

Clerk: Moldable Live Programming for Clojure

```
(ns nextjournal.clerk.px23
  {:nextjournal.clerk/toc true
   :nextjournal.clerk/visibility {:code :hide}}
  (require [nextjournal.clerk :as clerk]))

(nextjournal.clerk/html [:div.rounded-lg.p-4.text-xs.font-sans.bg-yellow-100.border
  " This is an early draft to be submitted to the " [:a {:href "https://2023.progra
```

Abstract

```
(clerk/html
  [:div.flex.flex-col.not-prose
    {:class "min-[860px]:flex-row"}
    [:div
      [:p.italic.leading.leading-snug
        "Clerk is an open source Clojure 'programmers assistant that builds upon the
      [:div.font-sans.flex-shrink-0.mt-6.text-sm
        {:class "min-[860px]:w-[165px] min-[860px]:ml-[40px] min-[860px]:text-xs min-
        [:a.hover:opacity-70 {:href "https://nextjournal.com"}]
        [:img.block.dark:hidden {:src "https://nextjournal.com/images/nextjournal-lo
        [:img.hidden.dark:block {:src "https://nextjournal.com/images/nextjournal-lo
        [:a.block.mt-2.hover:opacity-70 {:href "https://twitter.com/mkvlr"}] "Martin K
        [:a.block.mt-1.hover:opacity-70 {:href "https://twitter.com/unkai"}] "Philippa
        [:a.block.mt-1.hover:opacity-70 {:href "https://twitter.com/jackrusher"}] "Jac
```

Introduction: Literate Programming, Notebooks and Interactive Development

Knuth's *Literate Programming* emphasized the importance of focusing on human beings as consumers of computer programs. His original implementation involved authoring files that combine source code and documentation, which were then divided into two derived artifacts: source code for the computer and a typeset document in natural language to explain the program.

At the same time, other software was developed to target scientific use cases rather than program documentation. These systems, which prefigured modern computational notebooks, ranged from REPL-driven approaches like Macsyma and Mathematica to integrated WYSIWYG editors like Ron Avitzur's *Milo*, PARC's *Tioga* and *Camino Real*, and commercial software like *MathCAD*.

In contemporary data science and software engineering practice, we often see interfaces that combine these two approaches, like Jupyter and Observable. In these notebooks, a user can mix prose, code, and visualizations in a single document that provides the advantages of Knuth's Literate Programming with

those of a scientific computing environment. Unfortunately, most such systems require the programmer to use a browser-based editing environment (which alienates programmers with a strong investment in their own tooling) and custom file formats (which cause problems for integration with broader software engineering practices).

Although notebooks of this kind present an improvement on the programming experience of many languages, they often feel like a step backward to experienced Lisp programmers. In Lisp environments, it is common to be able to place the cursor after a single Lisp form and evaluate it in the context of a running program, providing finer granularity of control compared to the per-cell model of most notebooks. This workflow leads to a development style that these programmers are loath to lose.

That LISP users tend to prefer structured growth rather than step-wise refinement is not an effect of the programming system, since both methods are supported. I believe, however, that it is a natural consequence of the interactive development method, since programs in early stages of growth can be executed and programs in early stages of refinement cannot.

– Erik Sandewall

At the same time, though a number of Lisp environments have included graphical presentations of program objects, the default Clojure development experience relies on text-based representations of evaluation output and doesn't include the ability to embed widgets for direct manipulation of program state. Additionally, problems often arise when printing structurally large results, which can cause editor performance to degrade or lead to the truncation of output, and there's limited room for customization or support for requesting more data.

In comparison, interactive programming in Smalltalk-based systems has included GUI elements since the beginning, and work to further improve programmer experience along these lines has continued in Smalltalk-based systems like Self, Pharo, Glamorous Toolkit and Newspeak, which offer completely open and customizable integrated programming environments. Glamorous Toolkit, in particular, champions the idea of using easily constructed custom tools to improve productivity and reduce time spent on code archeology, which is also a big inspiration for what we'll present here.

Programming with Clerk

In such a future working relationship between human problem-solver and computer 'clerk', the capability of the computer for executing mathematical processes would be used whenever it was needed.

– Douglas Engelbart

Basic Interaction: Bring-Your-Own-Editor

Clerk combines Lisp-style interactive programming with the benefits of computational notebooks, literate programming, and moldable development, all without asking programmers to abandon their favorite tools or give up their existing software engineering practices. Its design stems partially from the difficult lessons we learned after years of unsuccessfully trying to get our *own team* to use an online browser-based notebook platform that we also developed.

