



C++ Toolkit

Denis Vakatov¹

Created: February 15, 2013; Updated: January 13, 2019.

2019 NOTICE

This chapter is no longer getting updated. For current, up-to-date information about C++ Toolkit see <https://ncbi.github.io/cxx-toolkit/>

Summary

The C++ Toolkit is a large body of C++ software that was built to support the medical literature and bioinformatics services that the National Center for Biotechnology Information (NCBI) makes available to the public. While the primary users of the Toolkit are within NCBI, the software is portable (Unix, Windows, Mac) and freely available with no restrictions on use.

Libraries and Applications

If you are a C++ developer you will find the portable nature of the libraries very useful in building cross-platform applications even if you do not have much interest in bioinformatics. Libraries such as those for CGI/Fast-CGI, networking, XmlWrapp, SQL database access, and serialization are quite general-purpose and can be used in a variety of applications outside the bioinformatics problem domain.

The Toolkit provides many general-purpose libraries, including:

- **CORELIB** - Provides a portable way to write C++ code and many useful facilities such as an application framework, argument processing, template utilities, threads, date/time, files, and strings.
- **CONNECT** - Networking and inter-process communication with IOSTREAM adaptors.
- **CGI** - CGI and Fast-CGI.
- **DBAPI** - SQL database access.
- **SERIAL** - Serialization using ASN.1, JSON, or XML.
- **GUI** - Portable wxWidgets and OpenGL based GUI and graphic libraries.
- **XmlWrapp** - XML parsing and handling, XSLT, XPath—this is an NCBI fork that adds some useful enhancements to the open-source `xmlwrapp` project.
- **JSONWRAPP** – same for JSON format.
- **UTIL** - Many generic facilities including compression, a diff API, floating point comparison, random number generation, thread pools, and UTF-8 conversion.

Libraries specific to bioinformatics are also provided, including:

- [ALGORITHM](#) - Sequence alignment algorithms.
- [BLAST](#) - An alignment engine.
- [OBJECT MANAGER](#) - Biological sequences (e.g., GenBank) retrieval and processing.

Applications are also provided, for example:

- [DATATOOL](#) - A data converter and C++ code generator for data storage classes.

The C++ Toolkit libraries and applications are in active development and are regularly built and tested on Unix, Windows, and Mac.

Typical Uses

Every day, thousands of people around the world use applications built on top of the C++ Toolkit. These include:

- [PubMed](#) - PubMed comprises more than 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher websites.
- [PubMed Central](#) - PMC is a free full-text archive of biomedical and life sciences journal literature at the National Institutes of Health's National Library of Medicine (NIH/NLM).
- [BLAST](#) - The Basic Local Alignment Search Tool is a tool for analyzing biological sequence information and is one of the most widely used bioinformatics programs.
- [Genome Workbench](#) - NCBI Genome Workbench is an integrated application for viewing and analyzing sequence data. With Genome Workbench, you can view data in publically available sequence databases at NCBI, and mix this data with your own private data. The Genome Workbench has an online counterpart, the [Sequence Viewer](#).

Installation

Before using the C++ Toolkit, you should ensure that your platform is supported by checking the [current release notes](#).

The first step in using the Toolkit is getting the source code—either by [downloading via FTP](#) or by [checking it out from Subversion](#).

Once you've gotten the source code, you can [configure and build the Toolkit](#) for your platform.

Because the C++ Toolkit supports a variety of platforms and compilers, the process of building the libraries involves determining the platform- and compiler-specific features as well as third-party packages. This is facilitated by running the platform-independent "configure" script on Unix and Mac, or building the "CONFIGURE" solution in C++ Visual Studio on Windows. Additional details on configuring and building can be found [online](#).

Successful builds result in immediately usable libraries and applications. Thus, downloading, configuring, and building comprise the installation process, and generally there is no need for a separate "install" step.

Public Releases

Typically, new [public versions](#) of the C++ Toolkit are released every year or two. The Release Notes are also updated when a new public version of the C++ Toolkit becomes available.

Documentation

An extensive C++ Toolkit book is available [online](#). There are also online browsers:

- [Doxygen](#)
- [LXR](#)
- [Subversion](#)
- [ASN.1](#)

Support

The software is provided on an as-is basis; however, the following mailing lists can be used:

- to receive [announcements](#) (read-only)
- for [general information](#)
- to contact [core developers](#)
- for help with the [Object Manager](#)
- to access [CVS logs](#) (read-only)
- to access [SVN logs](#) (read-only)

Related Resources

Many resources are available to C++ Toolkit users, and the best way to find them is either by using the sidebar on any page of the online [Toolkit Book](#), or from any of the following dedicated search pages:

- [Toolkit Book Search](#)
- [Source Code Search](#)
- [Combined Book and Source Code Search](#)

These pages all include a search tool and links to several source browsers, Subversion access to the source code, library symbol search tools, ASN.1 specifications, and more.

More generally, the [NCBI home page](#) has a wealth of information about NCBI resources, including the C++ Toolkit and many other tools.

For more information about the Toolkit, please see the online [NCBI C++ Toolkit Book](#). The [Introduction](#) provides a broad overview of the capabilities in the C++ Toolkit with links to other chapters that cover topics in more detail. The [Getting Started](#) chapter provides a description of how to obtain the C++ Toolkit, the layout of the source distribution tree, and how to get started. The online Toolkit book is intended to cover pretty much everything you need to know about the Toolkit. If you can't find answers to your questions there, please try one of the email lists above or email cpp-doc@ncbi.nlm.nih.gov.