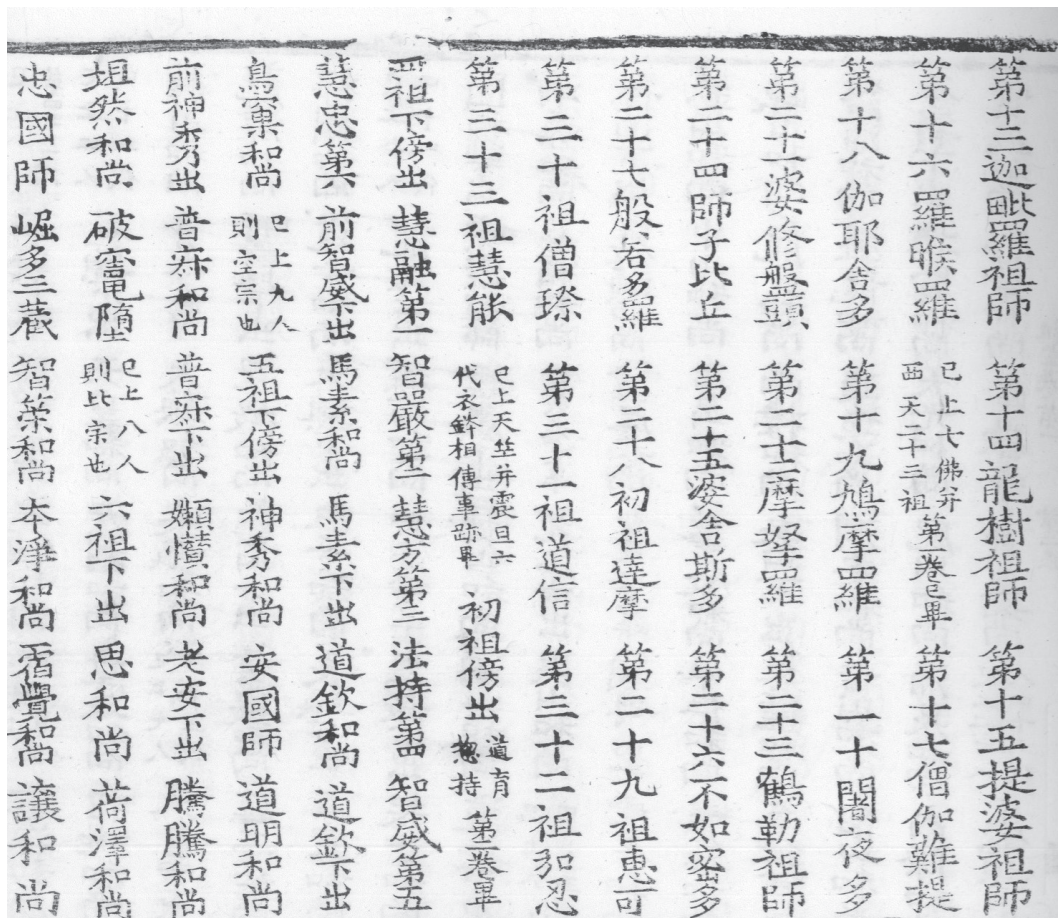


# センター研究年報2015

## CIEAS Research Report 2015



京都大学人文科学研究所附属  
 東アジア人文情報学研究センター  
 Center for Informatics in East Asian Studies  
 Institute for Research in Humanities, Kyoto University

特集 漢籍リポジトリ (ウィッテルン・クリスティアン編)  
 Special issue Kanseki Repository (ed: Christian Wittern)

## Preface

The topic of this volume is the “Kanseki Repository”, a digital collection of premodern Chinese texts. The work on these texts goes back more than 20 years, but the shape these texts and the software to interact with it take are largely the outcome of the research seminar “Fundamental topics in Digital Humanities” coordinated at the center since 2013. I would therefore especially thank the core members of this seminar, Ryūichi Kogachi, Bill Mak, Izumi Miyazaki, Tomoyuki Nagata, Hiroyuki Shirasu, Fuminori Tsuchiguchi, and Kōichi Yasuoka for the support during the seminar and the preparation of this report. It goes without saying that all errors are solely my own responsibility.

Christian Wittern

## まえがき

今回の特集のテーマは「漢籍リポジトリ」という漢籍の総合的なデジタル・コレクションである。このデジタル・テキストの作業そのものは20年以上前から始まったが、今の形で世に出せることは本センターで2013年から組織した「人文情報学の基礎研究」研究班の成果による。従って、研究班のコア・メンバーである古勝隆一、麥文彪（マク・ビル）、宮崎泉、永田知之、白須裕之、土口史記、安岡孝一に対して感謝の意を表す。いうまでもなく、誤りと誤植は全て編集者の責任である。

ウィッテルン・クリスティアン

## Contents

## 目次

1. Digital Philology meets premodern Chinese texts	
デジタル文献学が漢籍と出会う	1
2. Sources and organization of the Kanseki Repository	
漢籍リポジトリの情報源と編成	16
3. Using the Kanseki Repository	
漢籍リポジトリの使い方	30
4. Some use cases for the Kanseki Repository	
漢籍リポジトリの使用例	57

# 1 Digital Philology meets premodern Chinese texts

## デジタル文献学が漢籍と出会う

### Preamble

One of the challenges posed to us as scholars and editors by our times is to preserve the cultural heritage that has come down to us through the ages on its transition to digital formats. These new formats do away with many of the limitations of the print medium and open up exciting new possibilities. These can be realized in many different ways, however, based on our assumptions about what is the best way to achieve the desired result, what aspects of a work should be preserved and what aspects can be changed or discarded, and how access to the digital products should be organized.

Some of these assumptions may be due to a lack of awareness of the new freedoms allowed by today's technology. However, while most of these assumptions are not technical in themselves, they nonetheless involve decisions about technical implementation that at some point will influence the design of the interfaces and tools used to access the cultural heritage.

Here I will discuss the basic assumptions concerning the *Kanseki Repository* 漢籍リポジトリ (Kanripo). Firstly, this preservation should not be achieved by **fixing** the cultural artifact in display cases, like those of a museum, such that modern users can not readily interact with it, but rather by making it **available for exploration**, research and all kinds of interactions, imaginable and as yet unimaginable. Another assumption is that the process of preservation should be as open and transparent as possible, based on the ideal of a free, non-authoritarian discourse that embraces diversity.

These assumptions concern the design of the technical side of the system, as well as how decisions about this process are reached, what goals should be envi-

### まえおき

学者や編集者にとって今の時代における挑戦の一つは、幾多の時を経て私たちの手に届いた文化遺産をデジタル形式に保存することである。これらの新しい形式は紙メディアの持つ限界の多くを取り除き、刺激的で新たな可能性を開いてくれる。それらは数多くの方法で実現できるが、そうした方法は、望ましい結果を得るための最良の方法は何か、作品のどの側面は保存すべきでどの側面は変えたり切り捨てたりできるか、デジタルの産物へのアクセスはどのように準備すべきかについての私たちの暗黙の前提に基づいている。

これらの前提のいくつかは、今日の技術によって可能になった新たな自由についての認識不足によるものかもしれない。しかしながら、これらの前提のほとんどが、それ自体は技術的でない一方で、文化遺産にアクセスするためのインターフェースやツールの設計にいずれかの時点で影響を及ぼす技術的実装実現についての決定を必要としている。

漢籍リポジトリ（漢リポ）に関する基本的な前提をここで論じる。まず一つ目は、文化遺産を、現代のユーザーが容易にやりとりできない博物館のそのように、陳列ケースに固定することでこの保存を達成すべきではない。むしろそれを探求、研究、その他の想像し得る、また未だ想像し得ないありとあらゆるやりとりに供与することで達成すべきである。もう一つの前提は、この保存のプロセスが、多様性を包括する自由で非権威主義的な討議の理念に基づいて、できるだけオープンで透明であるべきということである。

これらの前提はシステムの技術的な側面の設計に関わっており、またそれだけでなく、この保存の方法についてどのように決定がなされるか、どのようなゴールを描

sioned, and so forth. All users should be involved and have a say, including future generations.

This is obviously not practicable, however, and since access to and use of the cultural heritage is fundamental, a key goal is to design a system that enables as many and as diverse usages as possible. This implies that there will be no central authority over the digital work, so the implementation will need to work in a way that does not require one. On the other hand, trust and reputation are important in allowing users to assess authoritativeness for themselves. Thus we will strive to develop a system that develops a strong and supportive community.

## Philology and digital text

The *Kanseki Repository* 漢籍リポジトリ (Kanripo) serves as a repository for premodern Chinese texts. These texts are provided in a way that facilitates their use as a reliable source for any kind of research for which they are needed. While the system will support a wide variety of different research approaches and methodologies, an awareness of the assumptions the repository is based on will help researchers use it more efficiently. These methodological assumptions are outlined below.

### Text and document

Scholars studying almost any aspect of premodern Chinese culture, depend on textual “witnesses” of things past. These witnesses are transmitted to us in the form of material objects, but in most cases we are interested not in the objects themselves, but rather the content they transmit. It is useful here to distinguish between the text as an abstract entity and the medium through which the text is transmitted. A text can be transmitted in the form of many different material objects, each with differences that might also extend to the content of the text. We thus need a model that can take account of these different aspects of an artifact.

One useful model to consider is that developed by the International Federation of Library Associations and Institutions (IFLA, see [6]). It describes the

くべきかななどにも関わっている。後世の人々も含めたすべてのユーザーがこれに関与し、発言権を持たなければならない。しかし、それは明らかに可能なことではなく、また文化遺産へのアクセスとその利用は基盤であるので、できるだけ多く、またできるだけ多様な用途を可能にするシステムを設計することが重要なゴールとなる。これは、デジタル作品においてはどのような中央権力も持たないこと、ゆえにそれを必要としない方法で実現されるべきであることを意味する。その一方で、信頼と評判はユーザー自身がそのデジタル作品の権威を評価するため重要である。それゆえに、しっかりとした、また好意的なユーザーコミュニティを培うシステムを開発することにも励むようにする。

## 文献学とデジタルテキスト

漢籍リポジトリ（漢リポ）は古典中国語文献のためのもの保管場所としての役目を果たしている。これらのテキストは、それを必要とするあらゆる種類の研究の信頼できる情報源として使える形で提供されている。このシステムは幅広い種類のさまざまな研究手法と方法論をサポートするが、それでもこのリポジトリが基盤とする前提を認識していることは、研究者がそれをより効率的に使う助けとなる。そうした方法論の前提をここに概説する。

### テキストとドキュメント

中国の古典文化を研究する学者は、そのほぼすべての側面において、過去の事柄のテキストでの「典拠」に依存している。そしてこれらの典拠は物質的な形をとって私たちのもとに伝えられてきた。しかし、ほとんどの場合、私たちはそうしたモノそのものではなく、むしろそれが伝える内容に興味を持つ。ここで、抽象的な実体であるテキストとそのテキストを伝えるための媒体とを区別することは有用である。一つのテキストは、それぞれがテキストの内容にまで及び得る違いを持った多くの異なるモノの形で伝えることができる。それゆえに、その有形物のこうした異なる側面を考慮に入れることができるモデルが必要である。

国際図書館連盟（[6] を参照）が開発したモデルは、検討に値する有用なモデルである。これは「書誌レコードの機能要件」（FRBR と略）を説明したもので、書誌



“Functional Requirements for Bibliographic Records” (abbreviated as FRBR) and is quite complex, describing all aspects of bibliographic data. Of relevance here is only the first set of terms that denote the “products of intellectual or artistic endeavor” as four related entities: **work**, **expression**, **manifestation**, and **item** (See Figure 1.1). These four terms are elaborated below.

- **Work** is a “distinct intellectual or artistic creation.” ([6], p. 17) For example, the *Analects* of Confucius or the *Dao De jing* are works. Works are independent of the rendering in a specific edition or language. Thus, an English translation is considered a representation of the work in the same way as the Chinese original. A **work** according to this model is an abstract entity that does not rely on a specific expression in words.
- **Expression** is “the specific rendering the work takes each time it is ‘realized’” ([6], p. 19). A specific formulation of the text of the *Analects* or a specific translation thereof, or a version of the text of the *Dao De jing* with the commentary by 王弼 *Wang Bi* are all examples of expressions. An expression is still independent of a specific edition, however.
- **Manifestation** is “the physical embodiment of an expression of a work.” ([6], p. 21). For textual works, this is usually a specific edition published by a specific publisher, for example D. C. Lau’s translation of the *Analects* published by Simon & Brown in 2012 or the edition of *Wang Bi*’s commentary included in the 四部叢刊 *Sibu congkan* of 1929.
- **Item** is “a single exemplar of a manifestation. The entity defined as item is a concrete entity.” ([6], p. 21). My own exemplar of D. C. Lau’s translation or my photocopy of Wang Bi’s *Sibu congkan* edition are examples of items.

As scholarly editors we work with **items** that carry evidence of the **work** through the **expression** that the

データのすべての側面を表したかなり複雑なものである。その中でここに関連するのは「グループ1」用語一式のみであり、著作、表現形、体現形、個別資料という四つの関係した実体で「知的あるいは芸術的活動の所産」を指している。これら四つの用語を以下に説明する（図1.1を参照）。

- 著作（Work）は「明確な知的あるいは芸術的産物である」（[6], p. 17）。例えば、孔子の『論語』、また『道德經』は著作である。著作は特定の版や言語への表現とは独立している。それゆえに、日本語翻訳は、中国語の原作がそうであると同様に、著作の表れとみなされる。このモデルによるところの著作は、言葉による特定の表現に依存しない抽象的な実体である。
- 表現形（Expression）は「著作が実現される度にそれが要する特定のレンダリングである」（[6], p. 19）。『論語』のテキストの特定の形態、またその特定の翻訳、あるいは王弼による注釈の付いた『道德經』のテキスト、これらはすべて表現形の例である。表現形もまだ特定の版とは独立している。
- 体現形（Manifestation）は「著作の表現の物質的な具体化である」（[6], p. 21）。テキストでの著作において、これは通常、特定の出版社から出版された特定の版のことである。1999年に岩波書店から出版された金谷治氏による『論語』の訳注の改定版、また1929年の『四部叢刊』に入っている王弼の注釈版がその例である。
- 個別資料（Item）は「体現形の一つの見本である。個別資料として定義された実体は具体的な実体である」（[6], p. 21）。金谷治氏の翻訳の私個人の所蔵本、また王弼の『四部叢刊』版の私個人のコピーが個別資料の例である。

私たちは学術分野の編集者として、著作の典拠を伝えている個別資料を使い、その個別資料が記録している表現

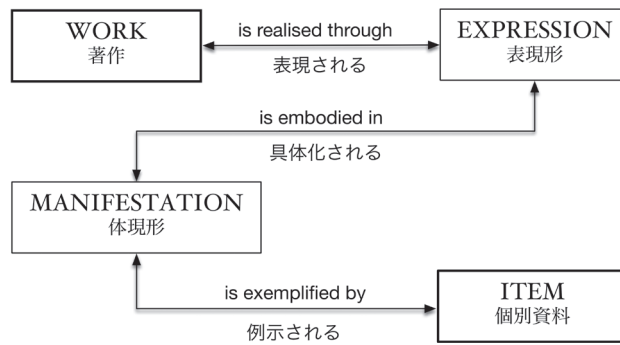


Figure 1.1: Entities of the first group and their relations

図1.1 「グループ1」における実体とそれらの関係

item records. If there are multiple textual witnesses of the **work**, we usually deal with multiple **manifestations** of this **expression**. On the other hand, in bibliographic discussions concerning specific texts and the way they have been transmitted historically, the focus is usually on **expressions** and **manifestations**, i. e., levels of abstraction above the level of the individual **item** but below the level of **work**.

In the Kanseki Repository, this model has been used to determine whether two “digital items” should be related to separate expressions, and thus have **separate entries** in the catalog, or merely to separate manifestations of the same expression, thereby representing two witnesses (editions) under the **same entry** in the catalog.

A useful simplification of this terminology might be to employ the more familiar, though slightly ambiguous term **text** when we speak of the “work”, and **document** for “item”. Multiple witnesses of a text (work) would then be multiple “documents”, which might or might not represent the same “expression/manifestation”. Since this second set of terms, while not as formal and precise as the FRBR set, is sufficiently expressive and also more familiar to philologists, I will use the terms **text** and **document** where possible.

So how should we understand the relationship between text and document? And what does this understanding mean for our task of producing scholarly reliable, philologically grounded digital editions of a text? Before exploring this further, yet another important clarification needs to be made.

## Record and interpretation

In a seminal article, the Swiss scholar Hans Zeller

形を通じて仕事をする。その著作に複数のテキストでの典拠があるならば、通常私たちはその表現形についての複数の体现形を使う。一方、特定のテキストとそれらが歴史的に伝えられてきた伝達の成り行きに関する文献学の議論において、焦点は通常、表現形と体现形におかれる。すなわち、抽象レベルは個別資料レベルの上に位置し、しかし著作レベルの下に位置する。

漢籍リポジトリにおいて、二つの「デジタル個別資料」を別々の表現形に関連づけるか、ゆえにそれらは目録に別々の項目を持つか、あるいは同じ表現形を別々の体现形に分けるだけにして目録には同じ項目の元で二つの典拠（版）を表すかを決めるのに、このモデルが使われてきている。

この専門用語を実用的に簡素化するならば、「著作」の代わりにテキスト、「個別資料」の代わりにドキュメントという、少し曖昧ではあってもよりよく知られている用語を使うことができる。そうであれば、テキスト（著作）の複数の典拠は複数の「ドキュメント」と対応するが、それらは同じ「表現形/体现形」と対応したりしなかったりする。これらの用語は、FRBRのそれより正式でも正確でもないが、十分に表現でき、文献学者によりよく知られている。それで、私は可能な限り、これらの用語、テキストとドキュメントを使うことにする。

では、テキストとドキュメントの関係をどのように理解したらよいただろうか。そしてこの理解は、学術的に信頼でき、文献学に基づいたテキストのデジタル版を作る課題においてどのような意味を持つのか？ このことを更に探求する前に、他の重要な事柄を明らかにする必要がある。

## 記録と解釈

スイスの学者ハンス・ツェラー [23] は、その影響力を



[23] emphasized the fact that all scholarly editing should make a clear distinction between the **record** of what has been transmitted in the “documents”, and the scholarly **interpretation** thereof. In this terminology, the **record** is the essentially objective evidence or record of a text in the form of a material document, while the **interpretation** is the way in which a specific reader of this document, in his function as editor of an edition, subjectively reads and understands the text, based on knowledge of the writing style and any other supporting evidence that is derived from knowledge about the text, but not necessarily part of this specific document itself. Equipped with such knowledge, an editor might be able to point out errors in the characters that make up the text, missing or misplaced pages, and additions or omissions in the textual content, and even attempt to produce a new edition (**manifestation**) of the text that improves on perceived shortcomings. Zeller stresses that we should not only be aware of this subjectivity and objectivity in editing, but also communicate it to the users of the editions we produce, to prevent users from having unreasonable expectations and trying avenues of analysis that are simply not supported by the edition. It also goes without saying that this is not a new distinction. As Zeller himself points out, in the European tradition it goes back to at least Aristarchus of Samothrace (*Ἀριστάρχος*; c. 220–c. 143 BC), an influential scholar of Homeric poetry, who established the first critical edition of Homer’s works, rejecting and marking lines he considered doubtful, but always making it clear what was based on his own judgment and what was evidence of the documents available to him.

While this distinction can be blurry at times, and some question the very notion of objectivity in the record, it still seems useful and it has informed the design of the *Kanseki Repository*, which arranges the editions of a text it represents into those that strive to faithfully reproduce a text according to some textual witness (**record**) and those that make alterations or additions to the text, for example by adding punctuation, normalizing characters, collating from other evidence, etc. (**interpretation**).

