DS and RSE Competences in relation to the Research Cycle

The RSE competencies can be structured into four major pillars as done in [@Goth2024]. It is a research pillar, indicating that RSEs have a deep knowledge how science works, at best enabling them to participate in research. software skills, that enable them to write proper software for science, and team skills, to work well in interdisciplinary teams. Their specialization on software makes RSEs an instance of RTPs(research technical professional) that work to support science.

1. Conceptualization

RSE Competencies:

- Understanding the research cycle (`RC`)
- Conducting and leading research (`NEW`)

Data Science Competencies:

- Understanding the Data Science Lifecycle and methods selection
- Awareness of data context, purpose, and interdisciplinary implications (ethics, economics, legal)

2. Design

RSE Competencies:

- Software Design
- Software re-use strategies (`SRU`)
- Creating documented code building blocks (`DOCBB`)
- Software behavior analysis (`MOD`)
- Building distributable software (`DIST`)
- Tool and environment configuration (`SWLC`, `SWREPOS`)
- Full-Stack Programming

Data Science Competencies:

- Data integration & feature engineering (ETL, pipelines, quality checks)
- Design of data workflows and modeling processes
- Visualization theory and editorial thinking (exploratory analysis)
- Algorithm selection and objective function definition (ML core)
- Designing responsible data usage frameworks (ethics, privacy)
- Data-oriented Programming (R, Python, Matlab etc.)

3. Implementation

RSE Competencies:

- Source control, testing, CI/CD (`SWREPOS`, `DIST`, `DOCBB`)
- Working in interdisciplinary teams (`TEAM`)
- Project and task management (`PM`)

Data Science Competencies:

- Deployment of data pipelines and operational models
- Tool usage (Python, R, Julia, ML libraries like scikit-learn, Dask)
- Executing experiments using machine learning, deep learning
- Applying project management and interdisciplinary communication

4. Analysis & Interpretation

RSE Competencies:

- Software behavior interpretation (`MOD`)
- Documentation of research results and workflows

Data Science Competencies:

- Explorative Data Analysis (EDA), multivariate visualizations
- Deep understanding of model inference and optimization strategies
- Time series analysis, pattern mining, and argumentation (Data Mining)
- Ethical analysis of bias, fairness, and model accountability

5. Dissemination

RSE Competencies:

- Software publication and citation (`SP`)
- Use of domain repositories (`DOMREP`)
- Teaching and communication (`TEACH`, `USERS`)

Data Science Competencies:

- Reporting results and dashboards
- FAIR principles and reproducibility practices
- Preparation of software and models for open science platforms
- Communicating findings across disciplinary and public boundaries

6. Reflexion and Improvements

RSE Competencies:

- Continuous integration and testing (`SWLC`, `MOD`)
- Feedback-informed iterative development
- Mentoring, community involvement, and ethics (`TEAM`, `USERS`)

Data Science Competencies:

- Critical reflection on model performance and bias
- Model tuning, reengineering, and lifecycle updates
- Reinforcement learning and emerging AI models
- Responsible innovation, economic awareness, data sovereignty
- Application of Data Science in real-world domain projects