Institute for Apprenticeships & Technical Education

ARTIFICIAL INTELLIGENCE (AI) DATA SPECIALIST

Details of standard

Occupation summary

This occupation is found in any sector or organisation that analyses high-volume or complex data sets using advanced computational methods, such as Agriculture, Environmental, Business, Leisure, Travel, Hospitality, Education, Public Services, Construction, Creative and Design, Media, Engineering, Technology, Manufacturing, Health, Science, Legal, Finance, Accountancy, Sales, Marketing, Procurement, Transport and Logistics

The broad purpose of the occupation is to discover and devise new data-driven AI solutions to automate and optimise business processes and to support, augment and enhance human decision-making. AI Data Specialists carry out applied research in order to create innovative data-driven artificial intelligence (AI) solutions to business problems within the constraints of a specific business context. They work with datasets that are too large, too complex, too varied or too fast, that render traditional approaches and techniques unsuitable or unfeasible.

AI Data Specialists champion AI and its applications within their organisation and promote adoption of novel tools and technologies, informed by current data governance frameworks and ethical best practices.

They deliver better value products and processes to the business by advancing the use of data, machine learning and artificial intelligence; using novel research to increase the quality and value of data within the organisation and across the industry. They communicate, internally and externally, with technology leaders and third parties.

In their daily work, an employee in this occupation interacts with a broad spectrum of people and collaborates with, and provides technical authority and insight to, a diverse business community of Senior Leaders Data Scientists, Data Engineers, Statisticians, Analysts, Research and Development Scientists and Academics. Their interactions extend to working externally alongside other organisations, such as local and international governments, businesses, policy regulators, academic research scientists and non-technical audiences. They will work independently and collaboratively as required, reporting to Heads of Data, Chief Architects, Company Directors, Product Managers and senior decision makers within any organisation.

An employee in this occupation will be responsible for initiating new projects in an agile environment, and collaboratively maintaining technical standards within AI solutions applied across the organisation and its customers. They lead research into AI and its potential application within the business. They collaborate

with and influence policy and operations teams to identify areas where AI solutions can create new business opportunities and efficiencies.

Typical job titles include:

Ai strategy manager | **Artificial intelligence engineer** | **Artificial intelligence specialist**

Director ai | **Machine learning engineer** | **Machine learning specialist**

Occupation duties

DUTY	CRITERIA FOR MEASURING PERFORMANCE	KSBS
Duty 1 Initiate new projects in an agile environment, and collaboratively maintain technical standards within AI solutions applied across the organisation and its customers.	Research prototypes are developed to organisational/customer requirements in line with industry standards.	K1 K5 K6 K13 K17 K21 K29 S4 S5 S11 S12 S21 S22 S23 S24 S28 B2 B5 B8
Duty 2 Critically evaluate and synthesise research findings in AI and related fields and translate into organisational context.	Research findings in AI and related fields are clearly articulated and documented, translating them into potential impacts, opportunities and threats for the organisation.	K1 K3 K7 K17 K18 K19 K21 K22 K26 S2 S3 S4 S5 S6 S11 S12 S24 S26 B1 B4 B8
Duty 3 Use the conclusions drawn from applied research in order to develop innovative, scalable data-driven AI solutions for business problems	New projects are initiated and maintained to organisational/customer requirements in line with industry standards.	K2 K5 K7 K14 K17 K18 K19 K26 K28 S2 S3 S5 S9 S11 S12 S24 S25 S26 B2 B4 B7
Duty 4 Contribute to the development and ethical and legal conduct of AI systems and processes, in line with organisational and regulatory requirements.	Effective solutions are delivered in accordance with principles of responsible research and innovation for automated AI decision-making systems. Governance frameworks are established which take into account legal and regulatory requirements including privacy issues.	K8 K9 K10 K11 K12 K24 K29 S6 S8 S12 S17 B1 B2 B3 B7
Duty 5 Investigate and devise the most efficient and effective architectures, to enable and maximise the use and impact of AI systems and solutions for the organisation.	Effective architectures are delivered in line with agreed timescales and to organisational requirements.	K2 K13 K15 K16 K19 K26 K29 S13 S14 S15 S16 S19 S25 B3 B7