持った論文の中で、すべての学術的な編集は「ドキュメント」の中で何が伝えられているかの記録とその学術的な解釈とを明確に区別しなければならないという事実を強調した。この専門用語において、記録とは、本質的に客観的な証拠、または物質的なドキュメントの形でテキストの記録である。一方、解釈とは、このドキュメントの特定の読者が、一つの版の編集者という役割において、書き方についての知識やその他、必ずしもこの特定のドキュメントの一部でなくてもこのテキストについての知識に由来している裏付け証拠に基づいて、テキストを主観的に読み、理解することである。このような知識を持って、編集者はテキストを構成している文字の誤り、乱丁や落丁、またテキストの内容の付加や抜けを指摘するかもしれないし、認知されている欠点を改善したテキストの新しい版（体現形）を作ることを試みるかもしれない。ツェラーは、編集作業におけるこの客観性と主観性を私たち自身で認識するだけでなく、私たちが作り出した版の利用者が過剰な期待を持ち、その版によっては不可能な分析方法を試みることはないように、それをその利用者にも伝えるべきだと力説する。言うまでもなく、この客観性と主観性の区別は新しいものではない。ツェラー自身が指摘するように、ヨーロッパの伝統において、これは少なくとも、ホメーロスの詩の学者であり影響力のあったサモトラケのアリスタルコス (*Ἀριστάρχος*; c. 220 – c. 143 BC) までさかのぼる。彼は、自らが疑わしいとみなす詩の行を却下し、また印をつけ、しかし常に何を基準にしてその判断を下したか、またどんなドキュメントの典拠を使用したかを明確にし、ホメーロスの作品の最初の校訂版を確立した。

この区別は時として曖昧であるかもしれず、また記録における客観の概念そのものを疑問視する者もいるが、それでもこの区別はまだ有用であるように思われ、またそれは漢籍リポジトリの設計に影響を与えている。この漢籍リポジトリは、それが表すテキストの版を、テキストでの典拠に従って忠実にテキストを再現することに尽力するもの（記録）と、例えば句読点を加えたり、文字を正規化したり、別の典拠を照合するなどしてテキストに変更や追加をするもの（解釈）に分類する。

## Additional requirements

Studying textual content (i. e., just text) in **isolation** severely limits understanding. In many cases it is desirable to have more context or supporting external evidence. Peter Shillingsburg [17] outlines the following requirements of a digital edition (slightly edited for clarity):

a. Digitized images of all the documents. This makes it possible to view any document side-by-side with any other document, from anywhere in the world, without traveling from Tokyo to Marburg and New York.

b. A table of variants to show how all the documentary texts differ from one another.

c. A textual history that explains the relationships between the variant documents and why we should care about them.

d. A transcription of at least one of the documents, so that the variants list can be used more easily. Or else transcriptions of all the documents so that readers can select and read any of them. This would also enable machine collation.

e. Edited copies of the transcriptions to correct errors, to create new editions based on individual **interpretation**. While the other editions of the **record** will preserve the text as a historical documentary text, such revisions will help readers avoid the distractions caused by scribal or compositorial errors.

As outlined above, a strict division is maintained between the “record”, i. e., a descriptive and neutral representation of a text, and the “interpretation” of that record, i. e., any edited version of, commentary to, or other enhancement of the text. For this purpose, anything that has been done before the activity of the modern editors, including previous modern editions of a text, is considered part of the record.

This should allow users to clearly distinguish at any moment between views or changes introduced within the tradition and those introduced later.

### The record: Two faces of digital text

What has been said so far could be applied in a similar

## 追加の必要事項

テキストでの内容（すなわちただのテキスト）のみを分離して研究することは、理解を厳しく制限することになる。多くの場合、より多くのコンテキスト、あるいは外部にある裏付け典拠を持つことが望ましい。ピーター・シリングスバーグ[17]は以下のようなデジタル版の必要事項を挙げた（明確にするために多少の手を入れた）：

a. すべてのドキュメントのデジタル化した画像。これが、例えば東京とマールブルクやニューヨークの間を移動する必要をなくし、あるドキュメントと別のドキュメントを、それらが世界中のどこにあらうとも、並べて見ることを可能にする。

b. すべてのドキュメンタリー・テキスト（記録）について、それぞれがどのように異なっているかを示すための異同の一覧表。

c. 異同のドキュメント間の関係や、なぜそうした関係に関心を持つ必要があるかを説明したテキストの伝達の歴史。

d. 異同のリストをより容易に使えるようにするために、少なくとも一つのドキュメントの翻刻。そうでなければ、読者がどのドキュメントも選択し、読めるようにするために、すべてのドキュメントの翻刻。これはまた機械による照合も可能とする。

e. 個々の解釈に基づいて新しい版を作るために、編集し誤りを訂正した翻刻のコピー。記録の別の版が歴史的なドキュメンタリー・テキストとしてテキストを保存するのに対し、このような改訂は読者が筆写や植字による誤りによって気がそらされるのを避ける助けをする。

上記のように、「記録」（すなわち記述的で中立的なテキストの表れ）と記録の「解釈」（すなわちテキストのあらゆる編集バージョン、テキストへの注釈、その他のテキストの改善）の間に厳しい分別が維持されている。この目的のために、以前出されたテキストの現代版も含め、現代の編集者の作業以前になされたものはすべて、記録の一部とみなされる。

これはユーザーがいつ何時でも、伝統の中で導入された見解や変更と後になって導入されたものとを明確に区別することを可能にする。

### 記録：デジタルテキストの二つの顔

ここまでで述べたことは印刷されたテキストにも同じよ





way to printed texts (Shillingsburg) and in some cases it has been written about exclusively with printed editions in mind (Zeller). But there are also issues that are specific to digital editions.

Every edition that enters the record can be represented in two distinct, but related digital forms: as **digital facsimile**<sup>1</sup> and as **digital transcription**. Both introduced by necessity alterations due to the medium used, but they should still be kept as close to the original as possible. Thus, a digital facsimile might be reduced to monochrome (black and white) images, with the white margins cut away and the contrast enhanced, but still considered a faithful representation of the original texts<sup>2</sup>. In the same way, the transcription of a text might strive to reproduce the text as closely as possible to the characters as seen on the facsimile (or more correctly, as read and interpreted by the typist or input program), without ever achieving a complete identity. (This means that if several variants of a character are available in the coded character set, the one closest to the character on the page is selected).

To be complete, the record also comprises descriptions of the available editions that are not yet part of the repository, ideally including all known editions.

Some conventions to easily distinguish these two types of texts are outlined below.

## The Interpretation

While the record tries to present the evidence in a form that is as neutral and unbiased as possible, it is in the interpretation that real scholarly editing takes place. The record is scrutinized and amended, commented on, etc. Such modifications and interpretations ensure the living evolution of the text that has continued ever since it was first fixed in writing.

One important use of this concept in the context of the Kanseki Repository is to provide an edition that employs modern character usage, rather than the characters used in the editions documented in the record. This allows us to cater to the needs of modern users but still provides a record of the existing textual tradition through reference to the editions in the “record”.

Premodern Chinese texts (古籍 *guji* in Chinese, 漢籍

うに適用でき (シリングスバーグ), また場合によっては印刷版だけを念頭において書かれたものもある (ツェラー)。しかしデジタル版特有の問題もある。

記録に入れられる版のそれぞれが, 二つの別個の, しかし関連しているデジタル形式で表される。デジタルファクシミリ<sup>1</sup>とデジタル翻刻である。両方とも, 使用する媒体により, 必要に迫られ変更を導入するが, それでもできるだけオリジナルに近い状態で維持されるべきである。ゆえに, デジタルファクシミリは, 白のマージンが切り取られコントラストが強調されたモノクロ (白黒) の画像にまで引き下げられているかもしれないが, それでもオリジナルテキストの忠実な表れであるとみなされる<sup>2</sup>。同様に, テキストの翻刻は, 完全な同一性を達成することは決してないままに, ファクシミリの中の文字にできるだけ近づけてテキストを表すように (より正確には, タイピストや入力プログラムによって読み解釈されるように) 励む。(これは, コード化された文字一式の中で, 一つの文字にいくつかの異体字がある場合, ページ上でその文字に最も近いものが選ばれることを意味する)。

完全を期すと, 記録はまた, リポジトリにはまだ含まれていない, 理想としては知られているすべての版を含めた, 利用可能な版の記述で構成されている。

テキストのこれら二つのタイプを容易に区別するための慣例を以下に概説する。

## 解釈

記録ができるだけ中立的で偏らない形で典拠を提供しようとする一方, 解釈の中でこそ真の学術的な編集が行われる。記録は細かに調べられ, 修正され, 注釈を付けられる。このような変更と解釈は, それが最初に文字の形で固定されて以来ずっとそうしてきたように, テキストがこれからも生き残り, 進化し続けることを保証する。

漢籍リポジトリのコンテキストにおけるこの概念の重要な効用の一つは, 記録の中に記された版で使用された文字ではなく, 現代の文字用法を使った版を提供することである。これにより私たちは現代のユーザーのニーズに応えることができ, そうしながらも「記録」の中の版を参照することで現存するテキストでの伝統の記録を提供する。

何世紀もの間, 美しい木版印刷の形で伝えられてきた

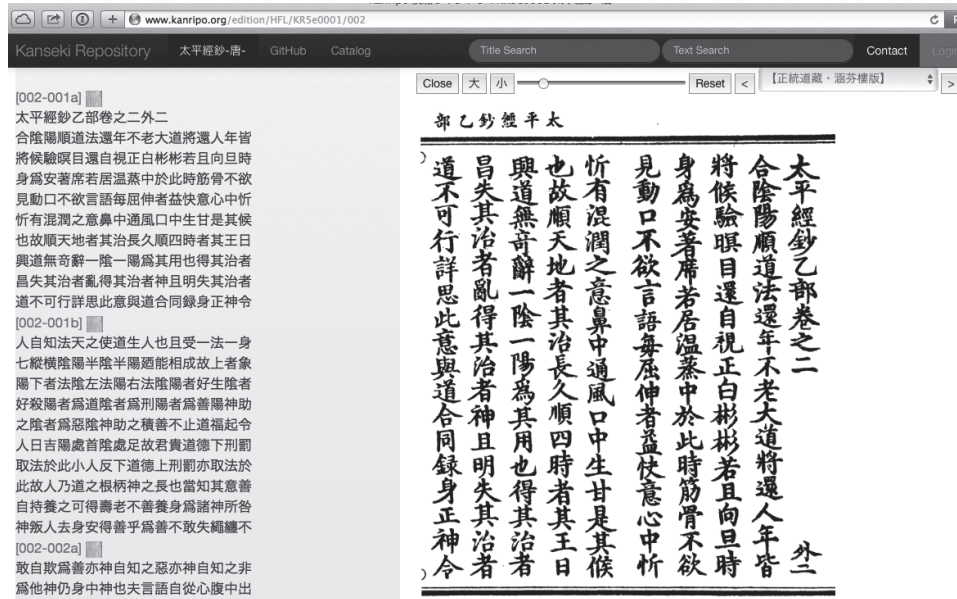


Figure 1.2: A digital facsimile of a page from the 太平經抄 (right) with the transcribed text (left)  
 図 1.2 太平經抄の中の一ページ分のデジタルファクシミリ (右) とその翻刻テキスト (左)

*kanseki* in Japanese - hereinafter I will use the Japanese designation) that were transmitted in the form of beautiful woodblock prints for many centuries reveal numerous traces of this kind of evolution, as their owners took possession of the texts as objects. As well as affixing collector's stamps, they also freely added marginal notes, reading marks, punctuation and more, thus making almost every copy of a work a unique item reflecting a specific trajectory of readership over the generations. Similarly, all users of the repository can create their own copies, marked with their individual readings and comments. As yet, we do not have the digital equivalent of a collector's stamp, however.

In addition to this ad hoc commenting on individual items, there is also the traditional exegesis, in which commentary, explication and sound-glosses were inserted interlinearly into the root text, albeit usually in smaller characters to set them off. The result was then published as a new edition. In other texts, the root text and the commentary might also be set apart by other typographical means, by indenting the first line of the commentary for example. It is thus a fact that outside of the Buddhist scriptures, almost all texts of any standing as a classic have been transmitted not as raw, unadulterated texts, but rather wrapped in layers of commentary and annotation. Within the Kanseki Repository, such 'texts with com-

漢籍は、その所有者がテキストを物として入手したことから、この種の変遷の数多くの形跡を示している。彼らは蔵書印を押しただけでなく、欄外注、訓点、句読点などを自由勝手に加えた。それゆえに、何世代にもわたる読者たちの軌跡を反映して、特定の著作一部一部のほぼすべてが唯一無二の個別資料になった。同様にして、リポジトリのユーザー全員が、個々の訓点や注釈を記し、それぞれ自分の一部を作ることができる。しかしながら、デジタルの蔵書印に相当するものはまだ存在しない。

こうした個々の個別資料にその場限りの注釈を付けることに加えて、注釈や詳説、また音注を、それらを際立たせるために通常は小さめの文字で、本文の間に差し込む割注もある。そしてその結果は新しい版として出版された。別のテキストでも、例えば注釈の最初の行をインデントするなど他の版式を使って、本文と注釈を区別した。ゆえに、仏教経典を除き、古典に数えられるテキストはほぼすべて、そのままの、混ぜ物なしのテキストとしてではなく、むしろ何層もの注釈や注解に包まれて伝えられてきた。漢籍リポジトリの中で、そのような「注釈の付いたテキスト」は通常、別々のテキストとして扱われ、目録に別々の項目を持つ。ここで使われる文献学的な意味において、これらは本文の解釈というよりむしろそれ自体が記録となる。



mentary' are usually treated as separate texts with their own entries in the catalog. Rather than an **interpretation** of the root text, in the philological sense used here, they become a **record** in their own right.

### Details of the technical implementation

The texts contained in the repository are presented with only the markup that is absolutely necessary for them to function within the supporting system. Every edition of the text is established separately and curated **alongside** other editions. The format for the texts is an extension to the Emacs Org-mode text format<sup>3</sup> and the technical means to handle the **alongside** is 'version control software' as outlined in the following section.

This approach is quite contrary to the current practice in Scholarly Editing and Digital Philology, which uses the XML based Text Encoding Initiative's (TEI, see [19]) textual model as the standard, and therefore requires an explanation.

First of all, it is important to note that there is not just **one** true rendition of an existing text in digital form. **Every** rendition will have to develop its own methodology, make decisions about how to represent certain distinct features of a source text and what to consider as ancillary baggage to be left behind during the transition to the digital medium. No one digital edition will therefore be the same as another, except where one is a direct **digital** copy of the other.

One important feature distinguishing such editions is of course the technology used. In Digital Philology, most of the work done since the turn of the century, is based on the text model developed by the Text Encoding Initiative (TEI) and syntactically based on XML, a technology that also drives most of the World Wide Web, among other things. With its sophisticated model of texts and many ways to record textual features of interest to the encoder, it is an important technology and its expressivity goes lightyears beyond the modest means of the syntax employed here.

The fact that TEI and XML are not used for the repository here should not be seen as a statement of their unsuitability for the task at hand. They are very suitable in fact, and can be used in conjunction with

### 実装の技術的な詳細

リポジトリに収められたテキストは、それらがサポートしているシステムの中で機能するのに絶対必要であるマークアップだけを付けて表示される。テキストの一版一版が別々に確立され、他の版と並べて提示される。このテキストの形式は Emacs Org-mode のテキスト形式<sup>3</sup>を発展させたものであり、並べて提示する技術的方法是、以下の項で概説する「版管理システム」を使っている。

このアプローチは、学術的な編集やデジタル文献学において、標準としてテキスト・エンコーディング・イニシアティブ (TEI: [19] を参照) の XML ベースのテキストモデルを使うという現在の慣行にかなり反しており、それゆえに説明を必要とする。

まず第一に、現存するデジタル形式のテキストの正しい再現方法は一つだけでないことに留意する必要がある。それぞれの再現について、それに見合った方法を開発しなければならず、またその元になるテキストの特定の特徴を表す方法やデジタル媒体への移行において何を不必要とみなして取り去るかを決めなければならない。それゆえに、デジタル版のどれ一つ取っても、それが別の版の直接のデジタルコピーである場合を除いて、他と同じものはない。

このような版を区別する一つの重要な特徴は、当然ながら、そこで使用されたテクノロジーである。デジタル文献学において今世紀に入ってからなされた仕事のほとんどは、TEI が開発したテキストモデルを基盤にしており、また他のさまざまなもの、中でもほとんどのウェブを動かしているテクノロジーである XML を構文的な基盤にしている。その洗練されたテキストモデル、およびエンコーダーに興味深いテキストの特徴を記録する数多くの方法ゆえに、これは重要なテクノロジーであり、その表現度はここで使われている構文の質素な方法をはるかに上回るものである。

このリポジトリに TEI と XML は使われていないが、それはこれらが当面の課題に適していないという表明であると理解してはならない。実のところ、これらはとても適しており、またとても効率的にここで提示している

what is presented here to great advantage. Alternatively, they can be used on their own in the usual way. In more than 20 years of working with these methodologies, however, I found that they are not always an ideal fit for every station on the path of a text from the reader (as editor) to the producer and back to the reader. There is also room for a simpler, but yet well defined textual model, which can be algorithmically “up-converted” to a more expressive syntax, if desired, to fit into a production pipeline. Additionally, it is useful to also provide a text in plain format as the final stage of the publication process.

While the simple format used here works well for establishing the text, and for reading, translating and annotating, it might prove inadequate for more sophisticated use cases, like linguistic annotations, or the markup of names, persons and locations<sup>4</sup>. In such cases, a conversion to XML representation of the texts is feasible.

The long and short of all this is that there is no single representation that is technically superior in all cases. Pragmatically, the one that best serves the needs of the text and the reader should be chosen at any given time.

This leads quite naturally to another observation: A digital text cannot be sufficiently appreciated by simply looking at the sequence of bits or bytes that represent the text; it is necessary to take into account the whole system that supports the presentation of the text to the reader — the application software for rendering the texts, the underlying operating system and its properties, the fonts used for rendering text, and the size, pixel density and luminosity of the screen (or paper!) on which the text is rendered all contribute to the system that presents the text to the user. With all of these items, there is a multitude of variation, depending on the purpose of viewing the text. Users of the Kanseki Repository are thus encouraged to experiment with this liberty and use what best fits their needs. Chapter 3 will provide further details of this.

### How texts and editions appear in the Kanseki Repository

As mentioned, a minor extension of the Emacs Org-

ものと共に使うことができる。あるいは、通常の方法で、これらだけで使うこともできる。しかしながら、20年以上もこうした方法論を使って仕事をしてきた中で、私はこれらが（編集者としての）読者から制作者へ、そして制作者から読者へのテキストの道中すべての中継点にいつも理想的に合うわけではないことを実感してきた。生産パイプラインに合わせて必要であればより表現度の高い構文へと機械的に「アップコンバート」できる、よりシンプルで、それでもしっかりと定義されたテキストモデルの入り込む余地はあると考える。これに加えて、出版プロセスの最終段階でプレーン形式のテキストもまた提供すると役立つであろう。

漢籍リポジトリで使用したシンプル形式はテキストを確立し、読み、訳し、また注解をつけるのにうまく機能する一方で、言語学的な注解や名前、人物、場所のマークアップといった、より洗練された用法には不十分<sup>4</sup>であることになるかもしれない。そのような場合、テキストのXML表現へのコンバージョンが適している。

要するに、すべての場合において技術的に優れている表現はひとつもないということである。実践的に、その時点でテキストと読者の必要を最もよく満たすものが選ばれるべきである。

このことはとても自然に別の観察へと私たちを導く：デジタルテキストは、そのテキストを表すビットやバイトの列をただ眺めるだけで十分に理解することはできない。つまり、読者へのテキストの提供をサポートするシステム全体を考慮に入れる必要がある。テキストをレンダリングするためのアプリケーション・ソフトウェア、基礎をなすオペレーティング・システムとその特徴、テキストをレンダリングするために使われたフォント、そしてテキストがレンダリングされたスクリーン（あるいは紙！）のサイズや解像度、明るさはすべて、ユーザーにテキストを提供するシステムに貢献している。これらすべての項目について、テキストを見る目的により、数多くのバリエーションがある。それゆえに、漢籍リポジトリのユーザーにはこの自由を試してみ、自分が必要とするものに一番合うものを使うことを奨めたい。第3章でこの更なる詳細について述べる。

### 漢籍リポジトリにおいてテキストと版はどのように表示されるか

前に述べた通り、ほぼプレーンテキストである Emacs



mode format, which is almost plain text, serves both as the first stage of preparing a text in the digital format, for example by typing it out, and as the final stage where the text is published and distributed. This is a format that makes the least assumptions about the receiving system by restraining it technically to a lowest common denominator, i. e., “plain text”<sup>5</sup>.

Kanripō implements the principles outlined above in the following way:

- A text “repository” is created for every text for holding all information related to the text, including links to facsimile renderings of pages and the content of the text itself<sup>6</sup>.
- The repository is named according to a convention that gives the text an appropriate place in the textual and conceptual universe of premodern Chinese texts. In principle, the conventions used here follow those for classification of traditional Chinese cataloging, as exemplified by such great works as the 四庫全書總目提要 *Siku quanshu congmu tiyao* or the bibliographical treatises of the dynastic histories, with a few alterations as outlined in the next chapter. The name of the repository serves as an identifier for the text and this identifier is used to uniquely and unambiguously refer to the text.
- The repository is created using a software called “git”<sup>7</sup>. This software was designed to help software developers collaborate across different time zones and locations. The details are unimportant, but the essential point is the concept of “branches”, which enables work on different versions of a text to be done simultaneously and allows the switching back and forth between multiple versions. The software is capable of dealing with as many different versions of a text as needed and offers many functions, such as efficient storage and retrieval, display of differences, etc.
- Every one of the editions in the “record” that has a digital representation in Kanripō will have a **branch** that represents this edition. Such a branch is given a name to identify it. We will use names that include only **uppercase** letters (and a hyphen

Org-mode 形式を少し拡張したものは、例えばそれをタイプすることなど、デジタル形式のテキストを準備する第一段階としても、またテキストが出版され配布されることの最終段階としてもその役目を果たす。これは受信システムについて、それを最小共通項、つまり「プレーンテキスト」<sup>5</sup>まで技術的に抑えることにより、最も少ない前提を作る形式である。

漢リポは上記した原則を以下の方法で実行する：

- テキスト「リポジトリ」<sup>6</sup>が、それぞれのテキストごとに作成される。これは、そのテキストに関するすべての情報を保持するため、そこにはページのファクシミリ・レンダリングへのリンクやテキストそのものの内容も含まれる。
- 漢籍という文字化し概念的な領域の中でテキストに適切な場所を与える慣例に従って、このリポジトリに名前が付けられる。ここで使われる慣例とは、原則として、『四庫全書総目提要』や正史の古籍誌のような偉大な作品によって例証される、伝統的な漢籍目録が用いられる分類法に従うが、幾つかの変更もあり、それらについては次の章で概説する。このリポジトリの名前はテキストの識別子の役割を果たし、この識別子はテキストを一意にそして明確に識別するために使われる。
- このリポジトリは「git」<sup>7</sup>と呼ばれるソフトウェアを使って作成される。このソフトウェアは異なる時間帯や場所の全域でソフトウェア・デベロッパーが共同して働く助けとなるために設計された。詳細は省略するが、要点は「ブランチ」の概念で、これは異なるテキストのバージョンで仕事を同時に行うことを可能にし、また複数のバージョン間で切り替えることも可能にする。このソフトウェアは必要なだけ多くの異なるテキストのバージョンを扱うことができ、効率的な保存と検索、違いの表示など多くの機能を提供する。
- 漢リポにデジタル表現を持つ「記録」の中の版の一つ一つが、その版を表すブランチを持つ。そのようなブランチにはそれを識別するための名前が付けられる。記録の中の版を表すブランチを指すのに大文字（及びハイフン「-」）だけを含めた名前を使う

“-”) to point to branches that represent editions in the record, such as WYG (for the reprint of the 文淵閣 *Wenyuan ge* edition of the 四庫全書 *Siku quanshu*) or CK-KZ (for the 考正 *Kaozheng* reprint of the 重刊道藏輯要 *Chongkan Daozang jiyao*). These branch names can be arbitrarily chosen, but they should be documented somewhere in the record.

