



**WA DATA SCIENCE
INNOVATION HUB**

DATA SCIENCE COMPETENCY FRAMEWORK

Background

The WA Data Science Innovation Hub (WADSIH) is committed to assisting data driven organisations to accelerate the uptake of data science. In doing so, a number of large Western Australian based organisations have identified a need for a well-articulated and industry driven data science competency framework to assist in effectively recruiting talent and progressing talent internally.

Whilst WADSIH is aware that a number of data science competency frameworks exist, we have received feedback from industry that these frameworks still have obvious competency gaps. In developing this competency framework, WADSIH is dedicated to partnering with all industries to build a consistent, extensive and relevant data science competency framework which can be used by all industries across WA.

Data Science Knowledge & Skill Competencies

The [Data Science in Western Australia](#) report found that organisations wanted data scientists who were competent across four domains:

- Statistics and/or mathematics;
- Data analysis;
- Communication; and
- Domain (or subject matter) knowledge.

Competency in these domains, which combines both technical and behavioural (or soft) skills, will become increasingly important as the sector grows and is impacted by technology, specifically automation and augmentation. And it is these competencies which WADSIH aims to expand on in this data science competency framework.

Defining Data Scientists

Data science is not a single discipline, but rather a diverse ecosystem of roles, methodologies and tools sitting across many fields of study. Machine learning, statistics, predication and data analysis are all types of data science, although not all are performed solely by 'data scientists'.

The Massachusetts Institute of Technology (MIT) have posited that 'data science is not a discipline but rather an umbrella term used to describe a complex process involving not one data scientist possessing all the necessary expertise, but a team of data scientists with non-overlapping complementary skills' (Irizarry, 2020). Their argument redefines data science based on industry need, rather than academic classification. Specifically, they note that 'data science includes several areas of expertise and we should not expect one individual to encompass all of these'.

Data Science Areas of Expertise

Broadly, data scientists can be categorised as practicing in one of three different areas of expertise:

- **Data Engineering:** Deals with hardware, efficient computing and data storage infrastructure.
- **Data Analytics:** Explores, quality assesses, develops data models, performs statistical inferences and develops prototypes.
- **Machine Learning (ML) / Artificial Intelligence (AI) Science:** Builds and assesses prediction algorithms and makes the solutions scalable and robust for many users.

The competencies, aligned with these areas of expertise are shown in Diagram 1.

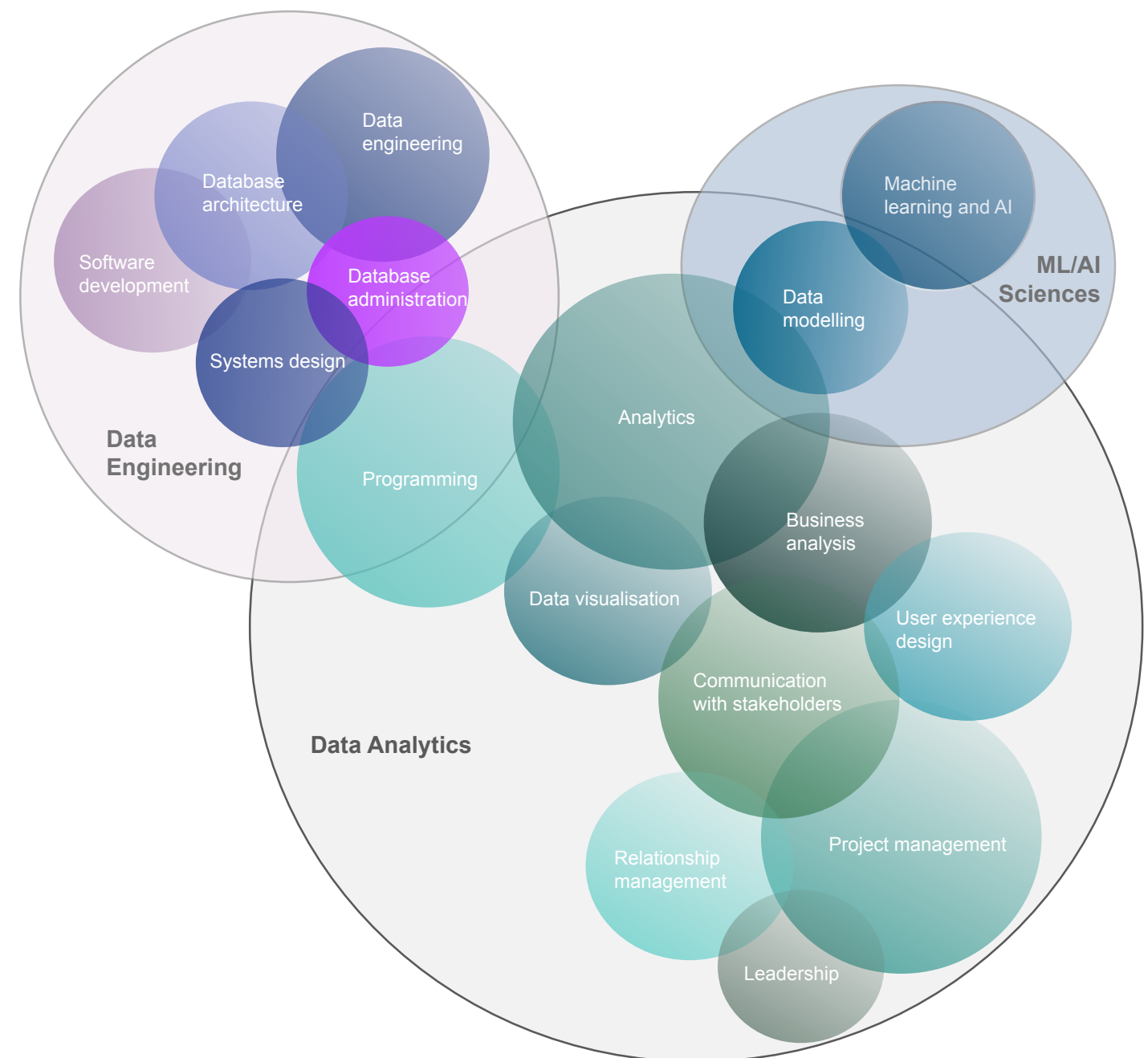


Diagram 1: Data Science areas of expertise

Data Science Knowledge & Skill Competencies

The **WADSIH Data Science Competency Framework (DSCF)** maps both knowledge and skill competencies.

The Harvard Data Science Review (*Fayyad & Hamatcu, 2020*) makes the following distinction between the two:

“Knowledge refers to information often acquired through formal education, books, or other media. Skills, on the other hand, refer to the ability to apply this knowledge, often gained through practice.”

Competency Frameworks

The DSCF uses three existing frameworks to capture competencies at different points along the Data Scientist Career Continuum.

Knowledge Competencies

Competencies at the Graduate & Data Scientist levels are mapped to the [Royal Statistical Society Standards](#), while those at the Senior & Principal Data Scientist levels are mapped to the [Science Council \(CSci\) Chartered Scientist Standards](#). See *Diagram 2 and 3*.

Skill Competencies

The [SFIA 7 Framework skills](#) and levels of responsibility are used to map skill competencies across the Data Scientist Career Continuum. See *Diagram 4*.



Diagram 2: Mapping knowledge competency frameworks

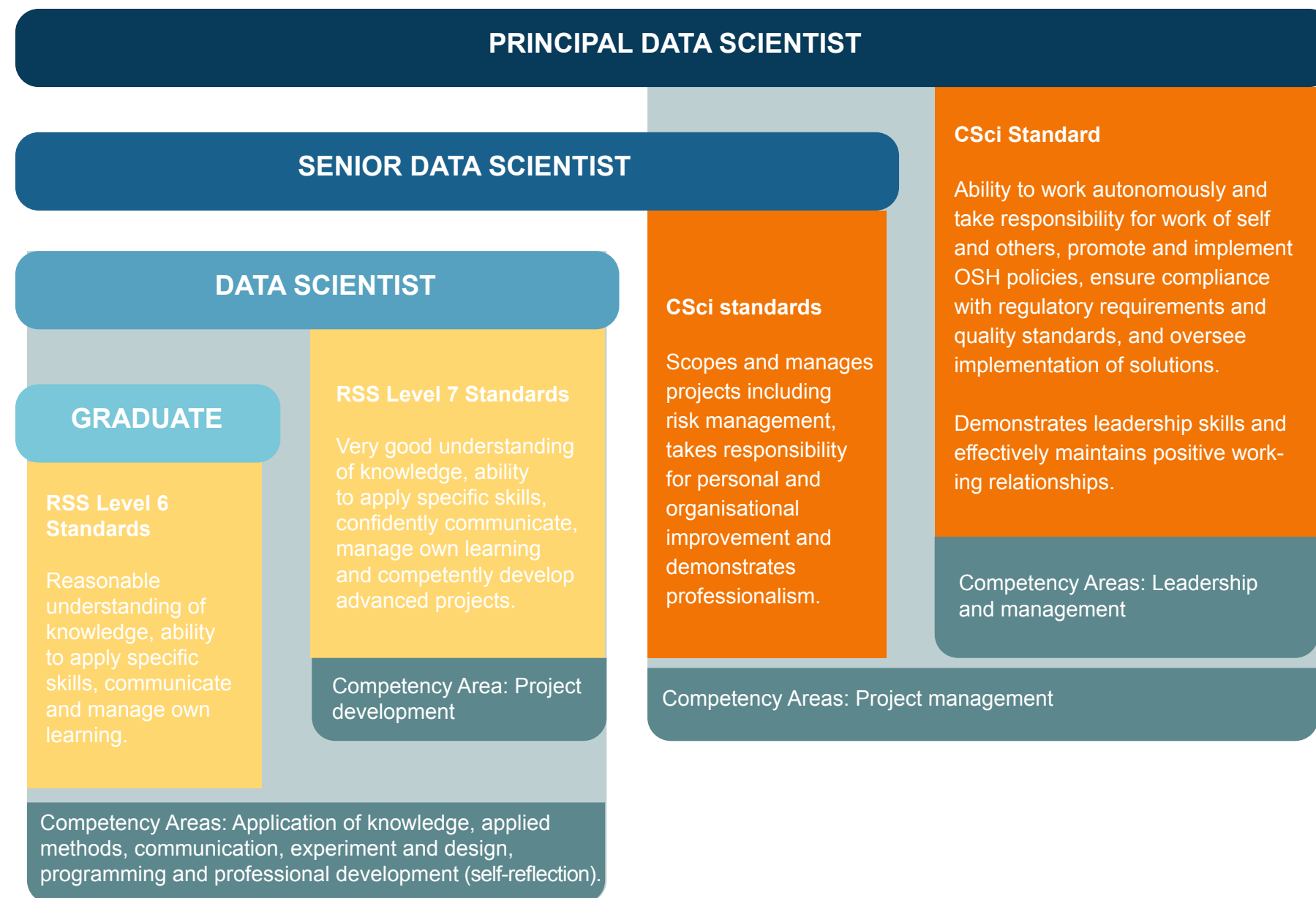


Diagram 3: Knowledge Competencies



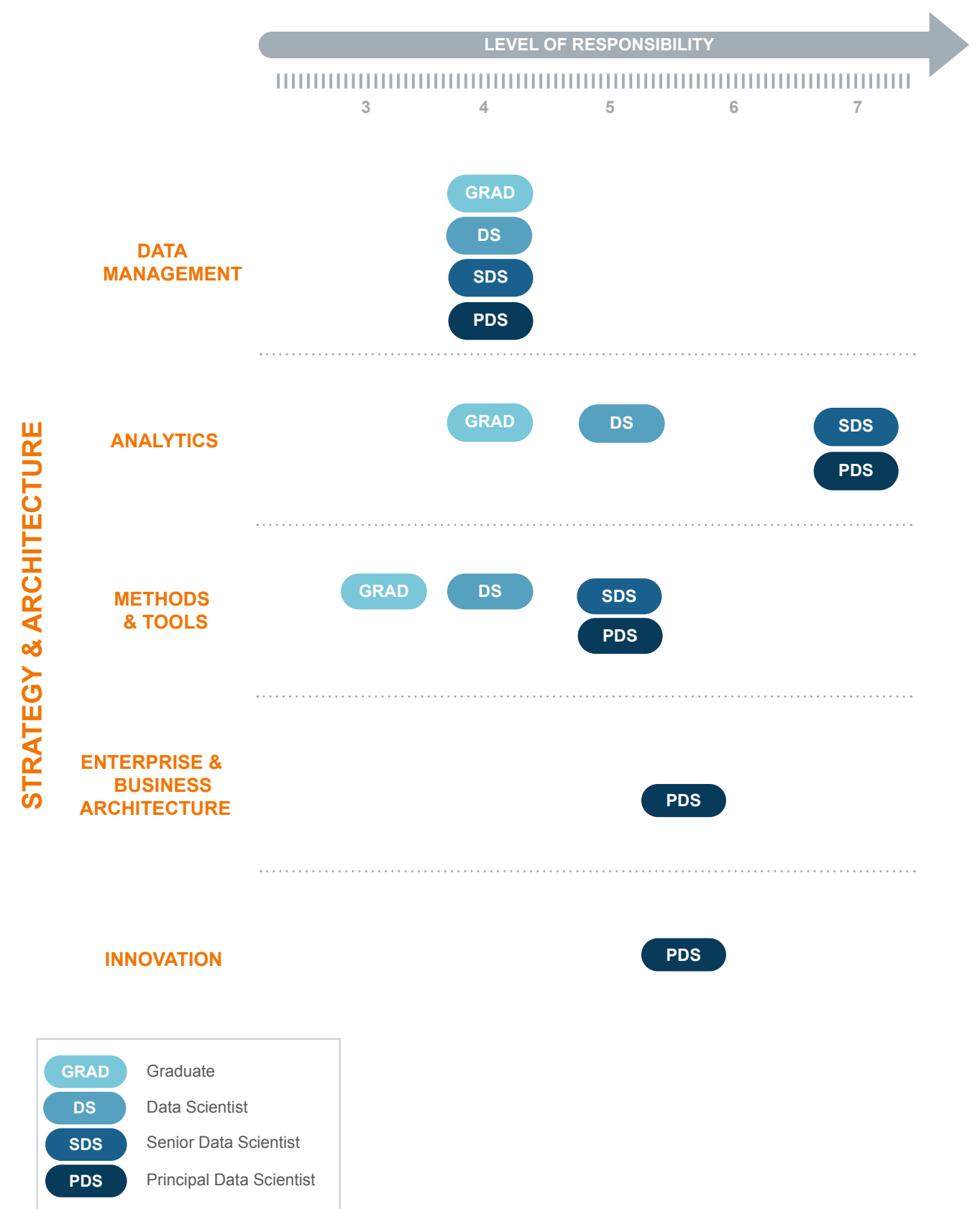
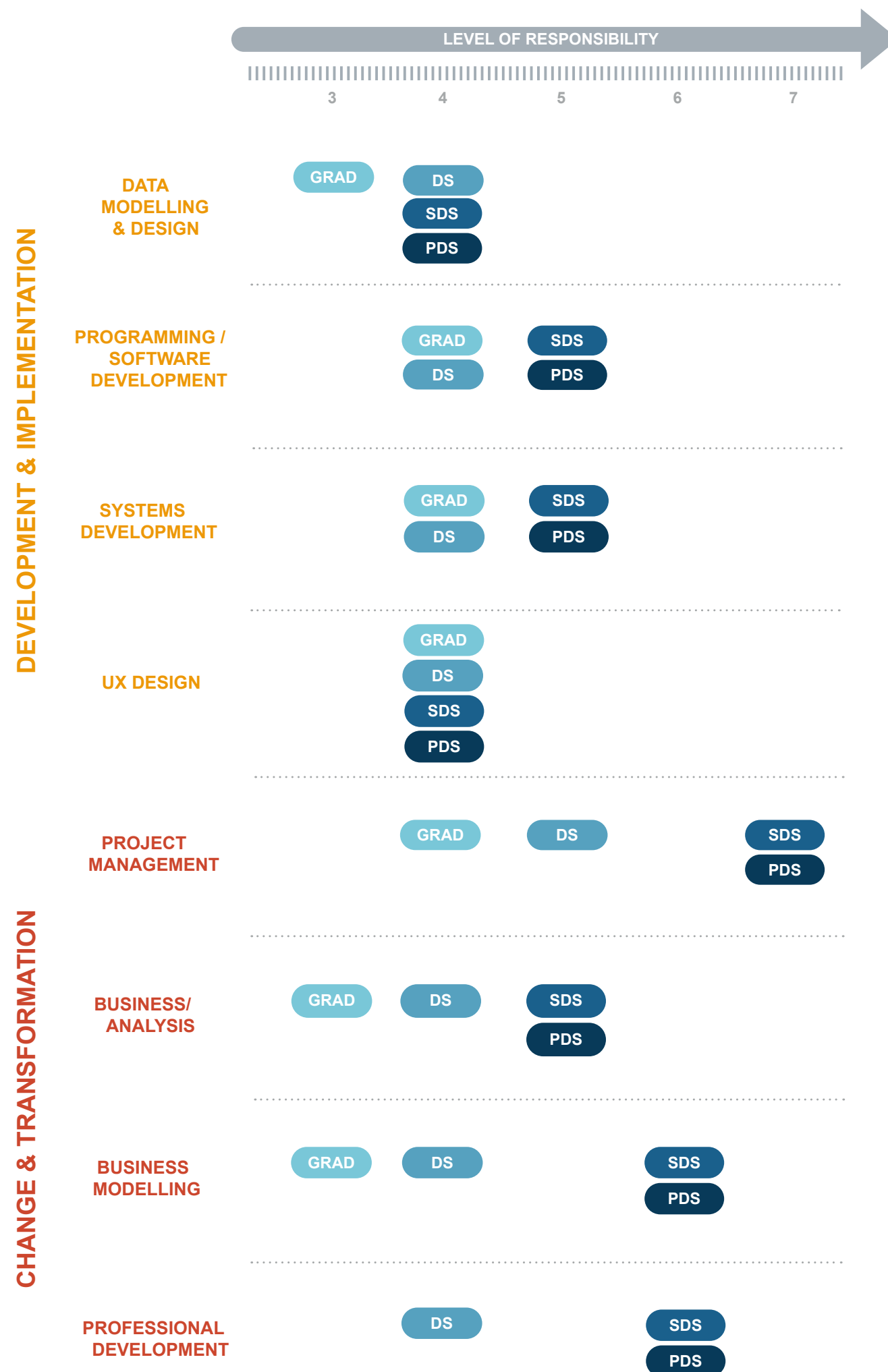


Diagram 4: Mapping SFIA 7 Framework Skills and Levels of Responsibility to levels on the Data Scientist Career Continuum

REFERENCES

Irizarry, R. A. (2020). The Role of Academia in Data Science Education . Harvard Data Science Review, 2(1). <https://doi.org/10.1162/99608f92.dd363929>

Fayyad, U., & Hamutcu, H. (2020). Toward Foundations for Data Science and Analytics: A Knowledge Framework for Professional Standards. Harvard Data Science Review, 2(2). <https://doi.org/10.1162/99608f92.1a99e67a>



**WA DATA SCIENCE
INNOVATION HUB**

Supported by



Curtin University