinburgh

Master Artificial Intelligence, Imperial College London

Master Artificial Intelligence, Queen Mary University London

Master Artificial Intelligence, University of Limerick (Dublin)

Master Artificial Intelligence, University of Aberdeen

Master Artificial Intelligence, King's College London

Master of Cyber Security and Artificial Intelligence, University of Sheffield

Master Artificial Intelligence, University of Leeds

Master Artificial Intelligence, University of Manchester

Master Artificial Intelligence, City, University of London

Master Artificial Intelligence, Heriot Watt University

Master Artificial Intelligence, University of Essex

Master Artificial Intelligence with Robotics, University of Hertfordshire

Master Artificial Intelligence, University of St Andrews

Master Artificial Intelligence, Cardiff University

Appendix C Needs Assessment Survey: Thailand

Current available positions in data science and artificial intelligence

Company Profile			Positions present in the company related to data science and AI															
Number of employees	Company name	Sector	Data Scientist	AI / data analyst	Machine learning Eng.	DevOps Eng.	Business Analyst	Quality assurance	AI architect	AI product manager	Software Eng. Manager	Software Eng.	Intelligence specialist	Security Engineer	Communication Manager	Risk Manager	Others	Total positions
1-10	Quantium Technology	Fintech										X						1
11-49	SmartContract Thailand	IT, Professional Service	X	X	X		X	X			X	X						7
	Intelligist Co.,Ltd.	-	X	X			X											3
	SOFTSION Co., Ltd.	-	X			X	X				X	X		X		X		7
	Q. Point Software Services Limited	Enterprise Software Development	X	X														2
	Thanadol & Kojchaboon Co., Ltd.	Healthcare		X			X							X				3
	BridgeAsia	Healthcare	X									X						2
	Deepscope	FinTech, WealthTech			X													1
	Energy Response	Commercial and industrial	X	X	X						X	X						5
	Microsoft Thailand Limited	Commercial Software Engineering	X		X	X			X	X	X	X						7

Company Profile			Positions present in the company related to data science and AI															
Number of employees	Company name	Sector	Data Scientist	AI / data analyst	Machine learning Eng.	DevOps Eng.	Business Analyst	Quality assurance	AI architect	AI product manager	Software Eng. Manager	Software Eng.	Intelligence specialist	Security Engineer	Communication Manager	Risk Manager	Others	Total positions
>=250	Nissan Motors Asia Pacific	Automotive					X	X			X			X	X	X		6
	Essilor Manufacturing (Thailand) Co.,Ltd.	Ophthalmic optics company	X	X		X	X				X	X						6
	Datapro computer system	-	X															1
	TEAM Consulting Engineering and Management PLC	Engineering Consultancy	X									X	X					3
	IBM	IT	X	X	X	X	X	X	X	X	X	X	X	X		X		13
	True Digital Group	Telecommunications	X	X			X						X					4
	Office of Agricultural Economics	Government, Agricultural	X	X	X													3
	Insurance Company	Insurance company		X			X											2
	Bank of Thailand	Central Bank	X				X					X		X				4
	Seagate	-										X						1
	PTTEP	-										X						1
	Innovation technology Co.	-															X	1
	CAT Telecom	ICT															X	1
	Total Companies		14	10	6	4	10	3	2	2	7	12	3	5	1	3	2	84

Employee training courses and student internships

Number of employees	Company name	Number of DS&AI employees				Enrol employees in short training courses					Providing internships to students completing a course in DS&AI?				
		1-3	4-10	11-25	>25	Yes	No	Maybe	How many	Duration	Yes	No	Maybe	How many	Duration
1-10	Quantium Technology	X				X			1-3	1 week	X			1-3	3 months
11-49	SmartContract Thailand	X					X					X			
	Intelligist Co.,Ltd.		X			X			>10	3 months	X			1-3	3 months
	SOFTSION Co., Ltd.	X				X			1-3	<3 months	X			1-3	>3 months
	Q. Point Software Services Limited	X					X				X			1-3	3 months
	Thanadol & Kojchaboon Co., Ltd.	X					X					X			
	BridgeAsia		X				X					X			
	Deepscope		X				X		1-3	2 weeks		X			
	Energy Response		X					X	1-3	1 week	X			1-3	2 months
>=250	Microsoft Thailand Limited				X	X			1-3	1 month	X			4-6	3 months
	Nissan Motors Asia Pacific				X			X	>10	3 days			X	>10	3 months
	Essilor Manufacturing (Thailand) Co.,Ltd.		X			X			4-6	3 months			X	4-6	6 months
	Datapro computer system		X					X			X			1-3	3 months
	TEAM Consulting Engineering and Management PLC	X						X	1-3	1 week			X	1-3	3 months

