





2nd **Edition** 2023







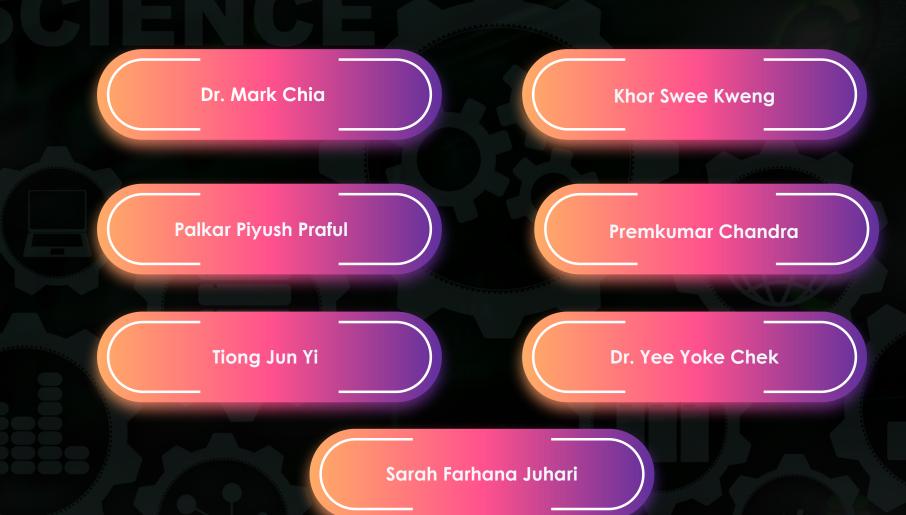




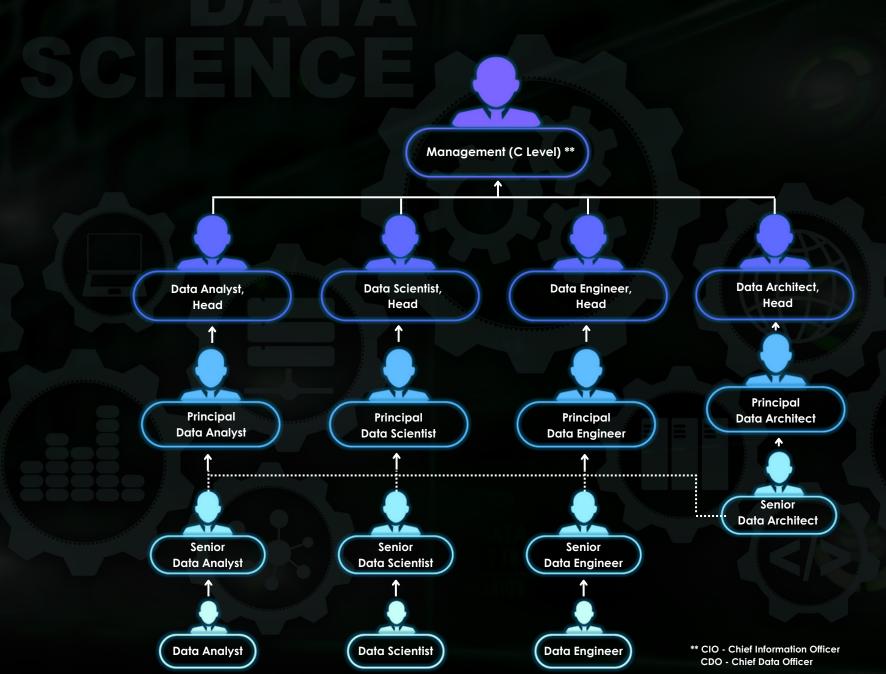


ACKNOWLEDGEMENT

We would like to thank all subject matter experts and validators who have contributed to the development and enhancement of the Industrial Skills Framework document Digital Technology - Data Science.



CAREER PATHWAY - DATA SCIENCE



Data Analyst



JOB DESCRIPTION

This section describes the key functions of a data analyst:

- Collect, clean and preprocess data from various sources, implement data processing workflows and ensure that data is suitable for analysis.
- Utilise mathematical and statistical methods to identify trends, patterns and insights.
- Create and present data insights in creative ways to facilitate the understanding of data across a range of technical and non-technical audiences.
- Monitor data trends and provide recommendations for process improvements and optimisations.
- Collaborate with different cross-functional teams to understand requirements and translate the requirements into data deliverables.

