

# Chapter 1

## Related work

The Nickel Language Server follows a history of previous research and development in the domain of modern language tooling and editor integration. Most importantly, it is part of a growing set of LSP integrations. As such, it is important to get a picture of the field of current LSP projects. This chapter will survey a varied range of popular language servers, compare common capabilities, and implementation approaches. Additionally, this part aims to recognize alternative approaches to the LSP, in the form of legacy protocols, extensible development platforms LSP extensions and the emerging Language Server Index Format.

### 1.1 Language Servers

The LSP project was announced (**lsp-announced?**) in 2016 to establish a common protocol over which language tooling could communicate with editors. This way it is able to fully concentrate on source analysis instead of integration with specific editors by building custom GUI elements and being restricted to editors extension interface.

At the time of writing the LSP is available in version 3.16 (**Spec?**). Microsoft's official website lists 172 implementations of the LSP(**implementations?**) for an equally impressive number of languages.

Assessing any majority of these is outside the scope of this work. Yet, a few implementations stand out due to their sophisticated architecture, features, popularity or closeness to the presented work.

#### 1.1.1 Considerable dimensions

To be able to compare and describe each project objectively and comprehensively, the focus will be on the following dimensions.

**Target Language** The complexity of implementing language servers is influenced severely by the targeted language. Feature rich languages naturally require more sophisticated solutions than simpler ones. Conversely, self-hosted and interpreted languages can often leverage existing tooling.

**Features** The LSP defines an extensive array of capabilities. The implementation of these protocol features is optional and servers and clients are able to communicate a set of *mutually supported* capabilities.

The Langserver.org<sup>1</sup> project identified six basic capabilities that are most widely supported:

1. Code completion
2. Hover,
3. Jump to def,
4. Workspace symbols,
5. Find references
6. Diagnostics

Yet, not all of these are applicable in every case and some LSP implementations reach for a much more complete coverage of the protocol.

**File Processing** Most language servers use very different methods of handling the source code they analyze. The means are mainly influenced by the complexity of the language. Distinctions appear in the way servers process *file indexes and changes* and how they respond to *requests*.

The LSP supports sending updates in form of diffs of atomic changes and complete transmission of changed files. The former requires methods of incremental parsing and analysis which are difficult by itself, but are able to process files much more quickly. An incremental approach makes use of an internal representation of the source code that allows to quickly derive analytic results from and can be updated efficiently. Additionally, to facilitate the parsing, it must be able to provide a parser with the right context to correctly parse a changed fragment of code. On the contrary, most language servers process file changes by re-indexing the entire file, updating their internal model much more broadly. This is a more approachable method. Yet, it is less performant since entire files need to be processed, which becomes more noticeable as file sizes and edit frequency increase.

For code analysis LSP implementers have to decide between *lazy* or *greedy* approaches for processing files and answering requests. Dominantly greedy implementations resolve most available information during the indexing of the file. The server can then facilitate this model to answer requests using mere lookups. This stands in contrast to lazy approaches where only minimal local information is resolved during the indexing. Requests invoke an ad-hoc resolution the results of which may be memoized for future requests. Lazy resolution is more prevalent in conjunction with incremental indexing, since it further reduces the work associated with file changes. This is essential in complex languages that would otherwise perform a great amount of redundant work.

---

<sup>1</sup><https://langserver.org>

### 1.1.2 Representative LSP Projects

Since the number of implementations of the LSP is continuously growing, this thesis will present a selected set of notable projects.

1. Three highly advanced and complete implementations that are the de-facto standard tooling for their respective language: *rust-analyzer* (**rust-analyzer?**), *ocaml-lsp/merlin* (**merlin?**) and the *Haskell Language Server* (**hls?**)
2. Two projects that provide compelling alternatives for existing specialized solutions: *Metals* (for Scala) (**metals?**), *Java LSP* (**java-lsp?**)
3. Language Servers for especially small user languages in terms of complexity and userbase, highlighting one of the many use cases for the LSP. *rnix-lsp* (**rnix-lsp?**), *frege-lsp*

#### 1.1.2.1 LSP as standard tooling

Today LSP based solutions serve as the go-to-method to implement language analysis tools. Especially emerging languages without existing integration with major IDEs make extensive use of the flexibility and reach provided by the LSP.

Most languages profit greatly from the possibility to leverage existing tooling for the language. For instance the Haskell language server facilitates a plugin system that allows it to integrate with existing tooling projects (**hls-plugin-search?**). Plugins provide capabilities for linting (**hls-hlint-plugin?**), formatting (**hls-floskell-plugin?**, **hls-ormolu-plugin?**) documentation (**hls-haddock-comments-plugin?**) and other code actions (**hls-tactics-plugin?**) across multiple compiler versions. While this requires HSL to be compiled with the same compiler version in use by the IDE, it also avoids large scale reimplementations of compiler features in an entirely different language.

Similarly, the Ocaml language service builds on top of existing infrastructure by relying on the Merlin project introduced in sec. ???. Here, the advantages of employing existing language components have been explored even before the LSP.

The rust-analyzer (**rust-analyzer?**) takes an intermediate approach. It does not reuse or modify the existing compiler, but instead implements analysis functionality based on low level components. This way the developers of rust-analyzer have greater freedom to adapt for more advanced features. For instance rust-analyzer implements an analysis optimized parser with support for incremental processing. Due to the complexity of the language, LSP requests are processed mainly lazily, with support for caching to ensure performance. While many parts of the language have been reimplemented with a language server context in mind, the analyzer did not however implement detailed linting or the rust-specific borrow checker. For these kinds of analysis, rust-analyzer falls back to calls to the rust build system. ### Honorable mentions

## 1.2 Alternative approaches

### 1.2.1 Platform plugins

**1.2.2 Legacy protocols****1.2.3 LSP Extensions****1.2.4 LSIF**