

CONCEPT PAPER

**A STRATEGY
FOR PRODUCING
DATA SCIENTISTS
BY THE PROVISION OF
EDUCATION & TRAINING
THROUGH THE TERTIARY
EDUCATIONAL
INSTITUTIONS
IN SRI LANKA**



Prepared By:
National Human
Resources Development
Council (NHRDC) of Sri Lanka

| | |
|--|---|
| Purpose | 3 |
| Rationale | 3 |
| Proposed Strategy | 3 |
| Components of the Strategy | 4 |
| Master’s and Phd programs in Data Science | 5 |
| Initiating Certification Programs | 5 |
| Bachelor’s degree programs in Data Science | 5 |
| Conclusion | 5 |
| COMMITTEE MEMBERS | 6 |

Purpose:

The purpose of this concept paper is to outline a workable strategy to provide the education and training required to produce a thousand data scientists within the next four years

Rationale:

2.5 quintillion bytes of data are created each day. Even a smart phone contributes to data proliferation.

Most of this extensive data is in the public domain and can be selectively mined for constructive use, i.e. in the area of human, social and economic development.

The need for effective usage of this data has spawned the new discipline of Data Science which is a multi-disciplinary amalgam of data inference, algorithm development and technology. It is focused towards solving problems of analytical complexity and is essentially about using data in creative ways to generate outcomes which have an economic and social value.

The professionals within this field are data scientists who leverage data to solve business problems. They use data to interpret, extrapolate prescribe and recommend appropriate actions for current and future strategies. A data scientist should be capable of original, creative and divergent thinking to invent new algorithms by integrating knowledge from different disciplines such as computer science, physics, statistics, finance, social science or applied mathematics.

At present there is a dearth of such professionals in the world. This issue was brought up for discussion at the World Economic Forum and attention was drawn to the fact that the US alone would require around 20,000 data scientists in the next few years. This assertion has been further confirmed by a study conducted by IBM. In this study future projections indicate that by 2020 the number of positions for data science and analytics (DSA) skills in the United States alone will surpass this number.

Since educators and trainers will have to respond to the increasing demand for the required skills, it will be necessary to initiate new programs for prospective future employees. Moreover suitable retraining programs will have to be established for those already in the work force so that they would be able to meet these demands.

Sri Lanka can be both a participant as well as a beneficiary if the educational institutions are able to produce data scientists in preparation for this envisioned and predicted future development.

Proposed Strategy

The strategy should provide a foundation for the creation and development of data scientists and big data engineers. The main concern is the provision of infrastructure and facilities essential for the propagation of data science. Furthermore, data from government organizations must be desensitized and made suitable for data scientists to use it in a constructive manner as well as for research.

The focus should be to attract the most suitable candidates among those seeking entry to tertiary and higher education, into data science and analytics. For those already employed in information

technology, opportunities should be created for channeling them to data science and analytics and provision made for retraining.

Since a cadre of data science trainers will be required, adequate measures will have to be taken to provide these trainers with the required skills through training programs.

The strategy should be to emphasize and publicize the benefits to organizations and the vast potential that exists for lucrative employment both in Sri Lanka and overseas in the field of data science.

Components of the Strategy

The strategy will require the formation of a team of experts to formulate policy, advise and plan strategy for the development of data science. This team will have to ascertain and study the diverse areas and factors that contribute to the development of data scientists in Sri Lanka. Whilst the focus of the team will be on the following, it will not prevent the team from directing their attention to other relevant concerns.

1. Identify current data science practitioners in industry, academia and other institutions who are able to support this initiative.
2. Conduct data science awareness sessions to educate important government and business organizations. These sessions would elucidate the benefits that could accrue to their respective organizations by the implementation of data science- based techniques.
3. Preparation of a directory of data sources belonging to government and business organizations. Big data sets from this directory should be made available to academic institutions for teaching, training and research. Furthermore they could be utilized effectively by data scientists for industry- based research and development. The outcomes from these initiatives could influence the future policy and plans of the respective establishments.
4. Preparing a suitable list of technologies, tools, computer languages, libraries, and knowledge sources that are commonly used in training data scientists.
5. Foster and promote data science projects to demonstrate the benefits of using data science technologies.
6. Harness Tertiary Educational Institutions to conduct programs for the development of data scientists.