Disclaimer:

This section lists the expected SFIA levels of responsibility for each professional skill against the various seniority levels for a data analyst (refer to table below). Information on SFIA can be found in Appendix A:

Skills	Data Analyst	Senior Data Analyst	Principle Data Analyst	Data analyst Head
Data Science (DATS)	2,3	4,5	6	7
Methods and Tools (MELT)	3	4	5	6
Machine Learning (MLNG)	2,3	4	5	6
Data Visualisation (VISL)	3	4	5	5
Business Situation Analysis (BUSA)	3	4	5	6
Requirements Definition and Management (REQM)	2,3	4	5	6

COMMON TOOLS

This section lists the common tools used by a data analyst:

- No-code tools
- A/B testing tools
- Programming tools
- Data preparation tools
- Data visualisation tools
- Model libraries (packages)
- Machine learning and Al tools
- Machine Learning Operations (MLOps)

Note:

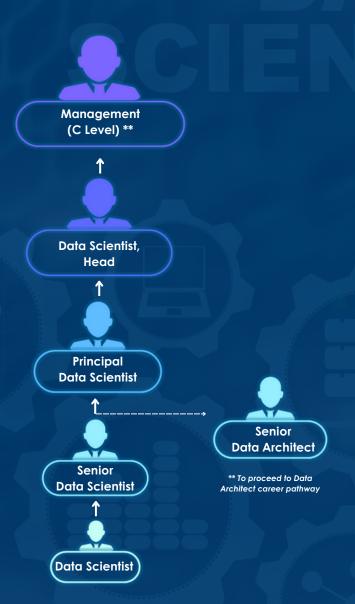
Refer Appendix B for examples of tools and Appendix C for examples of common certification.

COMMON QUALIFICATIONS

This section lists the qualifications recommended for a data analyst where at least one qualification is required:

- Professional certification or courses related to a data analyst.
- Diploma, degree or post-graduate course in a technical, quantitative, science or engineering field.
- Other equivalent degrees in related fields which cover knowledge and skills from data science.
- Relevant working experience in a related field.





JOB DESCRIPTION

This section describes the key functions of a data scientist:

- Collect, clean and preprocess data from various sources, implement data processing workflows and ensure that data is suitable for analysis.
- Utilise mathematical and statistical methods to identify trends, patterns and insights.
- Utilise advanced data science techniques to develop models using specialised programming languages and tools.
- To ensure developed models are regularly evaluated and updated to meet expected objectives.
- Create and present data insights in creative ways to facilitate the understanding of data across a range of technical and non-technical audiences.
- Monitor data trends and provide recommendations for process improvements and optimisations.
- Collaborate with different cross-functional teams to understand requirements and translate the requirements into data deliverables.

Disclaimer:

This section lists the expected SFIA levels of responsibility for each professional skill against the various seniority levels for a data scientist (refer to Table below). Information on SFIA can be found in Appendix A:

Skills	Data Scientist	Senior Data Scientist	Principle Data Scientist	Data Scientist Head
Data Science (DATS)	2,3	4,5	6	7
Methods and Tools (MELT)	3	4	5	6
Machine Learning (MLNG)	2,3	4	5	6
Data Visualisation (VISL)	3	4	5	5
Business Situation Analysis (BUSA)	3	4	5	6
Requirements Definition and Management (REQM)	2,3	4	5	6

COMMON TOOLS

This section lists the common tools used by a data scientist:

- No-code tools
- A/B testing tools
- Programming tools
- Version control tools
- Data preparation tools
- Data visualisation tools
- Deep learning frameworks
- Model libraries (packages)
- Machine learning and AI tools
- Machine Learning Operations (MLOps)

Note

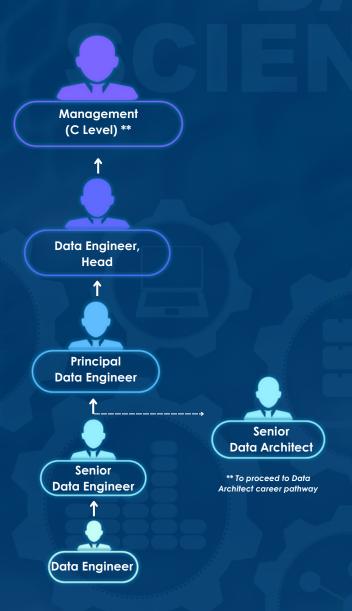
Refer Appendix B for examples of tools and Appendix C for examples of common certification.

