

The Alliance for Data Science Professionals
Certification Guidance and Process:
Data Science Professional

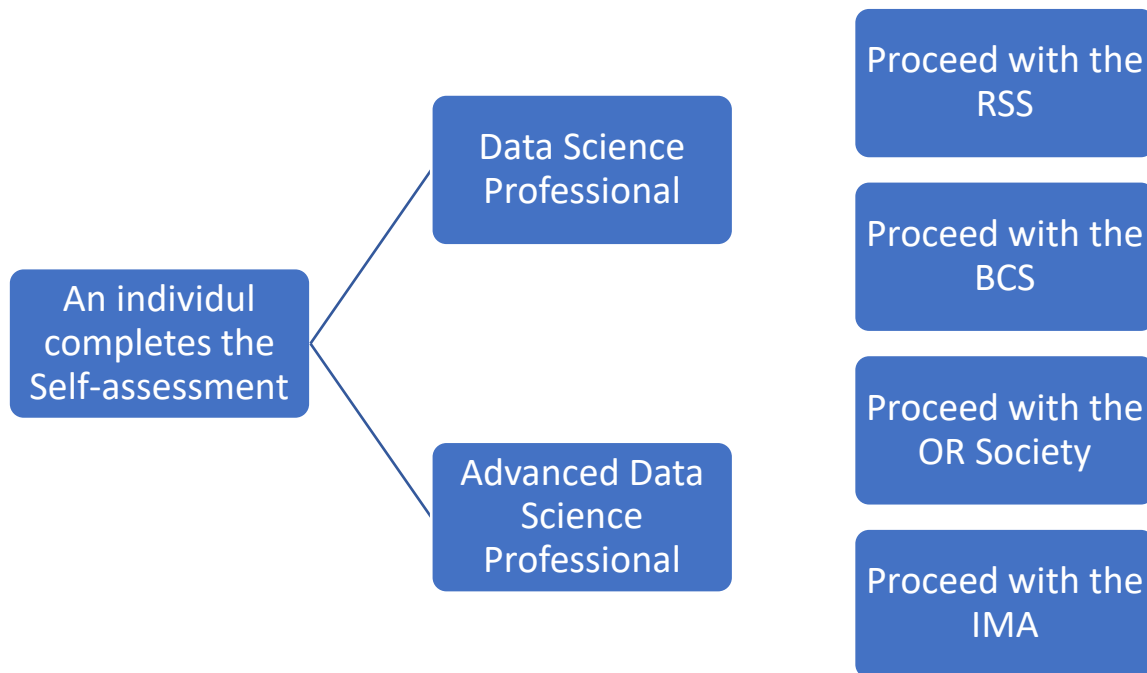
Introduction

All prospective applicants are advised to carefully read through this document before submitting an application with their selected Alliance member.

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

Please note:

Prior to beginning your formal application, all applicants are expected to have completed the self-assessment form available at the Alliance website. Completing this self-assessment will provide guidance as to which Alliance member and level of certificate may be appropriate, providing it is then supported by a completed application based on the original information provided.



Standards and Breadth of Knowledge

The data science standards have five main categories, with ethics* being a sixth category included within each of the other five. The categories are below.

**Please see the breadth of knowledge section for guidance on ethics*

Skill Area	Evidential Requirements
<p>A. Data Privacy and Stewardships</p> <p>This skill is about the security and protection of data including design, creation, storage, distribution and associated risk.</p>	1. Ensure the protection of personal and sensitive data.
	2. Managing loss of sensitive data
	3. Data Stewardship
<p>B. Definition, acquisition, engineering, architecture, storage and curation.</p> <p>This skill is about the collection, manipulation and secure storage of data safely and securely. Applying data management and analytical techniques.</p>	1. Data Collection and Management
	2. Data modelling
	3. Data Engineering
	4. Deployment
<p>C. Problem definition and communication with stakeholders</p> <p>This skill is about engaging stakeholders, demonstrating the ability to clearly define a problem and agree solutions.</p>	1. Problem definition
	2. Relationship management
<p>D. Problem solving, analysis, statistical modelling, visualisation.</p> <p>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.</p>	1. Identifying and applying appropriate solutions.
	2. Data Management
	3. Data Analysis
<p>E. Evaluation and Reflection</p> <p>This skill is about reflecting on performance and outcomes, identifying development needs and applying important principles associated with ethics and sustainability.</p> <p>Note: we expect items under Section E to be cross-cutting. Evidence of these should be embedded through the examples given in earlier sections.</p>	1. Project Evaluation

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:

- Applied understanding applied understanding and knowledge in 3 areas of skill from A –E.