ICTA can be used to proceed with this initiatives to provide 1000 data scientist by the year 2020

7. The following table provides the commitment made by the universities and educational institutes in enrolling prospective students towards meeting the objective of creating data scientists

Master's and Phd programs in Data Science

| Institution | 2018 | 2019 | 2020 | 2021 | 2022 |
|-----------------------------------|------|------|------|------|------|
| University of Colombo | | | | | |
| University of Sri Jayawardanepura | | | | | |
| University of Peradeniya | | | | | |
| SLIIT | | 20 | 30 | 40 | 60 |
| University of Kelaniya | 100 | 100 | 100 | 200 | 200 |
| Other Institutions | | | | | |

*University of Kelaniya, Postgraduate: 2018 - 100 (1st batch 100), 2019 - 200 (1st batch + 2nd batch), 2020 - 200 (2nd batch + 3rd batch)

Initiating Certification Programs

This would include Continuing Professional Development (CPD) initiatives with the objective of producing internationally accredited Data Science practitioners.

| Institution | 2018 | 2019 | 2020 | 2021 | 2022 |
|-----------------------|------|------|------|------|------|
| University of Colombo | 50 | 50 | 50* | 50* | 50* |

*subjected to change with demand.

Bachelor's degree programs in Data Science

| Institution | 2018 | 2019 | 2020 | 2021 | 2022 |
|-----------------------------------|------|------|------|------|------|
| University of Colombo | | | | | |
| University of Sri Jayawardanepura | | | | | |
| University of Peradeniya | | | | | |
| SLIIT | 30 | 40 | 50 | 60 | 100 |
| University of Kelaniya | | 50 | 65 | 65 | 65 |
| Other Institutions | | | | | |

To drive this strategy, the government should consider data science as a high priority area of interest by defining proper policies and providing the necessary administrative and financial support.

Conclusion

Once the concept is approved, a team of experts can be appointed to proceed with the development of data scientists in Sri Lanka. This team could commence working on the strategy immediately and facilitate the process of recruitment to the tertiary institutions. Simultaneously steps should be taken to disseminate the importance of data science analytics to public and private organizations as well as to demonstrate the economic and social benefits of this initiative.

COMMITTEE MEMBERS

Chairperson of the Committee: Dr.Prasad Samarasinghe, Managing Director, Lanka bell

| No. | Name | Designation | Institution |
|-----|---------------------------------|--|--|
| 1 | Dr Prasad samarasinghe | Managing Director | Lanka Bell 344, Galle Road Colombo |
| 2 | Dr.Sankalpa Gamwerige | CEO | Zone24X7 |
| 3 | Mr.Supun weerasinghe | Chief Executive | Dialog |
| 4 | Prof. Koliya Pulasinghe | Dean / Associate Professor. Faculty of Computing | Sri Lanka Institute of Information Technology, Colombo 03. |
| 5 | Dr. Ajith P. Madurapperuma | Head, Department of Electrical and Computer Engineering | Open University of Sri Lanka, Chairman, Sri Lanka Computer Emergency Response Team Member, Board of Directors, ICT Agency of Sri Lanka |
| 6 | Prof. K. R. Ranjith Mahanama | DEAN | Faculty of Science University of Colombo. |
| 7 | Mr.Lasitha Devendra | Dean | Faculty of Information Technology, Aquinas College of Higher Studies, Colombo |
| 8 | Mr.Felix Weerakkody | Visiting Lecture | Sri Lanka Institute of Information Technology, Colombo 03 |
| 9 | Dr.Sankalpa Gamwarige | CEO | Zone24X7 |
| 10 | Dr.Dilhari Attygalle | Senior Lecturer | Faculty of Science, University of Colombo |
| 11 | Rasika Jayathilake | Senior Lecturer | Faculty of Science, University of Colombo |
| 12 | Dr.Ruwan Werasinghe | Senior lecture | School of Computing, University of Colombo |
| 13 | Dr.Gamini Wijayawardhana | Dean | University of Kalaniya |
| 14 | S.Thanigaseelan | Assistant Director | NHRDC |

Guidance and Support: NHRDC