COMMON QUALIFICATIONS

This section lists the qualifications recommended for a data scientist where at least one qualification is required:

- Professional certification or courses related to a data scientist.
- Diploma, degree or post-graduate course in a technical, quantitative, science or engineering field.
- Other equivalent degrees in related fields which cover knowledge and skills from data science.
- Relevant working experience in a related field.





JOB DESCRIPTION

This section describes the key functions of a data engineer:

- Design, implement, and evaluate on-premises cloudbased, and hybrid data engineering solutions.
- Assist organisations in understanding their data assets by structuring and storing data for organisational use.
- Ensure data is appropriately stored and archived, in line with relevant legislation.
- Ensure data quality in the organisation by utilising appropriate tools and techniques.
- Apply data engineering standards and tools to create and maintain data pipelines.
- Collaborate with different cross-functional teams to understand requirements and translate the requirements into data deliverables.

Disclaimer:

This section lists the expected SFIA levels of responsibility for each professional skill against the various seniority levels for a data engineer (refer to Table below). Information on SFIA can be found in Appendix A:

Skills	Data Engineer	Senior Data Engineer	Principle Data Engineer	Data Engineer Head
Data Management (DATM)	4	4,5	5,6	6
Methods and Tools (MELT)	3	4	5	6
Data Modeling and Design (DTAN)	2,3	3,4	4,5	5
Database Design (DBDS)	3	3,4	4,5	5
Data Engineering (DENG)	2,3	4	5	6
Requirements Definition and Management (REQM)	2,3	4	5	6
Quality Management (QUMG)	3	4,5	5,6	7
Quality Assurance (QUAS)	3	4,5	5,6	7

COMMON TOOLS

This section lists the common tools used by a data engineer:

- Programming tools
- Orchestration tools
- Cloud technologies
- Version control tools
- Data modelling tools
- Data integration tools
- Big data technologies
- Database technologies
- Data quality and data profiling tools
- Data governance and metadata management tools

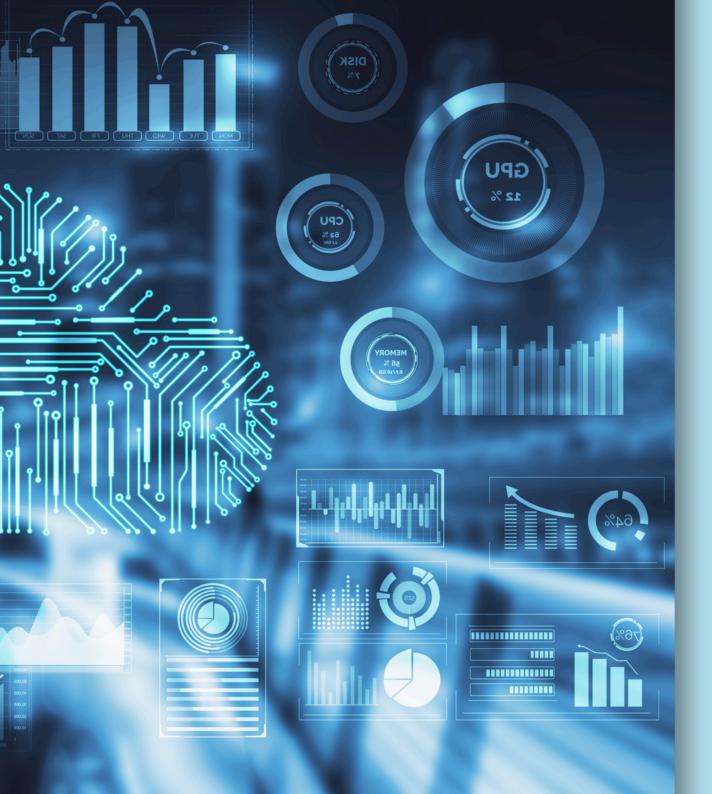
Note

Refer Appendix B for examples of tools and Appendix C for examples of common certification.