Duty 6 Develop innovative approaches to tackle known business problems that previously did not have a feasible solution within the constraints of a specific business context.	Innovative approaches are developed to meet industry standards utilising a full range of AI and related technologies to create and build solutions that can be used by strategic or operational users and can be further integrated into business systems.	K1 K7 K13 K29 S3 S6 S13 S27 B1 B3 B4 B6
Duty 7 Initiate and design scalable batch/real-time analytical solutions to business problems leveraging AI and related technologies such as, data science, machine learning and statistics and related technologies.	Solutions are designed and developed in line with agreed timescales and organisational and industry standards.	K5 K17 K18 K20 K23 K25 K27 S14 S18 S19 S20 S25 S26 S28 B2 B4 B7
Duty 8 Enhance awareness of the wider application of AI tools and technologies across the business so that opportunities for its use can be identified	The use of AI and its applications are championed within the organisation and novel tools and technologies are adopted.	K10 K11 K12 K14 K21 K24 K27 K28 S4 S7 S8 S11 S23 S27 B1 B5 B7 B8
Duty 9 Develop and architect new robust data sourcing and processing systems to serve the organisation.	New data sources are integrated into business processes in line with organisational change management processes. Analytics and statistical methods for data preparation and pre-processing are applied. Opportunities are identified to integrate data from silos both within and outside the organisation, to provide value added insights. These data pipelines should follow	K2 K4 K7 K8 K15 K16 K18 S1 S2 S4 S5 S6 S10 S13 S14 S15 S16 S25 B1 B3 B4

	organisational and general architecture best practice.	
Duty 10 Design technical roadmaps for data life-cycles ensuring appropriate support and business processes are in place.	Technical roadmaps are designed and maintained to organisational requirements. Clear plans for evolution of the technologies, and the relevant support and business processes, are in place.	K6 K9 K10 S1 S17 S18 S25 B1 B2 B4 B7
Duty 11 Create and optimise efficient mechanisms for accessing and analysing datasets that are too large, too complex, too varied or too fast, that render traditional approaches and techniques unsuitable or unfeasible, in order to deliver business outcomes	Bespoke problem-specific mechanisms that consider performance limitations are developed and tested to meet organisational/customer data access and analysis requirements.	K16 K18 K19 K20 K22 K23 S8 S9 S10 S20 S25 S26 B1 B2 B4
Duty 12 Identify best practice in AI data systems, data structures, data architecture and data warehousing technologies and provide technical oversight in order to meet business objectives.	Future business/domain opportunities are researched, identified and completed in line with organisational/customer requirements. Rigorous scientific methodology is followed at all stages of research activity, including communication of uncertainty in results of experiments and analysis.	K2 K4 K7 K8 K15 K16 K24 S1 S2 S4 S5 S6 S10 S13 S23 B1 B6 B7
Duty 13 Assess risks/limitations and quantify biases associated with applications of AI within given business contexts.	Risks are assessed according to organisational policy/customer requirements/industry standards.	K1 K3 K22 K24 K25 S10 S11 S12 S13 S21 S22 S28 B1 B4 B8
Duty 14 Provide technical authority for the business regarding emerging opportunities for AI.	Direction and guidance for the business is clearly	K8 K10 K11 K12 K17 K21 K28

	articulated to industry standards. Strategic opportunities are identified and new insights relevant to business goals are generated.	S4 S5 S6 S7 S18 S19 S23 S26 B6 B7 B8
Duty 15 Practice continuous self-learning to keep up to date with technological developments to enhance relevant skills and take responsibility for own professional development		K10 K17 S7 S11 B5 B8

KSBs

Knowledge

K1: How to use AI and machine learning methodologies such as data-mining, supervised/unsupervised machine learning, natural language processing, machine vision to meet business objectives

K2: How to apply modern data storage solutions, processing technologies and machine learning methods to maximise the impact to the organisation by drawing conclusions from applied research

K3: How to apply advanced statistical and mathematical methods to commercial projects

K4: How to extract data from systems and link data from multiple systems to meet business objectives

K5: How to design and deploy effective techniques of data analysis and research to meet the needs of the business and customers

K6: How data products can be delivered to engage the customer, organise information or solve a business problem using a range of methodologies, including iterative and incremental development and project management approaches

K7: How to solve problems and evaluate software solutions via analysis of test data and results from research, feasibility, acceptance and usability testing

K8: How to interpret organisational policies, standards and guidelines in relation to AI and data

K9: The current or future legal, ethical, professional and regulatory frameworks which affect the development, launch and ongoing delivery and iteration of data products and services.

K10: How own role fits with, and supports, organisational strategy and objectives

K11: The roles and impact of AI, data science and data engineering in industry and society

K12: The wider social context of AI, data science and related technologies, to assess business impact of current ethical issues such as workplace automation and misuse of data

K13: How to identify the compromises and trade-offs which must be made when translating theory into practice in the workplace

K14: The business value of a data product that can deliver the solution in line with business needs, quality standards and timescales

K15: The engineering principles used (general and software) to investigate and manage the design, development and deployment of new data products within the business

K16: Understand high-performance computer architectures and how to make effective use of these

K17: How to identify current industry trends across AI and data science and how to apply these

K18: The programming languages and techniques applicable to data engineering

K19: The principles and properties behind statistical and machine learning methods