When working with Clerk, a split-view is typically used with a code editor next to a browser showing Clerk’s representation of the same notebook, as seen in *Clerk side-by-side with Emacs*.

```
^{::: clerk / width : full}
(c clerk / html
  [:div#clerk-side-by-side-with-emacs.not-prose
   [:video {:loop true :controls true}
    [:source {:src "https://cdn.nextjournal.com/data/QmVYLx5SByNZi9hFnK2zx1K6Bz8F
    [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
    [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure:"] "Clerk side-by-side w
```

As shown here, our *notebooks* are just source files containing regular Clojure code. Block comments are treated as markdown text with added support for LaTeX, data visualization, and so on, while top-level forms are treated as code cells that show the result of their evaluation. This format allows us to use Clerk in the context of production code that resides in revision control. Because files decorated with these comment blocks are legal code without Clerk loaded, it they can be used in many contexts where traditional notebook-specific code cannot. This has led, among other things, to Clerk being used extensively to publish documentation for libraries that are then able to ship artifacts that have no dependency on Clerk itself.

Clerk’s audience is experienced Clojure developers who are familiar with interactive development. They are able to continue programming in their accustomed style, evaluating individual forms and inspecting intermediate results, but with the added ability to show! a namespace/file in Clerk. A visual representation of the file is then re-computed either:

- every time the file is saved, using an an optional file watcher; or alternatively,
- via an editor hot-key that can be bound to show the current document. (The authors generally prefer the hot-key over the file watcher, as it feels more direct and gives more control over when to show something in Clerk.)

Control and configuration of Clerk primarily occurs through evaluation of Clojure forms from within the programmer’s environment, rather than using outside control panels and settings. This integration with the programmer’s existing tooling eases adoption and allows advanced customization of the system through code.

Fast Feedback: Caching & Incremental Computation

To keep feedback loops short and avoid excess re-computation, Clerk uses dependency analysis to recompute only the minimum required subset of a file's forms. In addition, it optionally caches the results of long-running computations to disk to allow the user to continue work after a restart without recomputing potentially expensive operations. This caching behavior can be fine-tuned (or disabled) down to the level of individual forms.

TODO too low level/not focused on end-user experience?

Clerk begins by parsing and analyzing the code in a given file, then performs macro expansion and recursively traverses each form's dependencies, collecting them in a graph. For each top-level form, a hash is computed from the form and its dependencies. Next, Clerk evaluates each form unless it finds a cached value for that form. Because Clojure supports lazy evaluation of potentially infinite sequences, safeguards are in place to skip caching unreasonable values.

On-disk caches use a content-addressed store where each filename is derived from the hash of the file's contents using a base58-encoded multihash. Additionally, each file contains a pointer from the hash of the form to the result file, which allows us to indirect lookups to, for example, a remote storage service. This combination of immutability and indirection makes distributing sharing of the cache trivial.

While I did believe, and it has been true in practice, that the vast majority of an application could be functional, I also recognized that almost all programs would need some state. Even though the host interop would provide access to (plenty of) mutable state constructs, I didn't want state management to be the province of interop; after all, a point of Clojure was to encourage people to stop doing mutable, stateful OO. In particular I wanted a state solution that was much simpler than the inherently complex locks and mutexes approaches of the hosts for concurrency-safe state. And I wanted something that took advantage of the fact that Clojure programmers would be programming primarily with efficiently persistent immutable data.

It is idiomatic in Clojure to use boxed containers to manage mutable state. While there are several of these constructs in the language, in practice atoms are the most popular by far. An atom allows reading the current value inside it with `deref/@` and updating it's value with `swap!`.

When Clerk encounters an expression in which an atom's mutable value is being read using `deref`, it will try to compute a hash based on the value *inside* the atom at runtime, and extend the expression's static hash with it. This extension makes Clerk's caching work naturally with idiomatic use of mutable state, and frees programmers from needing to manually opt out of caching for those

expressions.

Semantic Differences from regular Clojure

Clojure uses a single-pass, whole-file compilation strategy in which each evaluated form is added to the state of the running system. One positive aspect of this approach is that manually evaluating a series of forms produces the same result as loading a file containing the same forms in the same order, which is a useful property when interactively building up a program.

A practical concern with this sort of "bottom-up" programming is that the state of the system can diverge from the state of the source file, as forms that have been deleted from the source file may still be present in the running system. This can lead to a situation where newly written code depends on values that will not exist the next time the program runs, leading to surprising errors. To help avoid this, Clerk defaults to showing an error unless it can resolve all referenced definitions in both the runtime and the source file.

It is our goal to match the semantics of Clojure as closely as possible but as a very dynamic language, there are limits to what Clerk's analysis can handle. Here are some of the things we currently do not support:

- Multiple definitions of the same var in a file
- Setting dynamic variables using `set!`
- Dynamically altering vars using `alter-var-root`
- Temporarily redefining vars using `with-redefs`

We have included a mechanism to override Clerk's error checking in cases where the user knows that one or more of these techniques are in use.

Presentation

Clerk uses a client/server architecture. The server runs in the JVM process that hosts the user's development environment. The client executes in a web browser running an embedded Clojure interpreter.

The process of conveying a value to the client is a *presentation*, a term taken from Common Lisp systems that support similar features. The process of presentation makes use of *viewers*, each of which is a map from well-known keys to quoted forms containing source code for Clojure functions that specify how the client should render data structures of a given type. When a viewer form is received on the client side, it is compiled into a function that will be then called on data later sent by the server.

When the `present` function is called on the server side, it defaults to performing a depth-first traversal of the data structure it receives, attaching appropriate viewers at each node of the tree. The resulting structure containing both data and viewers is then sent to the client.

To avoid overloading the browser or producing uselessly large output, Clerk’s built-in collection viewer carries an attribute to control the number of items initially displayed, allowing more data to be requested by the user on demand. Besides this simple limit, there’s a second global *budget* per result to limit the total number of items shown in deeply nested data structures. We’ve found this simple system to work fairly well in practice.

One benefit of using the browser for Clerk’s rendering layer is that it can produce static HTML pages for publication to the web. We could not resist the temptation to produce this document with Clerk, and have used that experience as an opportunity to improve the display of sidenotes.

It’s also possible to use Clerk’s presentation system in other contexts. We know of at least one case of a user leveraging Clerk’s presentation system to do in-process rendering without a browser.

Built-in Viewers

```
^{::: clerk / width : wide}
( clerk / html
  [: div .not-prose .overflow-hidden .rounded-lg
    [:img {:src "https://cdn.nextjournal.com/data/QmQLcS1D9ZLNQB8bz1TivBEL9AWttZdoP
```

Clerk comes with a set of built-in viewers for common situations. These include support for Clojure’s immutable data structures, HTML (including the hiccup variant that is often used in Clojure to represent HTML and SVG), data visualization, tables, LaTeX, source code, images, and grids, as well as a fallback viewer based on Clojure’s printer. The Book of Clerk gives a good overview of the available built-ins. Because Clerk’s client is running in the browser, we are able to benefit from the vast JS library ecosystem. For example we’re using Plotly and vega for plotting, CodeMirror for rendering code cells, and KaTeX for typesetting math.

Clerk’s built-in viewers try to suit themselves to typical Data Science use cases. By default, Clerk shows a code block’s result as-is with some added affordances like syntax coloring and expandability of large sub-structures that are collapsed by default.

```
^{::: clerk / width : wide}
( clerk / html
  [:div#showing-result-data.not-prose.overflow-hidden.rounded-lg
    [:img {:src "https://cdn.nextjournal.com/data/QmYJUox1pu3Yh3hiYgCn61TpKLmRuFPR
    [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
      [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure:"] "Showing result data "
```

Additional affordances are modes to auto-expand nested structures based on shape heuristics and expanding multiple sub-structures of the same level:

```

^{:clerk/width :wide}
(clerk/html
  [:div#expanding-multiple-sub-structures-at-once.not-prose.overflow-hidden.round
    [:video {:loop true :controls true}
      [:source {:src "https://cdn.nextjournal.com/data/QmWNN15jP8dujxKR71FCacTdASmca
      [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
        [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure: "] "Expanding multiple

```

Using the built-in clerk/table viewer, the same data structure can also be rendered as table. The table viewer is using heuristics to infer the makeup of the table, such as column headers, from the structure of the data:

```

^{:clerk/width :wide}
(clerk/html
  [:div#clerk-tables.not-prose.overflow-hidden.rounded-lg
    [:img {:src "https://cdn.nextjournal.com/data/QmXYYTUrqj5GuhTkKv7KnKUUX9bwPF
    [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
      [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure:"] "Clerk tables"]]])

```

Together with tables, plots make up for the most common Data Science use cases. Clerk comes with built-in support for the popular vega and Plotly plotting grammars. In the following figure, the same data, as shown in the above table example, is used to render a vega-lite plot using the built-in clerk/vl viewer:

```

^{:clerk/width :wide}
(clerk/html
  [:div#plotting-using-the-vega-grammar.not-prose.overflow-hidden.rounded-lg
    [:img {:src "https://cdn.nextjournal.com/data/QmYkCRqGcxSGH4EbjSsXrJ5fkHLJkiMD
    [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
      [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure:"] "Plotting using the v