	IBM				X		X					X			
	True Digital Group			X			X					X			
	Office of Agricultural Economics							X	4–6	1 month	X			1–3	3 months
	Insurance Company			X				X	1–3	1 week		X			
	Bank of Thailand				X	X			4–6	2 weeks			X	1–3	1 month
	Seagate	X				X			1–3	2 weeks	X			1–3	Other
	PTTEP			X			X		>10	Others		X		1–3	Other
	Innovation technology Co.	X				X			4–6	1 month			X		3 months
	CAT Telecom				X			X	4–6	1 month			X	1–3	2 months

Required knowledge for internships (score 1 – 10)

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
1–10	Quantium Technology	7	8	5	10	9	3	4	2	6	1
11–49	SmartContract Thailand	–	–	–	–	–	–	–	–	–	–
	Intelligist Co.,Ltd.	10	7	9	8	6	5	4	3	2	1
	SOFTSION Co., Ltd.	6	3	4	10	8	7	2	9	5	1
	Q. Point Software Services Limited	10	6	9	7	3	8	5	1	4	2
	Thanadol & Kojchaboon Co., Ltd.	–	–	–	–	–	–	–	–	–	–
	BridgeAsia	–	–	–	–	–	–	–	–	–	–

	Deepscope	9	8	7	7	8	6	7	4	2	3
	Energy Response	2	3	1	2	2	2	3	2	2	2
	Average: Small-size company (rounded)	**7**	**6**	**6**	**7**	**6**	**5**	**4**	**4**	**4**	**2**
>=250	Microsoft Thailand Limited	8	10	5	4	3	9	7	6	2	1
	Nissan Motors Asia Pacific	9	8	2	7	1	6	3	4	5	10
	Essilor Manufacturing (Thailand) Co.,Ltd.	8	10	7	5	4	9	2	6	3	1
	Datapro computer system	10	7	6	5	8	9	3	4	1	2
	TEAM Consulting Engineering and Management PLC	9	4	7	3	1	2	8	10	5	6
	IBM	-	-	-	-	-	-	-	-	-	-
	True Digital Group	-	-	-	-	-	-	-	-	-	-
	Office of Agricultural Economics	9	6	7	8	5	10	4	2	3	1
	Insurance Company	-	-	-	-	-	-	-	-	-	-
	Bank of Thailand	8	7	3	6	10	5	4	2	9	1
	Seagate	8	10	10	10	10	5	8	1	1	1
	PTTEP	8	8	8	1	6	8	8	9	8	8
	Innovation technology Co., Ltd.	6	7	10	10	6	8	5	5	10	10
	CAT Telecom	3	3	5	8	3	8	8	3	3	9
	Average: Large-size company (rounded)	**8**	**7**	**6**	**6**	**5**	**7**	**5**	**5**	**5**	**5**

Small-size companies (based on average score)

Top 5 (6-8 scores)	Average (4-5 scores)	Least interested (2 scores)
1. **Data modelling and management**	1. Statistics	1. Market analysis
2. **Data base**	2. Knowledge representation	
3. **Machine learning**	3. Natural language processing	
4. **Data engineering**	4. Management of information systems	

5. **Algorithms development**		

Large-size companies (based on average score)

Top 5 (6-8 scores)	Average (5 scores)
1. **Data modelling and management**	1. Algorithms development
2. **Statistics**	2. Knowledge representation
3. **Machine learning**	3. Natural language processing
4. **Data engineering**	4. Management of information systems
5. **Data base**	5. Market analysis

Demanded subject areas for professional development (score 1 – 10)

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
1-10	Quantium Technology	7	10	6	1	8	4	3	5	9	2
11-49	SmartContract Thailand	–	–	–	–	–	–	–	–	–	–
	Intelligist Co.,Ltd.	10	9	8	7	6	5	4	3	2	1
	SOFTSION Co., Ltd.	6	8	3	2	1	10	4	5	7	9
	Q. Point Software Services Limited	–	–	–	–	–	–	–	–	–	–
	Thanadol & Kojchaboon Co., Ltd.	–	–	–	–	–	–	–	–	–	–
	BridgeAsia	–	–	–	–	–	–	–	–	–	–
	Deepscope	9	9	4	1	8	2	6	7	5	6
	Energy Response	3	2	2	2	3	2	2	2	2	2