K20: How to collect, store, analyse and visualise data

K21: How AI and data science techniques support and enhance the work of other members of the team

K22: The relationship between mathematical principles and core techniques in AI and data science within the organisational context

K23: The use of different performance and accuracy metrics for model validation in AI projects

K24: Sources of error and bias, including how they may be affected by choice of dataset and methodologies applied

K25: Programming languages and modern machine learning libraries for commercially beneficial scientific analysis and simulation

K26: The scientific method and its application in research and business contexts, including experiment design and hypothesis testing

K27: The engineering principles used (general and software) to create new instruments and applications for data collection

K28: How to communicate concepts and present in a manner appropriate to diverse audiences, adapting communication techniques accordingly

K29: The need for accessibility for all users and diversity of user needs

Skills

S1: Use applied research and data modelling to design and refine the database & storage architectures to deliver secure, stable and scalable data products to the business

S2: Independently analyse test data, interpret results and evaluate the suitability of proposed solutions, considering current and future business requirements

S3: Critically evaluate arguments, assumptions, abstract concepts and data (that may be incomplete), to make recommendations and to enable a business solution or range of solutions to be achieved

S4: Communicate concepts and present in a manner appropriate to diverse audiences, adapting communication techniques accordingly

S5: Manage expectations and present user research insight, proposed solutions and/or test findings to clients and stakeholders.

S6: Provide direction and technical guidance for the business with regard to AI and data science opportunities

S7: Work autonomously and interact effectively within wide, multidisciplinary teams

S8: Coordinate, negotiate with and manage expectations of diverse stakeholders suppliers with conflicting priorities, interests and timescales

S9: Manipulate, analyse and visualise complex datasets

S10: Select datasets and methodologies most appropriate to the business problem

S11: Apply aspects of advanced maths and statistics relevant to AI and data science that deliver business outcomes

S12: Consider the associated regulatory, legal, ethical and governance issues when evaluating choices at each stage of the data process

S13: Identify appropriate resources and architectures for solving a computational problem within the workplace

S14: Work collaboratively with software engineers to ensure suitable testing and documentation processes are implemented.

S15: Develop, build and maintain the services and platforms that deliver AI and data science

S16: Define requirements for, and supervise implementation of, and use data management infrastructure, including enterprise, private and public cloud resources and services

S17: Consistently implement data curation and data quality controls

S18: Develop tools that visualise data systems and structures for monitoring and performance

S19: Use scalable infrastructures, high performance networks, infrastructure and services management and operation to generate effective business solutions.

S20: Design efficient algorithms for accessing and analysing large amounts of data, including Application Programming Interfaces (API) to different databases and data sets

S21: Identify and quantify different kinds of uncertainty in the outputs of data collection, experiments and analyses

S22: Apply scientific methods in a systematic process through experimental design, exploratory data analysis and hypothesis testing to facilitate business decision making

S23: Disseminate AI and data science practices across departments and in industry, promoting professional development and use of best practice

S24: Apply research methodology and project management techniques appropriate to the organisation and products

S25: Select and use programming languages and tools, and follow appropriate software development practices

S26: Select and apply the most effective/appropriate AI and data science techniques to solve complex business problems

S27: Analyse information, frame questions and conduct discussions with subject matter experts and assess existing data to scope new AI and data science requirements

S28: Undertakes independent, impartial decision-making respecting the opinions and views of others in complex, unpredictable and changing circumstances

Behaviours

B1: A strong work ethic and commitment in order to meet the standards required.

B2: Reliable, objective and capable of independent and team working

B3: Acts with integrity with respect to ethical, legal and regulatory ensuring the protection of personal data, safety and security

B4: Initiative and personal responsibility to overcome challenges and take ownership for business solutions

B5: Commitment to continuous professional development; maintaining their knowledge and skills in relation to AI developments that influence their work

B6: Is comfortable and confident interacting with people from technical and non-technical backgrounds. Presents data and conclusions in a truthful and appropriate manner

B7: Participates and shares best practice in their organisation, and the wider community around all aspects of AI data science

B8: Maintains awareness of trends and innovations in the subject area, utilising a range of academic literature, online sources, community interaction, conference attendance and other methods which can deliver business value

Qualifications

English and Maths

Apprentices without level 2 English and maths will need to achieve this level prior to taking the End-Point Assessment. For those with an education, health and care plan or a legacy statement, the apprenticeship's English and maths minimum requirement is Entry Level 3. A British Sign Language (BSL) qualification is an alternative to the English qualification for those whose primary language is BSL.

Additional details

Occupational Level:

7

Duration (months):

24

Review

This standard will be reviewed after three years.

Version log

VERSION	CHANGE DETAIL	EARLIEST START DATE	LATEST START DATE	LATEST END DATE
1.0	Approved for delivery	13/05/2020	Not set	Not set