- Other branches represent editions or versions of the text that do not form part of the record, but are added by editors or researchers for specific purposes, which might be the creation of a new critical edition, reading notes, translation, commentary, etc. If such an edition is based on an edition in the record, the referencing system of the base edition is carried over to allow better integration into Kanripo. These branches are usually named with a combination of uppercase and lowercase letters or lowercase letters only (such as **WYG-chris** or **trans-en**) to clearly distinguish them from the non-interpretative branches<sup>8</sup>.
- One of the editions in the record will be selected as a “base edition”. The page numbers and line markers of this base edition will be used as a reference framework in all editions of this text. This is not necessarily the philologically most valuable edition, but rather a pragmatic choice for the purpose of supplying a reference framework for the other editions in the “record”. This does allow the cross-referencing of locations within different editions across editions. All editions have a reference to the base edition as part of the information at the beginning of each file.
- Among the other branches, there are two which have a special meaning within the Kanseki Repository: the **master** and **\_data** branches.
- **master** is the branch which has the most up-to-date text as chosen by the editors of the Kanseki Repository. This might be the one in the base edition, but might also be a different one, if that seems to be a better starting point. It will usually also introduce (different) punctuation, character normalization, structural markup, and other characteristics. As this shows, the **master** branch rep-

ことにする。例えば、WYG (『四庫全書』の文淵閣、ピンインで *Wenyuan ge* 版の影印版)、あるいは CK-KZ (『重刊道藏輯要』、*Chongkan Daozang jiyao* の考正社 (Kaozheng) の影印版) である。これらのブランチ名は任意に選ぶことができるが、それらは記録のどこかに記載しておかなければならない。

- 他のブランチは記録の一部を成さないが、新しい批判校訂版、リーディング・ノート、訳、注釈などの作成といった特定の目的のために編集者か研究者によって加えられたテキストの版やバージョンを表す。もしそのような版が記録の中の版を基盤にしているならば、漢リポへよりよく統合するために、その基盤である版の参照システムが持ち越される。これらのブランチは通常、解釈的でないブランチ<sup>8</sup>と明確に区別するために、大文字と小文字の組合せ、あるいは小文字のみで名付けられる (例えば **WYG-chris** や **trans-en**)
- 記録の中の版の一つが「ベース版」として選ばれる。このベース版のページナンバーやライン・マーカはこのテキストのすべての版の参照用枠組みとして使われる。これは文献学的に最も価値のある版である必要はなく、むしろ「記録」中の他の版に参照用枠組みを提供するための実際的な選択となる。これはすべての版にわたって異なる版における位置を相互参照することを可能とする。すべての版がそれぞれのファイルの始めに情報の一部としてそのベース版が何かを表示している。
- 他のブランチの中で漢籍リポジトリにおいて特別の意味を持つブランチが二つある。**master** ブランチと **\_data** ブランチである。
- **master** は漢籍リポジトリの編集者がそれと決めた最新のテキストを表すブランチである。最新テキストは最もアクティブな編集作業の焦点となるテキストである。これはベース版のテキストかもしれないし、または、そちらの方が学術的な編集のよりよいスターティングポイントになると思われるのであれば、それとは異なるテキストであるかもしれない。それはまた通常、(異なった) 句読点、文字の正規



resents an **interpretation** of the text, not a record, hence the lowercase letters used for its name.

- `_data` is a branch that holds data used by the system. These data might pertain to one or more of the other branches. For example, information about which facsimile page belongs to which text location is maintained here.
- Within every branch of the repository, there will be a file called “Readme.org”. This file is **not** part of the textual content itself, but gives the minimum necessary information about the record (metadata), most notably the specific edition that is described in the branch in which the file `Readme.org` is located. For this reason, the content of this file will vary across different editions. By convention, the `Readme.org` file of the `master` branch, at least, will contain information about all the other branches as well to serve as a quick reference. More detailed information about the text will also be found in the Catalog.

### Texts in the Kanseki Repository

One of the distinguishing features of the Kanseki Repository compared to many other projects is that it primarily attempts to **document** as many editions of a text as feasible. It is therefore not based on one specific edition by default. For every edition, a transcribed version of the edition and a digital facsimile can be kept and these can be read side by side.

It should also be mentioned here that sometimes texts with the same title and author/editor can still be quite different. An example of this can be seen in the 雲笈七籤 *Yunji qiqian* (KR5d0055), which has 120 卷 *juan* in the Daozang, but only 40 *juan* in the Daozang jiyao. Should these be considered to be the same text, or rather different ones? Good arguments can be put for both points of view. Currently, the Kanseki Repository considers them as versions of the same text and will put them side by side as branches in the same repository, since this arrangement facilitates research on the differences of these versions.

In addition to documenting historical editions, many

化, 構造的マークアップ, その他の特徴を導入する。このことが示すように, `master` ブランチはテキストの記録ではなく, その解釈を表している。それゆえに, その名前には小文字が使われている。

- `_data` はシステムが使用するデータを保持するブランチである。これらのデータは他のブランチと関係しているかもしれない。例えば, どのファクシミリページがどのテキスト・ロケーションに属しているかの情報はここに保持されている。
- リポジトリの各ブランチの中に「`Readme.org`」というファイルがある。このファイルはテキスト内容そのものの一部ではないが, 記録 (メタデータ) についての必要最低限の情報, とりわけそのファイル (`Readme.org`) の置かれているブランチに記述されている特定の版についての情報を与えてくれる。この理由から, このファイルの内容はさまざまな版のそれぞれで異なる。慣例では, `master` ブランチの `Readme.org` ファイルは, 早見表としても使えるように, 他のすべてのブランチの情報を必ず保有している。テキストについてのより詳細な情報は目録にも見ることができる。

### 漢籍リポジトリにおける「テキスト」

他の多くのプロジェクトと比べ, 漢籍リポジトリの顕著な特徴の一つは, それがテキストのできるだけ多くの版を記録することを第一に努めていることである。したがって, それはデフォルトとして特定の版を基盤にしていない。それぞれの版ごとに, その版の翻刻されたバージョンとデジタルファクシミリが保持され, それらは並べて読むことができる。

時として, 同じタイトルで同じ著者 / 編集者のテキストであっても, かなりの違いがあることもここで述べておく。この例は『雲笈七籤』(KR5d0055)に見ることができる。これは『正統道藏』では120巻あるのに対し, 『道藏輯要』では40巻しかない。これらは同じテキストであるとみなすべきだろうか, あるいはむしろ異なるものとみなすべきだろうか? 双方の見解とも, すぐれた論拠を持っている。目下のところ, 漢籍リポジトリはそれらと同じテキストのバージョンとみなしており, 同じリポジトリの中でブランチとして並べて置いてある。このような配置がこれらのバージョンの違いについての研究を容易にするからである。

歴史的な版をドキュメントすることに加え, 例えば現

additional editions of a text can be created in the Kanseki Repository, for example by collating existing editions, and adding punctuation, further markup, annotations, translations, etc.

## Catalog

The catalog is used to maintain information about the texts. This is held in a format similar to the texts and distributed in the same way. Users of the repository can maintain their own versions of the catalog to accommodate specific analytic needs.

At the time of this writing, the Catalog can be accessed at [13] but is not yet integrated into the system.

## Notes

- 1 The use of the term “digital facsimile” might need some explication here. Due to the malleability of digital entities, a digital image is in many senses not “made similar” to the text-bearing object that served as its source. While there are some ways to make up for this, for example including dimensions and a color bar for calibration in photographs of documents, wherever software is involved there is always the possibility of unaccounted change. Some curators of digital text images therefore prefer the term “digital avatar”, to emphasize the fact that they represent recreated resemblances in the digital realm rather than simple reproductions. In the current context, however, I have chosen the term digital facsimile because I think it is less confusing and still sufficiently precise.
- 2 It goes without saying that for archival storage, the facsimiles should be kept in as high a resolution as possible and retain as much ancillary information as possible. A raw version as captured by the input device should also be included, as well as a record of how the image was reproduced. (A detailed description of how to digitize a text is beyond the scope of this introduction).
- 3 A detailed description of this format is available in the appendix. There is also a description online, which will be updated as necessary at [20]
- 4 Where exactly this line should be drawn is also up to the user, of course, and the text format used here does provide for simple annotations.
- 5 When I first arrived at the CBETA (Chinese Buddhist Electronic Text Association) project, everything was done using a plain text format. Over time, I injected a certain amount of XML into the workflow. However, there are still widely used plain text renditions in circu-

存版を照合することによって、また句読点、さらなるマークアップ、注解、訳などを追加することによって、漢籍リポジトリの中でテキストの追加版をたくさん作ることができる。

## 目録

目録はテキストについての情報を維持するのに使われる。これはテキストに似た形式で保持され、同じ方法で配布される。リポジトリのユーザーは特定の分析的な必要に合わせて自らの目録バージョンを維持することができる。

これを書いている今現在、目録は [13] においてアクセスできるが、まだシステムに統合されていない。

## 注

- 1 「デジタルファクシミリ」という用語の使用についてここでいくらかの説明が必要かもしれない。デジタル的な要素の持つ柔軟性ゆえに、デジタル画像は、多くの点において、その原本であるテキストが記載されている物と「似て作る」ことがなされていない。これを補う方法は、例えば、ドキュメントの写真中のキャリブレーションに寸法やカラーバーを含めるなどいくつかあるが、それでもソフトウェアを使う時にはいつでも、説明のつかない変更の可能性が常にある。それゆえに、デジタルテキスト画像のキュレーターの中には、「デジタルアバター」という用語の方を好む者もいる。単なる複製ではなくデジタル世界で再現された類似品を表していることを強調するためにである。しかしながら、このコンテキストにおいて、私はデジタルファクシミリの方を使うことにした。より混乱を招きにくいし、十分に正確だと考えるからである。
- 2 言うまでもなく、ファクシミリを正しく保存するためには、それをできるだけ高解像度で保存しなければならない。できるだけ多くの補助的情報を確保しなければならない。入力デバイスが取り込んだ raw 画像も、その画像がどのように再現されたかの記録と共に含めなければならない。(テキストをデジタル化する方法についての詳細にわたる説明はここでの範囲外である)
- 3 この形式の詳細にわたる説明は付録に掲載した。またオンラインにも説明があり、それは必要に応じて更新される：[24]
- 4 この線を厳密にどこで引くかもまた、当然ながらユーザー次第である。そしてここで使用しているテキスト形式はシンプルな注解を付けることができる。
- 5 私が最初に CBETA (中華電子仏典協会) のプロジェクトに参入した時、すべてはプレーンテキスト形式を使って行われていた。時間をかけて、私は特定の量の XML を仕事の流れに投入していった。しかしながら、自動化され





lation, published from the XML source of the texts by automated scripts.

- 6 “repository” is used here in the specific technical sense of a folder on a computer, the content of which is subject to a version control system.
- 7 git belongs to the category of version control software used widely by software developers. There are many introductions to git, but they are mostly geared towards programmers. A useful starting point for more information about how to use git is [7].
- 8 There is yet another form of branch names, which is specific to the Buddhist texts derived from the Taishō edition. Since this edition reports other textual witnesses in text-critical footnotes, these have been used to reconstitute the texts in separate branches. It has not yet been possible to actually tie them to the original text witnesses here, however, so the naming of the branches reflects this indirect way of constituting the text: T@SONG or 大 → 宋 are both ways to indicate a so-called Song edition (there are in fact several, but the Taishō editors grouped them together into one), without actually having a reference to the source available. It is of course hoped that some day these stopgap measures can be replaced with references to the actual witnesses of these editions.

たスクリプトによってテキストの XML ソースから出版されるプレーンテキストのレンダリングもまだ広く使われている。

- 6 ここでの「リポジトリ」はコンピューター上のフォルダーという特定の技術的な意味で使われており、その内容は版管理システムによって決まる。
- 7 git はソフトウェア・デベロッパーによって広く使われている班管理システムのカテゴリーに属する。git についての入門書は数多くあるが、それらのほとんどはプログラマー向きである。git の使い方について更なる情報を得るための有用なスターティングポイントは次のとおり：[7]
- 8 さらにまた別の形のブランチ名があるが、これらは大正新脩大藏經から派生した仏教のテキストに特定されるものである。この版は本文批評の脚注の中で他のテキストでの典拠を説明しているの、これらは漢リボの別のブランチでテキストを再構成するのに使われている。しかしながら、ここで別のテキスト典拠を元々のテキスト典拠に実際に結びつけることはまだ可能ではない。それで、ブランチの呼称が、テキストを構成するこの間接的なやり方を示している。例えば、原本へのリファレンスがなない時、T@SONG と 大 → 宋 は両方とも、いわゆる宋版を示す方法である（実のところ宋版とみなされている版はいくつかあるが、大正蔵の編集者たちはそれらを一つの版として分類した）。当然ながら、やがてはこのような当座しのぎの手段が宋版の実際の典拠のリファレンスに取って代わることが望まれている。

## 2 Sources and organisation of the Kanseki Repository

### 漢籍リポジトリの情報源と編成

#### Overview

This chapter will give an overview of the contents of the Kanseki Repository. It will first give some background about earlier phases of the development and then proceed to discuss the layout of the current iteration.

One of the most frequently raised questions about the repository concerns the source of the texts. The short answer is that about half of the texts available here have been collected from the Internet over a period of more than 20 years, while the other half is from projects I have been involved with over the past 25 years. All texts have been extensively curated, i.e., proofread, edited, linked to a digital facsimile in many cases, and converted to the format required by the repository. The digital facsimiles, where available, have been separately acquired and subsequently linked to the text. This process is still ongoing and the results have not yet been thoroughly proofread.

The representation of characters that cannot yet be represented in Unicode has been standardized and the process of finding substitute characters for a normalized version is still ongoing<sup>1</sup>.

As can be seen from this, the curating of the texts in the Repository is an ongoing, open-ended process. In fact, the texts are being moved to a public repository in their current state not so much because the curation process has reached a major milestone, but rather to enable wider participation in the process. Editing even a single text is a major undertaking that requires much time and effort. Doing this on the scale of ten thousand texts is impossible for any one individual or even a single institution. Hopefully, releasing the repository even in its current state will encourage more contributions and in time lead to the inclusion of

#### 概要

この章では漢籍リポジトリの内容についてその概要を紹介する。まずはじめに開発の初期段階についていくらか背景を述べ、その後で現在のバージョンの編成についての議論に進む。

リポジトリについて最もよく受ける質問の一つがテキストの出所に関するものである。これに手短かに答えると、リポジトリで提供しているテキストの約半分は20年以上の期間にわたりインターネットから集めたもので、残りの半分は過去25年にわたり私が関わってきたプロジェクトからのものである。テキストはすべて大幅に整理されている。すなわち校正され、編集され、多くの場合デジタルファクシミリにリンクされ、またリポジトリで必要とされるフォーマットに変換されている。デジタルファクシミリが提供されている場合、それはテキストとは別に入手され、その後でテキストにリンクされたものである。このプロセスはまだ進行中で、その結果はまだしっかりと校正されていないところもある。

ユニコードで（今はまだ）表すことができない外字の表現は統一化されており、また正規化したバージョンのための代用文字表は作成中である<sup>1</sup>。

ここからわかるように、レポジトリのテキストを整理することは、継続して行われるもので期限のないプロセスである。実のところ、テキストはその現状のままパブリックリポジトリに移動されている。それは整理のプロセスが一段落したからというよりむしろ、そうすることでこのプロセスにおいてより幅広い参加が可能になるからである。一つのテキストを編集することでさえ、多くの時間と労力を必要とする大仕事である。それを万単位 of テキストで行うのはどんな個人にも、あるいは一機関でさえ不可能なことである。それで、たとえ現状のままであってもこのレポジトリを公開することで、願わくばより多くの参与を促し、やがては重要なテキストすべて



all important texts with the relevant editions in a well curated form. Even if that day is years or, more likely, decades away, the work done here will hopefully serve as a foundation and building block for achieving that worthy goal.

## Phases of development

### (1) 1989 - 1996: Early history - mostly individual texts

The Kanseki Repository as it stands today started out as a personal collection of digital text files, created and maintained as a by-product of my research in Sinology and more specifically the Chinese sources of Chan/Zen Buddhism, through the collection of resources for facilitating my work.

The first text to enter the archive was the subject of my MA thesis, the 300+ poems attributed to 王梵志 *Wang Fanzhi* (currently text No. KR6s0055<sup>2</sup>), which I input by myself in the autumn of 1989. This was a relatively short text that could be entered in just a few weeks (including the time to create characters not available to the system). The next text, the collection 全唐詩 *Quan Tang shi* (KR4h0140), was more substantial, however, consisting of 900 juan. It was handed to me by a colleague and friend I came to know through a “Chinese Computing” workshop in Bremen in 1990. The stage was then set for the arrival of more texts, such as the 25 histories (many of which can be found in KR2a 正史類 *Zhengshi lei* section). At this point, the repository had already exceeded the capacity of my computer’s hard disk, forcing me to rely on an external drive. By this time I was doing research for my doctoral dissertation in Kyoto. I was attached to the International Research Institute for Zen Buddhism (IRIZ), where I worked on the Zen Knowledgebase project lead by Urs App. We produced a number of digital texts for this project that were subsequently released on CD-ROM and on the Internet in June 1995. These texts are now in section KR6q 禪宗類 *Chanjong lei* of the repository.

を関連した版と共によりしっかりと整理された形で含むようになりポジトリに繋がることを期待する。たとえその日が何年も先、あるいはおそらく何十年も先であったとしても、ここでなされた仕事はその価値あるゴールを達成するための土台や基本要素としての役割を果たすことを願っている。

## 開発の諸段階

### (1) 1989-1996: 初期の歴史—ほとんどが個人的なテキスト

現行の漢籍リポジトリは、私の中国学研究（より具体的に言えば禅仏教の中国語の原典の研究）の副産物として、私が仕事を容易にするために文献を作成し維持したデジタルテキストファイルの個人的なコレクションから始まったものである。

最初にアーカイブに入ったテキストは私の修士論文のテーマであり、1989年の秋に私が入力した王梵志による300以上の詩（現在のテキスト番号 KR6s0055<sup>2</sup>）であった。これは比較的短いテキストで、（システムで対応できない文字を作る時間も含めて）ほんの数週間で入力することができた。しかしその次のテキスト、『全唐詩』（KR4h0140）はより内容があり、900巻から成るものだった。これは1990年にブレーメンで行われた「Chinese Computing」ワークショップで知り合った友人でもある同僚から頼まれたものだった。これによって（その多くが KR2a 正史類に収められている）二十五史といったより多くのテキストの到来の準備ができた。この時点でレポジトリはすでに私のコンピューターのハードディスク容量を超えており、外付けドライブに頼らなければならなかった。私は、その頃にはすでに、京都で博士論文のための研究を行っており、また国際禅学研究所 (IRIZ) に所属して、そこでウルス・アップ率いる禅知識ベースプロジェクトに取り組んでいた。私たちはこのプロジェクトで多くのデジタルテキストを作成し、それらはその後1995年6月に CD-ROM とインターネット上で公開された。これらのテキストは現在、リポジトリの KR6q 禪宗類セクションに収められている。

**(2) 1996 - 2001: Buddhist canonical texts**

After the submission of my PhD thesis, I was given a CD-ROM of the Korean Tripiṭaka<sup>3</sup> and spent a full summer analyzing its content and converting it to a readable format. The WWW Database of Buddhist Texts [22] was one of the fruits of this work. This database later proved useful for the work of the Chinese Buddhist Electronic Text Association (CBETA), to which I contributed from February 1998 (until 2001 as resident and later as a consultant). Nearly 50% of the content of the Kanseki Repository originates from the texts produced by the CBETA team. The collaboration is ongoing and additions or changes in the CBETA texts are systematically carried over to the Kanseki Repository.