	Average: Small-size 1-49 (rounded)	**7**	**8**	**5**	**3**	**5**	**5**	**4**	**4**	**5**	**4**
>=250	Microsoft Thailand Limited	8	10	5	4	3	9	6	7	2	1
	Nissan Motors Asia Pacific	9	7	4	8	3	1	5	6	2	10
	Essilor Manufacturing (Thailand) Co.,Ltd.	10	8	7	6	3	1	4	5	9	2
	Datapro computer system	-	-	-	-	-	-	-	-	-	-
	TEAM Consulting Engineering and Management PLC	10	6	8	5	4	1	9	7	3	2
	IBM	-	-	-	-	-	-	-	-	-	-
	True Digital Group	-	-	-	-	-	-	-	-	-	-
	Office of Agricultural Economics	9	7	10	3	4	5	8	2	6	1
	Insurance Company	9	3	4	5	2	7	8	1	10	6
	Bank of Thailand	10	9	5	3	4	6	8	7	2	1
	Seagate	10	10	10	10	10	10	10	10	10	10
	PTTEP	8	8	8	1	8	8	8	5	8	8
	Innovation technology Co., Ltd.	1	4	4	9	5	8	5	5	8	6
	CAT Telecom	3	3	8	9	9	9	3	3	8	10
	Average: Large-size (rounded)	**8**	**7**	**7**	**6**	**5**	**6**	**7**	**5**	**6**	**5**

Small-size companies (based on average score)

Top 2 (7-8 scores)	Average (5 scores)	Least interested (3-4 scores)
1. **Machine learning**	1. Data engineering	1. Knowledge representation
2. **Data modelling and management**	2. Algorithms development	2. Natural language processing
	3. Statistics	3. Market analysis
	4. Management of information systems	4. Data base

Large-size companies (based on average score)

Top 4 (7-8 scores)	Average (6 scores)	Least interested (5 scores)
1. **Data modelling and management**	1. Data base	1. Algorithms development
2. **Machine learning**	2. Statistics	2. Natural language processing
3. **Data engineering**	3. Management of information systems	3. Market analysis
4. **Knowledge representation**		

Appendix D Needs Assessment Survey: Indonesia

Current available positions in data science and artificial intelligence

Company Profile			Positions present in the company related to data science and AI															
Number of employees	Company name	Sector	Data Scientist	AI / data analyst	Machine learning Eng.	DevOps Eng.	Business Analyst	Quality Assurance Manager	AI architect	AI product manager	Software Eng. Manager	Software Eng.	Intelligence specialist	Security Engineer	Risk Manager	App Developer and DB Admin	Automation Eng.	Others	Total positions
1-10	FR Softnet (Pvt) Ltd	Information Tech.	X	X									X						3
11-49	PT. Lunata Technologies	-										X							1
	SMKS Islamic Technology Marinah Al-Hidayah Medan	-	X				X	X											3
	UPT TIK (Teknologi Informasi dan Komunikasi) - Universitas Syiah Kuala	ICT				X						X		X					3
	Sangkuriang	IT, custom software development	X								X							1	3
50-249	BALAI VETERINER MEDAN KEMENTERIAN PERTANIAN	-	X									X							2
	Oracle Indonesia	Software & Hardware Vendor	X	X									X						3
	PT Prosa Solusi Cerdas	-			X														1
	GDP Labs	Information Tech.	X		X	X				X	X	X		X					7
	eFishery	Aquaculture	X	X		X	X				X	X			X			X	8

Company Profile			Positions present in the company related to data science and AI																
Number of employees	Company name	Sector	Data Scientist	AI / data analyst	Machine learning Eng.	DevOps Eng.	Business Analyst	Quality Assurance Manager	AI architect	AI product manager	Software Eng. Manager	Software Eng.	Intelligence specialist	Security Engineer	Risk Manager	App Developer and DB Admin	Automation Eng.	Others	Total positions
>= 250	PT Perkebunan Nusantara IV	Agriculture Palm Oil					1									X			2
	PT Dua Empat Tujuh	ICT	X		X	X	X	X	X	X		X		X					9
	PT Aplikanusa Lintasarta	Telecommunication		X		X	X	X				X		X	X				7
	BPJS Ketenagakerjaan	-		X								X			X				3
	PT Solusi Bangun Andalas	Cement Industry		X	X		X	X	X	X			X		X				8
	PT. Bank sumut	-		X	X					X	X								4
	Tokopedia (Mobile engineer)	IT	X			X					X	X			X				5
	SMEC Group (CTO)	Eye Hospital	X			X		X			X	X		X	X				7
	SMEC Group (IT-HRD)	Eye Hospital	X			X		X			X	X		X	X				7
	PT Bank Mandiri (Persero) Tbk.	-		X		X	X	X	X	X	X	X	X		X		X		11
	PT. TELKOM INDONESIA	-	X	X			X												3
	Tokopedia (Software engineer)	Technology	X	X	X							X							4
	Bukalapak	E-commerce	X	X	X				X	X									5
	GDP Labs	-	X	X	X	X				X		X						X	7
	Airy	-	X	X							X	X							4
	PT Airy Nest Indonesia	Hospitality; Information Tech.	X	X							X	X							4
	Total Companies		**17**	**14**	**8**	**10**	**8**	**7**	**4**	**7**	**10**	**16**	**4**	**6**	**8**	**1**	**1**	**3**	**124**

Employee training courses and student internships

Number of employees	Company name	Number of DS&AI employees				Enrol employees in short duration courses					Providing internships to students completing a course in DS&AI?				
		1-3	4-10	11-25	>25	Yes	No	Maybe	How many?	Duration	Yes	No	Maybe	How many?	Duration
1-10	FR Softnet (Pvt) Ltd	X				X			1-3	1 month	X			1-3	Other
11-49	PT. Lunata Technologies	X				X							X	1-3	1 month
	SMKS Islamic Technology Marinah Al-Hidayah Medan	X				X			1-3	1 week	X			1-3	3 months
	UPT TIK - Universitas Syiah Kuala	X				X						X			
	Sangkuriang	X				X			1-3	1 week			X	1-3	2 months
50-249	Balai Veterinier Medan Kementerian Pertanian	X				X			1-3	3 days			X	1-3	2 weeks
	Oracle Indonesia		X				X					X			
	PT Prosa Solusi Cerdas				X	X			4-6	1 week	X			1-3	3 months
	GDP Labs			X				X	7-10	Other	X			4-6	3 months
	eFishery		X					X	1-3	2 weeks			X	1-3	3 months
>=250	PT Perkebunan Nusantara IV						X		4-6	1 month			X	1-3	3 months
	PT Dua Empat Tujuh			X				X	1-3	1 week	X			1-3	1 month
	PT Aplikanusa Lintasarta		X				X		1-3	2 weeks	X			1-3	3 months
	BPJS Ketenagakerjaan		X					X					X	1-3	3 months
	PT Solusi Bangun Andalas		X				X		1-3	1 week			X	1-3	1 month
	PT. Bank sumut						X				X			1-3	3 months

Number of employees	Company name	Number of DS&AI employees				Enrol employees in short duration courses					Providing internships to students completing a course in DS&AI?				
		1-3	4-10	11-25	>25	Yes	No	Maybe	How many?	Duration	Yes	No	Maybe	How many?	Duration
	Tokopedia (Mobile eng.)				X			X	>10	1 month			X	>10	3 months
	SMEC Group (CTO)	X					X		4-6	1 month	X			1-3	1 month
	SMEC Group (IT-HRD)	X					X					X			
	PT Bank Mandiri (Persero) Tbk.				X	X			7-10	1 month	X			1-3	2 months
	PT. Telkom Indonesia				X			X	>10	1 week	X			>10	Depends
	Tokopedia (Soft. eng.)				X	X			1-3	1 week	X			>10	3 months
	Bukalapak				X			X	1-3	Other	X			1-3	3 months
	GDP Labs			X				X	>10	1 week	X			>10	2 months
	Airy		X					X	4-6	3 days	X			1-3	Other
	PT Airy Nest Indonesia		X			X			1-3	1 month	X			1-3	3 months

Required knowledge for internships (score 1 - 10)

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
1-10	FR Softnet (Pvt) Ltd	3	4	5	3	4	5	5	4	3	7
11-49	PT. Lunata Technologies	10	6	5	7	9	4	8	3	2	1

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
	SMKS Islamic Technology Marinah Al-Hidayah Medan	8	10	9	6	7	2	4	1	5	3
	UPT TIK - Universitas Syiah Kuala	-	-	-	-	-	-	-	-	-	-
	Sangkuriang	8	7	8	10	8	9	8	9	10	7
	Average: Small-size company : 1-49 (rounded)	**7**	**7**	**7**	**7**	**7**	**5**	**6**	**4**	**5**	**5**
50-249	Balai Veterinier Medan Kementerian Pertanian	6	4	7	9	2	8	5	3	10	1
	Oracle Indonesia	-	-	-	-	-	-	-	-	-	-
	PT Prosa Solusi Cerdas	10	9	8	6	5	4	3	7	2	1
	GDP Labs	10	10	10	5	5	8	10	5	5	8
	eFishery	8	7	8	8	6	8	6	4	9	9
	Average: Medium-size company: (rounded)	**9**	**8**	**8**	**7**	**5**	**7**	**6**	**5**	**7**	**5**
>=250	PT Perkebunan Nusantara IV	6	5	3	10	8	1	7	4	9	2
	PT Dua Empat Tujuh	9	10	6	5	4	1	7	8	3	2
	PT Aplikanusa Lintasarta	8	10	4	1	9	5	2	7	3	6
	BPJS Ketenagakerjaan	8	7	9	10	4	1	5	6	3	2
	PT Solusi Bangun Andalas	3	7	9	4	2	5	6	1	8	10
	PT. Bank sumut	8	10	6	4	3	1	9	7	2	5
	Tokopedia (Mobile engineer)	6	5	9	10	8	4	2	1	3	7
	SMEC Group (CTO)	7	5	10	9	8	6	3	4	2	1
	SMEC Group (IT-HRD)	-	-	-	-	-	-	-	-	-	-

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
	PT Bank Mandiri (Persero) Tbk.	4	1	10	8	9	2	5	3	6	7
	PT. Telkom Indonesia	8	1	9	5	3	4	2	7	6	10
	Tokopedia (Software engineer)	8	9	6	5	10	4	1	7	2	3
	Bukalapak	8	9	10	4	7	6	2	5	1	3
	GDP Labs	3	10	4	2	8	9	7	6	1	5
	Airy	7	8	6	6	5	8	6	6	5	7
	PT Airy Nest Indonesia	7	8	9	10	5	6	4	3	1	2
	Average: Large size company (rounded)	**7**	**7**	**7**	**6**	**6**	**4**	**5**	**5**	**4**	**5**

Small-size companies: 1-49 (based on average score)

Top 5 (7 scores)	Average (5-6 scores)	Least interested (4 scores)
1. **Data modelling and management**	1. Knowledge representation	1. Natural language processing
2. **Machine learning**	2. Statistics	
3. **Data engineering**	3. Management of information systems	
4. **Data base**	4. Market analysis	
5. **Algorithms development**		

Medium-size companies: 50-249 (based on average score)

Top 3 (8-9 scores)	Average (6-7 scores)	Least interested (5 scores)
1. **Data modelling and management**	1. Data base	1. Algorithms development
2. **Machine learning**	2. Statistics	2. Natural language processing
3. **Data engineering**	3. Management of information systems	3. Market analysis
	4. Knowledge representation	

Large-size companies (based on average score)

Top 3 (7 scores)	Average (5-6 scores)	Least interested (4 scores)
1. **Data modelling and management**	1. Data base	1. Statistics
2. **Machine learning**	2. Algorithms development	2. Management of information systems
3. **Data engineering**	3. Knowledge representation	
	4. Natural language processing	
	5. Market analysis	

Demanded subject areas for professional development (score 1 - 10)

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
1-10	FR Softnet (Pvt) Ltd	4	5	4	5	4	5	7	6	4	6
11-49	PT. Lunata Technologies	-	-	-	-	-	-	-	-	-	-

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
	SMKS Islamic Technology Marinah Al-Hidayah Medan	8	7	9	10	2	3	5	4	6	1
	UPT TIK - Universitas Syiah Kuala	-	-	-	-	-	-	-	-	-	-
	Sangkuriang	10	8	9	10	8	10	8	8	9	8
	Average: Small-size company (rounded)	**7**	**7**	**7**	**8**	**5**	**6**	**7**	**6**	**6**	**5**
50-249	Balai Veterinier Medan Kementerian Pertanian	8	5	7	9	2	6	3	4	10	1
	Oracle Indonesia	-	-	-	-	-	-	-	-	-	-
	PT Prosa Solusi Cerdas	10	9	8	7	6	5	4	3	2	1
	GDP Labs	9	7	9	4	9	4	4	4	9	9
	eFishery	9	8	8	8	6	10	2	5	6	6
	Average: Medium-size company (rounded)	**9**	**7**	**8**	**7**	**6**	**6**	**3**	**4**	**7**	**4**
>=250	PT Perkebunan Nusantara IV	6	5	3	9	7	1	8	4	10	2
	PT Dua Empat Tujuh	10	9	6	4	5	1	7	8	3	2
	PT Aplikanusa Lintasarta	8	10	4	6	7	1	3	9	5	2
	BPJS Ketenagakerjaan	-	-	-	-	-	-	-	-	-	-
	PT Solusi Bangun Andalas	4	8	10	6	2	3	7	1	5	9
	PT. Bank sumut	-	-	-	-	-	-	-	-	-	-
	Tokopedia (Mobile engineer)	10	9	7	6	5	2	8	1	4	3
	SMEC Group (CTO)	10	5	8	9	7	6	3	4	2	1
	SMEC Group (IT-HRD)	-	-	-	-	-	-	-	-	-	-
	PT Bank Mandiri (Persero) Tbk.	5	3	10	9	8	1	7	4	6	2

Number of employees	Company name	Data modelling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
	PT. Telkom Indonesia	8	1	7	5	4	3	2	6	9	10
	Tokopedia (Software engineer)	8	9	6	5	10	4	1	7	2	3
	Bukalapak	9	7	10	1	4	3	5	8	2	6
	GDP Labs	4	8	3	2	7	1	5	9	10	6
	Airy	10	10	10	8	8	8	10	8	10	10
	PT Airy Nest Indonesia	7	9	10	6	5	4	3	2	8	1
	Average: Large-size company (rounded)	**8**	**7**	**7**	**6**	**6**	**3**	**5**	**5**	**6**	**4**

Small-size companies: 1-49 (based on average score)