```

While these viewers may cover the bulk of Data Science use cases, additional built-in viewers like clerk/image, clerk/katex, and more can be added to enrich the output of a Clerk notebook. It is important to note that Clerk's viewers work in a way that encourages composition. Multiple viewers can be combined to suit a specific use case such as the following example showing a table that combines bird species' names, with exemplary images and geo-spatial plots:

```

^{:clerk/width :wide}
(clerk/html
  [:div#combining-multiple-built-in-viewers.not-prose.overflow-hidden.rounded-lg
    [:img {:src "https://cdn.nextjournal.com/data/QmbkLu55j4wv7QhkvNUchifBvVxzYxHZV
    [:div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
      [:div.mx-auto.max-w-prose.px-8 [:strong.mr-1 "Figure:"] "Combining multiple ,

```

Moldable Viewer API

Clerk’s viewers are an ordered (and thus prioritized) collection of plain Clojure hash maps. Clerk interprets the following optional keys in each viewer map:

- `:pred` is a predicate function that tests whether this viewer should be used for a given data structure
- `:transform-fn` is an optional function run on the server side to transform data before sending it to the client. It receives a map argument with the original value under a key. Additional keys carry the path, the viewer stack, and the budget (for elision)
- `:render-fn` is a quoted form that will be sent to the browser, where it will be compiled into a function that will be called to display the data
- `:page-size` is a number that indicates how many items to send in each chunk during elision/pagination

Viewers can also be explicitly selected by wrapping a value in the `clerk/with-viewer` function, which produces a presentation for that value using that viewer. Alternatively, viewers can be selected by placing a Clojure metadata declaration before a form. Because of the way Clojure handles compilation, metadata in this position is ultimately ignored in the generated code. So far as we know, this is a novel mechanism for out of band signaling to a specialized Clojure parser. TODO add example viewer source and metadata declaration here

The process of selecting viewers happens programmatically on the server side, thus using the programmer’s already existing interactive programming environment as a user interface.

Sync

To help with creating interactive tools using Clerk, it also supports bidirectional sync of state between the client and server Clojure environments. If a Clojure atom on the server is annotated with metadata indicating it is sync, Clerk will create a corresponding var in the client environment. Both of these atoms will be automatically instrumented with an update watcher that broadcasts a *diff* to the other side.

In addition, a server-side change will trigger a refresh of the currently active document, which will then re-calculate the minimum subset of the document that is dependent on that atom’s value. This allows us to use Clerk for small local-first apps, as shown in the Regex Dictionary Example.

Prose-oriented Documents

The first and primary use case for Clerk was adding prose, visualizations, and interactivity to Clojure namespaces. However, when writing documents that are mainly prose, but would benefit from *some* computational elements, it is rather tedious to write everything in comment blocks. To make this easier, Clerk

can also operate on markdown files with “code-fenced” source code blocks. All Clojure source blocks in such a file are evaluated and replaced in the generated document with their result.

This format is very similar to other markdown-based notebooks, like R Mark-down, but specifically tailored to Clojure. We used this approach to write this paper, the source for which is located on Github.

Examples of Moldable Development with Clerk

In addition to the sorts of traditional data science use cases that one might expect from something that has “notebook” features, we intend Clerk to be a general purpose programmer’s assistant that allows the rapid construction of tiny interfaces during daily work. Here are a few samples of tools and documentation created in this manner.

Augmenting table names

This example illustrates an approach we used to make working with a legacy DB2 database easier. The database’s column names are made up of largely human-unreadable 8 character sequences:

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#as400-column-names.not-prose.overflow-hidden.rounded-lg
   [ :img {:src "https://cdn.nextjournal.com/data/QmWnzjc5c9qpUUaLoK3ytZk4Zs1AzDp2
     [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
       [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:"] "AS/400 Column Names "
```

We were able to automatically translate these names using a metaschema extracted from the database. This allowed us to create a viewer that maps those 8-character names to human-readable (German-only) names (which we can then translate to English names). In typical Lisp fashion, we go on and inspect a query interactively. We can use the translated names in the table even print them but one quickly sees the limit of plain-text printing:

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#inspecting-a-query-using-the-repl.not-prose.overflow-hidden.rounded-lg
   [ :video {:loop true :controls true}
     [ :source {:src "https://cdn.nextjournal.com/data/QmbGFKpEXLGyqngHe7q1dqAsEAW
       [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
         [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:"] "Inspecting A Query U
```