**(3) 2001 - 2012**

With my move in 2001 to Kyoto University and the Documentation and Information Center for Chinese Studies (DICCS) (since 2009 the Center for Informatics and East Asian Studies (CIEAS)) of the Institute for Research in Humanities, the range of texts for investigation and curation was broadened further. Several research projects were instrumental here:

**Toward an Overall Inheritance and Development of Kanji Culture (2003-2008)**

As part of this project, which under the leadership of 高田時雄 Takata Tokio was selected as a Center of Excellence (COE) project by the Japan Society for the Promotion of Science (JSPS), enabling a particular focus on historical texts that dealt with events during the Tang dynasty. This resulted in extensive annotation of sections of the 資治通鑑 *Zizhi tongjian* (KR2b0007), but work has also been done on the two official histories of the Tang, among other things<sup>4</sup>.

**Ceremony and Punishment in East Asia (2006-2010)**

This project, which was also sponsored by the JSPS (project leader 富谷至 Tomiya Itaru, project #18102003) resulted in a few texts related to institutional history, especially 唐六典 *Tang liudian*

**(2) 1996-2001: 仏教経典**

博士論文を提出した後、私は高麗大藏經<sup>3</sup>のCD-ROMを受け取り、ひと夏かけてその内容を分析し、それを読みやすい形式に変換した。仏教テキストのWWWデータベース [22] はこの仕事の成果の一つである。このデータベースは後に、私が1998年2月から(2001年までは常駐として、またその後はコンサルタントとして) 参与している中華電子仏典協会 (CBETA) の仕事の役に立つことが判明した。漢籍リポジトリの内容の半分近くはCBETA チームが作成したテキストからきている。この共同作業は現在も進行しており、CBETA テキストへの追加や変更は体系的に漢籍リポジトリにも反映されている。

**(3) 2001-2012:**

2001年に私は京都大学の漢字情報研究センター (DICCS) (2009年より東アジア人文情報学研究センター CIEAS) へ移動し、それにともなって、その研究範囲は広がり、これまでとは異なるテキストも調査とキュレーションのための対象となった。漢籍リポジトリに貢献したいいくつかの研究プロジェクトをここに挙げる:

**漢字文化の全き継承と発展のために (2003-2008)**

高田時雄率いるこのプロジェクトは日本学術振興会 (JSPS) のセンター・オブ・エクセレンス (COE) プロジェクトに選ばれ、の一部は唐代の出来事を扱う歴史的テキストに特別な焦点を置くことを可能にした。その結果が『資治通鑑』 (KR2b0007) の唐紀の膨大な注解であるが、他の仕事に加えて、唐の二つの正史についての仕事もまた行われている<sup>4</sup>。

**東アジアにおける儀礼と刑罰 (2006-2010)**

これもまた JSPS の科学研究所助成金を受けたプロジェクト (研究代表: 富谷至, プロジェクト番号: #18102003) で、制度史に関するいくつかのテキスト、とりわけ『唐六典』 (KR2i0001) を整理し、また方法と手順の更なる



(KR2I0001). and facilitated the further development of methods and procedures.

#### 道藏輯要 Daozang jiyao project (since 2005)

This is a major ongoing project aimed at analyzing the early Qing compilation of the Daoist texts 道藏輯要 *Daozang jiyao* (DZJY). It was started and initially lead by the late Monica Esposito (1962–2011), with active contributions from Mugitani Kunio 麦谷邦夫 and others; it was sponsored by the JSPS (project #20242001, 2008–2011) and two grants from the Chiang Ching-kuo Foundation for Scholarly Exchange (projects #RG003-P-05 (2006–2009) and #RG006-P-09 (2010–2013)). As part of this project, the entire text of the DZJY was input and proofread. In addition, a digital version of the text of the 正統道藏 Zhengtong Daozang (ZTDZ) was acquired, and a set of high-quality digital images was integrated to bring the collection to its present form. We also received another digital version, curated and proofread from the 道教學術資訊網站 *Daojiao xueshu zixun wang zhan* ([26]), but this is not yet thoroughly incorporated into the repository.

#### Chan database project (since 2008)

This project was conceived as a collaboration between the late John McRae (1947–2011), Christoph Anderl, and Christian Wittern. We have not yet received any substantial funding, but we exchanged data (within the scope of the project) and prepared an enhanced version of the 祖堂集 *Zutang ji* (KR6q0002), as well as of some editions of the Chan texts 景德傳燈錄 *Jingde chuandeng lu* (KR6q0003) and 五燈會元 *Wudeng huiyuan* (KR6q0012). A part of this work was made possible with funds from CIEAS.

#### (4) 2013 – 2016 Fundamental topics in Digital Humanities

Several times during the above mentioned phases, efforts were made to create an overarching framework and common format to better enable comprehensive research on these texts. After 2010, the germ of an idea had taken shape and I started running some tests on how such a framework could be realized. The essential need is a systematic arrangement of the texts

開発を促した。

#### 『道藏輯要』プロジェクト（2005年より）

これは現在も進行中の大きなプロジェクトで、道教文献の清代初期に編纂された叢書『道藏輯要』を分析することを目的としている。これは麦谷邦夫からの積極的な支持を得ながら故モニカ・エスポジト（1962–2011）が開始し、また当初は指揮したプロジェクトで、JSPSの助成を受け（プロジェクト番号：#20242001, 2008–2011）、また台湾蔣經國基金からも二つの助成を受けた（プロジェクト番号：#RG003-P-05（2006–2009）と#RG006-P-09（2010–2013））。このプロジェクトの一環として『道藏輯要』のすべてのテキストが入力され校正された。これに加えて、『正統道藏』のテキストのデジタルバージョンが取得され、このコレクションを現在の形に持っていくために高品質のデジタル画像一式が組み込まれた。私たちはまた『道教學術資訊網站』（[26]）から整理され校正された別のデジタルバージョンを受け取ったが、これはまだリポジトリに完全に組み込まれていない。

#### 禅データベースプロジェクト（2008年より）

このプロジェクトは故ジョン・マクレイ（1947–2011）とクリストフ・アンデル、クリスティアン・ウィッテルンとの共同研究として始められた。私たちはまだしっかりとした財政的助成を得ていないが、（このプロジェクトの範囲内で）データを交換し、『景德傳燈錄』（KR6q0003）や『五燈會元』（KR6q0012）などいくつかの禅宗のテキスト、および『祖堂集』（KR6q0002）の強化バージョンを準備した。この仕事の一部は東アジア人文情報学研究センターからの助成によって可能となった。

#### (4) 2013–2016: 人文情報学の基礎研究

上記の諸段階において、テキストの総合的な研究をより行いやすくするために、すべてを包含する枠組みや共通の形式を作る努力が幾度かなされた。2010年以降、ある考えの芽生えが形をなして、私はそのような枠組みをどのようにしたら実現できるかのテストを始めた。道教（KR5）や仏教（KR6）のテキスト用追加カテゴリーを補強し文献の古典的な四部分類に基づいた、共通した、

in a common, but very simple, format, based on the classical four sections of literature, but enhanced with additional categories for Daoist (KR5) and Buddhist (KR6) texts, as well as an analytical catalog and tools for efficiently interacting with the text corpus.

In April 2013, a research team was formed and began to meet regularly [25] to discuss methodological and practical issues related to this work, to ensure that the results would have the broadest possible application. All this has led to the current collection of around 10,000 texts. In addition to the texts and collections already mentioned, these are drawn mainly from the following collections:

- ・四庫全書 *Siku quanshu*
- ・四部叢刊 *Sibu congkan*
- ・維基文庫 *Wikisource.zh*

#### 四庫全書 *Siku quanshu* (SKQS)

The Siku Quanshu with more than 3400 texts is the greatest single collection in the Kanseki Repository after the CBETA collection (which has roughly 1,100 texts more). Unlike the CBETA collection, however, the SKQS has not been acquired as a single entity, but rather pieced together over time. It may be useful to recount how this was done and describe the current state of the collection, which is still far from complete.

In the early years of the 21st century, more and more pre-modern Chinese texts became available online. In early 2007 I discovered a potential treasure trove of works<sup>5</sup>, which inspired me to attempt a systematic collation of the texts that belong to the SKQS.

Some years later, in addition to the transcribed texts, digital facsimiles became available at the Internet Archive. These turned out to be from the 文瀾閣 *Wenlan ge* copy of the SKQS, located today at the Zhejiang Library ( 浙江圖書館 ) in Hangzhou. Unfortunately, this edition, which was originally held in the Wenlan Pavillon on Gushan Island of the West Lake in Hangzhou, had been scattered during the years of the Taiping wars and only later in the 19th century, through the efforts of the brothers 丁丙 Ding Bing (1829-1887) and 丁申 Ding Shen (1832-1899) could a part of this collection be restored and moved to the

しかもとてもシンプルな形式でのテキストを系統的に配置すること、および全文のテキストと効率的にやりとりするための分析的なカタログとツールを持つことが不可欠であった。

2013年4月、「人文情報学の基礎研究」研究班が結成され、最大限に幅広い応用を持つ成果が得られるように、この仕事に関連した方法論的および実際的な問題を話し合うための定期的なミーティングが持たれ始めた [25]。こうした努力のすべてが約一万のテキストを持つ現在のコレクションをもたらした。これらのテキストは、すでに言及したテキストやコレクションに加え、主に以下のコレクションから取られたものである：

- ・四庫全書 *Siku quanshu*
- ・四部叢刊 *Sibu congkan*
- ・維基文庫 *Wikisource.zh*

#### 『四庫全書』

3,400以上のテキストから成る『四庫全書』は、漢籍リポジトリの中では（それより1,100ほど多いテキストを持つ）CBETA コレクションに次ぐ一大叢書である。しかし、CBETA コレクションとは異なり、『四庫全書』は一つの実体として獲得されておらず、むしろ長い期間の間に継ぎ合わされたものである。どのようにそれが行われたかを振り返り、完成には程遠いこのコレクションの現状を説明することは有用だと思われる。

21世紀の初頭にはより多くの漢籍がオンラインで入手できるようになった。2007年の初め、私は仕事における貴重な宝となり得るものを見つけ<sup>5</sup>、それに励みを得て『四庫全書』に属するテキストの系統的な照合を試みた。

その何年か後、翻刻されたテキストに加えて、デジタルファクシミリもインターネット・アーカイブで提供されるようになった。これは今日、杭州市の浙江図書館に収蔵されている『四庫全書』の文瀾閣版からのものであった。残念ながら、もともと杭州市西湖の孤山にある文瀾閣に収蔵されていたこの版は、太平天国の乱の期間に離散してしまい、19世紀の最後になって初めて再発見された。丁丙 (1829-1887) と丁申 (1832-1899) 兄弟の努力の成果で、この文瀾閣版の一部が修復され、省の図書館に移された。この原本そのものの不完全な状態に加えて、デジタルファクシミリが目録に載せられインターネット・アーカイブで提供される方法が、それらを順序立



Zhejiang Library. In addition to the problematic state of the source itself, the way the digital facsimiles are cataloged and made available in the Internet Archive makes it very difficult to arrange them in order, align them to the transcribed text, or assess their completeness. It was fortunate, therefore, that in early 2015 the opportunity arose to purchase a complete set of digital facsimiles of the 文淵閣 *Wenyuan ge* edition. These were scanned from a reprint by the Taiwan branch of Commercial Press (商務印書館 *Shangwu yinshuguan*), the same source used as the reference for the curation of the transcribed versions of the texts. These facsimiles made it much easier to produce a version that aligned the textual representation and the digital facsimiles. At the time of this writing, this work has been completed for about 2900 texts, but since the easiest of the texts were prepared first, the remaining texts will require considerably more effort in editing and alignment.

#### 四部叢刊 *Sibu congkan* (SBCK)

Since the texts in the SKQS are sometimes heavily edited and not always the best possible source, where possible texts have also been aligned to the SBCK. This is a much smaller collection of fewer than 400 texts, of which some 100 are not in the SKQS, thereby providing additional material. Again, digital facsimiles could be purchased, but the work to establish the text and align it to the facsimiles is time consuming and ongoing. Thus, only very few texts have so far been processed.

#### 維基文庫 *Wikisource.zh*

This is the Chinese version of the Wikisource website, which has a few hundred texts. While these texts do not indicate the edition used, they come in modern, well punctuated editions and can thus be used in addition to the sources mentioned, especially in the **master** branch.

#### Current state

At the time of this writing, almost 9,000 or 90% of the texts in the catalog are released and freely available

てて配置し、翻刻テキストと並べ、あるいはそれらの完成度を評価することをとても困難にした。それゆえに、2015年の初めに文淵閣版のデジタルファクシミリの全セットを購入する機会が訪れたことはとても幸運であった。これらは商務印書館の台湾支社の影印版からスキャンしたもので、テキストの翻刻と整理もこの版をリファレンスとして使ったので、原本と同じものである。これらのファクシミリによって、テキストでの表現とデジタルファクシミリを並べたバージョンを作成する作業がより簡単になった。これを書いている今現在、この作業において約2,900のテキストが完成されている。しかし、比較的簡単なテキストから作業が手がけられたので、残りのテキストは編集と配列にかなり多くの労力が必要とされる。

#### 『四部叢刊』

『四庫全書』の中のテキストは時として大幅に編集されており常に最善の原本であるとは限らないので、それが可能な場合、テキストは『四部叢刊』とも並べてある。これは400足らずのテキストを持ったずっと小さなコレクションで、そのうちの100ほどは『四庫全書』に入っておらず、そのため追加資料を提供している。これもまた、そのデジタルファクシミリを購入することができたが、テキストを確立しそれをファクシミリと並べる仕事は時間がかかり、また継続して行われるものである。それで、これまでのところ、ほんのいくつかのテキストが処理されただけである。

#### 維基文庫 *Wikisource.zh*

これはウィキソースの中国語バージョンであり、数百のテキストを持っている。これらのテキストは使用した版を明示していないが、その一方で現代的でしっかりと句読点の付いた版であり、それゆえに、先に言及した原本に加えて、とりわけマスターブランチで使うことができる。

#### 現在の状況

現時点、漢籍リポジトリの目録中のほぼ9,000（あるいはほぼ90%）のテキストが公開され、GitHub ウェブサ

in the @kanripo account of the GitHub website [12]. Most of these texts are not yet vetted according to the strict philological methodology outlined above. They are placed here partly in the hope that they will nevertheless prove useful as they are, but more importantly because the software interfaces described in the next chapter will allow users to work independently on texts of interest, without interference from the editors of the Kanseki Repository.

## Content of the Kanseki Repository

The content in the Kanseki Repository is organized similarly to a traditional Chinese library, using the four divisions of literature first used to organize the imperial library, which later served as the principle for classification of other collections as well. Since this traditional classification scheme has an inbuilt bias against Daoist and Buddhist texts, which play an important role in the Kanseki Repository, these two sections have been elevated to stand alone as top-level divisions, thus increasing the overall number of top-level divisions to six.

The bibliographic treatise 隋書 *Sui shu* KR2a0023, compiled under the leadership of 魏徵 *Wei Zheng* (581–643), was also organized in this way at a time when the enormous production of Buddhist and Daoist texts pushed the limits of the traditional classification scheme. These two sections were later moved to the end of the 子部 *Zi bu*, but in most cases only a very small selection of Buddhist and Daoist texts were ever deemed worthy of admission to the libraries. The current selection thus revives this classification scheme of the early Tang, as shown in Table 2.1.

Table 2.1: The six top categories in the Kanseki Repository  
表 2.1 漢籍リポジトリの六部分類

KR1	經部 <i>Jing bu</i>	Confucian Classics (incl. music, dictionaries and elementary learning)
KR2	史部 <i>Shi bu</i>	Historiography and politics
KR3	子部 <i>Zi bu</i>	Masters, philosophers and treatises
KR4	集部 <i>Ji bu</i>	Anthologies (Poetry and Collected Writings)
KR5	道部 <i>Dao bu</i>	Daoist texts
KR6	佛部 <i>Fo bu</i>	Buddhist texts

The appropriate place in this classification has to be determined for every text in the Kanseki Repository.

イトの @kanripo アカウントで自由に閲覧できる [12]。これらのテキストのほとんどは、上で概要を述べた厳密な文献学的方法論に従っては未だ処理されていない。こうしたテキストを公開したのは、一つにはそれらがそのままの状態でも役立つことを立証したいということもあるが、それより何より、次の章で説明するソフトウェア・インターフェースがユーザーが漢籍リポジトリの編集者からの干渉を受けることなく興味のあるテキストについて独立して仕事することを可能にするからである。

## 漢籍リポジトリの内容

漢籍リポジトリの内容は、初め中国の宮廷の蔵書所を整理するのに使われ、その後は他のコレクションの分類の原理ともなった文献の四部分類を使った伝統的な中国の図書館に似せて整理されている。この伝統的な分類表は、漢籍リポジトリにおいて重要な役割を持つ道教と仏教の経典に対して不利な偏りを内包しているため、これら二つのセクションは独立させて、トップレベルの区分に配置し、それゆえにトップレベルの区分の総数は六つに増やしてある。

魏徵 (581–643) の指揮の元で編纂された『隋書』(KR2a0023) の経籍志もまた、膨大な量の仏教と道教の経典が新しく作成或は翻訳されたことで、伝統的な分類表の限界を破った当時、この方法で整理された。これら二つのセクションは後になって子部の終わりに移されたが、ほとんどの場合、仏教と道教のテキストはそのほんの少しのセレクションのみが図書館に収蔵するに値するとみなされた。それゆえに、表 2.1 に示すように、現在のセレクションはこの唐代初期の分類表を復活させている。

この分類における適切な場所は、漢籍リポジトリの各テキストごとに決定されなければならない。ほとんどの



In most cases, this follows the precedents set by previous text collections, with some important differences, as outlined below.

Another question concerns the exact delimitation of a text. Where are the boundaries between different texts within a collection located exactly? And there may be sets of pages that could be considered a single text or several texts. The problem can sometimes be seen in anthologies of texts, especially when an anthology is itself part of a larger anthology. The ZTDZ, for example, contains the work 修真十書 *Xiuzhen shishu* “Ten texts of practices for obtaining the way”, an anthology of 10 texts of very different length and widely varying characteristics, written by different authors in different times. Some catalogs consider it to be one text, thus assigning it a single entry. Since research may require a focus on only a subset of these texts, for the purposes of the Kanseki Repository these texts have been considered separately, with each assigned a different number (KR5a0264 to KR5a0275, making, in fact, 12 entries). This decision was based on the notion that it is always desirable for a catalog to be as fine-grained as possible. On the other hand, this also means that the *Xiuzhen shishu* does not exist as a single textual entity in the Kanseki Repository.

### The four divisions of Chinese literature

As explained, the first four top-level divisions of the Kanseki Repository correspond to the divisions that are commonly used in many libraries and collections, such as the 四庫全書 *Siku quanshu* (SKQS) and the 四部叢刊 *Sibu congkan* (SBCK). When it comes to further divisions below this top level, there is considerable more variation, though. While the SBCK does not have any further subdivisions, the SKQS has up to two more levels — the 類 *lei* and the 屬 *shu*. For example, the third second-level class is 禮類 *Li lei*, the ‘Ritual Books,’ which is further subdivided into the 周禮之屬 *Zhouli zhi shu*, the ‘Rites of Zhou,’ and five more third-level classes. Other second-level entries, like the 易類 *Yi lei*, 書類 *Shu lei*, and 詩類 *Shi lei*, do not have any further subdivision.

In order to standardize classification depth, the hierarchy of the Kanseki Repository has been limited to

場合、その前のテキストコレクションが決めた前例に従うが、以下に概説するようにいくつかの重要な違いがある。

また別の問題はテキストの正確な範囲に関してである。すなわち、コレクション内の異なるテキスト間の境界線は正確にはどこに引くのか。また幾組かのページは一つのテキストとしてみなすべきか、あるいは複数のテキストとしてみなすべきか。この問題は時々テキストの叢書、とりわけその選集がより大きな選集の一部である時に見られる。例えば、『正統道藏』には、異なる著者によって異なる時代に書かれ、大きく異なる長さや多様な性質を持つ十の文献である『修真十書』が入っている。研究にはこれらのテキストの部分集合のみに焦点をあてる必要があるかもしれないので、漢籍リポジトリの方針として、これらのテキストは別々に扱われ、それぞれが異なる番号を割当てられている (KR5a0264から KR5a0275まで、実のところ12の項目となっている)。目録はきめ細かく分類するに越したことはないという考えに基づいてこの決定がなされている。その一方で、これはまた、『修真十書』が漢籍リポジトリの中では一つのテキストでの実体として存在しないことも意味している。

### 漢籍の四部分類

先に説明した通り、漢籍リポジトリのトップレベルの最初の四つは、『四庫全書』や『四部叢刊』など、多くの図書館やコレクションで一般的に使われている分類に相当する。このトップレベルの更に下の区分となると、かなり多くのバリエーションがある。『四部叢刊』は更なる下位区分を持たないが、その一方で『四庫全書』は更に2レベルまで持っている。それらは類と屬である。例えば、第二レベルクラスの三つ目は禮類で、これは更に周禮之屬「周の儀礼」とその他五つの第三レベルクラスに細分されている。易類、書類、詩類など他の第二レベルエントリーは下位区分を持たない。

分類の深度を標準化するために、漢籍リポジトリの階層はただ一つの下位区分のみに制限されている。それゆ

just one subdivision. Thus, all books in the Rites section have been grouped at the same level. Initially, the sequential order of the subdivisions on the 屬 *shu* level is preserved, but as the repository expands and more texts are added in the future, this order will become distorted, because newly added texts will always have to be added to the end of the relevant class.

In most cases, the divisions on the 類 *lei* level are as seen in the SKQS. The only exception to this is the 別集類 *Bieji lei*, “Collections of individual authors”. Since there is such a large number of texts in this category, the subdivisions, which here reflect the time of the text creation by dynastic period as they are found in the SKQS, have been moved up to the 類 *lei* level as shown in Tables 2.2 and 2.3.

As for the arrangement of the texts within a class, our basic guideline is that texts are grouped by topic and usually by date or assumed date of creation, in such a way that related topics can be found together. In cases where several topical groups exist within one class, as in the example of the 禮類 *Li lei* (ritual books) above, the groupings have been preserved so that related texts still occur together. There is no overarching chronological arrangement, since the original chronological arrangement was within the subdivision, and this order gets preserved.

Furthermore, root texts are usually placed before commentaries and treatises that deal with them. Since most non-Buddhist texts in the Chinese tradition have been transmitted intermingled with commentaries or groups of commentaries, for some important texts with many commentaries a separate modern version containing only the root text has been created. Examples include the 周易 *Zhou yi* (KR1a0001) ‘Book of Changes’ and the 尚書 *Shang shu* (KR1b0001) ‘Book of Documents’. In these cases, the characters 正文 *Zhen-gwen* ‘root text’ have been added to the title. This has been harmonized in the Kanseki Repository as shown in Table 2.3.

Amongst other changes that have been necessary, all texts of the sections of the 釋家類 *Shijia lei* and the 道家類 *Daojia lei* in the SKQS have been moved to the corresponding sections of the newly created 佛部 *Fo bu* and 道部 *Dao bu*, as detailed in the tables 2.4 and

えに、儀式セクションの本はすべて同じレベルとして分類されている。最初のうち、属レベルにある下位区分の一連の序列は維持されるが、今後リポジトリが拡大し、より多くのテキストが加えられるにつれ、新たに加えられるテキストは常に対応するクラスの最後に加えられるので、この序列がゆがめられることになる。

ほとんどの場合、類レベルにある下位区分は『四庫全書』に見られる通りである。唯一の例外は、別集類「個別著者のコレクション」である。このカテゴリーには膨大な数のテキストがあるので、表 2. 2 と表 2. 3 に示すように、下位区分（『四庫全書』に見られるように、王朝時代ごとに分けられたテキストの作成時期を表している）は類レベルまで繰り上げられている。

一つのクラス内のテキストの配列については、関連したトピックが一緒に見つけられるようなやり方で、つまりテキストをトピックごとに、それから通常は作成日、あるいは想定作成日ごとに分類するというのが私たちの基本的なガイドラインである。上記の禮類「儀式の本」の例のように、一つのクラス内に複数のトピックのグループが存在する場合、それでも関連したテキストが一緒に出てくるように、その分類は維持されている。すべてを包含する年代順別の配列はなされていないが、元々の年代順別の配列が下位区分内にあり、この順序は維持される。

更に、原文は通常、それらを扱う注釈や論文より前に配置される。中国の伝統では仏教以外のテキストのほとんどが注釈や注釈のグループと入り混じって伝えられてきたので、多くの注釈を持つ重要なテキストについては、原文のみを含めた別の現代バージョンが作成されている。この例として周易 (KR1a0001) や尚書 (KR1b0001) が挙げられる。このような場合、「正文」(原文) という文字が題名に加えられている。表 2. 3 に示すように、漢籍レポジトリではこれは正規化されている。

その他にも必要とされていた変更がある。四庫全書の中の釋家類と道家類のセクションにあるすべてのテキストが新たに作られた佛部と道部の中にある対応するセクションへと移された。これは表 2. 4 と表 2. 5 に詳しく示されている。

## 2.5.

**Daoist texts**

The Daoist material is mainly compiled from the collections 正統道藏 Zhengtong daoang (ZTDZ) and 道藏輯要 Daozang jiyao (DZJY). Titles from the 中華道藏 Zhonghua daoang (ZHDZ) that are drawn from the Dunhuang collection have also been added, but these texts are not yet in the repository.

In arranging the material in this section, the seven-part division that can be seen in the ZTDZ of the early Ming has been preserved (corresponding to subsections KR5a to KR5g). This is followed by the *Supplement to the Daoist Canon*, which was added to the Daozang towards the end of the Ming dynasty. Wherever possible the texts in the DZJY have been positioned as additional editions with the corresponding texts of the ZTDZ. For the texts not in the ZTDZ, an additional subsection 清代道教文獻 *Qingdai daojiang wenxian*, “Daoist texts from the Qing dynasty” has been added. Additional Qing material not in the DZJY is placed here as well. This is followed by a subsection for Daoist texts excavated in Dunhuang, but no texts have yet been added to this subsection.

**Buddhist texts**

The Buddhist texts are mostly collected and converted from the fine collection prepared by 中華電子佛典協會 CBETA (see [3], [2]) and follow the general structure developed there. However, since CBETA does not include the parts of the 大正新脩大藏經 *Taishō shinshū daizōkyō* ‘Taishō Tripitaka’ that contain commentaries by Japanese authors or texts pertaining to Japanese Buddhist sects, these parts have been added anew into the overall classification.

**List of text witnesses**

Table 2.8 shows a list of text witnesses and their designation in the Kanseki Repository<sup>6</sup>. It is to be expected that this list will continue to grow. Details are found in the catalog KR-Catalog [13].

**道教の文献**

道教の資料は主に『正統道藏』と『道藏輯要』のコレクションから編集されている。敦煌文献から取られた『中華道藏』からのタイトルもまた加えられたが、これらのテキストはまだレポジトリに入っていない。

このセクションの資料を配列するにあたり、明代初期の『正統道藏』に見られる7区分が維持されている（下位区分 KR5a から KR5g までに相当）。これに続くのが、明代の終わり近くで道藏に加えられた「續道藏」である。『道藏輯要』の中のテキストは、それが可能な限り、追加版として位置づけられ、これに対応する『正統道藏』のテキストが付けられている。『正統道藏』にないテキストについては、追加の下位区分である「清代道教文獻」が加えられている。『道藏輯要』にない追加の清の資料もここに置かれている。これに続くのが敦煌で発掘された道教のテキストのための下位区分であるが、この下位区分にはまだ何のテキストも加えられていない。

**仏教の文献**

仏教のテキストはそのほとんどが中華電子佛典協會 CBETA ([3], [2] を参照) が準備したみごとなコレクションから集められ、変換されたもので、そこで開発された一般構造を踏襲している。しかしながら、CBETA は『大正新脩大藏經』の日本人著者による注釈や日本の仏教宗派に関する文献が入ったいくつかの部分を含めていないので、新たにこれらの部分を全体的な分類に加えた。

**テキストでの典拠のリスト**

表 2. 8 にテキストでの典拠とそれらの漢籍レポジトリ内での名称のリストを示す<sup>6</sup>。このリストが増え続けるだろうことは予想される。詳細は KR-Catalog に見ることができる [13]。

## Research

The *raison d'être* of Kanripo is to enable and facilitate any kind of research that can make use of the texts contained. The texts in this digital library can be accessed through different interfaces, including machine and software interfaces. Another purpose of Kanripo is to render the texts and the entire library programmable. For this reason, every repository needs to have a **master branch**, which is essentially the “best” text, i. e., the text that would be published as the **copy text** of a critical textual edition. A program trying to access the library in Kanripo should normally work on the **master branch**, unless there is a special interest in some other particular version or other versions in general.

### Notes

- 1 All information pertaining to these characters is available in the Kanseki Repository, in the folder KR-Gaiji.
- 2 This is the serial number within the repository, composed of the repository identifier KR, the section identifier 6s, and a number indicating the sequential position of the text in the section, in this case 55 (though the number is padded out to four digits, hence 0055).
- 3 This was the version described in [4], completed in November 1996.
- 4 The results of this project can still be accessed at [18], but integration into the Kanseki Repository is ongoing.
- 5 The link to the site, <http://203.72.198.245>, became inactive after a few months, but some parts of the site's content can still be found using the Internet Archive's Wayback Machine, at [11]. I am very grateful for this effort, but since I did not take notes of who produced these texts, unfortunately the producers have to go without proper recognition here.
- 6 This includes only the direct witnesses. The indirect witnesses derived from the text-critical footnotes of the Taishō Tripitaka have been omitted, a list of these can be found online.

## 研究

漢リポの存在意義は、それが擁するテキストを活用できるような種類の研究も、それを可能にし、容易にすることである。このデジタル図書館のテキストは、機械やソフトウェアのインターフェースを含めさまざまなインターフェースを通じてアクセスできる。漢リポのまた別の目的は、リポジトリ内の個々のテキストやその全体をプログラム可能にすることである。この理由から、それぞれのリポジトリすべてが、基本的には「ベストな」テキスト、すなわち本文批評によって得られた版を出版され得るテキストを **master ブランチ** とすべきである。漢リポのライブラリーにアクセスしようと試みるプログラムは通常、別の特定のバージョン、あるいは他のバージョン一般に特別な興味を持たないかぎり、**master ブランチ** 上で動作するはずである。

### 注

- 1 これらの文字に関するすべての情報が漢籍リポジトリの =KR-Gaiji= というフォルダーの中で提供されている。
- 2 これはリポジトリ内のシリアル番号であり、リポジトリ識別子 KR、セクション識別子 6s、そしてこの場合は 55 (番号は 4 桁に引き伸ばされ、ゆえに 0055) であるテキストのセクション内でのシーケンシャル・ポジションを示す番号で構成されている。
- 3 これが 1996 年 11 月に完了した、ここに記述されているバージョンである：[4]。
- 4 このプロジェクトの成果にはまだアクセスできる ([18]) が、漢籍リポジトリへの組み込みは現在進行中である。
- 5 このサイトへのリンク (<http://203.72.198.245>) は数ヶ月後に非アクティブになったが、サイトの内容のいくらかの部分はインターネット・アーカイブの Wayback Machine ([11]) を使ってまだ見つけることができる。
- 6 これには直接的な典拠のみ含めている。大正新脩大藏經の本文批評の脚注に由来する間接的な典拠は省略されている。(これらのリストはオンラインで見ることができる)



Table 2.2: The subdivisions of the SKQS 集部 *Ji bu*  
 表 2.2 『四庫全書』集部の類

SK4	集部	<i>Jibu</i>	Anthologies and belles-lettres
4a	楚辭類	<i>Chuci</i>	Poetry of the South
4b	別集類	<i>Bieji</i>	Collections of individual authors
4b1	漢至五代	<i>Han zhi Wudai</i>	Han to Five Dynasties periods
4b2	北宋建隆至靖康	<i>Bei Song</i>	Northern Song period (Jianlong to Jingkang reigns)
4b3	南宋建炎至德佑	<i>Nan Song</i>	Southern Song period (Jiayuan to Deyou reigns)
4b4	金至元	<i>Jin zhi Yuan</i>	Jin and Yuan empires
4b5	明洪武開至崇禎	<i>Ming</i>	Ming period (Hongwu to Chongzhen reigns)
4b6	清代	<i>Qing dai</i>	Qing period
4c	總集類	<i>Zongji</i>	Anthologies and collections of many authors
4d	詩文評類	<i>Shi-wenping</i>	Critique on poetry and prose
4e	詞曲類	<i>Ciqu</i>	Poetry and arias
4e1	詞集之屬	<i>Ciji</i>	Collected poems
4e2	詞選之屬	<i>Cixuan</i>	Selected poems
4e3	詞話之屬	<i>Cihua</i>	Explanations to poems
4e4	詞譜韻之屬	<i>Cipuyun</i>	Rhymes and musical notes to poems
4e5	南北曲之屬	<i>Nanbeiqu</i>	Arias from the North and the South

Table 2.3: The subdivisions of the Kanseki Repository 集部 *Ji bu*  
 表 2.3 漢籍リポジトリ集部の類

KR4	集部	<i>Jibu</i>	Anthologies and belles-lettres
KR4a	楚辭類	<i>Chuci</i>	Poetry of the South
KR4b	別集類-漢至六朝	<i>Bieji Han zhi Liuchao</i>	Individual collections Han to Six Dynasties
KR4c	別集類-唐	<i>Bieji Tang</i>	Individual collections Tang
KR4d	別集類-宋	<i>Bieji Song</i>	Individual collections Song
KR4e	別集類-明	<i>Bieji Ming</i>	Individual collections Ming
KR4f	別集類-清	<i>Bieji Qing</i>	Individual collections Qing
KR4g	別集類-近代	<i>Bieji Jindai</i>	Individual collections after Qing
KR4h	總集類	<i>Zongji</i>	Collective anthologies
KR4i	詩文評類	<i>Shi-wenping</i>	Critique on poetry and prose
KR4j	詞曲類	<i>Ciqu</i>	Poetry and arias

Table 2.4: Texts moved from the 釋家類 *Shijia lei* 'Texts of the Buddhists'  
 of the SKQS to the 佛部 *Fo bu* of the Kanseki Repository  
 表 2.4 『四庫全書』の釋家類から漢籍リポジトリの佛部へ移動されたテキスト

SKQS ID	KR ID	Title
SK3m2133	KR4d0007	法藏碎金錄-宋-晁迥
SK3m2134	KR4d0008	道院集要-宋-晁迥
SK3m2138	KR6q0012	五燈會元-宋-普濟
SK3m2135	KR6q0040	禪林僧寶傳-宋-惠洪
SK3m2140	KR6r0013	佛祖歷代通載-元-念常
SK3m2139	KR6r0014	釋氏稽古略-元-覺岸
SK3m2132	KR6r0054	宋高僧傳-宋-贊寧
SK3m2137	KR6r0092	羅湖野錄-宋-曉瑩
SK3m2128	KR6r0137	弘明集-梁-僧祐
SK3m2129	KR6r0138	廣弘明集-唐-道宣
SK3m2136	KR6r0157	林間錄-宋-慧洪
SK3m2130	KR6s0002	法苑珠林-唐-道世
SK3m2131	KR6s0093	開元釋教錄-唐-智昇

Table 2.5: Texts moved from the 道家類 *Daojia lei* 'Texts of the Daoists' of the Siku quanshu to the 道部 *Dao bu* of the Kanseki Repository

表 2.5 『四庫全書』の道家類から漢籍リポジトリの道部へ移動されたテキスト

SKQS ID	KR ID	Title
SK3n2141	KR5a0114	黃帝陰符經解-蹇昌辰
SK3n2142	KR5a0110	黃帝陰符經講義-夏元鼎
SK3n2143	KR5c0065	道德真經註 (一)-漢後期-河上公
SK3n2144	KR5c0078	道德真經指歸-嚴遵
SK3n2145	KR5c0073	道德真經註 (二)-魏-王弼
SK3n2146	KR5c0074	道德真經註 (三)-蘇轍
SK3n2147	KR5c0076	道德寶章-宋-葛長庚
SK3n2148	KR5c0091	道德真經註 (四)-吳澄
SK3n2149	KR5h0056	老子翼-明-焦竑
SK3n2150	KR5c0384	御定道德經註-清-世祖
SK3n2151	KR5c0385	老子說畧-清-張爾岐
SK3n2152	KR5c0077	道德經註-清-徐大椿
SK3n2153	KR5c0048	無上妙道文始真經-周-尹喜
SK3n2154	KR5c0125	釋文-周-列禦寇
SK3n2155	KR5c0121	沖虛至德真經解-江適
SK3n2156	KR5c0138	莊子注-晉-郭象
SK3n2157	KR5c0136	南華真經新傳-王雱
SK3n2158	KR5c0128	南華真經口義-林希逸
SK3n2159	KR5c0127	南華真經義海纂微-褚伯秀
SK3n2160	KR5h0057	莊子翼-明-焦竑
SK3n2161	KR5c0118	文子-周-辛鈞
SK3n2162	KR5c0119	文子續義-宋-杜道堅
SK3n2163	KR5a0306	列仙傳-劉向
SK3n2164	KR5d0019	周易參同契分章通真義-彭曉
SK3n2165	KR5d0018	周易參同契 (三)-朱熹
SK3n2166	KR5d0025	周易參同契解-陳顯微
SK3n2167	KR5d0024	周易參同契釋疑-俞琰
SK3n2168	KR5d0023	周易參同契分章註-元-陳致虛
SK3n2169	KR5d0026	古文參同契集解-明-蔣一彪
SK3n2170	KR5f0019	抱朴子內篇-葛洪
SK3n2171	KR5c0317	神仙傳-晉-葛洪
SK3n2172	KR5d0036	真誥-陶弘景
SK3n2173	KR5d0037	亢倉子-唐-王士源
SK3n2174	KR5d0052	玄真子 (外篇)-張志和
SK3n2176	KR5d0051	無能子--
SK3n2177	KR5a0307	續仙傳-南唐-沈汾
SK3n2178	KR5d0055	雲笈七籤-唐-張君房
SK3n2179	KR5a0142	紫陽真人悟真篇註疏-張伯端
SK3n2180	KR5d0013	古文龍虎經註疏-王道
SK3n2181	KR5d0028	易外別傳-宋-俞琰
SK3n2182	KR5d0029	席上腐談-宋-俞琰
SK3n2183	KR5g0239	道藏目錄詳註-明-白雲齋

Table 2.6: The subdivisions of the Kanseki Repository 道部 Dao bu

表 2.6 漢籍リポジトリ道部の類

KR5	道部	Daobu
KR5a	洞真類	<i>Dongzhen lei</i>
KR5b	洞玄類	<i>Dongxuan lei</i>
KR5c	洞神類	<i>Dongshen lei</i>
KR5d	太玄類	<i>Taixuan lei</i>
KR5e	太平類	<i>Taiping lei</i>
KR5f	太清類	<i>Taiqing lei</i>
KR5g	正一類	<i>Zhengyi lei</i>
KR5h	續道藏類	<i>Xudaozang lei</i>
KR5i	清代道教文獻	<i>Qingdai daojiao wenxian</i>
KR5j	敦煌道教文獻	<i>Dunhuang daojiao wenxian</i>



Table 2.7: The subdivisions of the Kanseki Repository 佛部 Fo bu  
表 2.7 漢籍リポジトリ佛部の類

KR6	佛部	<i>Fobu</i>
KR6a	阿含類	<i>Ahan lei</i>
KR6b	本緣類	<i>Benyuan lei</i>
KR6c	般若類	<i>Banruo lei</i>
KR6d	法華類	<i>Fahua lei</i>
KR6e	華嚴類	<i>Huayuan lei</i>
KR6f	寶積類	<i>Baoji lei</i>
KR6g	涅槃類	<i>Niepan lei</i>
KR6h	大集類	<i>Daji lei</i>
KR6i	經集類	<i>Jingji lei</i>
KR6j	密教類	<i>Mijiao lei</i>
KR6k	律類	<i>Lü lei</i>
KR6m	中觀類	<i>Zhongguan lei</i>
KR6l	毘曇類	<i>Bitan lei</i>
KR6n	瑜伽類	<i>Yuqie lei</i>
KR6o	論集類	<i>Lunji lei</i>
KR6p	淨土宗類	<i>Jingtu lei</i>
KR6q	禪宗類	<i>Chanzong lei</i>
KR6r	史傳類	<i>Shizhuan lei</i>
KR6s	事彙類	<i>Shihui lei</i>
KR6t	續諸宗 (日本)	<i>Xu zhuzong (Riben)</i>
KR6u	敦煌文獻類	<i>Dunhuang wenxian lei</i>
KR6v	新編類	<i>Xinbian lei</i>

Table 2.8: Frequently seen witnesses and their abbreviations used in the Kanseki Repository

表 2.8 漢籍リポジトリによく使われている版本とその記号

WYG	【四庫全書・文淵閣】	Wenyuan edition of the SKQS
SBCK	【四部叢刊】	1919 edition of the SBCK
ZTDZ	【正統道藏・三家本】	Reprint of the Zhentong Daozang
YP-C	【原版道藏輯要】	First edition of the Daozang jiyao
CK-KZ	【重刊道藏輯要】	New edition of the DZJY (Kaozheng reprint)
CBETA	【電子佛典集成】	CBETA collection of Buddhist texts
TKD	【高麗藏・東國影印版】	Tripitaka Koreana, Dongguk University reprint
T	【大】	Taishū Shinshū Daizōkyō
X	【新纂大日本續藏經】	Manji Zokuzokyō
F	【房山石徑】	Fangshan Shijing
J	【嘉興】	Ming canon (Jiaxing edition)
K	【麗】	Tripitaka Koreana
DCS	【東禪寺】	Edition of the Dongchan Tempel in Fuzhou
L	【乾隆大藏經】	Qing Canon (Qianlong edition)
P	【永樂北藏】	Ming canon (Northern Yongle edition)
GOZAN	【日本五山版】	Gozan editions (Japan)
C	【中華大藏經】	Zhonghua canon
G	【佛教大藏經】	Fojiao canon
S	【宋藏遺珍】	Remains of the Song canon

## 3 Using the Kanseki Repository

### 漢籍リポジトリの使い方

#### Overview

The texts of the Kanseki Repository are maintained in the @kanripo user account on the Web site GitHub.com. This is the raw material for all other uses of the texts and the texts there are freely available to everybody. In this part, the software used to interact with this repository is discussed.

