

Documentation - Data Patterns

**The proposal period for 2022 internships is now open
Submit your final proposal to Lorraine Chapman before Friday 18th March 2022**

Contribute to our suite of documentation guides by producing a manual on the Data Patterns capabilities available for use with the HPC Systems Platform. Find out more about the [HPC Systems Summer Internship Program](#).

Project Description

Data Patterns is a library or ECL Bundle that provides data profiling and research tools to an ECL programmer. It can be used in three ways:

- Inside ECL Watch (the easiest)
- Using the [Data Patterns Bundle](#)
- Using the Std.DataPatterns module in ECL

The candidate can use the free version of XMLMind Editor (or the XML Editor of their choice) and will submit pull requests to our GitHub Repository.

Completion of this project involves:

A new stand-alone Data Patterns book with the following sections:

- Overview
- Using Data Patterns in ECL Watch
- Using the DataPatterns bundle
- Using Std.DataPatterns module methods

We will provide some sample data files to process. These files will be made to best demonstrate some of the capabilities of the product. For example, one file might have an unusual skew or interesting MIN/MAX values.

An includible module (chapter) in DocBook XML format

This chapter will be included in two books: The Standard Library Reference and the new Data Patterns book

Poster - Showcasing deliverables and their value to the HPC Systems Open Source Project and Community

If you are interested in this project, please contact [Jim DeFabia](#).

By the mid term review we would expect you to have:

-

Mentor	Jim DeFabia Backup Mentor: Greg Panagiotatos
Skills needed	<ul style="list-style-type: none">•
Deliverables	Midterm <ul style="list-style-type: none">• End of project <ul style="list-style-type: none">• A new stand-alone Data Patterns book• An includible module (chapter) in DocBook XML format• Poster - Showcasing deliverables and their value
Other resources	<ul style="list-style-type: none">• HPC Systems website• HPC Systems Documentation Suite• HPC Systems GitHub Repository• Data Patterns GitHub Repository• Presentation about DataPatterns