

# Available Projects

**The proposal period for 2022 internships is now closed**  
**The proposal period for 2023 internships will open in November 2022**

Browse the list below. One you have selected a project, we recommend that you contact the mentor to discuss your ideas. Full details of how the application process works are available in this blog. Also, see our Proposal Guidelines and an example to help you prepare your proposal submission. If you are new to HPCC Systems, find out more about here and view our tips on how to prepare for your project and internship.

The projects listed here are available as student work experience opportunities with HPCC Systems as part of our summer intern program. Submissions are welcome from high school through to PhD. Find out about our Class of 2021 and their projects.

Contact [Lorraine Chapman](#) with your questions about the program.

1. [Additional Embedded Languages in ECL](#)  
Clojure, Haskell, Julia, MariaDB, MatLab, MongoDB (**Already taken**), ODBC, Postgres, SAS, Scala, SQL, or suggest one!
2. [Additional external data stores](#)  
Ceph or suggest one!
3. [Address Cleaner Plugin Optimizations](#)
4. [Build Process Improvements -Ninja, Jenkins X and Azure](#)
5. [Build System as Code](#)
6. [Cloud specific projects](#)  
HPCC Cluster Storage Support With Container Storage Interface (CSI)  
Baseline test suite for an HPCC Systems cluster on Kubernetes - **No longer available**  
Explore Kubernetes Federation Support  
HPCC Systems Kubernetes Cloud Namespace hierarchy  
HPCC Systems machine learning cluster on Kubernetes  
Enhance K8S Ingress with Istio Gateway and API Gateway  
Istio and other Service Mesh with HPCC Systems cluster  
Jenkins X **Already taken**  
Performance test suite for an HPCC Systems cluster on Kubernetes
7. [Develop an Automated ECL Watch Test Suite](#)
8. [Marketing/Documentation Projects](#)  
Technology/Marketing and Branding - **Already taken**  
Documentation - [Data Patterns](#) - **Already taken**  
These projects are suitable for students who would like to get involved with the HPCC Systems Open Source Project by contributing in a way that does not involve coding in a programming language.
9. [ECL Code Documentation Generator Improvements](#) - **No longer available**  
Make major improvements to the ECL Code Documentation Generator (ECLDoc), written in Python.
10. [Investigate Test Frameworks and Best Practises for HPCC Systems Cloud Native](#)
11. [Investigate Third Party Environments Working with HPCC Systems](#)
12. [Locking engine to replace DALI - Investigative project](#)  
Research, test and do a POC of a 3rd party inter-machine/process locking engine, for example ZooKeeper, HashiCorp's Consul or other suitable contenders.
13. [Machine Learning Algorithms on the HPCC Platform](#)  
[Anomaly Detection Algorithms](#) - **No longer available**  
[Applying the Causality Toolkit to Real World Datasets](#) - **Already taken**
14. [Natural Language Processing - All places for NLP projects under offer](#)  
[Enhance the English Dictionary](#)  
[Enhance the full English phrase parser](#)  
[Enhance the performance of the ECL NLP Plugin](#)  
[Create an NLP dictionary for Portuguese](#) - **Already taken**  
[Create an NLP dictionary for Chinese](#)  
[Create an NLP dictionary for Spanish](#)  
[Create an NLP dictionary for the Kurdish language](#) - **Already taken**  
[Build knowledge base for the human body](#) - **Already taken**  
[Build an OCR cleanup analyzer](#)  
[Create a phrase parser for Portuguese](#) (Follow-on project- Requires the completion of the Portuguese dictionary project)  
[Create a phrase parser for Chinese](#) (Follow-on project- Requires the completion of the Chinese dictionary project)  
[Create a phrase parser for Spanish](#) (Follow-on project- Requires the completion of the Spanish dictionary project)  
[Sentiment Analysis in Chinese](#)  
[Sentiment Analysis in English](#)  
[Sentiment Analysis in Portuguese](#)  
[Sentiment Analysis in Spanish](#)  
[Sentiment Analysis in the Kurdish Language](#)
15. [Performance Testing - Bare Metal vs Cloud Native](#)
16. [Provide SELinux Policies for the HPCC-Platform installation on Linux environments](#)  
Build SELinux domains for hpccsystems-platform services.
17. [Replace existing socket-based message passing interface with an open source package](#)  
Explore if using a different message layer (open-source package, such as ZeroMQ) offers improved performance, robustness and code maintainability

These projects are still under development and more details will be added soon. If you want to know more about any of these projects, view the associated JIRA issue and please contact [Lorraine Chapman](#) or the mentor of the project:

1. [Implement reference dafilesrv in other languages](#)
2. [Implement a Reverse activity](#)
3. [Incorporating self test code into a bundle](#)

4. Provide test code for bundles with no self test - **Already taken**