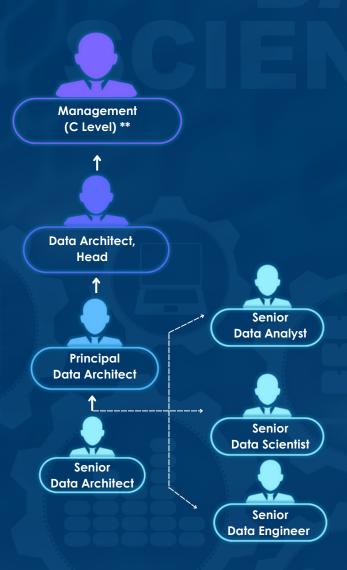
COMMON QUALIFICATIONS

This section lists the qualifications recommended for a data engineer where at least one qualification is required:

- Professional certification or courses related to a data engineer.
- Diploma, degree or post-graduate course in a technical, quantitative, science or engineering field.
- Other equivalent degrees in related fields which cover knowledge and skills from data science.
- Relevant working experience in a related field.



Data Architect



JOB DESCRIPTION

This section describes the key functions of a data architect:

- Design and implement systems compatible with onpremise, cloud, and / or hybrid computing architectures using data architecture standards.
- Ensure that existing and planned solution components are compatible with relevant architectures, strategies, policies, standards and processes.
- Contribute to create and review a systems capability strategy to meet the business's strategic requirements.
- Help organisations understand their data assets by designing and implementing systems.
- Ensure the organisation has appropriate systems to store and archive data in line with relevant legislation.
- Collaborate with different cross-functional teams to understand requirements and translate the requirements into data deliverables.

Disclaimer:

This section lists the expected SFIA levels of responsibility for each professional skill against the various seniority levels for a data architect (refer to Table below). Information on SFIA can be found in Appendix A:

Skills	Senior Data Engineer	Principle Data Engineer	Data Engineer Head
Enterprise and Business architect (SPTL)	5	6	7
Methods and tools (MELT)	4	5	6
Data modeling and design (DTAN)	3,4	4,5	5
Database design (DBDS)	3,4	4,5	5
Feasibility assessment (FEAS)	4	5	6
Requirements definition and management (REQM)	4	5	6
Solution architect (ARCH)	4	5	6
Systems design (DESN)	4	5	6

COMMON TOOLS

This section lists the common tools used by a data architect:

- Programming tools
- Orchestration tools
- Cloud technologies
- Version control tools
- Data modelling tools
- Data integration tools
- Big data technologies
- Database technologies
- Data quality and data profiling tools
- Data governance and metadata management tools

Note:

Refer Appendix B for examples of tools and Appendix C for examples of common certification.

COMMON QUALIFICATIONS

This section lists the qualifications recommended for a data architect where at least one qualification is required:

- Professional certification or courses related to a data architect.
- Diploma, degree or post-graduate course in a technical, quantitative, science or engineering field.
- Other equivalent degrees in related fields which cover knowledge and skills from data science.
- Relevant working experience in a related field.

SOFT SKILLS

This section lists the soft skills required by a data architect:

- Problem-solving
- Critical and analytical thinking
- Creativity
- Strong communication
- Time management
- Teamwork collaboration skills
- Proactive
- Agile and fast learner
- Golden circle*
- Project management
- Business acumen
- Strategic thinking
- *Encourages individuals and organisations to start with a clear understanding of their purpose and values (Why), then define their unique approach (How), and finally, communicate the tangible results (What).
 Intellectual curiosity.

