The Bodin Project

In 1962, Kenneth Douglas McRae noted that Jean Bodin's "concept of sovereignty, his theory of climate, and his advocacy of religious toleration have today become commonplaces in practically all the histories of political thought." McRae offered this observation in his introduction to a facsimile of the English edition of 1606. That facsimile, the first and only full version since 1606, has since gone out of print. As a result, the work remains available in English only through electronic reproductions available online. This is a striking fact, given that the dissemination of works has become easier and less costly, scholarship has increased exponentially, and research into the less-known areas of modern political thought has exposed new, previously neglected territories. Scholars have known for a long time that Bodin had shaped the political thought of such figures as Hobbes, Locke, and Rousseau. The topics that animated Bodin's inquiry during the religious civil wars of the late sixteenth century have remained current, and his ideas and language have continued to exert a considerable influence, yet this most important of his works remains inaccessible to most readers, because of its size, scope, and publication history.

Bodin first published the French version, in 1576, and then a Latin version some ten years later.¹ These two versions are different in some important respects (the title of the Latin edition advertises it as "much richer than before"). When Richard Knolles decided to produce an English translation, he used both the French and the Latin. As a result, his Six Bookes of a Commonweale, is unlike either the French or the Latin. As McRae noted, prior to his publication of the facsimile in 1962, no complete edition of this book existed "since the Latin edition of 1641, none in French since 1629, and none in English since the Knolles translation of 1606." The only English edition in print today is severely abridged (4 chapters, 126 pages).

The size of the book--whose French edition numbers 1102 pages, while Knolles's comes to 792 quarto pages--would render the cost of the publication of even a single version prohibitive. A parallel edition would simply be out of the question. By way of comparison, consider that the recently-published annotated version of Hobbes's Leviathan, a much shorter and popular work, and in only two rather than three languages (English and Latin), was issued in three volumes totaling 1400 pages.² The text alone of the three editions of the Six livres is estimated at over 1,500,000 words. The book's complexity adds yet another layer of difficulty, since Bodin uses Greek, Latin, or Hebrew terms, and often alludes to persons, places, events, and sources that are unintelligible to the modern reader. Notes explaining these references and translating foreign or obscure terms would add significantly to the size of a critical edition.

The first step towards a modern critical edition of this work must therefore be the creation of an electronic edition based on the three existing versions. Such an edition will solve many of the aforementioned problems, as it would begin by forming an electronic version of the French, Latin, and English texts, which would be checked for accuracy and would thus provide the foundation for future work. Once checked, these versions would be set alongside one another, so as to indicate not simply how the text is translated (originally by Bodin himself, from French into Latin, and then into English by Knolles), but also--and most importantly--how one edition differs from another. This parallel

---


presentation would provide the basis for a master version of the text, which would in turn provide the source for a modern translation and critical edition. The electronic version of the texts would allow users to search terms and track them, by pointing to them in one version and seeing them highlighted in the others. It would also allow annotators to include references to sources cited by Bodin, provide explanations of relevant historical events and allusions to other sources, and thus exploit fully the extraordinary comparative richness of Bodin’s treatise.

In preparation for this project, I attended a National Endowment for the Humanities workshop at Tufts University, in 2012, entitled “Working with Text in the Digital Age.” During the 2012-2013 academic year, I applied for and received grants from the Earhart Foundation and the Office of the Provost at Tufts University, to begin work on the *Six livres*. In June 2013, I assembled a team of six Tufts undergraduate students and one alumnus who is currently pursuing a PhD in political theory at Duke University. We began by consulting the Perseus staff about their experiences in extracting digital text from print editions. At the same time, I coordinated with librarians from Tufts’ Tisch Library to acquire copies of Bodin’s French and Latin, which we digitized and made publicly available in collaboration with the Internet Archive, at the Boston Public Library.³ Our next step was to experiment with various optical character recognition (OCR) programs, in order to see which could address the various challenges presented by early modern French and Latin, such as ligatures, variations in spelling, etc., best. Having settled on a program and compiled a dictionary based in part on working dictionaries previously used by Perseus staff, we began the process of digitization. Not knowing quite what to expect, we set the relatively modest goal of extracting Book I in all three versions over the summer. Through careful monitoring of the workflow and training of the dictionaries, we were in fact able to extract the entire Latin text, as well as approximately 30% of the French. Because of Tufts’ participation in the Text Creation Partnership ⁴ that provides Early English Books Online ⁵ with the text version of some of its holdings, we were able to obtain an XML version of Knolles’s English edition.

It is worth noting that if the extraction of editable text from scans or photographs is a challenging task, the attempt to apply OCR software to early modern texts is all the more so. In addition to the difficulties introduced by the traits of early printed books, which frequently feature smudges and stains that obscure the type, and whose ink is sometimes too faint to read (see example 1) and at other times so bold as to bleed through (see example 2), the text itself is often an insurmountable barrier.

The most frequent of the many problems associated with sixteenth century type is the confusion introduced by the "long s" (ſ) which can only be distinguished from the letter "f" when both are printed


⁴ http://www.textcreationpartnership.org/

⁵ http://eebo.chadwyck.com/about/about.htm
very clearly. In many cases, knowledge of the language in question will suffice to show which of the two was intended, but that is not always the case, since there are several instances in which either letter yields a valid word (e.g., "fail" and "sail"). Training the dictionaries to recognize the difference between these letters and the various combinations of other ligatures and signs was one of the biggest challenges we faced, but we have developed a technique that has been yielding highly satisfactory results.

A further layer of difficulty was added by the fact that the vast majority of existing Latin dictionaries contain Ancient Latin, which differs considerably from its early modern variants. This means that we cannot rely on these dictionaries to check our work automatically, but must train them to recognize the variants as valid. Thus, to use the most obvious example, our otherwise excellent Latin dictionary took exception to the very title of the work, since knowing only "respublica," it would not recognize Bodin's "Republica."

To solve these problems and begin the construction of an XML file with initial markup, we tested our version of Book I of the Latin through Morpheus (a morphological parsing and lemmatizing tool produced by the Perseus Project), which can identify errors and mark them for correction. Book I contains 82,459 word tokens, of which Morpheus could not parse 14,589 words stemming form 4,684 distinct tokens. Excluding some 1,361 ampersands leaves 3,323 instances that need to be checked. We have begun to review these in an XML editor, marking the instances that need to be corrected, and expect to produce a highly accurate version of Book I soon. The updated list of corrections will be applied to the remainder of the Latin, thereby reducing the number of errors that need to be corrected manually.

Bodin's use of Hebrew and Greek terms in the original presents yet another difficulty. Occasionally, these are clear and legible, but very often they are not, either because of abbreviations and ligatures (Ex. 4), or because of typographical errors (Ex. 3) and smudges (Ex. 4).
The difficulties these pose are illustrated by the most up-to-date scholarly edition of a subset of the *Six livres*, edited by one of the foremost Anglophone authorities on Bodin's political thought. At the very start of that edition, the second part of the Greek in Ex. 4, above--admittedly hard to read--is misspelled.  

To solve the problems arising from Greek and Hebrew quotations, as well as from references or allusions to various sources, the Tufts Bodin Project joined the Dynamic Variorum Editions group. Along with colleagues from Canada (Mount Allison University), the United Kingdom (King's College; Imperial College), and the United States (Northeastern University) we are participating in a project entitled "The Dynamic Variorum Editions Toolkit," aimed at providing enhanced support for measuring and exploring the influence of Greco-Roman culture in large semi-structured digital collections across languages and cultures. The toolkit will be invaluable for the Bodin project in enabling us to identify the Greek, Hebrew, and Latin references, trace them to their original sources, and thus enhance the variorum edition's annotation significantly. The DVE group applied for a grant from the Digging into Data 3 challenge, but our application was unsuccessful; we expect to re-apply in the next round.

Having cleaned up the first few chapters, in the spring of 2014 we began to align the Latin text to Knolles's English using a platform based on the Perseids model that is being developed by the Perseus team specifically for the Bodin Project. In this platform, the users select a source text on one side of the screen and the translation on the other, they proceed to select a passage in each, and then save the link. The annotations resulting from the alignment process form the basis of the parallel edition.

