File Format Guide

This guide outlines the required file formats and naming conventions for uploading SFW and Sector files to the application.

📋 Table of Contents

🎯 SFW File Format

File Naming Convention

- **Pattern:** `SFW_{SECTOR}.xlsx`
- **Examples:**
    - `SFW_HR.xlsx`
    - `SFW_FS.xlsx`
    - `SFW_IT.xlsx`

Important Rules:

- Must start with `SFW_`
- Sector code must be in **UPPERCASE**
- No spaces or special characters in sector code
- Supported file types: `.xlsx` only

Sheet Structure

- **Sheet Name:** Must match the filename pattern (e.g., `SFW_HR`)
- **Number of Sheets:** Exactly **one sheet** only
- Remove any extra sheets before uploading

Required Columns

Your SFW file must contain **exactly** these 10 columns with the specified data types:

Column Name	Data Type	Description	Example
`TSC_CCS_Type`	Text	Type classification	"Core Skill"
`TSC_CCS Code`	Text	Unique code identifier	"CC-001"
`Sector`	Text	Sector name	"Human Resources"
`TSC_CCS Category`	Text	Category classification	"Leadership"
`TSC_CCS Title`	Text	Skill title	"Team Management"
`TSC_CCS Description`	Text	Detailed description	"Ability to lead and manage teams effectively"
`Proficiency Level`	Whole Number	Numeric level (1-5)	3
`Proficiency Description`	Text	Level description	"Intermediate"
`Knowledge / Ability Classification`	Text	Classification type	"Knowledge"
`Knowledge / Ability Items`	Text	Specific items	"Team building, Conflict resolution"

Data Requirements

- **No missing values** in any column
- **Proficiency Level** must be whole numbers only (no decimals or empty cells)
- All text columns should contain meaningful data (not just spaces)
- File must contain at least one data row

🏢 Sector File Format

File Naming Convention

- **Pattern:** `{SECTOR}_{Full_Sector_Name}_sector_course_listing_curated.xlsx`
- **Examples:**
  - `HR_Human_Resources_sector_course_listing_curated.xlsx`
  - `FS_Financial_Services_sector_course_listing_curated.xlsx`
  - `IT_Information_Technology_sector_course_listing_curated.xlsx`

Important Rules:

- Sector code must be in **UPPERCASE**
- Full sector name can contain underscores for spaces

- Must end with `_sector_course_listing_curated.xlsx`
- Only `.xlsx` files are supported

Sheet Structure

- **Sheet Name:** Must be the sector code only (e.g., `HR`, `FS`, `IT`)
- **Number of Sheets:** Exactly **one sheet** only

Required Columns

Your Sector file must contain **exactly** these 5 columns:

Column Name	Data Type	Description	Can Be Empty?	Example
`Course Reference Number`	Text	Unique course ID	❌ No	"CRS-001"
`Skill Title`	Text	Associated skills	✅ Yes	"Leadership" or "['Python', 'SQL']"
`Course Title`	Text	Course name	❌ No	"Advanced Leadership Skills"
`About This Course`	Text	Course description	✅ Yes	"This course covers..."
`What You'll Learn`	Text	Learning outcomes	✅ Yes	"Students will learn..."

Special Note: Skill Title Format

The `Skill Title` column can contain two different formats:

1. **Single Skill (String):** `"Leadership"`
2. **Multiple Skills (JSON List):** `"['Python', 'SQL', 'Data Analysis']"`

Mixed Format Detection:

- If your file contains **both formats**, the system will automatically run preprocessing
- This helps standardize the skill data for better analysis
- Example of mixed format:

```
Row 1: "Leadership"
Row 2: "['Python', 'SQL']"
Row 3: "Communication"
Row 4: "['Excel', 'PowerPoint', 'Word']"
```

💾 Load Checkpoint Function

The Load Checkpoint feature allows you to resume your work from where you left off.

How to Use Checkpoints

Step 1: Save Your Progress

- Upload and validate your SFW and Sector files
- System automatically saves a checkpoint after successful validation

Step 2: Load Previous Work

- Click "Load Checkpoint" on the home page
- Choose from your saved checkpoints (shows date/time)
- Your files and progress will be restored instantly

What Gets Saved

- Your uploaded SFW and Sector files
- File validation status
- Any processing results

Important Notes

- Checkpoints are cleared when you close your browser
- Keep backup copies of your original files

⚠️ Common Issues & Solutions

File Naming Issues

Problem: Filename validation errors

```
❌ "sfw_hr.xlsx" → Sector code must be uppercase
❌ "SFW HR.xlsx" → No spaces allowed
❌ "HR_sector_courses.xlsx" → Missing required suffix
```

Solutions:

```
✅ "SFW_HR.xlsx"
✅ "HR_Human_Resources_sector_course_listing_curated.xlsx"
```

Column Issues

Problem: Missing or extra columns

❌ `Missing required columns: TSC_CCS_Type, Sector`
❌ `Unexpected columns: Extra_Column, Notes`

Solution: Use exactly the columns specified in the format guide

Data Type Issues

Problem: Incorrect data types

❌ `Proficiency Level contains text: "Level 3"`
❌ `Proficiency Level has empty cells`

Solution:

- Use whole numbers only: ③ not `"Level 3"`
- Fill all empty cells in required columns

Sheet Structure Issues

Problem: Multiple sheets or wrong sheet names

❌ `File has 3 sheets: SFW_HR, Sheet1, Sheet2`
❌ `Sheet named "Data" instead of "SFW_HR"`

Solution:

- Keep only one sheet
- Name the sheet exactly as specified

📞 Support

If you encounter issues not covered in this guide:

1. Check the error messages carefully - they often contain specific solutions

2. Verify your file matches the exact format requirements

3. Try downloading and re-uploading your file

4. Contact jerry_yang_from.tp@tech.gov.sg with specific error messages

---

Last Updated: May 2025
Version: 2.0