TALENT ENGINEERING ACADEMY

Data Quality, Taxonomies and Record Hygiene

PRACTICE WORKBOOK

Build reliable field definitions, resolve duplicates carefully and maintain useful skills and role vocabularies.

PATHWAY	Recruitment AI, Data and Systems
LEVEL	Practitioner
EDITION	2026.09
UPDATED	2026-09-24

Free learning. Practical work. Better recruitment.

An education initiative by Vitae.

WORKBOOK GUIDE

Build something useful as you learn.

This workbook turns the course into a practical piece of work you can review, refine and adapt. Print it, annotate it digitally, or use the prompts in your preferred notes tool.

YOUR COURSE DELIVERABLE

A data-quality rule set and taxonomy maintenance plan.

How to use this workbook

01 Work through the course lessons in order, then complete the matching workbook prompts.

02 Use fictional or properly authorised information. Do not add candidate or client confidential data.

03 Compare your decisions with the self-review guidance. Revise assumptions, gaps and unclear ownership.

04 Keep the finished workbook as evidence of your learning and as a starting point for real work.

BEFORE YOU START

Talent Engineering Fundamentals, or equivalent recruitment experience.

Course page: https://talentengineering.org/courses/data-quality-taxonomies

PART 1

Plan your course deliverable.

Use these planning pages to define what you are producing before you work through the lesson exercises. Return to them as your thinking develops.

WORKSHEET 01

Critical fields and quality dimensions

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

WORKSHEET 02

Vocabulary terms, synonyms and scope

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

WORKSHEET 03

Duplicate evidence and merge authority

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

WORKSHEET 04

Validation and exception queue

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

WORKSHEET 05

Freshness and correction rules

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

WORKSHEET 06

Quality measures and maintenance owner

Capture the evidence, decisions, assumptions, owners and next actions that make this section usable.

YOUR WORKING NOTES

PART 2

Apply each lesson.

Each exercise is designed to turn the lesson into evidence of practice. Complete the task, then use the guidance to challenge your first answer.

LESSON 01 / 04

Prioritise quality by decision impact

YOUR PRACTICE TASK

Rank ten fictional fields by consequence and define checks for the top four.

YOUR RESPONSE

FICTIONAL WORKED EXAMPLE

A vacancy has a complete salary field but the value is monthly while reports assume annual. The quality rule checks units and meaning, not just whether a cell is filled.

SELF-REVIEW GUIDANCE

Include semantic checks such as currency and time basis. Explain when each value can reasonably be known and how missing information is routed for follow-up.

LESSON 02 / 04

Build taxonomies that support retrieval

YOUR PRACTICE TASK

Create a ten-term vocabulary with synonyms and two examples of terms that must remain distinct.

YOUR RESPONSE

FICTIONAL WORKED EXAMPLE

PLC programming and automation engineering overlap but are not interchangeable. The vocabulary links related concepts while recording specific evidence and proficiency context separately.

SELF-REVIEW GUIDANCE

Define inclusion examples and ambiguous cases. Avoid treating a keyword match as verified competence. Show how an unfamiliar term can be reviewed without immediately creating a duplicate label.

LESSON 03 / 04

Resolve duplicates without losing meaning

YOUR PRACTICE TASK

Write a merge decision tree for three fictional pairs, including one uncertain identity.

YOUR RESPONSE

FICTIONAL WORKED EXAMPLE

Two records share a name but have different employment histories. The reviewer keeps them separate pending verification. Another confirmed duplicate retains the stricter contact preference and both source histories after merging.

SELF-REVIEW GUIDANCE

Include evidence needed, who can approve, field conflict rules and downstream checks. Uncertain identity should remain unresolved rather than being forced into a clean-looking database.

LESSON 04 / 04

Make hygiene a maintained workflow

YOUR PRACTICE TASK

Design a weekly hygiene routine and an import failure response.

YOUR RESPONSE

FICTIONAL WORKED EXAMPLE

An import repeatedly converts unknown availability into immediately available. The team stops the import, fixes the mapping and reviews affected records instead of asking recruiters to correct them indefinitely by hand.

SELF-REVIEW GUIDANCE

Include prevention, exception ownership and validation after repair. A useful routine focuses on consequential defects and does not reward staff for making uncertain fields look complete.

FINAL REVIEW

Strengthen the work before you use it.

Review the completed workbook as one connected piece of work. The strongest deliverables are internally consistent, explicit about limitations and clear about who decides what happens next.

- [ ] The work distinguishes facts, evidence, assumptions and open questions.
- [ ] Decisions have named owners, dates and review points.
- [ ] Measures connect to the outcome rather than activity alone.
- [ ] Risks, accessibility needs, privacy considerations and limitations are visible.
- [ ] The next action is specific enough for another person to follow.
- [ ] The finished deliverable can be explained in plain language.

FINAL CHANGES AND NEXT ACTIONS

FURTHER READING

Continue with trusted sources.

These sources support further study. A reference does not imply endorsement. Check the current official rules and obtain appropriate professional advice for real legal, financial or regulated decisions.

GOV.UK: Responsible AI in recruitment

UK guidance on procuring and deploying recruitment AI.

https://www.gov.uk/government/publications/responsible-ai-in-recruitment-guide/responsible-ai-in-recruitment

NIST: AI Risk Management Framework

Voluntary framework for organising AI risks and controls.

https://www.nist.gov/itl/ai-risk-management-framework

ABOUT THIS WORKBOOK

Original educational scenarios. Practical work is self-directed and is not independently assessed. Adapt any method to your context with appropriate review.

Talent Engineering Academy is an education initiative by Vitae.