

Report of Validation Panel

Date of Meeting: 12th March 2013

Named Award: Higher Diploma
Programme Title: Higher Diploma in Data Science & Analytics
Award Type: Higher Diploma
Award Class: Major Award
NFQ Level: 8
Intakes Commencing: April 2013
ECTS/ACCS Credits: 60

PANEL MEMBERS

| Name / Function / Institution |
|---|
| Mr Matt Cotterell, Head of School, Mechanical, Electrical & Process Engineering, Cork Institute of Technology(Chair) |
| Dr Pat Doody, Director, Centre for Innovation in Distributed Systems, Institute of Technology, Tralee |
| Dr Ailish Hannigan, Professor of Biomedical Statistics, Graduate Entry Medical School, University of Limerick |
| Mr Barry Guiney, Chief Operations Officer, Bluemetrix, Northpoint House, Northpoint Business Park, Mallow Road, Cork |
| Dr Catherine Frehill, Module Moderator, Office of the Registrar and Vice President for Academic Affairs, Cork Institute of Technology |

PROPOSING TEAM MEMBERS

| Name / Function / Department |
|---|
| Dr Áine Ní Shé, Head, Department of Mathematics |
| Mr Jim O'Dwyer, Head, Department of Computing (apologies) |
| Dr Hugh McGlynn, Head, School of Science & Informatics |
| Mr Tim Horgan, Head, Cloud Computing Centre of Excellence |
| Mr Aengus Daly, Lecturer, Department of Mathematics |
| Dr Seán Lacey, Lecturer, Department of Mathematics |
| Ms Aisling O'Driscoll, Lecturer, Department of Computing |
| Dr Ted Scully, Lecturer, Department of Computing |
| Ms Catherine Palmer, Lecturer, Department of Mathematics |
| Ms Linda O'Sullivan, Lecturer, Department of Computing |
| Mr Larkin Cunningham, Lecturer, Department of Computing |
| Dr Declan O'Connor, Lecturer, Department of Mathematics (apologies) |
| Dr Paul Walsh, Lecturer, Department of Computing (apologies) |

BACKGROUND TO THE PROPOSED PROGRAMME

The proposal seeks validation for a one-year Higher Diploma in Science in Data Science & Analytics which has been developed in response to the HEA call for proposals for ICT skills conversion courses announced in November 2012. The proposers made a submission to the HEA in December 2012 and the HEA awarded funding to CIT for 20 places on the programme. The sponsor department for the programme is the Department of Mathematics and the programme development has had significant input from the Department of Computing and collaboration with CIT Extended Campus. Through this collaboration a thorough and detailed consultation took place with relevant industrial partners to discuss programme design and development, recruitment and implementation, and the provision of internship opportunities. This is the first programme to be developed by the Department of Mathematics and in the future the Department plans to develop further offerings in several areas of expertise such as Mathematics for Education.

FINDINGS OF THE PANEL

*NOTE: In this report, the term “Requirement” is used to indicate an action or amendment which in the view of the Panel **must** be undertaken prior to commencement of the Programme. The term “Recommendation” indicates an item to which the Institute/Academic Council/Course Board should give serious consideration for implementation at an early stage and which should be the subject of ongoing monitoring.*

The Panel has considered the documentation provided and has discussed the programme with the proposers. Based on this, the Panel has arrived at a number of Findings, Requirements and Recommendations as follows.

1. Programme-Level Findings

1.1 NEED FOR THE PROGRAMME

Validation Criterion: Is there a convincing need for the programme with a viable level of applications?

Overall Finding: Yes

1.2 AWARD

Validation Criterion: Are the level and type of the proposed award appropriate?

Overall Finding: Yes

1.3 LEARNING EXPERIENCE

Validation Criterion: Is the learning experience of an appropriate level, standard and quality overall?

Overall Finding: Yes, subject to certain Requirements and Recommendations

The proposed Programme Outcomes as presented to the Panel are attached as Appendix 1. Findings, requirements and recommendations concerning individual modules are recorded in Section 3 below.

1.4 PROGRAMME STRUCTURE

Validation Criterion: Is the programme structure logical and well designed (including procedures for access, transfer and progression)?

Overall Finding: Yes

The Semester Schedules as proposed are in Appendix 2.

Recommendations: The panel recommends that the previous mathematical attainment of applicants be considered in the selection of candidates for admission to the programme.

1.5 PROGRAMME MANAGEMENT

Validation Criterion: Are the programme management structures adequate?

Overall Finding: Yes

1.6 RESOURCE REQUIREMENTS

Validation Criterion: Are the resource requirements reasonable?

Overall Finding: Yes

The Panel was assured on behalf of the President and Head of School that appropriate resources in terms of staffing and facilities will be put in place when the programme is validated.

1.7 IMPACT ON THE INSTITUTE

Validation Criterion: Will the impact of the programme on the Institute be positive?

Overall Finding: Yes

Findings: The panel thought that the rapid response to the HEA ICT Skills Conversion call was to be commended and welcomed the proposal as there is an identified national need for the programme.

2. Module-Level Findings

As this is CIT's first offering in the Data Science and Analytics field, all the modules are newly developed. These modules may be delivered as free choice electives to full-time students of other Level 8 and 9 programmes across CIT.

The Panel was informed that the new draft modules have been the subject of internal scrutiny by the CIT module moderator (Dr Catherine Frehill).

In exercising its brief to consider the overall standard and appropriateness of modules, the Panel wishes to add the following findings, requirements and recommendations.

2.1 ALL MODULES

2.1.1 Requirement: Any revisions to Module Descriptors or Semester Schedules made to address the recommendations and requirements in this require sign-off from the CIT Module Moderator and the Registrar's Office prior to approval by the CIT Academic Council.

2.1.2 Requirement: The appropriate Reassessment Requirement should be selected for each module.

2.1.3 Comment: During the panel discussion the problem of the availability of suitable data sets was highlighted. Due to the nature of large data sets (messy and missing data), ethical approval and various issues of confidentiality the timescale to obtain these data sets should not be underestimated.

2.2. Various Modules: Resource Listings

Requirement: Reading lists including supplementary texts should be expanded and updated, in order to ensure currency of content.

2.3. Various Modules: Workload

Recommendation: The part-time mode of delivery should be completed for all modules.

2.4 Module: DATA8001 Data Science and Analytics

Recommendation: Topics including technical writing, ethics, data privacy and security, data protection legislation should be included in this module.

2.5. Module: DATA8003 Text Analysis & Data Visuals

Recommendation: The panel recommends the modules should be retitled as the proposed title is too specific. Suggested titles discussed at the meeting included *Unstructured Data Visualisation* and *Data Analysis & Visualisation*, with the panel favouring the latter.

2.6. Module: STAT8006 Applied Stats & Probability

Recommendation: A new section should be included in the indicative content of the module to include Statistical Packages.

2.7. Modules: DATA8004 Data Mining & Knowledge Discovery and STAT8007 Statistical Methods for Big Data

Requirement: The learning outcomes and indicative content for DATA8004 needs to be reduced to a manageable workload for the learner. The panel recommends that topics such as data reduction techniques, clustering and classification methods, logistic regression models and other data mining techniques be rebalanced into STAT8007 with a corresponding reduction in focus on experimental design and traditional hypothesis testing in this module

2.8 Module: DATA8006 Project

Requirement: The proposed project module was found to be too generic in nature and the panel requires that the module be written specifically for the Data Science and Analysis programme.

3. Other Findings

None.

4. Conclusion

Based on the above findings, the Panel has arrived at the following Conclusions:

- The Programme meets the required standards for an award in the Science field of study at Level 8 of the National Framework of Qualifications.
- The Programme meets the criteria for validation of a new programme adopted by the Academic Council of Cork Institute of Technology.

The Panel therefore recommends that the Programme be validated for five academic years, or until the next programmatic review, whichever is soonest, subject to implementation of the Requirements above, and with due regard to the Recommendations made.

| Implementation of Requirements and Recommendations |
|--|
| Requiring Registrar's Office Sign-Off: |
| 2.1.1 Requirement: Completed |
| 2.1.2 Requirement: Completed |
| 2.2 Requirement: Completed |
| 2.3. Requirement: Completed |
| 2.4 Recommendation DATA8001: Completed |
| 2.5 Recommendation DATA8003: Completed |
| 2.6 Recommendation STAT8006: Completed |
| 2.7 Requirement DATA8004 and STAT8007: Completed |
| 2.8 Requirement DATA8006: Completed |

Higher Diploma in Science in Data Science & Analytics

Implementation Report
Department of Mathematics
3 May 2013

The proposers of the *Higher Diploma in Science in Data Science & Analytics* wish to thank all members of the validation panel for the extremely beneficial contribution which they have made to the final stages of the development of the programme. It was clear at the meeting that the panel had reviewed and considered the documentation in the utmost detail. The discussion which ensued at the meeting has further informed the development of the programme, and has also created further opportunities for collaboration with educational and industrial partners.

We also wish to acknowledge the support of our colleagues in the Office of the Registrar & Vice President for Academic Affairs throughout the programme development process.

The programme development team is satisfied that the validation report is fair and accurate, and we are happy that the implementation of its various requirements and recommendations will further enhance the programme.

The following report outlines how these requirements and the recommendations have been addressed.

1.4 PROGRAMME STRUCTURE

Recommendations: The panel recommends that the previous mathematical attainment of applicants be considered in the selection of candidates for admission to the programme.

Response: Level 8 degrees in cognate disciplines (e.g. Mathematical Sciences, Engineering, Physics) were favoured in the shortlisting process. In many cases, supplementary information was sought, for example, the level at which Mathematics was taken for Leaving Certificate, and the grade attained in same. Due account was also taken of selection of related subjects for Leaving Certificate level, e.g. Applied Mathematics, Physics. A short diagnostic Maths test was given in the first week of delivery, and Mathematics learning support (1.5 hours per week) has been made available to students since Week 2.

2.1 ALL MODULES

2.1.1 Requirement: Any revisions to Module Descriptors or Semester Schedules made to address the recommendations and requirements in this require sign-off from the CIT Module Moderator and the Registrar's Office prior to approval by the CIT Academic Council.

Response: All modules have been revised and rechecked in the light of the requirements and recommendations in the panel's report, and the documentation has been submitted to the Module Moderator.

2.1.2 Requirement: The appropriate Reassessment Requirement should be selected for each module.

Response: complete

2.1.3 Comment: During the panel discussion the problem of the availability of suitable data sets was highlighted. Due to the nature of large data sets (messy and missing data), ethical approval and various issues of confidentiality the timescale to obtain these data sets should not be underestimated.

Response: Noted, and agreed. The programme team has already commenced work on this task.

2.2. Various Modules: Resource Listings

Requirement: Reading lists including supplementary texts should be expanded and updated, in order to ensure currency of content.

Response: completed

2.3. Various Modules: Workload

Recommendation: The part-time mode of delivery should be completed for all modules.

Response: completed

2.4 Module: DATA8001 Data Science and Analytics

Recommendation: Topics including technical writing, ethics, data privacy and security, data protection legislation should be included in this module.

Response: completed

2.5. Module: DATA8003 Text Analysis & Data Visuals

Recommendation: The panel recommends the modules should be retitled as the proposed title is too specific. Suggested titles discussed at the meeting included *Unstructured Data Visualisation* and *Data Analysis & Visualisation*, with the panel favouring the latter.

Response: completed

2.6. Module: STAT8006 Applied Stats & Probability

Recommendation: A new section should be included in the indicative content of the module to include Statistical Packages.

Response: completed

2.7. Modules: DATA8004 Data Mining & Knowledge Discovery and STAT8007 Statistical Methods for Big Data

Requirement: The learning outcomes and indicative content for DATA8004 needs to be reduced to a manageable workload for the learner. The panel recommends that topics such as data reduction techniques, clustering and classification methods, logistic regression models and other data mining techniques be rebalanced into STAT8007 with a corresponding reduction in focus on experimental design and traditional hypothesis testing in this module

Response: completed

2.8 Module: DATA8006 Project

Requirement: The proposed project module was found to be too generic in nature and the panel requires that the module be written specifically for the Data Science and Analysis programme.

Response: completed. In addressing this particular requirement, the programme team was anxious that, while meeting the requirements of the panel, the descriptor should remain flexible enough to allow each project to have a particular focus/emphasis, depending on the interests and ability of the learner, be that towards the Statistics/Mathematics field, or the Computer Science field.