Appendix A

Skills Framework for the Information Age (SFIA) is a widely used framework for describing skills and competencies within the IT and digital professions. SFIA defines seven levels of responsibility, ranging from the lowest level (Level 1) to the highest (Level 7) as can be seen in Table A. Further information on SFIA can be found at: https://sfia-online.org/

Level	Responsibilty
1	Follow
2	Assist
3	Apply
4	Enable
5	Ensure and Advise
6	Initiate and Influence
7	Set Strategy, Inspire, and Mobilise

Table 2: SFIA levels of responsibility

<u>Appendix B</u>

This section comprises a compilation of prevalent tools and illustrative instances essential for the field of data science:

COMMON TOOLS	EXAMPLES
No - Code Tools	Spreadsheet (Microsoft Excel, Google Sheets)
Programming Tools	Python, R, SAS
Data Visualisation Tools	Tableau, Power BI, ggplot2, Qlik D3.js
Product Analytics Tools	Google Analytics, Mixpanel, Amplitude
Model Libraries (Packages) Tools	NumPy, Pandas, Matplotlib, seaborn, scikit-learn, statsmodels, LightGBM, XGBoost, CatBoost, NetworkX, SciPy, Keras
Deep Learning Frameworks Tools	TensorFlow, PyTorch, Caffe, Theano, FastAl
MLOps / DS Tools	DataRobot, SAS Viya, AutoML, Comet ML
Data Preparation Tools	Tableau Prep, SAS EG
A/B Testing Tools	Optimizely, VMO, Google Optimize

COMMON TOOLS	EXAMPLES
Orchestration Tools	Kubernetes, Nomad, OpenShift, Cloudify, Docker Swarm, AWS Fargate, Rancher, Ansible, Marathon, Apache Airflow
Version Control Tools	Mercurial, Helix Core, GitLab, Apache Subversion, Bitbucket, Team Foundation Server, Bazzar, Monotone, IBM Rational ClearCase
Big Data Technologies Tools	Hadoop, MongoDB, Tableau, Spark, Cloudera, HPCC, Stats iQ, Oracle Big Data Appliance, Xplenty
Database Technologies Tools	Amazon RDS, IBM Db2, Oracle RDBMS, Redis, DbVisualizer, Informix, Airflow, Microsoft
Clouds Technologies Tools	AWS, Puppet, Microsoft AZURE, AtomSphere, Terraform, Apache, CloudStack, Cloudability, BigQuery, Informatica
Data Modeling Tools	Erwin Data Modeler, Power Designer, IBM InfoSphere, Data Architecture, SqIDBM
Data Intergration and ETL Tools	Informatica, Talend, Qlik, Azure, TIBCO, AWS, Oracle Data Integrator, Hevo Data
Data Warehousing Tools	Amazon Redshift, Pie, Cloudera, MarkLogic, Amazon DynamoDB, Snowflake

COMMON TOOLS	EXAMPLES
Data Visualisation Tools	Zoho Analytics, Pie, Microsoft Power BI, Looker
Data Governance and Metadata Management Tools	Collibra, Ataccama, Alation, OvalEdge, Octopai
Data Quality and Data Profiling Tools	Data Ladder, SAP, Data quality software, Datactics, Cloudingo, Demand Tools
Machine Learning and Al Tools	TensorFlow, PyTorch, scikit-learn, Keras, XG Boost, IBM Watson

Disclaimer:

Appendix C

This section outlines the essential certifications required for a career in data science:

- 1. Certified Analytics Professional (CAP)
- 2. Open Certified Data Scientist (Open CDS)
- 3. SAS Certified Big Data Professional
- 4. Microsoft Certified Azure Data Scientist Associate
- 5. AS Certified Advanced Analytics Professional
- 6. IBM Data Science Professional Certificate
 - Data analysis and visualisation
 - Data science methodology
 - Databases, Python and SQL
 - Machine learning
 - Open source tools
 - A capstone in applied data science
- 7. Google Data Machine Learning
- 8. SAS Certified Data Scientist
- 9. Dell EMC Data Science Track (Data Science Associate and Data Science Specialist)
- 10. Cloudera Certified Associate Data Analyst
- 11. Microsoft Certified: Power BI Data Analyst Associate
- 12. Oracle Business Intelligence
- 13. MongoDB Associate Developer
- 14. Google Certified Professional Data Engineer
- 15. IBM Certified Data Architect
- 16. TOGAF (The Open Group Architecture Framework) Certification
- 17. Certified Information Systems Architect (CISA)

Disclaimer: