The Alliance for Data Science Professionals Guidance and Process:

Data Science Professional

Introduction

In this document we outline guidance notes and the process for how an individual can apply for **Data Science Professional** and the process and standards against which they are assessed.

All prospective applicants are advised to carefully read through this document before submitting an application.

Standards and Breadth of Knowledge

The data science standards have five main categories which are detailed below.

Skill Area	Evidential Requirements
A. Data Privacy and Stewardship This skill is about the security and protection of data, including design, creation, storage, distribution and associated risk.	1. Ensure the protection of personal and sensitive data.
	2. Managing loss of sensitive data
	3, Data Stewardship and Standards
B. Definition, acquisition, engineering, architecture, storage and curation. This skill is about the collection, manipulation and secure storage of data safely and securely. Applying data management and analytical techniques.	1. Data Collection and Management
	2. Data Engineering
	3. Deployment
	1. Problem definition

C. Problem definition and communication with stakeholders *This skill is about engaging stakeholders, demonstrating the ability to clearly define a problem and agree on solutions.*	*2. Relationship management*
D. Problem solving, analysis, statistical modelling, visualisation. *This skill is about identifying and presenting solutions using a range of methods, tools and techniques. Demonstrating an ability to analyse a problem and define and present options.*	*1. Identifying and applying appropriate solutions.* *2. Data Modelling* *3. Data Analysis*

Section E as a cross-cutting consideration to be evidenced throughout.

E. Evaluation and Reflection *This skill is about reflecting on performance and outcomes, identifying development needs and applying important principles associated with ethics and sustainability.* **Note:** *we expect items under Section E to be cross-cutting. Evidence of these should be embedded through the examples given in earlier sections.*	*1. Project Evaluation* *2. Ethical behaviour* *3. Sustainability and Best Practices* *4. Reflective Practice and Ongoing Development*

See appendix 1 for an expanded version, including types of suggested evidence and related skills.

When applying for the Data Science Professional, it is expected that all applicants:

- deliver **Applied** level of competence for **section E** and at least **two** other sections.
- Demonstrate a **limited** knowledge of the other **two** sections.

Level Descriptors

Level	Description
Limited	Has knowledge and understanding of facts, procedures and ideas in the field of work. Can interpret relevant information and ideas. Is aware of a range of information that is relevant to the area of work.
General	Has factual, procedural and theoretical knowledge and understanding. Can interpret and evaluate relevant information and ideas. Is aware of the area of work. Is aware of different perspectives or approaches within the area of work.
Applied	Has practical, theoretical or technical knowledge and understanding the field of work to address problems that are well defined but complex and non-routine. Can analyse, interpret and evaluate relevant information and ideas. Is aware of the nature of approximate scope of the area of work. Has an informed awareness of different perspectives or approaches within the area of work.
Deep	Has advanced practical, conceptual or technological knowledge and understanding of the field of work to create ways forward in contexts where there are many interacting factors. Understands different perspectives, approaches or schools of thought and the theories that underpin them. Can critically analyse, interpret and evaluate complex information, concepts and ideas.

Breadth of knowledge: Data Science Professional

Ethics & Efficacy

By working within the field of Data Science, it is important that all professionals have a clear understanding of the ethics which underpins the: collection, management, use and communication of the data and results they work with. It is equally important that a Data Scientist takes responsibility for the assurance of the models they build. Assurance covers both the efficacy of the application and the ethical natures of its design and implementation. As such, these attributes are not something that can, or should, be assessed as one standalone criteria. Therefore, when completing this application, you should wherever possible include your knowledge and working practices relating to the appropriate **ethical** considerations such as:

- data: collection, validity for use in the intended purpose, permission for usage, storage, security
- model: development, testing (e.g. fairness, bias, error rates) usage (how could the model and results be used for an unintended purpose?) and transparency
- communication: explanation of why the science is required; the results achieved and how can misinterpretation of the results be minimised?

- Relevant laws and permissions of usage for data (including legal rights of individuals, privacy and anonymity)

And **efficacy** considerations such as:

- Quality assurance of code and data
- Validation of model fit
- Robustness of the model and software implementation
- Ongoing monitoring of model implementation

It is important to note that the list is not exhaustive. It is here to serve as a guide to help you show the assessors you are aware of the professional expectations of those who work in this field. You should include any other areas of ethical and efficacy considerations you feel are important with your area of expertise.

Levels of Competence
Whilst there are two levels of certification associated with the Data Science standards. The standards will remain generic statements that can apply across a wide range of roles within the data science field. We intend for the standards to be agnostic of a practitioner's choice of tools.

The distinguishing features that define the levels are associated with the application of the standards and therefore levels of competence.

The distinction will be related to the following:

- **Responsibility** – the higher the level of registration the greater level of responsibility and accountability.
- **Decision making** – the level of authority to make decisions and the impact across the organisation.
- **Complexity** – this can be delivered within 2 spheres of complexity:
  - *Technical Complexity* – specifically associated with the technical skills applied.
  - *Organisational Complexity* – associated with skills and decision-making responsibilities that would apply across, and potentially beyond, the organisation.
- **Business impact** – relating to how far-reaching actions apply and impact and understanding who and how they impact.

Data Science Professional

For those wanting to apply for the Data Science Professional, we expect an individual to provide suitable evidence within their application that they are accountable and responsible for their own activities.

Applicants should demonstrate:
- Accountable for their own work.
- Works under general direction and understands when issues need escalating.
- Able to deliver technically complex solutions.

They should also demonstrate some of:
- Have responsibility for a function within the Team/Department.
- Have limited decision-making authority within their given area of expertise.
- Act as technical lead at Team level.
- Consider the impact across the Team and Business, of actions undertaken based on their decisions.

Requirements & Flexibility

Dependent on the experience or current role, areas of strength and weakness within the standards may differ, for example:
- As people progress towards Leadership/Management positions so their responsibility will increase, but technical complexity may diminish.
- Others may not progress towards Leadership roles but will develop highly complex and valuable technical skills.

Therefore, it is not essential that an applicant meets all the criteria at the required level, but that on balance, the totality of their evidence for each section meets the required level.

Similarly, different roles at the same level will have differing levels of competence with the Skill areas defined in the standard. For example, a Data Engineer may have strong evidence against Skills Area B but less developed evidence against skills area D.

Taking this into consideration, it is expected that all applicants at either level can display an appropriate level of competence for skill area E. However, sections A, B, C and D may be weighted differently depending on the area of specialism.

At the Data Science Professional level, an applicant would be expected to deliver applied level of competence for section E and two other sections. They should also demonstrate a limited knowledge of the other two sections.

To explain the levels and how the evidence might differ, below are a couple of examples:

Data Security

when presenting their evidence with regard to data security issues they may draw on specific examples of where they have influenced, helped develop or implemented a policy to ensure that the organisation's practices are commensurate with data security requirements. They should also be able to demonstrate sufficient understanding of appropriate practical responses to data security issues, to be able to provide oversight and governance of others' practical work.

Modelling

when presenting their evidence with regard to Data Modelling they may draw on specific examples of their role in determining the tools and techniques that the organisation may employ and why these tools and techniques were selected.

Application Process

Applicants applying to be a certified Data Science Professional must do so via the competency-based route.

Please be aware that whilst different Alliance members may have differences in how an application is assessed, the information requested and assessed remains the same.

The competency-based route

An individual applies, meeting all of the requirements of this route to the appropriate Alliance member → The application is reviewed by the assessors → The individual is notified of the outcome

Those individuals applying via the competency route will need to complete all sections of the application to an appropriate level.

- Personal information (if not known already)
- Academic/training history (including copies of transcripts and certificates)**
- Competency-based and responsibility-based statements
- Details of experience within a data science role.
- A completed CPD document

Assessment of Applications

By providing both information and evidence within the five sections below, applications will be passed to the relative assessors, whereby they will be able to clearly identify where and how you have met standards and as to if they have been met at the level of responsibility, awareness and understanding required.

Section 1	Personal Information
Section 2	Academic/training history
Section 3	Competency & Responsibility-based questions
Section 4	Work experience
Section 5	CPD

Whilst it is expected that most of those individuals applying at this level would have met a number of the standards via their academic and professional training, Assessors will be able to identify any skills and knowledge gaps that have then been met via a mixture of work experience, competencies and CPD. For a full list of suggested evidence in addition to academic and professional training, please see an extended view of the Data Science Standards in appendix 1 and a copy of the application form at appendix 2.