I have practical, theoretical or technical knowledge and understanding within the field of work to address problems that are well defined but complex and non-routine. I 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..

- and limited understanding and knowledge in the remaining two areas

I have limited conceptual 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.

Level Descriptors

Skill Level	Proficiency	Experience
Limited	Limited conceptual knowledge	Minimal experience – read about it, some education and some practice with tools Some exposure in educational training setting
General	General conceptual knowledge Theoretical knowledge applied in education or training context	Performs in education or training setting Completed formal education, including Capstone type project
Applied	Applied knowledge	Performs with supervision or mentoring
Deep	In-depth knowledge	Mastered the current state-of-the-art and is able to perform without supervision
Expert	Expert knowledge	Advances the state-of-the-art

Breadth of knowledge: Data Science Professional

Ethics

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. As such, ethics is not something that can, or should, be assessed as one standalone criteria. Rather, when completing this application you should wherever 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)

It is important to note that this 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 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. We intend for our standard to be inclusive of practitioners who use low-/no-code solutions to perform their work and practitioners who approach their work primarily programmatically.

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

Applicants should demonstrate that they:

- Have personal responsibility for their own work.
- Apply technical skills in delivering outcomes.

They should also demonstrate some of the following:

- Have responsibility for activities of a section or team.
- Have decision-making authority at the Section or team level.
- Act as an advisor/consultant and departmental level.
- Understands how their individual practice impacts other departments.

Taking this approach, individuals would provide evidence against the skills determined in the standard, demonstrating their level of application against the criteria above.

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 higher 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 Advanced Data Science Professional level, an applicant would be expected to deliver deep level of competence for section E two other sections. They should also demonstrate a general 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 Advanced 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

For those that have been educated to degree level in a non-accredited degree at least some formal training within data science either as part of their degree or achieved by other methods. In addition to this applicants must meet all other criteria listed below:

- have at least two-years' qualifying work experience
- at least one year's evidence of CPD
- Can evidence they meet the competencies and level of responsibilities of the Data Science Certificate via the chosen assessment method of the awarding society.



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 at least two-years of qualifying work experience
- A completed CPD document

Application form

The Application form for the Data Science Certificate is typically split into five sections, with section 3 usually only required for those applying via the competency-based route. Please be aware that the information you provide and how this information is assessed may vary between different Alliance members so do ensure you have reviewed the requirements of the member body you are applying with.

Section 1	Personal Information	Name, address, memberships etc...	This section may be auto-completed
Section 2	Academic/training history	Certificates Transcripts Name of institution/training provider Dates Relative modules	Failure to provide evidence of both certificates and transcripts may result in applications being rejected
Section 3	Competency & Responsibility-based questions	Competency statements Personal statements	In addition to information provided within an application, you may be asked for references, others may cover this at an interview.
Section 4	Work experience	Brief details of work experience from the last two years, including tasks undertaking, the level of responsibility during your time there	
Section 5	CPD	A completed document highlighting how you maintaining your professional development in accordance to the CPD guidelines of the appropriate membership body	

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.

Example application

The competency-based Route:

This individual has completed a non-accredited degree

Section 1	Personal Information	Joe Bloggs
Section 2	Academic/training history	Degree – MSc in Maths & Statistics Certificates - attached Transcripts - attached Name of institution/training provider – The Open University Dates – qualified May 2018 Relative modules – M348 Applied Statistical Modelling'
Section 3	Competency & Responsibility-based questions	Need advice from group and assessors for this part
Section 4	Work experience	Employer 1 as data analyst August 2019 – Present Level of responsibility – Whilst I had responsibility for some daily tasks, I worked as part of an analytics team that analysed and cleaned data for use by our research team Brief description of my role: XXXXX
Section 5	CPD	Please see a CPD activity summary attached from the last year, including 60 hours of CPD.

This is a basic example only

Please note that based on the information provided and other certificates the individual may be applying for at the time of submitting this application, the reviewing body may wish to request more information that can include, but is not limited to:

- Shorter application questions, followed by an interview
- Longer application questions only
- A mixture of both
- Personal statements
- Short assessment/accredited certificate

Appendix

Skill Area	Evidential Requirements	Types of evidence	Related skills
<p>A. Data Privacy and Stewardships</p> <p>This skill is about the security and protection of data including design, creation, storage, distribution and associated risk.</p>	<p>1. Ensure the protection of personal and sensitive data.</p>	<p>i. Assess risks and enact data protection policies and procedures</p> <p>ii. Ensure safe and secure management of sensitive data, models and infrastructures</p> <p>iii. Apply appropriate data controls, such as encryption, (pseudo)anonymisation, and synthetic data.</p> <p>iv. Risk management around environment and infrastructure</p>	<p>Level 7 AI Data Specialist</p> <p>S1, B3, Duty 13</p> <p>DDat Data Scientist</p> <p>Supporting strategic and operational decision making</p> <p>Level 7 Operational Research Specialist</p> <p>K18, S14</p>
	<p>2. Managing loss of sensitive data</p>	<p>i. Act with integrity, giving due regard for legal and regulatory requirements.</p> <p>ii. Be aware of the actions that should be taken to respond to potential loss of data in line with organisational, legal and regulatory procedures.</p>	<p>SFIA</p> <p>CORE</p> <p>OpenCDS</p> <p>CBS07</p>
	<p>3. Data Stewardship</p>	<p>i. Incorporates the FAIR Guiding Principles for scientific data management and stewardship into practices, where appropriate and practicable.</p> <p>ii. Identify opportunities for efficient and creative reuse of data.</p>	
<p>B. Definition, acquisition, engineering,</p>	<p>1. Data Collection and Management</p>	<p>i. Sourcing and accessing data appropriate for the problem.</p>	<p>Level 7 AI Data Specialist</p>

<p>architecture, storage and curation.</p> <p>This skill is about the collection, manipulation and secure storage of data safely and securely. Applying data management and analytical techniques.</p>		<p>ii. Critically analyse the availability of appropriate data and resources to meet project requirements.</p> <p>iii. Critically evaluate and synthesise data.</p> <p>iv. Ensure data provenance</p> <p>v. Identifying data characteristics (volume, velocity and variety)</p> <p>vi. Familiarity or experience with tabular and non-tabular data (e.g. unstructured and streaming data).</p>	<p>K14, S17, K27</p> <p>DDat Data Scientist</p> <p>Sourcing, accessing and manipulating</p> <p>Level 7 Operational Research Specialist</p> <p>S3, K6, K24</p> <p>SFIA</p> <p>DATM, INAS, STPL, RLMT, KNOW, MEAS, AVMT, STMG, CPMG</p> <p>OpenCDS</p> <p>DSS04</p>
	2. Data modelling	<p>i. Identify appropriate solutions, including statistical and machine learning approaches.</p> <p>ii. Identify and evaluate appropriate evaluation metrics, including computational performance and accuracy.</p> <p>iii. manipulating data with due regard for differences in characteristics.</p> <p>iv. Application of the scientific method in delivering solutions</p> <p>v. Engage with the latest developments across industry and academia and incorporate these into your solutions.</p>	
	3. Data Engineering	<p>i. Conducting exploratory data analysis</p>	

		<ul style="list-style-type: none"> ii. Take a systematic approach to data curation and the application of data quality controls. iii. Identify infrastructure requirements for data storage and analysis. iv. Identify the most appropriate solutions (e.g. cloud vs on-premise) in response to business and project needs 	
	4. Deployment	<ul style="list-style-type: none"> 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. 	
<p>C. Problem definition and communication with stakeholders</p> <p>This skill is about engaging stakeholders, demonstrating</p>	1. Problem definition	<ul style="list-style-type: none"> 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. 	<p>Level 7 AI Data Specialist</p> <p>S5, S8, K1</p> <p>DDat Data Scientist</p> <p>Supporting Strategic and operational decision making</p>

<p>the ability to clearly define a problem and agree solutions.</p>		<p>v. Recognise and quantify biases and identify solutions to manage and mitigate these.</p> <p>vi. Assess risk.</p> <p>vii. Sector/domain knowledge and knowledge of how data science can deliver value to these sectors/domains</p>	<p>Level 7 Operational Research Specialist</p> <p>K1, K2, K3, K17, B1, B4, S1, S13, K10, K12, S9, S10</p> <p>SFIA</p> <p>LEDA, PEMT, ETDL, DLMG, ARCH, REQM, QUMG</p>
	<p>2. Relationship management</p>	<p>i. Communicate in an effective manner for diverse audiences, including technical colleagues, subject matter experts and leadership.</p> <p>ii. Effectively manage the expectations of diverse stakeholders with conflicting priorities to mediate equitable solutions.</p> <p>iii. Use relevant communication techniques (written, oral or visual), appropriate for the audience.</p> <p>iv. Build appropriate and effective business relationships.</p>	<p>OpenCDS</p> <p>CBS04</p>
<p>D. Problem solving, analysis, statistical modelling, visualisation.</p> <p>This skill is about identifying and</p>	<p>1. Identifying and applying appropriate solutions.</p>	<p>i. Identify viable solutions based on requirements and data available.</p> <p>ii. Identify and provide guidance to technical and non-technical stakeholders on the most appropriate solution.</p>	<p>Level 7 AI Data Specialist</p> <p>S12, B3, S10, K7, S2, K4, S1, S17, S16, K2, S13, K3, K19, S11, S18, K20, S9, S21</p>

<p>presenting solutions using a range of methods, tools and techniques. Demonstrating an ability to analyse a problem and define and present options.</p>		<p>iii. Apply appropriate technical and project management methodologies appropriate for the organisation and project.</p>	<p>DDat Data Scientist</p> <p>Exploring and visualising data.</p> <p>Level 7</p>
	<p>2. Data Management</p>	<p>i. Sourcing and accessing data appropriate for the problem.</p> <p>ii. Constructing data sets, potentially drawing from multiple disparate sources</p> <p>iii. Perform data profile and characterisation to understand the surface properties of the data</p> <p>iv. Handling missing data, through enforcing inclusion/exclusion criteria and imputation methods.</p> <p>v. Incorporates the FAIR Guiding Principles for scientific data management and stewardship into practices, where appropriate and practicable.</p>	<p>Operational Research Specialist</p> <p>B8, S15, K4, K5, K6, K7, K8, K9, K13, S3, S4, S6, S7, S9</p> <p>SFIA</p> <p>ARCH, DTAN, DBDS, BUAN, DESN, PROG, BPTS, TEST, INAN, VISL, INCA, ICPM, DBAD</p> <p>OpenCDS</p> <p>DSS08, DSS04, DSS05</p>
	<p>3. Data Analysis</p>	<p>i. Apply appropriate solutions, including statistical and machine learning approaches.</p> <p>ii. Use appropriate analysis platforms and tools.</p> <p>iii. Adopt a systematic approach to exploratory data analysis to embrace and manage ambiguity and uncertainty</p>	

		<p>iii. Critically analyse data and analytical results.</p> <p>iv. Adopt appropriate methods to visualise data and communicate complex findings.</p> <p>v. Ensure high technical standards, in line with software development best practices; for example, software testing, version control, Continuous Integration and Continuous Delivery.</p> <p>vi. Apply automation to promote reproducibility analyses</p>	
<p>E. Evaluation and Reflection</p> <p>This skill is about reflecting on performance and outcomes, identifying development needs and applying important principles associated with ethics and sustainability.</p> <p>Note: we expect items under Section E to be</p>	<p>1. Project Evaluation</p>	<p>i. Ongoing monitoring of project performance and outcomes.</p> <p>ii. Identify and feed forward lessons learned.</p> <p>iii. Participate and lead collaborative project evaluations, e.g. retrospectives</p>	<p>Level 7 AI Data Specialist</p> <p>B3, S22, Duty 3, Duty 6, K7, S18, K6, S22, S14, K29</p> <p>DDat Data Scientist</p> <p>Taking and interdisciplinary focus, Adhering to the data science ethics framework, supporting strategic decision making.</p> <p>Level 7 Operational Research Specialist</p>

cross-cutting. Evidence of these should be embedded through the examples given in earlier sections.			S10, S9, SFIA BPRE, INOV, OpenCDS DEC02
--	--	--	--