The term Kanseki Repository used thus far is mainly intended as a designation for the texts made accessible on GitHub. (for technically minded readers, this can be understood as a 'back-end' of some sort); non-technical readers can imagine it as the stack room in a library where librarians will go to fetch the books ordered by the readers. Most users will encounter these texts not firstly as 'librarians', but as 'readers', therefore the perspective presented here will be mostly that of a reader. However, academics are not only readers, but also writers. Books are not only used in public libraries, but also purchased and kept in a personal library. So the dividing line is not as fixed as it might seem.

Furthermore, readers do have different ways to work with the texts in the Kanseki Repository. Again using a metaphor from the way books are used, the texts can be either read in a 'reading room', or borrowed from the library and taken back to a 'study'. Depending on the purpose and context, a reader can also switch back between these uses; in a way of speaking bringing his own books back to the reading room.

In this chapter, 'stack room' corresponds to the repository on GitHub, 'reading room' to the web application at kanripo.org and 'study' to the Mandoku application, which allows texts to be owned and worked on.

#### 概要

漢籍リポジトリのテキストはウェブサイト github.com の @kanripo ユーザーアカウントの中に維持されている。これはテキストのその他の用途すべてに使われる素材であり、そこにあるテキストは誰でも自由に利用できる。ここではこのリポジトリを利用するために使われるソフトウェアについて論じる。

これまで使ってきた漢籍リポジトリという用語は GitHub 上でアクセスできるようにしたテキストを指す名称である。技術に詳しい読者には、GitHub をバックエンドのようなものと理解してもらうことができる。技術に詳しくない読者は、図書館の司書が読者から指定された本を取りに行く書庫のようなものだと思ってもらえばよい。ほとんどのユーザーはこれらのテキストに「司書」としてではなく「読者」としてアクセスするので、ここで紹介する見地のほとんどは読者のそれである。しかしながら、研究者たちは読者であるだけでなく、著者でもある。そしてまた、本は公的な図書館で使われているだけでなく、購入され個人的な蔵書として保管されてもいる。それで、この境界線はそう思われているほど固定していない。

更に、読者はさまざまな方法を使って漢籍リポジトリのテキストを利用できる。再び本の喩えを使うならば、テキストは図書館の「閲覧室」で読むこともできれば、図書館から借り出して「書斎」に持ち帰ることもできる。また自ら所有する本を「閲覧室」に持ち込むなど、読者は目的やコンテキストに合わせて、これらの用途を切り替えることができる

この章において「書庫」は GitHub 上のリポジトリに対応し、「閲覧室」は kanripo.org におけるウェブアプリケーションに、「書斎」はテキストを所有しそれに手を加えることを可能にするマンドクというソフトウェア・アプリケーションに対応する。





## Kanripo.org

Thus, the reading room is the web application that can be visited by pointing a browser to the web address <http://www.kanripo.org>. A home page similar to the one in Figure 3.1 will appear.

From this page, there are three ways to access the texts that are stored in the stack room:

- (1) Browsing the catalog
- (2) Searching for titles
- (3) Searching within the texts

### Browsing the catalog

Catalog might be a slightly fancy expression for what is essentially a list of texts<sup>1</sup>.

The top page of the catalog (Figure 3.2) shows a table of contents to the left and a content area (yellow background) to the right. Initially the top level headings of the catalog (部 *bu*, KR1 to KR6) and the second level headings (類 *lei*, KR1a to KR6v) are visible, currently comprising 85 items in total. Users can now browse the content by clicking on these serial numbers and immediately access the corresponding section of the catalog.

Alternatively, the top level headings are also available on the left side, providing convenient and quick access to the sections at all times. The main classification follows the structure of the Repository as outlined previously. Since most users, at least initially, will be more familiar with other collections, the main constituting collections of the Repository are also available in the lower left part. Clicking on a title will bring up a listing for this specific collection. The serial numbers are now those of the collection in question, effectively providing a view of a subset of the Repository containing only the texts that are part of the collection in view. These serial numbers are only for the orientation on the catalog page; all other pages will only display the serial numbers as used in the Kanseki Repository beginning with “KR”.

As an example, browsing the catalog of the 道藏輯要 *Daozang jiyao* will bring up a screen similar to Fig-

## Kanripo.org

このように、「閲覧室」はブラウザをウェブアドレス <http://www.kanripo.org> に指定することでアクセスできるウェブアプリケーションである。図 3. 1 のようなホームページが表示される。

このページから「書庫」に保管されているテキストにアクセスするには 3 通りの方法がある：

- (1) 目録の閲覧
- (2) 題名で検索
- (3) テキスト内を検索

### 目録の閲覧

実質的にはテキストのリスト<sup>1</sup>にすぎないものを目録と呼ぶのは少々行き過ぎた表現かもしれない。

図 3. 2 の目録のトップページ（ホームページの「目録」をクリックすると現れるページ）には、左側に目次、右側に（黄の背景で）内容範囲が表示される。まず、目録のトップレベルの見出し（部, KR1 から KR6）と第二レベルの見出し（類, KR1a から KR6v）は表示される。現在、全 85 項目から成っている。ユーザーはこれらのシリアル番号をクリックすることでこの目録の対応セクションに即座にアクセスし、内容を閲覧できる。

また、トップレベルの見出しは左側にも掲載されており、各セクションへの素早く簡単なアクセスを常に提供している。主となる分類は前に概説した通り、このリポジトリの構造に従っている。ほとんどのユーザーは、少なくとも最初のうちは、コレクションの他の並べ方によりなじんでいるので、このリポジトリの主なコレクションを左側（漢リポ分類の下）に異なる分類法で提供している。題名をクリックすることで特定の目録のリストを表示できる。シリアル番号は特定のコレクションの番号で、事実上リポジトリの部分集合、すなわちコレクションに含まれるすべてのテキストリストを一覧で表示している。これらのシリアル番号は目録ページにテキストを並べるためだけのもので、他のすべてのページでは漢籍リポジトリで使用している「KR」で始まるシリアル番号のみを表示している。

一つの例として、『道藏輯要』をクリックすると図 3. 3 のような画面が表示される。これはテキストの題名を

ure 3.3. This gives the titles of texts, followed by the text's dynasty of origin and the principal author (or other representative person, depending on the type of text), as can be seen for example in the entry for JY003<sup>2</sup> “太上洞玄靈寶無量度人上品經法-南宋-陳春榮”. The information here is given as a best estimate, mostly based on traditional cataloging. As work on the catalog proceeds, this will be updated with the latest information, and all the details, including the reasons for updates, will be available in the catalog. The different parts, “太上洞玄靈寶無量度人上品經法”, “南宋”, and “陳春榮” are separated by a hyphen (“-”). If the dynasty or name is not available for some reason, this will be skipped over and the hyphen character will stand by itself, as for example in the case of JY019, “太上中道妙法蓮花經--”.

Clicking on the text, in this case JY019, will bring up the landing page of the text, as shown in Figure 3.4.

This happens to be the text with the number KR5h0001, the first text of section KR5h 續道藏. This page gives an overview of the text, including available versions and a table of contents. If no detailed overview of the inner divisions of a text is available for longer texts, the 卷 *juan* are enumerated and listed here. In the Repository, a *juan* is the basic unit in which a text is stored and accessed. Any content before the first *juan*, such as prefaces, introductions, etc. is placed in the file with number ending in “0.” (For technical reasons, this file exists even if there is no such preliminary content.)

To start reading the text, simply click on *juan* 1 in the list; this will bring up a screen similar to Figure 3.5. If no other version is specified, this will open the **master** version of the text. As previously explained, this is the most up-to-date and best curated version; it does not necessarily represent one specific documentary edition of the text, but rather the text that has been edited by the editors of the Kanseki Repository. Other versions can be accessed via the links at the top left of the page, on white background. There are five versions in this case, two versions of the Daozang, two editions of the DZJY (all these editions are “documentary”) and one more edition of the DZJY, fea-

示しており、その後にテキストの出所の王朝、そして主な著者（あるいは、テキストの種類により、代表者）が続く。これは、例えば、JY003<sup>2</sup>「太上洞玄靈寶無量度人上品經法-南宋-陳春榮」のエントリーに見ることができる。ここで提供されている情報はほとんどが伝統的な目録法に基づいた最良の推定である。目録の仕事が進むにつれ、これも最新情報に更新され、更新理由も含めたすべての詳細が目録中で提供されることになる。「太上洞玄靈寶無量度人上品經法」「南宋」「陳春榮」という異なる部分はハイフン（「-」）で区切られている。何らかの理由で王朝や名前が不明な場合、JY019「太上中道妙法蓮花經--」の場合のように、不明部分は飛ばし、ハイフンのみを単独で表示する。

図 3.4 に示されているように、テキスト、この場合は JY019、をクリックするとテキストのランディングページが表示される。


この JY019 の場合はたまたま KR5h0001 番、つまり KR5h 續道藏セクションの一番目のテキストである。このページは閲覧可能なバージョンや目次も含めたテキストの概要を表示している。長いテキストについて、その内部区分の詳細な概要がない場合、ここには巻が一つ一つ列挙される。このリポジトリでは、巻が基本単位となり、そこにテキストが保管され、またアクセスされる。前書きや序文など一番目の巻の前のコンテンツはすべて「0」で終わる番号のファイルに置かれる。（技術的な理由から、このような先行するコンテンツがない場合にもこのファイルは存在する）

テキストを読み始めるには、ただリスト中の巻の「1」をクリックすればよく、図 3.5 のような画面が表示される。他のバージョンが指定されていないかぎり、このクリックでテキストの master バージョンが開けられる。前にも説明した通り、これが最新で最もよく整理されたバージョンである。これは必ずしも一つの特定のドキュメンタリー版を代表しておらず、それよりも漢籍リポジトリの編集者によって編集されたテキストであることを意味する。他のバージョンにはページの左上、白背景上にあるリンクからアクセスできる。ここでの場合、五つのバージョンがある。『正統道藏』の 2 バージョン、『道藏輯要』の 2 版（これらの 4 版が「ドキュメンタリー」版）、そして『道藏輯要』のもう一つの版、つまり京都での『道藏輯要』プロジェクトの一環として作成さ



turing punctuation created as part of the DZJY project in Kyoto. (This latter edition is interpretative.) At the bottom left of the page is a table of contents, in this case simply giving numbers for each of the 10 卷 *juan* of the text.

The area with the yellow background features the text to be read. The line breaks follow the selected edition, as do the page numbers, which are given following the pattern 001-001a, meaning *juan* 1, page 1, recto. The exact pattern depends on the edition, but is always in the order of “unit within the text” (*juan* in this case), page number and a letter indicating a subdivision of a page or something similar. Next to the page number is an icon, a small image of a woodblock printed page, which indicates that the digital text is accompanied by a scanned image of the corresponding physical page. Clicking on the page number will bring up a screen similar to Figure 3.7.

The right side of the screen here shows the digital facsimile of the requested page, while the left side shows the text as before. The size of the facsimile can be adjusted using the buttons above the image, + to enlarge the view and - to reduce it. The view can also be adjusted using the slider to the right of the - button. Clicking on  (reset) will return the view to its default size. On the right there is a drop-down menu showing the name of the currently displayed version, 【正統道藏・涵芬樓版】 in this case. Clicking the name of the edition allows the user to select a different version, if available (Figure 3.6). The buttons labeled “<” and “>” to the left and right of the drop-down menu can be used to display the previous and next facsimile page, respectively. Clicking on X will hide the digital facsimile, thereby restoring the screen to its previous state.


## Search

### Title search

There are two types of search in Kanripo—title searches and content searches. A title search searches the full strings of characters used to reference a text, including the dynasty and author. The result of a title

れた句読点をつけた版である（この最後の版は解釈版）。ページの左下には目次があり、ここでの場合、ただテキスト10巻のそれぞれの番号が示されているだけである。

黄の背景部分は読むべきテキストを表示している。改行は選択した版のそれを踏襲しており、改頁についても同様で、「第1巻、第1ページの表」を意味する001-001aのパターンに従って表示されている。厳密なパターンはそれぞれの版により異なるが、常に「テキスト内の小単位」（この場合は巻）、ページ番号とそのページの区分を示す文字あるいはそれに相当するものの順に従っている。ページ番号の隣には木版印刷ページの小さな画像であるアイコンがあり、それはそのデジタルテキストが対応する実際のページのデジタルファクシミリを伴っていることを示している。ページ番号をクリックすると図3.7のような画面が表示される。

ここで画面の右側にはリクエストしたページのデジタルファクシミリが表示され、左側には先ほどと同じテキストが表示される。ファクシミリのサイズは、表示を拡大するには「+」、縮小するには「-」と画像の上のボタンを使って調整できる。表示はまた、「-」ボタンの右にあるスライダーを使って調整できるをクリックすると表示がそのデフォルトサイズに戻る。この右側にはドロップダウンメニューがあり、現在表示されているバージョンの名前を示している。ここでの場合、【正統道藏・涵芬樓版】である。この版名をクリックすることでユーザーは、（それが提供されている場合には）異なるバージョンを選択できる（図3.6）。ドロップダウンメニューの左と右にある「<」と「>」のボタンはそれぞれ前のファクシミリページと次のファクシミリページを表示するのに使うことができる。「X」をクリックするとデジタルファクシミリが非表示になり、よって画面が以前の状態に戻る。

## 検索

### 題名で検索

漢リポには題名検索とテキスト検索という2種類の検索がある。題名検索はテキストを言及するのに使われた、王朝や著者も含めた文字列すべてを検索する。題名検索の結果として目録から該当する題名が表示される。この

search is one or more titles from the catalog. The text or texts can then be accessed as described above for browsing the catalog.

### Search within texts

The search field on the right is used to enter a search term for a full-text search. At the moment only simple searches are supported, but more sophisticated methods will become available. An example of a full-text search result is shown in Figure 3.8.

The top part with the red background gives a summary of the results, specifically the total number of matches and the matches shown on the current page — “1 to 20” in this case. The main part of the screen, with yellow background, shows the details of these first 20 results, using a so-called “keyword in context” (KWIC) format. The text number, title (sometimes truncated), and location within the text are displayed for each match. This location functions as a hyperlink, such that clicking it will open the text at the indicated location with the search term highlighted. The third item is an excerpt of the text at the location of the match, including a fixed number of characters before the search string (three in this example) and a number of characters after the match. The total length of this line is fixed; in this case there are 10 characters including the search term. The line displayed here very closely resembles the index that is used internally to locate the matches. This is useful to know, because it means that searches for strings of more than 10 characters will return no results. In practice, it is recommended that search terms be 2 to 5 or 6 characters in length. Searches for a single character will be rejected by the system, because they tend to overload the server.

The search function has been specifically designed for the Kanseki Repository. While building the index that is internally used for executing the search, all editions of a text are taken into account. Those index entries that are not from the **master** branch are marked in the index display with 異本 *yiben* ‘Different version’. In addition to that, the portions of a text that are printed as notes in smaller characters are considered as a separate sequence of characters, outside of the se-

テキストは、上記の目録を閲覧する項で説明したようにしてアクセスできる。

### テキスト内を検索

右側の検索フィールドは全文検索のための検索用語を入力するのに使われる。現時点では、単純検索のみをサポートしているが、より洗練されたメソッドが利用可能となるであろう。全文検索の結果の例が図 3.8 に示されている。

赤の背景を持つトップ部分には結果のサマリー、具体的には一致結果の合計数および現ページに表示されている一致結果の数（ここでの場合「1 から20まで」）が表示される。黄の背景を持つ画面のメイン部分には、いわゆる「文脈付き索引」(KWIC) フォーマットを使って、この検索結果の最初の20の詳細が表示される。各検索結果について、そのテキスト番号、題名（時として切り詰められている）、テキスト内での位置が表示される。この部分はハイパーリンクになっており、これをクリックすると、その指定した位置で検索した用語がハイライトされたテキストが開く。三つ目の項目は一致箇所の抜粋で、それには検索語の前部分の決まった文字数（この例では3文字）と後ろ部分の数文字も含まれる。この行の長さは決められており、ここでの場合、検索語も含めて10文字である。ここに表示された行は一致箇所を探すのに内部で使われたインデックスととてもよく似ている。これを知っておくことは有用である。というのは、10文字以上の文字列の検索は何の結果も出さないということだからである。実際のところ、検索語の長さは2文字から5、6文字までであることが推奨されている。1文字だけを検索することはシステムによって却下される。サーバーをオーバーロードしがちだからだ。

この検索機能は漢リポのために特に作られたインデックスに基づいている。このインデックスを作成するのに、全ての版本が持つ文字がインデックスに導入された。マスター版以外の検索結果には該当する行の最後に「異本」の表示がある。これに加えて、小さな文字で注釈として印刷されているテキストの部分は、本文のシーケンスの外にある別の文字シーケンスとみなされる。注釈の中から引き出された検索結果はそのインデックス表示に「夾註」の印が付けられている。



quence of the main text. Matches that are drawn from the text in notes are marked as 夾註 *jiazhu*, 'intralinear notes' in the index display.

Search is central to research, because it is used to locate new material, but it also serves as a tool for analysis to better understand the cultural tradition as represented by the collection. There are two main ways that searches can be refined to present the material in a desirable way: (1) changing the sort order, i. e., the specific way that results are presented to the user; and (2) applying filters to display only the subset of matches that is of interest for the purpose at hand.

### Changes to the sort order

The top left of the results page with the white background has links that allow the sort order to be changed to any of the following:

- “By text number”: This sorts the search results by the text number. Since the numbering follows the traditional classification, this effectively places together texts that have been considered to be closely related.
- “By date”: The date here is to be understood as the date of creation of the major part of the text. Dating texts is a very complex undertaking and in many cases there are conflicting views. The solution used here is to arrange all texts into a list. The order of the texts in this list will be used to determine the sorting<sup>3</sup>.
- “By search term”; This is the sort order that is displayed by default if no other search order is specified. It simply arranges the results according to the characters following the highlighted search term. This should ideally be the lexical order using the radical/stroke count-based ordering of the 康熙字典 *Kangxi zidian*. However, for technical reasons this is only approximately achieved, since the underlying character codes on which the sorting is based have been defined in several blocks, not continuously. In addition, some variant characters that are represented by images because they
- ・ テキスト番号順：これは検索結果をテキスト番号順に並べ替える。この番号付けは伝統的な分類を踏襲しているので、実質的にこれは密接に関連しているとみなされているテキストを一緒に並べることになる。
- ・ 年代順：ここでの年代はテキストの大部分が作成された年代とする。テキストに年代を付けることはとても複雑な作業で、多くの場合、相反する見解が存在する。ここではすべてのテキストをリストにして並べるという解決法を使った。このリストのテキストの順番は並べ替えを決定するのに使われる<sup>3</sup>。
- ・ 検索語順：これは、他の並べ替え順序が指定されていない時、デフォルトとして表示される並び替え順序である。ハイライトされた検索語に続く文字に従って、ただ検索結果を並べるだけのことである。これは、理想的には、『康熙字典』の部首 / 画数に基づいた順序付けを使った辞書式の順序でなければならない。しかしながら、技術的な理由で、並び替えが基礎としている基底文字コードがいくつかのブロックごとに定義されており、継続してではないことから、これはおおよそでしか達成されていない。またこれに加えて、文字コードセットにはなく、それで画像として表示されるいくつかの異体文字が、そ

are not available in the character code set will sort before the other characters in that position.

- ・“By preceding characters”. In this case the order (subject to the same limitations as in the previous case) is based on the characters immediately preceding the search term, taking the characters into account in reverse order of reading, starting from the search term.

The other links further down the left side of the screen, grouped into several sections, all serve as filters to drill down to specific parts of the result set:

- ・朝代 *Chaodai* ('Dynasty') breaks down the results by dynasty and lists the most frequent matches. The numbers in parentheses are the number of hits for the specific dynasty.
- ・部 *Bu* ('Section') breaks down the results according to the six fold classification, thus providing a quick overview of where the most matches are found.
- ・部／類 *Bu lei* ('Subsection') breaks down the results by subsection. The six most frequent matches are displayed. The top-level section names are included for reference to enable easier orientation.
- ・文献 *Wenxian* ('Text') gives the top six texts, i.e., the texts with the most matches, along with the number of matches.

### List of texts with matches

In addition to this KWIC display, the search results can also be viewed in a completely different way — according to text. This view can be activated by clicking “Show results by text” at the right of the area with red background. The results are then displayed in order from the texts with the highest number of matches, but the order can be changed using the links on the left.