Appendix 1 – Professional Standards

Skill Area	Evidential Requirements	Types of evidence
A. Data Privacy and Stewardship This skill is about the security and protection of data, including design, creation, storage, distribution and associated risk.	1. Ensure the protection of personal and sensitive data.	i. Assess risks and enact data protection policies and procedures. ii. Ensure safe and secure management of sensitive data, models and infrastructures iii. Apply appropriate data controls, such as encryption, (pseudo)anonymisation, and synthetic data. iv. Risk management around environment and infrastructure
	2. Managing loss of sensitive data	i. Act with integrity, giving due regard to legal and regulatory requirements. ii. Be aware of the actions that should be taken to respond to potential data loss in line with organisational, legal and regulatory procedures.
	3, Data Stewardship and Standards	i. Incorporates the FAIR Guiding Principles for scientific data management and stewardship into practices, where appropriate and practicable. ii. Identify opportunities for efficient and creative reuse of data. iii. Understand the relationship between technical standards and regulation/governance, and their benefits for interoperability and knowledge sharing.
B. Definition, acquisition, engineering, architecture, storage and curation. This skill is about the collection, manipulation and secure storage of data safely and securely. Applying data	1. Data Collection and Management	i. Sourcing and accessing data appropriate for the problem. ii. Critically analyse the availability of appropriate data and resources to meet project requirements. iii. Critically evaluate and synthesise data. iv. Ensure data provenance processes are followed v. Identifying data characteristics (volume, velocity and variety) vi. Identify infrastructure requirements for data storage and analysis. vii. Familiarity or experience with tabular and non-tabular data (e.g. unstructured and streaming data).

management and analytical techniques.	2. Data Engineering	i. Sourcing and accessing data appropriate for the problem. ii. Constructing data sets, potentially drawing from multiple disparate sources using data linkage iii. Perform data profiling and characterisation to understand the surface properties of the data iv. Handling missing data, through principled inclusion/exclusion criteria and imputation methods. v. Take a systematic approach to data curation and the application of data quality controls. vi. Identify the most appropriate solutions (e.g. cloud vs on-premise) in response to business and project needs
	3. Deployment	i. Plan the deployment of data products with their end-users. ii. Develop monitoring and maintenance processes. iii. Deliver secure, stable and scalable data products to meet the needs of the organisation, e.g. Application Programming Interfaces (APIs), derivative datasets, dashboards, reports iv. Design and deliver data products that meet appropriate accessibility standards for their users.
C. Problem definition and communication with stakeholders This skill is about engaging stakeholders, demonstrating the ability to clearly define a problem and agree on solutions.	1. Problem definition	i. Identify and elicit project requirements ii. Determine success criteria and frame these in the context of the business. iii. Clearly articulate the problem statement. iv. Identify and critically evaluate assumptions. v. Recognise and quantify biases and identify solutions to manage and mitigate these. vi. Assess risk. vii. Sector/domain knowledge and knowledge of how data science can deliver value to these sectors/domains
	2. Relationship management	i. Communicate in an effective manner for diverse audiences, including technical colleagues, subject matter experts and leadership. ii. Effectively manage the expectations of diverse stakeholders with conflicting priorities to mediate equitable solutions.

		iii. Use relevant communication techniques (written, oral or visual), appropriate for the audience. iv. Build appropriate and effective business relationships. v. Experience in the human factors considerations with respect to data-driven solutions.
D. Problem solving, analysis, statistical modelling, visualisation. This skill is about identifying and presenting solutions using a range of methods, tools and techniques. Demonstrating an ability to analyse a problem and define and present options.	1. Identifying and applying appropriate solutions.	i. Identify viable solutions based on requirements and data available. ii. Identify and provide guidance to technical and non-technical stakeholders on the most appropriate solution. iii. Apply appropriate technical and project management methodologies appropriate for the organisation and project.
	2. Data Modelling	i. Identify appropriate solutions, including statistical and machine learning approaches. ii. Identify and evaluate appropriate evaluation metrics, including computational performance and accuracy. iii. manipulating data with due regard for differences in characteristics.
	3. Data Analysis	i. Apply appropriate solutions, including statistical and machine learning approaches. ii. Use appropriate analysis platforms and tools. iii. Adopt a systematic approach to exploratory data analysis to embrace and manage ambiguity and uncertainty iv. Critically analyse data and analytical results. v. Adopt appropriate methods to visualise data and communicate complex findings.

Section E as a cross-cutting consideration to be evidenced throughout.