APPENDIX 1 – Proposed Programme Outcomes

Programme Outcomes

On successful completion of this programme the learner will be able to :

| | | |
|------------|---------------------------------------|--|
| PO1 | Knowledge - Breadth | Demonstrate detailed knowledge and understanding of areas of Mathematics, Statistics, Computer Science and Business Intelligence relevant to the Data Analyst. |
| PO2 | Knowledge - Kind | Demonstrate understanding of the terminology, defining concepts and theories underlying the Data Science and Analytics field; demonstrate knowledge of the advanced methods and technologies for acquiring, interpreting and analysing big data, with a critical understanding of the appropriate contexts for their use; relate current issues in Data Science to society; understand current knowledge of the Data Science field, including current limits of theoretical and applied knowledge. |
| PO3 | Skill - Range | Demonstrate mastery of relevant skills and tools in Statistics, Mathematics, Computer Science and Business Intelligence; use these to solve complex problems involving big data sets; interpret and apply appropriate and referenced literature and other information sources; work independently within defined time and resource boundaries; communicate scientific information in a variety of forms to specialist and non-specialist audiences. |
| PO4 | Skill - Selectivity | Formulate and test hypotheses; design experiments; appreciate current limits of knowledge in the Data Science field and respond appropriately; think independently and make effective decisions; contribute fully to the day-to-day operations of the Data Science work setting. |
| PO5 | Competence - Context | Apply data analysis skills and technologies in a range of contexts in order to critically interpret existing knowledge and apply in new situations; make and report appropriate decisions in a responsible and ethical manner. |
| PO6 | Competence - Role | Act effectively under guidance in a peer relationship with qualified practitioners; participate constructively in a complex interdisciplinary team environment; plan for effective project implementation; reflect on own practices. |
| PO7 | Competence - Learning to Learn | Learn to act in variable and unfamiliar learning contexts; identify learning needs and undertake continuous learning in the Data Science field; assimilate and apply new learning. |
| PO8 | Competence - Insight | Demonstrate an understanding of the wider social, political, business and economic contexts of Data Science, including an appreciation of the philosophical and ethical issues involved. |

Appendix 2 – Proposed Semester Schedules

Semester 1

| Mandatory | | | | | | | | |
|-----------|--|--------------|----------|---------|------------------------|------------------------|-------------|------------|
| Mod Code | Module Title | Co-ordinator | Level | Credits | FT Hours Contact Hours | PT Hours Contact Hours | Course Work | Final Exam |
| DATA8001 | Data Science and Analytics (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 3.00 | 40.0% | 60% |
| MATH8009 | Maths Methods and Modelling (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 0.00 | 50.0% | 50% |
| DATA8002 | Data Management Systems (Draft) | JIM O DWYER | Advanced | 5.0 | 4.00 | 4.00 | 50.0% | 50% |
| DATA8003 | Text Analysis & Data Visuals (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 0.00 | 100.0% | 0% |
| COMP8042 | Analytical and Scientific Prog (Draft) | JIM O DWYER | Advanced | 5.0 | 4.00 | 4.00 | 100.0% | 0% |
| STAT8006 | Applied Stats & Probability (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 0.00 | 50.0% | 50% |

Semester 2

| Mandatory | | | | | | | | |
|-----------|---|--------------|----------|---------|------------------------|------------------------|-------------|------------|
| Mod Code | Module Title | Co-ordinator | Level | Credits | FT Hours Contact Hours | PT Hours Contact Hours | Course Work | Final Exam |
| DATA8004 | DataMining & KnowledgeDiscovery (Draft) | JIM O DWYER | Advanced | 5.0 | 4.00 | 4.00 | 50.0% | 50% |
| STAT8007 | Statistical Meth for Big Data (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 3.00 | 100.0% | 0% |
| DATA8005 | Distributed Data Management (Draft) | JIM O DWYER | Advanced | 5.0 | 4.00 | 4.00 | 100.0% | 0% |
| DATA8006 | Data Science Analytics Project (Draft) | AINE NI SHE | Advanced | 10.0 | 0.50 | 0.00 | 100.0% | 0% |

| Elective | | | | | | | | |
|----------|------------------------------------|--------------|----------|---------|------------------------|------------------------|-------------|------------|
| Mod Code | Module Title | Co-ordinator | Level | Credits | FT Hours Contact Hours | PT Hours Contact Hours | Course Work | Final Exam |
| COMP8043 | Machine Learning (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 0.00 | 100.0% | 0% |
| MATH8010 | Time Series & M-V Analysis (Draft) | AINE NI SHE | Advanced | 5.0 | 4.00 | 3.00 | 100.0% | 0% |