Top 5 (7-8 scores)	Average (6 scores)	Least interested (5 scores)
1. **Data base**	1. Statistics	1. Algorithms development
2. **Data modelling and management**	2. Natural language processing	2. Market analysis
3. **Machine learning**	3. Management of information systems	
4. **Data engineering**		
5. **Knowledge representation**		

Medium-size companies (based on average score)

Top 2 (8-9 scores)	Average (6-7 scores)	Least interested (3-4 scores)
1. **Data modelling and management**	1. Machine learning	1. Natural language processing
2. **Data engineering**	2. Data base	2. Market analysis

	3. Management of information systems 4. Algorithms development 5. Statistics	3. Knowledge representation

Large-size companies (based on average score)

Top 3 (7-8 scores)	Average (5-6 scores)	Least interested (3-4 scores)
1. Data modelling and management **2. Machine learning** **3. Data engineering**	1. Data base 2. Algorithms development 3. Management of information systems 4. Knowledge representation 5. Natural language processing	1. Market analysis 2. Statistics

Appendix E Needs Assessment Survey: Sri Lanka

Current available positions in data science and artificial intelligence

Company Profile			Positions present in the company related to data science and AI													
Number of employees	Company name	Sector	Data Scientist	AI / data analyst	Machine learning	DevOps Eng.	Business Analyst	Quality Assurance	AI architect	AI product	Software Eng.	Software Eng.	Intelligence	Security Engineer	Others	Total positions
11-49	Trabeya Pvt Ltd	ICT	X	X	X	X	X	X								6
	Linear Squared Private Limited	AI/Data Science	X	X	X		X		X	X		X				7
	Paraqum Technologies (Pvt) Ltd	Service Integration		X	X	X	X	X	X	X	X	X	X	X		11
	Axiata Digital Labs (Pvt) Ltd	-	X													1
50-249	VizuaMatix	-	X													1
	n*able pvt ltd	Electronics Design Services and Network Product Development		X	X											2
	Hsenid Mobile Solutions	IT/Telco	X	X	X	X			X							5
	SimCentric Technologies	Mobile Application													X	1
>=250	DIALOG AXIATA PLC	IT			X											1
	V S Information Systems (Pvt) Ltd.	ICT	X	X	X		X									4
	Total		6	6	7	3	4	2	3	2	1	2	1	1	1	39

Employee training courses and student internships

Number of employees	Company name	Number of DS&AI employees				Enroll employees in short duration courses					Providing internships to students completing a course in DS&AI?				
		1-3	4-10	11-25	>25	Yes	No	Maybe	How many?	Duration	Yes	No	Maybe	How many?	Duration
11-49	Trabeya Pvt Ltd			X				X	1-3	1 week	X			1-3	6 months
	Linear Squared Private Limited		X					X	1-3	1 month	X			1-3	3 months
	Paraqum Technologies (Pvt) Ltd	X				X			1-3	1 month	X			1-3	Other
	Axiata Digital Labs (Pvt) Ltd			X			X		1-3	Other	X			7-10	3 months
50-249	VizuaMatix		X			X			1-3	1 month	X			4-6	6 months
	n*able pvt ltd		X					X	1-3	1 month	X			1-3	3 months
	Hsenid Mobile Solutions	X						X	> 10	Other	X			1-3	Other
	SimCentric Technologies	X				X			1-3	3 days			X	7-10	Other
>=250	DIALOG AXIATA PLC			X		X			1-3	Other	X			1-3	Other
	V S Information Systems (Pvt) Ltd.		X			X			1-3	3 days	X			1-3	Other

Required knowledge for internships (score 1 – 10)

Number of employees	Company name	Data modeling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
11-49	Trabeya Pvt Ltd	4	10	9	5	7	8	6	3	2	1
	Linear Squared Private Limited	7	9	8	7	9	9	3	6	3	6
	Paraqum Technologies (Pvt) Ltd	8	8	8	7	8	8	8	1	1	8
	Axiata Digital Labs (Pvt) Ltd	8	8	10	8	9	7	8	8	5	5
	Average: Small-size company(rounded)	**7**	**9**	**9**	**7**	**8**	**8**	**6**	**5**	**3**	**5**
50-249	VizuaMatix	10	8	6	5	7	9	3	1	4	2
	n*able pvt ltd	10	8	9	5	3	7	6	4	2	1
	Hsenid Mobile Solutions	9	9	9	8	8	8	8	6	5	5
	SimCentric Technologies	1	7	1	1	10	1	1	1	1	1
	Average: Medium-size company (rounded)	**8**	**8**	**6**	**5**	**7**	**6**	**5**	**3**	**3**	**2**
>=250	DIALOG AXIATA PLC	8	8	8	7	8	8	8	6	8	8
	V S Information Systems (Pvt) Ltd.	4	6	6	5	5	6	6	5	5	7
	Average: Large-size company (rounded)	**6**	**7**	**7**	**6**	**7**	**7**	**7**	**6**	**7**	**8**