の位置にある他の文字よりも先に並べられる。

- ・検索語の前の字で：この場合、（前の場合と同じ制約に準じている）順序は、検索語から逆に数えて、そのすぐ前の文字を基にしている。

画面の左側でこれらの更に下にある、いくつかのセクションにグループ分けされた他のリンクはすべて、結果を特定の部分に絞り込むためのフィルターの役目を果たす：

- ・朝代（王朝）は結果を王朝別にし、最も頻度の高い検索結果を列挙する。数字はそれぞれの王朝の検索ヒット数である。
- ・部（セクション）は結果を漢リポの分類別にし、それによって最も検索結果の多い部分はどれかの手取り早い概観を提供する。
- ・部／類（サブセクション）は結果を類別にする。最も頻度の高い検索結果を上から六つ列挙する。簡単に位置付けできるよう参考とするために、該当類の上位の部名も含まれている。
- ・文献（テキスト）は最も頻度の高い検索結果を持つ上位六つのテキストを、その検索ヒット数と共に表示する。

### 検索結果によるテキストリスト

この KWIC デisplayに加えて、検索結果はまったく異なる方法、テキストによる方法でも表示できる。この表示は、赤の背景を持つ部分の右側にある「テキスト毎まとめて表示」をクリックすることで作動させることができる。すると結果が、最も検索結果数の多いテキストの順に表示されるが、上記した通り、この順序は左のリンクを使うことで変えることができる。



## Preparation for more advanced use

What has been explained so far is the kind of interaction with the material that is possible for any visitor to the [kanripo.org](https://kanripo.org) website without any preparation.

For more advanced usages, many web sites offer the possibility to sign up for an account that allows personal settings and data to be preserved between sessions. The [kanripo.org](https://kanripo.org) web site does not maintain accounts by itself, but instead asks users to log in using an account with the website [github.com](https://github.com). No personal information whatsoever is permanently stored at [kanripo.org](https://kanripo.org), although some transient data is cached while the site is used. Using GitHub for this purpose has two main advantages: (1) It allows users to transparently take ownership and control of their data independently of the web site; and (2) Since the repository itself is hosted on GitHub, using a GitHub account for this purpose is the most logical way to enable settings and data to be shared with other sites where relevant. (Note that multiple web sites and services can provide access to the Kanseki Repository.)

GitHub is a commercial company based in the U.S. There is a Japanese branch as well, but so far the web site is available in English only. Accounts are free of charge if the data placed into the accounts is available for public access without limitations. It is also possible to have private repositories, but these require an account with a subscription fee.

Creating an account is quite straightforward; simply visit <https://github.com> and follow the instructions. For use with the Kanseki Repository a free account will do.

Once an account is ready, it can be accessed from [kanripo.org](https://kanripo.org) using the “Login” link on the red background at the top right of every page of the web site. If you are already logged in to GitHub when you click this link on Kanripo for the first time, GitHub will require you to authorize access from Kanripo. This is necessary, because Kanripo will interact on your behalf with GitHub. This authentication is carried out using the OAuth2 protocol, which is currently thought

## より高度な用法の準備

これまでに説明してきた資料とのインタラクションは、何の準備もなく漢リポのウェブサイトへアクセスしたユーザーでも可能なたぐいのものである。

より高度な用法のために、多くのウェブサイトは個別の設定やデータがセッション間でも保持されるようにアカウントにサインアップする選択を提供している。[kanripo.org](https://kanripo.org) ウェブサイトはそれ自体ではアカウントを維持しないが、その代わりに、ユーザーに [github.com](https://github.com) ウェブサイトのアカウントを使ってログインすることを求める。一時的なデータがサイトを使っている間キャッシュされることはあるが、どのようなユーザーの情報も [kanripo.org](https://kanripo.org) に永久保管されることはない。この目的のために GitHub を使うことには二つの主な利点がある：(1)ユーザーが自らのデータをこのウェブサイトから独立して透明な形で所有しコントロールすることを可能にする；また(2)リポジトリ自体が GitHub によってホストされているので、この目的のために GitHub アカウントを使うことは、設定やデータを関連する他のサイトと共有させる最も論理的な方法である。(複数のウェブサイトやサービスが漢籍リポジトリへのアクセスを提供できることに留意すること)

GitHub はアメリカに拠点を置く営利会社である。日本支社もあるが、今までのところウェブサイトは英語のみで提供されている。アカウント内のデータが制限なしのパブリック・アクセスを提供しているのであれば、そのアカウントは無料である。プライベートのリポジトリを持つことも可能であるが、それには利用料金が課せられるアカウントが必要とされる。

アカウントを作成するのはとても簡単である。ただ <https://github.com> にアクセスし、指示に従うだけのことである<sup>4</sup>。漢籍リポジトリで使うためには、無料のアカウントで十分である。

ひとたびアカウントを準備すれば、あとは [kanripo.org](https://kanripo.org) から、どのページでもトップ右側に表示されている赤の背景を持つ「サインイン」リンクを使ってアクセスできる。この漢リポ上のリンクを初めてクリックする時、GitHub にすでにログインしている状態ならば、GitHub は漢リポからのアクセスを認証するよう要求してくる。これは、漢リポがユーザーに代わって GitHub とインタラクションするために必要なことである。この認証は、ウェブ間のインタラクションを可能にする現在のところ

to be the most secure and safe way of enabling interaction between web sites. As you also can see on this page, the permission you have to give to Kanripo is very specific — allowing access only to public and private repositories (Figure 3.9). No other information associated with your account will be accessible to Kanripo. It is also important to know that, if necessary, this authorization can be revoked at any time from your GitHub account. After you click on the green “Authorize application” button on the GitHub page, you will be returned to the [kanripo.org](http://kanripo.org) website. At the same time, Kanripo will receive a token from GitHub for authenticating further interactions between the two websites.

One of the first things Kanripo will do with this token is to create a new repository under your GitHub account with the name `KR-Workspace`. Kanripo will use this repository to store the data you create while working on Kanripo (see Figure 3.10). Having this data under your own GitHub account will allow you to access the data, either directly yourself or via a software application, provided there is appropriate authentication.

Once you are logged in, the “Login” button changes to display your GitHub user name. Clicking this button now takes you to a profile page that summarizes your account information. (Although it appears on [kanripo.org](http://kanripo.org), as shown here, the data actually resides in your workspace on GitHub).

## Advanced uses of the Kanseki Repository

Interaction with GitHub enables information about users and their work to be saved across visits and it also allows customization of some settings. Furthermore, texts of special interest can be copied to a user’s account (“to fork a repository” in the language of GitHub) and edited there. Kanripo.org is smart enough to use a text from the user in preference to the text in the common `@kanripo` account if available. A user can, for example, also correct misprints and report them to the editors of the Kanseki Repository using

最も安全で安定した方法とされている OAuth2 プロトコールを使って実行される。このページを見てもらえばわかる通り、漢リポに与えるべき許可はとても明確で、パブリックとプライベートのリポジトリへのアクセスのみである (図 3.9)。漢リポはユーザーのアカウントに関連した他のどの情報にもアクセスできない。また、これは重要なポイントだが、必要であれば GitHub のあなたのアカウントからいつでもこの認証を取り消すことができる。GitHub ページ上の緑の「Authorize application」ボタンをクリックすると、漢リポウェブサイトに戻る。これと同時に、漢リポは GitHub からこれら二つのウェブサイトの更なるインタラクションを認証するトークンを受け取る。

漢リポがこのトークンで最初に行う事柄の一つは、ユーザーの GitHub アカウントに `KR-Workspace` という名前で新しいリポジトリを作成することである。漢リポ上で作業する間に作成するデータを保存するために、漢リポはこのリポジトリを使う (図 3.10 を参照)。こうしてデータを自分の GitHub アカウント下に持つことは、(適切な認証を持つことを条件として) ユーザーがこのデータに直接、あるいはソフトウェア・アプリケーションを通じてアクセスすることを可能にする。

ひとたびログインをすれば、この「Login」ボタンが変わって、あなたの GitHub ユーザー名が表示される。このボタンをクリックすると、あなたのアカウント情報をまとめたプロフィールページが表示される。(ここで示されているように、これは [kanripo.org](http://kanripo.org) 上に表示されるが、データは実のところ GitHub にあるあなたのワークスペースに存在する)

## 漢籍リポジトリの高度な用法

GitHub との連携によって、ユーザーとその仕事についての情報をアクセス間にも保持することを可能にし、またいくつかの設定をカスタマイズすることを許可する。更に、特に興味のあるテキストをユーザーのアカウントにコピーし (GitHub 用語では「to fork a repository (リポジトリをフォークする)」)、そこで編集することも可能にする。kanripo.org は、ユーザーのテキストが存在する時には、`@kanripo` アカウントの標準テキストよりそちらを優先して使う賢さを持ち合わせている。この方法を使い、例えば、ユーザーは誤字の訂正をし、それ





this method. How this is done is explained further below.

Web browsers have the advantage of being readily available without the need to install additional software and they are familiar to everyone, but they also have limitations. It becomes tedious to edit or even read a text in the browser for long periods of time. For this purpose, a specialized piece of software called Mandoku has been developed, which will also be introduced below.

### Kanripo.org

Here we introduce some of the features of the Kanripo.org web site that become available to users who are logged in. The first allows users to easily collect a list of texts of special interest; the second shows how to customize the “sort by date” function mentioned above.

#### Lists of texts

In many cases, the focus of a specific research question lies in a number of specific texts that require highly refined or specialized searching. When performing searches that generate lists of texts, the result screen will display an additional line at the top, as shown in Figure 3.11.

As can be seen, the current search term is already suggested as a filename for the list of texts that result from this search, but of course it can be changed to any other name. All texts in the list have a checkbox which is unselected by default, but clicking the line “Toggle selection” at the left will select all the texts. The “Save text list” button saves the list of selected texts to the KR-Workspace on GitHub. (In GitHub terminology, the list is “added” and “committed”). To use the list as a filter for searches on Kanripo.org, however, the list has to be loaded. This can be done by selecting “Immediately load the text list”. If the list is not immediately loaded, or on a subsequent visit to the site, the text list can also be loaded from your profile, as seen in Figure 3.10. Here, we see a number of available text lists, one of which is even loaded. Figure 3.12 shows the results of a filtered search.

を漢籍リポジトリの編集者に報告できる。これがどのように行われるかは後で説明する。

ウェブブラウザには、追加のソフトウェアをインストールする必要なく即座に利用でき、誰にでもなじみがあるという利点があるが、限界もある。ブラウザ上でテキストを編集すること、また読むことでさえ、それが長時間になると面倒になる。この目的のために、マンドクと呼ばれる特化されたソフトウェアが開発された。これについても後で紹介する。

### kanripo.org

ログインしたユーザーに提供されている kanripo.org ウェブサイトの機能をいくつかここで紹介する。一つ目はユーザーが特に興味のあるテキストリストを簡単に集めることを可能にする。二つ目は上記した「年代順」機能をカスタマイズする方法を示す。

#### テキストリスト

多くの場合、専門性の高い研究課題は、とりわけ検索が必要な幾つかの特定のテキストに焦点を絞ることになる。検索を行ないテキストリストが作成されると、図 3.11 に示されるように、結果の画面のトップ部分に追加の一行が表示される。

これを見てもらえばわかる通り、この検索で作成されたテキストリストのためのファイル名として、現在の検索語がすでに提案されている。もちろん、これは他のどんな名前に変えることもできる。リスト中のすべてのテキストにはそれぞれチェックボックスがあり、それらはデフォルトとして「未選択」になっているが、左にある「選択を切り替え」をクリックすることですべてのテキストを「選択」に切り替えることができる。「テキストリストの保存」ボタンは選択されたテキストリストを GitHub 上の KR-Workspace に保存する。(GitHub 用語では、リストが「added (追加された)」そして「committed (コミットされた)」)。しかしながら、kanripo.org 上の検索フィルターとしてリストを使うためには、それがロードされていなければならない。これは「テキストリストを即座にロードする」を選択することで実行できる。リストが即座に、あるいはサイトへの次のアクセスでロードされなければ、図 3.10 に示されるように、このテキストリストはユーザーのプロファイルからもロードできる。ここでは、いくつかの提供されているテキストリ

### Changing the list for “ordering by date”

As already explained, the “order by date” feature effectively sorts texts according to a list of texts arranged in a specific, predetermined order. This list can be inspected by selecting “Order by date” on the profile page. This will open the file for editing on the GitHub site, where the sorting order can be changed as necessary. Don’t forget to save the changes by clicking on the green “Commit changes” button at the very bottom of this page. From now on, the search results of this user will be sorted in this order, according to the changed file.

### Forking a text

As mentioned above, it is possible to “check out” texts of interest from the Kanseki Repository and copy them to a private “library.” This action is known as “forking” the text, in the sense of a fork in a roadway from which roads lead in two separate directions. By “forking” a text, a user can edit the text independently of the text as it still exists in the Kanseki Repository. This is conceptually different from creating a branch, which adds a new version to a text. Forking makes a complete copy of all branches; it takes a snapshot of the state of all branches and records them together with the location of their origin. It is important to note that forking is a specific action of the GitHub system. After forking the new text will be available for editing in the user’s account on GitHub. Texts displayed in the kanripo.org website have a link labeled “GitHub”. This link leads directly to the corresponding page on GitHub where the text can be forked. If the text has been forked already, the link will put the page in editing mode instead. Thus, this link can be used when a user wishes to make a correction of the text, for example.

There is also a different way to copy a text, referred to as “cloning.” In this case, the text is copied to the local machine from which cloning is initiated. Both the text in the @kanripo account on GitHub and the

ストが見られ、その内の一つはロードもされている。図 3.12 にフィルター付き検索の結果が示されている。

### 「年代順」にリストを変更

すでに説明した通り、「年代順」機能は、特定のあらかじめ決められた順序に整理されたテキストリストに従って、テキストを効率的に並び替える。このリストはプロフィールページの「年代順」を選択することで検査できる。これが GitHub サイトで編集するためのファイルを開き、この時、必要であれば並び替え順序を変えることができる。このページの一番下にある緑の「Commit changes (変更をコミットする)」ボタンをクリックしてこの変更を保存することを忘れてはならない。それ以降、このユーザーの検索結果は変更されたファイルに従って、その順序で並び替えられることになる。

### テキストをフォークする

上記した通り、漢籍リポジトリから興味のあるテキストを「借り出」して、それらを個人の「図書館」にコピーすることが可能である。この行為は、その場所から道が二つの異なる方向へと向かっている分岐点という意味で、テキストの「フォーク (fork)」として知られている。テキストを「フォーク」することで、そのテキストは漢籍リポジトリにもいまだ残存するので、ユーザーはそれとは独立してテキストを編集できる。これは、テキストに新しいバージョンを加えるというブランチ作成とは概念的に異なる。フォークすることですべてのブランチの完全なコピーを作ることになる。それはすべてのブランチの状態のスナップショットを取り込み、それらをそれぞれの元の位置と共に記録する。これは重要なポイントだが、フォークは GitHub システムに特有なアクションである。フォーク後、この新しいテキストは GitHub 上のユーザーのアカウントで編集できる。kanripo.org ウェブサイトに表示されたテキストは「GitHub」とラベル付けされたリンクを持つ。このリンクが GitHub 上の対応ページに直通しており、そこでテキストはフォークされる。テキストがすでにフォークされている場合、リンクはその代わりにそのページを編集モードにする。それゆえに、このリンクは例えばユーザーがテキストの訂正を行ないたい時に使うことができる。

「クローン (clone)」と呼ばれる、テキストをコピーするまた別の方法もある。この場合、テキストはクローンが開始されたローカルマシンにコピーされる。GitHub 上の @kanripo アカウントにあるテキストも、ユーザーア



“fork” of a text in the user account can be cloned. In the latter case, users can also “push” updates made locally back to the version of the text in their online accounts. Note that a text in the @kanripo account can not be updated directly from a user account, since pushing is not allowed. Users wishing to push changes to the Kanseki Repository will have to fork the text and then make use of a different mechanism, called a “pull request”, which will be explained later. A necessary condition for a pull request is the existence of a fork, so users who want to make a pull request will need to fork the text first.

## Mandoku

For intensive work with texts, such as close reading or translation, we recommend the use of Mandoku. This application allows users to download texts of interest and also edit them offline (while not connected to the Internet).

Mandoku is based on Emacs, a very sophisticated text editor that has been in active development for over 30 years. Some of the concepts and the terminology to describe them predates current software and might seem unfamiliar. It will take some time to get used to it, but that time is well invested.

Here we describe how to install Mandoku on Windows, Mac OS X and Linux and then give an overview of some of the program’s functions.

### Requirements:

- On all platforms, git and python are required for the installation process and software updates, and to run the program Emacs is also needed, of course.
- A working network connection.
- About 400 MB of free space on the harddisk used for installation<sup>4</sup>.

アカウントにあるテキストの「フォーク」も、両方ともクローンすることができる。後者の場合、ユーザーはローカルで作ったアップデートをそれらのオンラインアカウントのテキストバージョンへと「プッシュ (push)」することもできる。@kanripo アカウントのテキストは、プッシュすることが許可されていないので、ユーザーアカウントから直接アップデートできないことに留意すること。変更を漢籍リポジトリへとプッシュすることを望むユーザーは、テキストをフォークし、その後で「プルリクエスト (pull request)」と呼ばれる異なる手段を使わなければならない。この手段については後ほど説明する。プルリクエストに必要な条件はフォークが存在することであり、それゆえに、プルリクエストを行いたいユーザーはまずテキストをフォークする必要がある。

## マンドク

クローズリーディングや翻訳といった、集中的にかつ長時間にわたりテキストを扱う時にはマンドクの使用を勧める。このアプリケーションはユーザーが興味のあるテキストをダウンロードし、それらをオフラインで（インターネットに接続していない間も）編集することを可能にする。

マンドクはその基盤として、30年以上も開発が継続されており、たいへん洗練されたテキストエディターである Emacs を使っている。これらを説明する概念や専門用語のいくつかは現在のソフトウェアより前にさかのぼるもので、なじみのないものかもしれない。

ここではマンドクをウィンドウズや Mac OS X, Linux にインストールする方法を説明する。そしてこのプログラムのいくつかの機能について概要を述べる。

### 要件:

- すべてのプラットフォームに対して、インストールのプロセスとソフトウェアアップデートのための git と python が必要。またプログラムを実行するのに Emacs が必要なのももちろんのことである。
- 使用できるネットワークコネクション。
- インストールのための約400MB のディスクスペース<sup>5</sup>。

## Installation<sup>5</sup>

### Windows

There are three ways to install Mandoku on computers running the MS Windows operating system:

- (1) Using a “bundle” that contains all necessary files and programs.
- (2) Installing all components separately and adding the Mandoku package for Emacs manually.
- (3) Installing the `gnupack` distribution and then proceeding with the relevant instructions for Linux<sup>6</sup>.

For users as yet unfamiliar with Emacs, the first method is recommended. Users who already have a relatively recent version of Emacs (24.4 or newer is recommended) are best served by the second method. Method (3) provides an environment that is well suited to working with Japanese and Chinese, but the exchange of files with other Windows programs requires some rather advanced technical knowledge, so it is not recommended for users who are not already familiar with this kind of distribution. Here below, I will mainly discuss the first method.

A distribution “bundle” with all necessary files to start using Mandoku is available from the Mandoku website.

### Installing the installation bundle for Windows

The bundle for Windows has been tested on Windows 8 and 10. The bundle may also work on earlier versions of Windows, but it has not been tested on them. The Emacs application included in the Mandoku bundle below is version 24.5 from [21]. This version is known to work with input methods for Chinese and Japanese.

1. The Mandoku bundle includes a version of the `git` program, which is used to download and up-

## インストール<sup>6</sup>

### ウィンドウズ

マイクロソフトウィンドウズのオペレーティングシステムを使っているコンピューターにマンドクをインストール方法は三つある：

- (1) 必要なファイルやプログラムすべてを含む「バンドル」を使用する。
- (2) すべてのコンポーネントを別々にインストールし、Emacs のためのマンドクのパッケージを手動で加える。
- (3) `gnupack` のディストリビューション<sup>7</sup>をインストールし、Linux の適切な指示に従って進める。

まだ Emacs になじみのないユーザーには一つ目の方法を勧める。すでに Emacs の比較的最近のバージョン (24.4以降を推奨) を持っているユーザーは、二つ目の方法が最適である。三つ目の方法は日本語や中国語で働くのに最適な環境を提供するが、他のウィンドウズプログラムとのファイル交換にはかなり高度な専門知識が必要とされる。それでこの手のディストリビューションになじみのないユーザーには勧めない。ここでは主に一つ目の方法について論じる。

マンドクを使い始めるのに必要なファイルすべての入ったディストリビューション「バンドル」はマンドクのウェブサイトから入手できる。

### ウィンドウズ用インストールバンドルをインストールする

ウィンドウズ用バンドルはウィンドウズ 8 と 10 上でテストされている。このバンドルはウィンドウズの以前のバージョンでも動作するかもしれないが、テストはしていない。以下のマンドクバンドルに含まれている Emacs アプリケーションは NTemacs プロジェクト ([21]) からの 24.5 バージョンである。このバージョンは中国語や日本語の入力メソッドで動作することが確認されている。

1. マンドクバンドルは、必要なファイルのダウンロードやアップデートに使われる `git` プログラムの



date necessary files.