---


7 [http://sites.tufts.edu/dynamicvariorum/home](http://sites.tufts.edu/dynamicvariorum/home)


9 See, e.g., [http://services.perseus.tufts.edu/berti_demo/berti_demo.html#urn:cite:perseus:lci.1.1](http://services.perseus.tufts.edu/berti_demo/berti_demo.html#urn:cite:perseus:lci.1.1)
The same tool can be used to annotate the text by providing links to Bodin’s sources. We began by adding the ability to link to some of the more frequent ones available on Perseus, but the list can expand to include other resources available online.

A beta version of the interface resulting from alignment is already available: ¹⁰

¹⁰ http://perseids.org/sites/bodin/
The interface works well with translation tools, such as the Alpheios add-ons that allow users to parse and translate Greek and Latin words without leaving the page:

Example of alignment using Perseids: Here, the English (middle column) is aligned with the French (right-hand column), as a result of having clicked on the blue highlighted sentence beginning "Or la definition ...". The highlighting in the Latin column is switched off.

Example of the Alpheios Latin tool in action: double-clicking on the term “perueniri” at C1 of the Latin (left-hand column), produces a pop-up box with the definition and grammatical information.
The user interface and alignment of I.i used for the beta version were put together during the spring semester, in the context of a political theory methods seminar. Over the course of the academic year, I presented the project to the NEH seminar entitled "Working with Text in a Digital Age" and the Tufts Digital Humanities Symposium. At the end of the spring semester, we held a workshop during which we discussed the project with Professor Ann Blair, one of the foremost authorities on Bodin in the United States. In June, I will participate in a conference on Bodin at Oxford University. Meanwhile, beginning at the end of May, a group of three undergraduates and one graduate student, will continue the alignment. Our goal is to have Book I fully aligned and at least partially annotated by the end of the summer. As we progress, we expect to add such features as links to maps through gazetteers, visual representations of the frequency with which places are mentioned in the work, as well as representations of the frequency with which the Six livres was printed, owned, and cited, so as to add a new dimension to the story of its reception and influence. After all, this was not just a book that interested Hobbes and Rousseau. One of the Latin copies currently housed in the greater Boston area belonged to King James I of England and another to John Adams. Other annotations will include links to biographies of individuals mentioned in the text, as well as links to information about historical events. Perhaps the greatest challenge with respect to annotation, however, is the identification of Bodin's sources. Kenneth McRae, whose research into Knolles' edition remains definitive, identified more than 4,850 sources in the Six Bookes, and there may well be more.

It is important to note that, though a large project in its own right, the digital variorum edition of Bodin is but the starting point of a broader, two-pronged effort aimed at the expansion of the Perseus Digital Library to include texts in the history of political thought and the dissemination of digital tools for the humanities to students and faculty members at Tufts and beyond. The project was founded on the idea that interested users should be welcome as participants and collaborators at any point. Incorporating these users has thus been a critical part of the process. Although many of the students who joined the project knew Latin and, some of them, French, one knew neither language. Moreover, none of the students had any experience with OCR and XML markup. All of them have since become familiar with the latter, and those who lacked the language skills necessary for corrections were nevertheless able to participate in the process through collaborative use of continuously updated lexica. Each stage of the process will add new skills, and will equip the participants with the foundation for future work in the digital humanities. Although daunting, the scope of the Bodin Project is also promising in that it provides a firm foundation on which to continue to train students in the history of ideas, political thought, and the application of digital technologies to the humanities.

No less important is the dissemination of these tools to interested students and scholars beyond Tufts. Much like the tools that we are developing, the edition of Bodin will not simply be available to interested students and scholars with access to the internet, it will also be open for submissions and contributions. Once fully operational, the Perseids platform will allow registered users to submit emendations and annotations. These will be directed to a board of experts who will be able to approve or reject them, and those approved will be incorporated into the edition. Every step of the process will be recorded, so that users can identify their contributions, and the integrity of the text can be maintained. Tufts students have already experimented with shorter texts, including texts that were influential at one point but have received little scholarly attention because they have been inaccessible. An inordinate amount of the scholarship on the history of political thought revolves around the same figures and questions, and not always because these figures and questions are too interesting to ignore. The tools and resources that we have been developing are opening the way for some exciting research into the history of political thought that can broaden the discipline's horizons and do so with students as full participants in the process.