
Preface to the Second Edition

Bioinformatics is a fast moving subject, so while updating this book for the second edition it has been interesting to consider which aspects of the field have changed and which have remained the same. As the commissioning of a second edition indicates, bioinformatics skills continue to be in high demand, as the acquisition of complex multivariate data sets becomes the norm across many areas of biology. Although many new tools have emerged since the first edition of this book was published, and many existing tools have been further developed, there is still a need for the development of new tools, and the creation of bespoke software that brings existing tools together, increasingly across the different 'omics' disciplines.

So the need to build bioinformatics solutions remains. The general approach to building these solutions also remains the same, typically involving a database element, programming, some quantitative analysis, and possibly the creation of a web front end. We have therefore kept this structure within the book. The core tools around which we wrote the first edition—MySQL, Perl, and R—are all still popular within the bioinformatics community. The choice of programming language was a tough one, with Python and Ruby both gaining ground in the bioinformatics community. Ultimately, however, we decided that Perl was still the right choice for this book as it remains the single most widely used language in bioinformatics, and has the benefit of a thriving ecosystem that gives a head start when developing most types of bioinformatics applications. However, in recognition of the increased diversity of programming languages used in the community, we have added new material explaining how to get started in Python, Ruby, and Java to help you experiment with these languages so you can make up your own mind which best suits your needs. Having said that, we should note that all these languages are underpinned by the same fundamental programming concepts, so time spent learning Perl will give you a head start in any of the other languages.

Something else we have witnessed over recent years is a gradual increase in the size and complexity of bioinformatics projects. More complex projects call for more systematic coding practices, such as those proposed by the discipline of software engineering. We have added a new chapter dedicated to this subject, covering version control, documentation, user-centred design, and unit testing. These techniques can help streamline the development of your software—especially

when working as part of a team—and make it more reliable, easier to use, and easier to maintain and extend.

The technical area that has changed the most since the first edition is web development. Perl CGI's dominant position in bioinformatics web development has been usurped by web frameworks that are far more powerful while, at the same, make the development and debugging of web applications much easier. At the same time, the arrival of HTML5—the most significant update to the core web standard for over a decade—together with increasingly sophisticated JavaScript libraries such as jQuery, has ushered in a new era of more engaging web applications. Chapter 5 has been substantially revised to take account of these latest developments.

We sincerely hope that this book provides the overview of bioinformatics software development that you are looking for, and that it gives you the knowledge and confidence to go ahead and build your own bioinformatics solutions. We very much welcome your comments and questions—the best way to find us is via www.bixsolutions.net.

*Conrad, Darren, and Ian
November 2013*