With Clerk, we were able to render the output as a graphical table without the limitations of plain text. Further, we can use the Viewer API to extend the table viewer’s headings to show the translated metaschema names (plus showing the original 8 character names in a de-emphasized way so that they aren’t lost). We

can go further still, showing the original German names when move the mouse over the headings:

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#augmented-table-headings .not-prose .overflow-hidden .rounded-lg
   [ :video {:loop true :controls true}
     [ :source {:src "https://cdn.nextjournal.com/data/QmVZsXsX2wcYYc758yHkZjijW2H"}
       [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
         [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:"] "Augmented Table Headings"] ] ] ] ] ]
```

Rich documentation features

This example illustrates the use of Clerk to create rich documentation for clojure2d's colors package. They used Clerk's Viewer API to implement custom viewers to visualize colors, gradients and color spaces, then publish that documentation on the web by generating a static website directly from the source code of the library.

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#custom-viewers-for-clojure2ds-colors-library .not-prose .overflow-hidden .rounded-lg
   [ :img {:src "https://cdn.nextjournal.com/data/QmQgTLi8qfzrBRTkaAGfWQ4RceM4v3f"}
     [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
       [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:"] "Custom Viewers for Clojure2D's Colors Library"] ] ] ] ]
```

Regex Dictionary

Built as a showcase for Clerk's sync feature, this example allows entering a regex into a text input and get dictionary matches as result while you type:

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#interactive-regex-dictionary .not-prose .overflow-hidden .rounded-lg
   [ :video {:loop true :controls true}
     [ :source {:src "https://cdn.nextjournal.com/data/QmTwZWw4FQT6snxT8RkKt5P7Vx"}
       [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
         [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:"] "Interactive Regex Dictionary"] ] ] ] ] ]
```

It is built using a Clojure atom containing the text input's current value that is synced between the client and server. As you type into the input, the atom's content will be updated and synced. Consequently, printing the atom's content in your editor will show the input's current value:

```
^{::: clerk/width :wide}
(c clerk/html
 [ :div#printing-the-value-of-a-synced-clojure-atom .not-prose .overflow-hidden .rounded-lg ] ] ]
```

```
[ :img { :src "https://cdn.nextjournal.com/data/QmNS2jigrDn2WdS7AVa4qMiWtwZovJm
[ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
[ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:" ] "Printing the value o
```

Lurk: Interactive Lucene-powered Log Search

Also building on Clerk's sync feature, this interactive log search uses Lucene on the JVM side to index and search a large number of log entries. In addition to using query input, logs can also be filtered by timeframe via an interactive chart. It is worth noting that this example has a completely custom user interface styling (nothing left of Clerk's default styling) via Clerk's CSS customization options.

```
^{::: clerk/width :wide}
( clerk/html
  [ :div#interactive-log-search.not-prose.overflow-hidden.rounded-lg
    [ :video { :loop true :controls true }
      [ :source { :src "https://cdn.nextjournal.com/data/QmRtGb5aByKD6i5SsxfS1JCJPKpC1
    [ :div.bg-slate-100.dark:bg-slate-800.dark:text-white.text-xs.font-sans.py-4
      [ :div.mx-auto.max-w-prose.px-8 [ :strong.mr-1 "Figure:" ] "Interactive Log Sear
```

Experience

Our experience as the developers and users of Clerk has been surprisingly positive, but we're heavily biased. We've chosen a few quotes from Clerk's user base to give a sense of how it has been received in the community:

[Clerk] is making the training of junior Clojure programmers a massive pleasure! [...]

It helps us to bypass what would otherwise be a lot of distracting UI programming. Set up your env, make a namespace, hit a keybind, hey presto, your code is running in a browser.

– Robert Stuttford

I'm using Clerk to visualize statistics properties from a simulation in a model checker [...] it's basically a wrapper over TLA+ [...]

Amazing that Clerk just lets you focus on what really matters and nothing else!

– Paulo Feodrippe

I just wanted to express some gratitude for Clerk. It's been a game changer for me in terms of understanding problems and communicating that understanding to other people.

– Jeffrey Simon

WORK IN PROGRESS

Open Toolkit

TODO discuss open library toolkit approach

Related & Future Work

Related:

- org-mode
- Mathematica

Future Work

- Viewers: lets :pred function opt into more context
- Open up caching
- Use distributed cache more
- Make caching more granular, also allow caching functions?
- Clerk printer to fix REPL printing problem

Conclusion