E. Evaluation and Reflection This skill is about reflecting on performance and outcomes, identifying development needs and applying important principles associated with ethics and sustainability. Note: we expect items under Section E to be cross-cutting. Evidence of	1. Project Evaluation	i. Ongoing monitoring of project performance and outcomes. ii. Identify and feed forward lessons learned. iii. Participate and lead collaborative project evaluations, e.g. retrospectives
	2. Ethical behaviour	i. Identify and manage the risks of erroneous and biased data ii. Acting with integrity with respect to legal and regulatory requirements. iii. Upholding principles of ethical and safe use of data and AI technologies iv. Implementing data use procedures to ensure sensitive data is only used for its agreed purpose

these should be embedded through the examples given in earlier sections.		v. Implement data retention strategies in line with regulatory and legal requirements.
	3. Sustainability and Best Practices	i. Evidence of incorporating the principles of open science and/or reproducible research within the organisation, and perhaps beyond. ii. A familiarity with programmatic approaches to undertaking data science work. iii. Application of the scientific method in delivering solutions iv. Ensure high technical standards, in line with software development best practices; for example, software testing, version control, Continuous Integration and Continuous Delivery. v. Apply automation to promote reproducibility analyses
	4. Reflective Practice and Ongoing Development	i. Learning from experience through self-assessment of one's own responses to practice situations. ii. Identify learning opportunities to maintain knowledge and skills in their area of data science. iii. Taking ownership for ongoing professional development. iv. Contributing to knowledge sharing across their organisation and/or the wider community. v. Contribution to the management and empowerment of the broader team. vi. Engage with the latest developments across industry and academia and incorporate these into your solutions.

Appendix 2 – Copy of application form

Data Science Professional (DSP)

To apply for DSP certification, please complete and submit this form, with your CV, through our DSP application process.

Surname		First name(s)	
Title (Mr/Mrs/Ms etc)		BCS membership no.	

1. Details of Relevant Qualification

You will be required to provide authentication of these qualifications, either through a copy of the certificate or provision of a URL through which they can be checked.

Type of Qualification	Master's degree	Bachelor's Degree with Honours	Apprenticeship	HND	Other
*if other please specify					
Course title (as shown on certificate)					
Institution					
Country					
Start date	Graduation date				
Mode of study:	Full time		Part time		Distance learning	Sandwich	

Verification

You'll need to provide an authenticated copy of your certificate **or** the URL of a public register or other method where we can verify your certification.

Please complete *one* of the following fields:

I've included an authenticated copy of my certificate with my application (mark with an 'x')		

My certification can be verified here:	*Enter URL*

Note: you are only required to provide evidence of competence against section E and two other sections of the criteria.

3. I have provided evidence against the following areas of the ADSP criteria

Please select the 2 sections for which you have provided evidence (choose only two) by marking with an 'x'.

Section E has already been pre-selected.

Section A - Data Privacy and Stewardships	
Section B - Definition, acquisition, engineering, architecture, storage, and curation.	
Section C - Problem definition and communication with stakeholders.	
Section D - Problem solving, analysis, modelling, visualisation.	
Section E - Evaluation and Reflection.	X

4. Breadth of knowledge

You are required to provide evidence of a limited understanding of the 2 sections for which you have not demonstrated competence.

Please include details of relevant academic qualifications and other training and development you've undertaken.

5. Experience

In the following sections, we're looking for evidence that you've worked in one or more challenging, multifaceted roles where you've had personal responsibility for your work, apply technical skills in delivering outcomes, with an understanding of how your practice impacts other departments within the organisation.

Here's a useful guide to completing your experience statements.

Note: In the following sections please annotate with an 'x' the section for which you are providing evidence.

Section	A	B	C	D	

Note: be sure to include ethical issues that you have had to consider, resolve or act upon within your evidence.

Section	A		B		C		D	

Note: be sure to include ethical issues that you have had to consider, resolve or act upon within your evidence.

Section E

E1 – Project Evaluation

E2 – Ethical Behaviour

E3 – Sustainability and Best Practice

E4 – Reflective Practice and Ongoing Development

How we use your data at BCS

We'll store your basic personal information, such as your name and email address, so that we can process your application and communicate with you about your registration. This may include contact from our assessor (during the application process), welcome communications, information about accessing and getting the most from your registration, and information about your renewal and revalidation.

We'll always keep your information safely and never pass it to a third party without your permission. Full details of our data protection and privacy policies are available online at bcs.org/privacy.

Marketing preferences
We're involved in a wide range of activities in the BCS Group, driven by our royal charter and our purpose to make IT good for society. If you'd like to know more about these, please log in to MyBCS and tell us your marketing preferences.

6. Register of Certified Data Science Professionals Would you like to appear on the Register for Data Science Professionals (No contact details are displayed.)	Yes	No