Small-size companies (based on average score)

Top 4 (8-9 scores)	Average (6-7 scores)	Least interested (3-5 scores)
1. **Machine learning**	1. Data modeling and management	1. Natural language processing
2. **Data engineering**	2. Data base	2. Market analysis

3. **Algorithms development** 4. **Statistics**	3. Knowledge representation	3. Management of information systems

Medium-size companies (based on average score)

Top 3 (7-8 scores)	Average (5-6 scores)	Least interested (2-3 scores)
1. **Data modelling and management** 2. **Machine learning** 3. **Algorithms development**	1. Data engineering 2. Statistics 3. Data base 4. Knowledge representation	1. Natural language processing 2. Management of information systems 3. Market analysis

Large-size companies (based on average score)

Top 1 (8 scores)	Average (7 scores)	Least interested (6 scores)
1. **Market analysis**	1. Machine learning 2. Data engineering 3. Algorithms development 4. Statistics 5. Knowledge representation 6. Management of information systems	1. Data modeling and management 2. Data base 3. Natural language processing

Demanded subject areas for professional development (score 1 – 10)

Number of employees	Company name	Data modeling and management	Machine learning	Data engineering	Data base	Algorithms development	Statistics	Knowledge representation	Natural language processing	Management of information systems	Market analysis
11-49	Trabeya Pvt Ltd	9	6	5	10	3	4	8	7	2	1
	Linear Squared Private Limited	4	4	7	4	4	4	7	4	4	4
	Paraqum Technologies (Pvt) Ltd	9	9	9	1	9	1	6	1	9	8
	Axiata Digital Labs (Pvt) Ltd	9	9	6	4	6	8	7	8	7	7
	Average: Small-size (rounded)	**8**	**7**	**7**	**5**	**6**	**4**	**7**	**5**	**6**	**5**
50-249	VizuaMatix	10	6	8	5	7	9	3	1	4	2
	n*able pvt ltd	7	8	9	4	3	10	5	6	2	1
	Hsenid Mobile Solutions	9	9	9	9	8	6	7	8	9	6
	SimCentric Technologies	1	10	1	1	1	1	1	10	1	1
	Average: Medium-Size (rounded)	**7**	**8**	**7**	**5**	**5**	**7**	**4**	**6**	**4**	**3**
>=250	DIALOG AXIATA PLC	8	8	8	7	8	8	6	8	8	8
	V S Information Systems (Pvt) Ltd.	7	7	7	4	7	4	8	6	5	4
	Average: Large-size (rounded)	**8**	**8**	**8**	**6**	**8**	**6**	**7**	**7**	**7**	**6**

Small-size companies (based on average score)

Top 4 (7-8 scores)	Average (6 scores)	Least interested (4-5 scores)
1. **Data modelling and management**	1. Algorithms development	1. Data base
2. **Machine learning**	2. Management of information systems	2. Natural language processing
3. **Data engineering**		3. Market analysis
4. **Knowledge representation**		4. Statistics

Medium-size companies (based on average score)

Top 4 (7-8 scores)	Average (5-6 scores)	Least interested (3-4 scores)
1. **Machine learning**	1. Natural language processing	1. Knowledge representation
2. **Data modelling and management**	2. Data base	2. Management of information systems
3. **Data engineering**	3. Algorithms development	3. Market analysis
4. **Statistics**		

Large-size companies (based on average score)

Top 4 (8 scores)	Average (7 scores)	Least interested (6 scores)
1. **Data modelling and management**	1. Knowledge representation	1. Data base
2. **Machine learning**	2. Natural language processing	3. Statistics
3. **Data engineering**	3. Management of information systems	4. Market analysis
4. **Algorithms development**		

Appendix F: Focus Group Discussions

Focus Group Thailand

A focus group discussion was conducted jointly by the three partner universities in Thailand: (AIT, KKU and WU), on Thursday 1st August 2019, during 13:30-15:30 via online (zoom teleconference). The discussion group composed of 4 representatives from ICT industrial sector, 3 students, and 8 academics, as shown below.

