BLASTPHYME TECHNICAL SPECIFICATIONS

Current as of application version 1.5.0.12
Last updated January 24, 2019

CONTENTS

1 INTRODUCTION

1.1 SCOPE

This document includes but is not limited to the programming languages, software platforms, standards, and technologies that the application employs, and specifications for its interactions with third-party systems (i.e.: NCBI GenBank and BLAST).

Functional requirements detailing what functionality the application is expected to provide are beyond the scope of this document. The "BlastPhyMe Functional Requirements" document should be referred to for those requirements.

2 CORE PLATFORMS

The BlastPhyMe application is coded in the C# programming language on the Microsoft .NET Framework platform, version 4.0.

http://msdn.microsoft.com/en-us/library/zw4w595w(v=vs.100).aspx

The database engine that BlastPhyMe interacts with for storing data is the Microsoft SQL Server 2014 Express LocalDB engine.

http://msdn.microsoft.com/en-ca/library/hh510202(v=sql.120).aspx

For communicating with the NCBI BLASTN web service, BlastPhyMe uses the .NET Bio open source library:

- Website: https://github.com/dotnetbio/bio
- License: https://github.com/dotnetbio/bio/blob/master/License

3 SYSTEM ARCHITECTURE

BlastPhyMe has been designed to the standard of an *n*-Tier application. The application code is separated between three distinct layers:
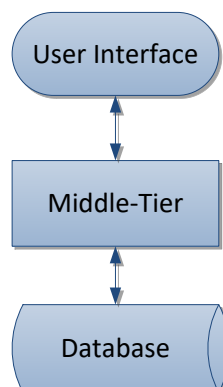
3.1 USER INTERFACE

The User Interface layer includes all code necessary to display visual interfaces for the user to interact with. User Interface code does not interact directly with the Database layer or third-party systems (i.e.: NCBI) and is abstracted from both by objects within the Middle-Tier layer.

3.2 MIDDLE-TIER

The Middle-Tier layer includes all code necessary to transfer data between the User Interface and Database layer. Code within the Middle-Tier interprets the database architecture into objects that are exposed for the User Interface to interact with. The Middle-Tier is responsible for all direct communication with the database, and all interaction with third-party systems (i.e.: NCBI) and products (i.e.: .NET Bio) is contained within the Middle-Tier layer.

3.3 DATABASE

The Database layer comprises all of the architecture necessary to store data for the BlastPhyMe application as well as code to manipulate that data within the database. The BlastPhyMe database exposes stored procedures to handle all data collection and modification processes performed by the Middle-Tier layer.

3.3.1 BLASTPHYME DATA FILE (*.BPMD)

BlastPhyMe is capable of exporting data from its database to a "data file", distinguished by the ".bpmd" file extension. A BlastPhyMe data file is an XML document compressed via the GZip compression algorithm.

4 THIRD-PARTY SYSTEMS

4.1 NCBI GENBANK NUCLEOTIDE DATABASE

BlastPhyMe communicates with NCBI's GenBank nucleotide database to search for and download GenBank records. This communication is performed via HTTP requests of the E-utilities web services hosted by NCBI:

http://www.ncbi.nlm.nih.gov/books/NBK25499/

BlastPhyMe is programmed to comply with the E-utilities Usage Guidelines and Requirements:

http://www.ncbi.nlm.nih.gov/books/n/helpeutils/chapter2/

4.2 NCBI BLAST WEB SERVICE

BlastPhyMe communicates with NCBI's BLAST web service using the aforementioned .NET Bio libraries, which make use of the QBlast URL API:

http://www.ncbi.nlm.nih.gov/blast/Doc/urlapi.html

BlastPhyMe is programmed to comply with the BLAST web service Usage Guidelines:

http://blast.st-va.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=DeveloperInfo

4.3 CODEML.EXE WITHIN PAML

For performing phylogeny tree analysis, BlastPhyMe uses the codeml.exe application from within the PAML suite of applications:

http://abacus.gene.ucl.ac.uk/software/paml.html