Representatives from ICT industrial sector:

1. Mr. Chinnawat Devahastin Na Ayudhya
   Senior Analyst, Data Analytics Group, Information Technology Group
   Bank of Thailand (BOT), Sector: Banking, Size: 250 or more
2. Mr. Chana Supatsorn
   Deputy CEO, Intelligist Company Limited,
   Sector: Big Data Platform Provider, Size: 20-50
3. Dr. Kan Ouivirach
   Software Engineer Manager
   Pronto Group Ltd.
   Sector: Software House, Size: 100-150
4. Mr.Tanin Sammanee
   CEO, Deepscope, FinTech Startup Company,
   Sector: Startup and FinTech, Size:

Students:

1. Mr. Akraradet Sinsamersuk, Computer Science master programme, AIT
2. Miss Sasipim Prasartsri, IT master programme, KKU
3. Miss Thohirah Husaini, Software Engineering master programme, WU

Academics:

1. Dr. Chutiporn Anutariya, ICT Department, AIT[8]
2. Dr. Wararat Songpan, CS Department, KKU
3. Dr.Chitsutha Soomlek, CS Department, KKU
4. Dr. Suratsavadee Korkua, Department of Electrical Engineering, WU
5. Dr. Chirawat Wattanapanich, Department of Computer Engineering, WU
6. Dr. Thimaporn Phetkaew, Department of Software Engineering, WU
7. Dr. Putthiporn Thanathamathee, Department of Software Engineering, WU
8. Dr. Siraporn Sakphrom, Department of Electrical Engineering, WU

[8] Also acting as a focus group moderator.

Focus Group Indonesia

A focus group discussion was conducted jointly by the three partner universities in Indonesia (ITB, USU and Unsyiah), on Thursday 20st August 2019, during 15:00–16:30 via online (zoom teleconference). The discussion group composed of 7 representatives from ICT industrial sector, 3 students, and 7 academics, as shown below.

Representatives from ICT industrial sector:

4. Ali Septiandri – airyrooms
5. Arie Sutiono – airyrooms
6. Samsu – airyrooms
7. Muhammad Ghifary – Bukalapak
8. Juanda Lokman – ions tech
9. Budi Prasetyo – Mandike
10. Teguh Eko Budiarto – Prosa

Students:

1. Denny syaputra – Unsyiah
2. Teuku putera – Unsyiah
3. Adine Belinda – USU
4. Haslin Lawalata – USU

Academics:

1. Nugraha P. Utama – ITB[9]
2. Ayu Purwarianti – ITB
3. Masayu Leylia Khodra – ITB
4. Taufik Abidin – Unsyiah
5. Opim Salim Sitompul – USU
6. Emerson Sinulingga – USU
7. Erna Budhiarti Nababan – USU
8. Maya Silvi – USU
9. Elvi Zamzami – USU

Focus Group Sri Lanka

A focus group discussion was conducted by the University of Peradeniya (UoP) and the University of Sri Jayewardenepura, Sri Lanka on Saturday 10th August 2019, during 9.00 – 12.30, with the participation of 3 ICT companies, 3 students and 13 academics, as shown below.

Representatives from ICT industrial sector:

5. Mr. Pradeep Indrajith
   Principal Data Analyst - Business Operations

[9] Also acting as a focus group moderator.

IFS Sri Lanka
Sector: Software vendor, Size: over 700

6. Mr. Shanika Amarasoma
   Principal Architect-Data Engineering
   Axiata Digital Labs (Pvt) Ltd
   Sector: innovative software service provider, Size: over 400
7. Ms. Nuzla Ismail
   Data Scientist
   Axiata Digital Labs (Pvt) Ltd
8. Mr. Hansa Perera
   Data Scientist and Machine Learning Researcher
   Zone 24x7 (Pvt) Ltd
   Sector: technology solutions provider, Size: 201-500

Students:

1. Mr. Kayanan Manickavasagar, PhD candidate in Statistics, UoP
2. Mr. Jarashanth Selvarajah, M.Phil. candidate in Computer Science, UoP
3. Ms. Amirthavarshini Mahadevan, M.Sc. candidate in Computer Science, UoP

Academics:

1. Prof. Saluka R. Kodituwakku – Professor in Statistics & Computer Science, UoP
2. Prof. Pushpakanthie Wijekoon – Senior Professor in Statistics, UoP
3. Prof. Roshan D. Yapa – Professor in Statistics & Computer Science, UoP
4. Dr. Chitraka Wickramarachchi, Dept. of Statistics, University of Sri Jayewardenepura
5. Dr. Ruwan D. Nawarathna, Dept. of Statistics and Computer Science, UoP
6. Dr. Amalka Pinidiyaarachchi, Dept. of Statistics and Computer Science, UoP
7. Dr. Hakim Usoof, Dept. of Statistics and Computer Science, UoP
8. Dr. Lakshika Nawarathna, Dept. of Statistics and Computer Science, UoP
9. Dr. Hemalika Abeysundara, Dept. of Statistics and Computer Science, UoP
10. Mr. Prabhath Gunathilake, Dept. of Statistics and Computer Science, UoP
11. Dr. Erunika O. Dayartna, Dept. of Statistics and Computer Science, UoP
12. Dr. Shameen Jinadasa, Director , International Affairs Office, UoP
13. Dr. Nanda Gunawardena, Former Director , International Affairs Office, UoP