2. Download the Mandoku installation bundle `mandoku-2016-03.zip` (250 MB).
3. Double-click the archive file then move the resulting `krp` folder to a root directory of your computer, e. g., `C:` or `D:`, as shown in Figure 3.13. The folder can also be moved to a USB memory stick or external hard disk. Do not copy it to your user folder because it may not work correctly there.
4. Inside the folder `krp`, navigate to the `bin` folder then click on the file “`start-mandoku`”, which is a batch file that configures the operating environment and launches Emacs. This file can also be accessed via a shortcut placed on the Desktop, so you can easily find it again next time. To create such a shortcut, right-click on the file “`start-mandoku`” and select “Send-to”, then “Desktop (create shortcut)”. You will then be able to use that shortcut to start Emacs (and Mandoku) next time. Alternatively, you can also pin the shortcut to the taskbar for convenient access.

### Installing Mandoku on a Macintosh computer

There is no need for an installation bundle for computers running Mac OS X 10.9 or newer, since everything except Emacs is already available in the standard installation of the operating system. To install Mandoku, use version 24.4 or later of Emacs. (The version containing the patch by Yamamoto Mitsuharu 山本光晴 of Chiba University, available at [15], is recommended, because it supports the whole range of Unicode.) Simply download a zip archive from this link, click on the downloaded file to unpack it, then drag the file Emacs to your Applications folder.

Once you have Emacs installed, you need to install and activate the Mandoku package. You can do this in different ways; two of these are outlined here. Use whichever method seems easier to you.

バージョンを含んでいる。

2. マンドクのインストールバンドル `mandoku-2016-03.zip` (250MB) をダウンロードする。
3. アrchiveファイルをダブルクリックし、図 3.13に示されているように、展開した `krp` フォルダをあなたのコンピュータのルートディレクトリ、例えば `C:` や `D:` に移動する。フォルダを USB メモリや外付けハードディスクに移動させることもできる。ユーザーフォルダにコピーしてはならない。そこでは正しく動作しないかもしれないからである。
4. `krp` フォルダの中で `bin` フォルダへナビゲートし、その後、動作環境を設定し Emacs を起動するバッチファイルである「`start-mandoku`」ファイルをクリックする。このファイルはデスクトップに置かれたショートカットからアクセスすることもできる。これで次回もまた簡単にファイルを見つけることができる。このようなショートカットを作成するには「`start-mandoku`」ファイルを右クリックし、「送る」を選択、そして、「デスクトップ (ショートカットの作成)」を選択する。そうすれば、次回から Emacs (とマンドク) をスタートするためにこのショートカットを使うことができる。あるいは、簡単にアクセスするために、このショートカットをタスクバーにピン留めすることもできる。

### マンドクを Mac OS X にインストールする

Mac OS X 10.9以降の入っているコンピュータにインストールバンドルは必要ない。このオペレーティングシステムの標準インストールに Emacs 以外はすべて入っているからである。マンドクをインストールするには Emacs のバージョン24.4以降を使うこと。([15] から入手できる千葉大学の山本光晴によるパッチを含むバージョンを推奨する。このバージョンはユニコードの全範囲をサポートしているからである)。このリンクから zip アrchiveをダウンロードし、ダウンロードしたファイルをクリックして展開し、そのファイル Emacs をあなたのアプリケーションフォルダにドラッグするだけである。

Emacs をインストールしたなら、今度はマンドクのパッケージをインストールしアクティベートする必要がある。これはさまざまな方法で行うことができるが、ここではその内の二つを概説する。どちらでもやりやすい

### Activating Mandoku: Method 1

One simple way of activating Mandoku is to download the file `activate-mandoku.el` [1], open it in Emacs and then execute it by opening the menu item “Emacs Lisp” and clicking on “Evaluate Buffer.” Emacs will now start to install mandoku and all other packages that are required by it.

### Activating Mandoku: Method 2

Open Emacs and find the “\*scratch\*” buffer. This buffer might not be listed on the Buffer menu, but you can always find it in the Buffer list, which is available on the Buffer menu under “List all buffers”. Once in the \*scratch\* buffer, copy the following code and paste it there:

```
(progn (require 'package)
      (add-to-list 'package-archives
                  '("melpa" . "http://melpa.org/packages/"))
      (package-refresh-contents)
      (package-initialize)
      (package-install 'mandoku))
```

Next, move the cursor to the very end of the buffer, after the last closing parenthesis, “)”, and finally press “Control-x” followed by “Control-e.” This should get things going: Emacs will install Mandoku and other packages that are required.

### Installing Mandoku on a Linux computer

To use Mandoku, you will need to install Emacs version 24.4 or later. How this is done depends on your system. Usually there will be a package manager that can be used for this task<sup>7</sup>.

Once Emacs is installed in a sufficiently recent version, simply proceed with adding the Mandoku package as detailed above for Macintosh-based computers using either Method 1 or Method 2.

と思う方を使って欲しい。

### マンドクをアクティベートする：方法1

マンドクをアクティベートする一つの簡単な方法は、[1] のテキストファイルをダウンロードし、それを Emacs で開け、それからメニュー項目「Emacs Lisp」を開けて「Evaluate Buffer」をクリックして実行することである。すると Emacs はマンドクとそれが必要とする他のパッケージすべてをインストールし始める。

### マンドクをアクティベートする：方法2

Emacs を開けて「scratch」バッファを見つける。このバッファはバッファメニューには入っていないかもしれないが、バッファメニューの「List all buffers」から入手できるバッファリストには常に入っている。\*scratch\* バッファに入ったら、以下のコードをコピーし、そこにペーストする：

次に、カーソルをこのバッファの一番最後、最後のとじカッコ「)」の後ろに持って行き、「Control-x」続いて「Control-e」を押す。これであまく進むはずである。Emacs はマンドクと必要な他のパッケージをインストールする。

### Linux コンピューターにマンドクをインストールする

マンドクを使うためには Emacs のバージョン24.4以降をインストールする必要がある。どのようにしてそれを行うかはシステムによって異なる。通常、このタスクのために使うパッケージマネージャーがあるはずである<sup>8</sup>。

十分に新しいバージョンの Emacs をインストールしたなら、Mac OS X のコンピューター用に上記した詳細に従って方法1か方法2を使い、マンドクパッケージを加える。



## Starting Mandoku for the first time on Macintosh or Linux

With the bundle for Windows, Emacs is set up to automatically start Mandoku when the program starts. On other systems, however, Mandoku has to be launched by the user. To start using Mandoku for the first time, you will have to manually ask it to show the catalog. This is done by typing “M-x.” (On Mac OS X, “M” here usually means the “Command” key, but in some configurations it could also be the key labelled “Option” or “alt”. Just try to press each of these keys at the same time as “x” until you see “M-x” appear at the bottom of the Emacs application window). Once you release the keys, you will see a prompt at the bottom of the Emacs window; this is called the minibuffer. At the prompt, now type “mandoku-show-catalog”, followed by the Enter key. This will initialize the mandoku package. This process will also ensure that the catalog will load right away the next time, so that you will not need to go through this process again.

Mandoku will now ask you where you want to place the files related to use of the Kanseki Repository. This will mostly be texts, but also other data files, data you produce while working with texts, etc. The program will suggest “~/krp” as the default location, i. e., the folder “krp” in your home directory. You can either accept this by pressing Enter, or specify a different path. From this point onwards the process is the same for all systems, as described in the next section.

## Connecting to GitHub and installing the workspace

For some time Emacs will go on downloading and installing more packages that are needed for operation. Some messages that report on the progress of these activities will appear on the screen. Eventually, when download and installation is complete, Emacs will connect to GitHub and download your workspace from there (if it exists).

Emacs will now ask you for your GitHub account name and password. This will be used to ask the GitHub site to create an authentication “token” for

## Macintosh か Linux 上でマンドクを初めて起動する

Windows 用のバンドルでは、プログラムが起動するとマンドクが自動的に起動するように Emacs は設定されている。しかし、その他のシステムでは、マンドクはユーザーが起動しなければならない。初めてマンドクをいはじめる時には、目録を表示するよう手動でマンドクに求めなければならない。これは「M-x」とタイプすることで実行できる。(Mac OS X ではこの「M」は通常「command」キーに対応するが、環境設定によっては「option」か「alt」キーであることもある。Emacs のアプリケーションウィンドウの下部に「M-x」が出るまでそれぞれのキーを「x」と同時に押すことを勧める)。ひとたびキーを離すと Emacs ウィンドウの下部にプロンプトが現れる。これはミニバッファと呼ばれる。このプロンプトに「mandoku-show-catalog」とタイプし、「Enter」キーを押す。これでマンドクパッケージが初期化される。このプロセスはまた、次回も目録が即座にロードされることを確実にする。これで再びこのプロセスを踏む必要がなくなる。

するとマンドクは、漢籍リポジトリの使用に関連したファイルをどこに置きたいか聞いてくる。これはほとんどがテキストであるが、他のデータファイルもある。テキストで働いている間に作成したデータなどである。プログラムはデフォルトのロケーションとして「~/krp」、すなわちホームディレクトリの「krp」フォルダーを提案してくる。「Enter」を押してこれを受け入れることもできれば、異なるパスを指定することもできる。これ以降のプロセスはすべてのシステムで同じで、それは次の項で説明する。

## GitHub に接続し workspace をインストールする

しばらくの間、Emacs はオペレーションに必要な他のパッケージのダウンロードとインストールを続ける。これらのアクティビティの進行を報告するメッセージがスクリーンに現れる。そして最終的にダウンロードとインストールが終了すると、Emacs は GitHub に接続し、そこからあなたの workspace を（それが存在する場合）ダウンロードする。

この時、Emacs はあなたの GitHub ユーザー名とパスワードを聞いてくる。これは GitHub サイトにああなたの認証「トークン」を作ることを求めるために使われる。

you. This token is stored and used for subsequent access from the particular computer you are using. Note that you will need to enter the password only once and it will not be stored; only the token will be stored. The startup process might take a while, but this is only necessary for the first time. If you do not want to log in to GitHub now, you can skip this step and start using Mandoku right away. Using your GitHub credentials will be necessary later when you want to download texts from the Kanseki Repository for local use.

### Layout of the krp folder on the computer

Mandoku will store all of the files that are created, edited or downloaded during its operation in a folder hierarchy that usually has the name `krp`. On Windows, this is generally placed in the root directory, while on other systems it is usually in the user's home directory. In this section, the layout and meaning of this folder will be described.

Table 3.1: The folders in `krp`  
表 3.1 `krp` 中のフォルダー

name	edit	description	備考
KR-Gaiji	part	List of non-system characters	外字表と画像
KR-Workspace	yes	Workspace shared with website	ワークスペース、ウェブサイトと共用
images	no	Facsimile images	デジタル・ファクシミリ
index	no	Index for local files	ローカル・テキストのインデックス
meta	part	Catalog files	目録
system	no	Some files used by mandoku	マンドクのシステムファイル
temp	yes	Temporary files, can be deleted occasionally	臨時的な物、削除可能
text	yes	Texts downloaded for local use	KR からコピーされたテキスト
work	yes	Additional files not from KR	ユーザーのファイル
bin	no	Windows only: Emacs, git, python and other programs	ウィンドウズのみ: Emacs, git など

Table 3.1 shows a typical list of the folders present in the `krp` folder, together with a short description. Their status (whether and to what extent they are editable) is indicated in the second column, “edit”. Some of the folders are used only by the system, so users should not edit their contents directly, to prevent problems arising from accidentally deleting or corrupting a file. Thus, edit is set to “no”. Other folders contain both files used by the system as well as files that a user can edit if necessary. The `meta` and `KR-Gaiji` folders are examples of this.

Users will usually edit some or all of the files in the `KR-Workspace`, `text`, and `work` files. The `KR-Work-`

このトークンは保管され、あなたの使用する特定のコンピュータからのそれ以降のアクセスに使われる。パスワードは一度入力するだけでよく、それは保管されず、トークンが保管されることに留意すること。この起動プロセスには少々時間がかかるかもしれないが、これは初回のみ必要なだけである。今の段階では GitHub にログインしたくなければ、このステップを省略し、すぐにマンドクを使い始めてもよい。後になって漢籍リポジトリからローカルコンピュータで使用するためにテキストをダウンロードする時には、あなたの GitHub 信用情報を使う必要が出てくる。

### コンピュータ上の krp フォルダーのレイアウト

マンドクは通常、そのオペレーションの間に作成され、編集され、ダウンロードされたファイルのすべてを `krp` という名前を持つフォルダー階層に保存する。Windows では、これは一般的にルートディレクトリーに置かれるが、他のシステムでは通常、ユーザーのホームディレクトリーに置かれる。この節ではこのフォルダーのレイアウトと意味を説明する。

表 3.1 は `krp` フォルダーの中に置かれているフォルダーの典型的なリストを短い説明と共に示している。2 列目「edit」にはこれらのフォルダーの状態（編集可能か、そうであればその程度）が表示されている。システムだけが使用するためのフォルダーがいくつかあり、ユーザーはそれらのコンテンツを直接編集してはならない。ファイルを図らずも削除してしまったり破損したりすることで発生する問題を防ぐためである。それゆえに、「edit」は「no」に設定してある。他のフォルダーはシステムが使うファイルと必要であればユーザーが編集できるファイルの両方を含んでいる。`meta` と `KR-Gaiji` フォルダーがその例である。

ユーザーは通常、`KR-Workspace` や `text`, `work` の中





space is a copy of the folder of the same name on GitHub and usually these two folders will be kept synchronized<sup>8</sup>. text is where the text files downloaded from the Kanseki Repository are stored. The texts contained in this folder can be edited by users as needed. work, on the other hand, is a folder in which users can place their own files, to make them available to the system. This is an advanced usage that is described in the online documentation.

## Using Mandoku

Mandoku is a package that extends the functionality of the editor Emacs. This means that all of the core functions of Emacs, as well as any other extension users have installed are all available while using Mandoku. Altogether, this puts a very powerful system at your service. Using Mandoku might seem daunting at first and it will take time to get used to the way things are done in Emacs, but it is definitely worth the effort and time. This is not the place for a general introduction to Emacs, however, so here below discussion will be limited to introducing some of the main functions available in Mandoku.

## Browsing

As on the web site, Mandoku allows the user to browse the contents of the Repository. This can be done easily from the catalog<sup>9</sup> by clicking on the link “Kanseki Repository 漢籍リポジトリ” or by positioning the cursor on any of the underlined, blue characters making up this phrase and then pressing the “Enter” key. This will display the familiar list of six main divisions in the catalog and by activating any of the links the desired subsection of the catalog can be viewed. A list of texts will be displayed, from which texts can be accessed for reading, as shown in Figure 3.14.

It is important to note here that the text has been downloaded from the server to a temporary location to allow browsing of the text. If the text is going to be read, it should be downloaded and made available in the private local library of the user. This is the meaning of the second line of the text shown in Figure 3.14. “# Don't edit this file. If you want to edit, press

のいくつかの、あるいはすべてのファイルを編集する。KR-Workspace は GitHub 上にある同じ名前のフォルダのコピーであり、通常、これら二つは同期している<sup>9</sup>。text は、漢籍リポジトリからダウンロードされたテキストファイルが保管される場所である。このフォルダに含まれるテキストは、ユーザーが必要に応じて編集できる。一方、work は、システムに提供するために、ユーザーが自らのファイルを置くことのできるフォルダである。これはオンライン資料の中で解説されている高度な用法である。

## マンドクの活用

マンドクはエディター Emacs の機能性を拡張するパッケージである。これは、マンドクを使っている間、Emacs の主要機能すべて、およびユーザーがインストールしたその他の拡張機能、これらのすべてを利用できるということである。総じて、たいへんパワフルなシステムを自由に使うことを可能にする。マンドクを使うことは最初は難しげに思えるかもしれないし、Emacs のやり方に慣れるのにも時間がかかるが、時間と労力をかけるに必ず値する。しかしながら、ここは Emacs の概要を紹介する場所ではないので、以下ではマンドクが提供している主な機能をいくつか紹介するに留める。

## 閲覧

マンドクは、ユーザーがウェブサイト上で行うようにリポジトリの内容をブラウズすることを可能にする。これは目録<sup>10</sup>から「漢籍リポジトリ」のリンクをクリックするか、あるいはカーソルをこの語句を構成する下線の付いた青文字に置いて「Enter」キーを押すことで簡単に実行できる。これにより目録中の 6 部分類が表示される。そしてこれらのリンクのどれか一つをアクティベートすることで、目録中の望みの小区分を表示できる。図 3.14 に示されているように、テキストのリストが表示され、そこから読みたいテキストにアクセスできる。

これは重要なポイントだが、テキストのブラウズを可能にするために、テキストはサーバーから一時的な場所にダウンロードされている。テキストを読むためには、それがユーザーのプライベートなローカルライブラリにダウンロードされ、そこで提供されていなければならない。これが図 3.14 に示されているテキストの 2 行目の意味である。「このファイルを編集しないでください。編集したい場合には、Control-c d を押してまずそれを

Control-c d to download it first.”

To download the file, do as follows. Press the “Control” and “c” keys together, release both and then press “d”. This will initiate the download process. At this point, your GitHub credentials (username and password) are needed. If you have entered them previously you should not need to enter anything here, because a token to authenticate access to your GitHub account will already have been generated.

Once the download is completed, Mandoku will prompt you with the question “Fork repository and add remote?” As explained above, “fork” is a GitHub term for a copy (“clone”) of a text in the GitHub user account. If you answer “yes” to this question, a fork will be created and a reference to this forked text will be added to the text that was cloned. Such a reference is called a “remote”. Answering “yes”, results in the situation illustrated in Figure 3.15. Altogether there are now three distinct instances of the text. The first is owned by the user @kanripo and a second by the current user (@cwittern in this case). Both of these copies are located on GitHub, i. e., in the “cloud”. The third copy is on the computer of the current user, cwittern. The location from which a text has been cloned is usually recorded as a remote reference named “Origin”. Answering “yes” to the above question generated a fork, resulting in another remote, this time under the name “cwittern”. That is, the name of the GitHub user serves as the name of the remote in this case. If the answer to the question had been “no”, no fork would have been created, and no additional remote added. Thus the text in @cwittern’s user account on GitHub would not exist and there would be just two copies of the text.

The situation can also be confirmed from within Mandoku. The Emacs package **Magit**, which is installed together with Mandoku, is used to interact with **git**. It is configured to be activated by pressing “Control-x” followed by **g**. (This displays an overview screen.) Pressing “y” now displays the branch manager<sup>10</sup>. If **Magit** is called while viewing newly downloaded text, the display appears similar to Figure 3.16. As can be seen, all available branches for the remotes are

ダウンロードしてください」

ファイルをダウンロードするには次の指示にしたがうこと。「Control」キーと「c」キーを同時に押し、両方を離し、それから「d」を押す。これがダウンロードのプロセスを開始させる。この時点で、あなたのGitHub 信用情報（ユーザー名とパスワード）が必要とされる。これらを以前すでに入力している場合、ここでは何も入力する必要はない。あなたのGitHub アカウントにアクセスする認証のトークンはすでに作成されているからである。

ダウンロードが完了すると、マンドクは「Fork repository and add remote? (リポジトリをフォークし、リモートを追加しますか?)」と聞いてくる。上で説明したように、「フォーク」はGitHub用語でGitHub ユーザーアカウントにあるテキストをコピー（「クローン」）することである。この問いに「yes」と答えたなら、フォークが作成され、このフォークされたテキストのリファレンスがクローンされたテキストに追加される。このようなりファレンスは「リモート」と呼ばれる。「Yes」と答えると図3.15に図解されるような状況になる。これでテキストは全部で三つ存在することになる。一つ目はユーザー @kanripo が所有し、二つ目は現在のユーザー（ここでの場合は@cwittern）が所有する。これらのコピーは両方ともGitHub上、すなわち「cloud」の中に位置する。三つ目のコピーは現在のユーザー、cwitternのコンピューター上にある。テキストがそこからクローンされたという場所は、通常、「Origin」と名付けられたリモートリファレンスとして記録される。上記の問いに「yes」と答えるとフォークが作成され、今度は「cwittern」の名前の下に別のリモートを生み出される。つまり、GitHub ユーザーの名前がこの場合リモートの名前となる。問いへの答えが「no」であった場合、フォークは作成されず、何のリモートも追加されない。それで、GitHub上の@cwitternのユーザーアカウントの中のテキストは存在せず、テキストは2コピーのみとなる。

この状況はまたマンドクの中からも確認できる。マンドクと一緒にインストールされた Emacs パッケージ **Magit** は **git** と対話するのに使われる。これは「Control-x」続いて「g」を押すことでアクティベートするように設定されている。（これはオーバービュースクリーンを表示する）。「y」を押すと、今度はブランチマネージャーを表示する<sup>11</sup>。新しくダウンロードしたテキストを読んでいる間に **Magit** が呼び出されると、図3.16のようなディスプレイが現れる。見てもらえばわかる通り、



displayed. The cursor keys can be used to move to any of the branches to view it. This retrieves and displays the text of this branch.

### Searching for titles

A special function key, “F7” has been set up for searching titles. After pressing this key, a prompt will be displayed at the bottom of the screen: “Mandoku | Search for title containing: “. Simply enter some characters and press Enter. A search will be conducted of all titles containing the entered characters. The list of titles displayed is the same as on the web site and the names of dynasties and authors are similarly included.

From the list of titles displayed in the search results, the desired one can be selected by moving the cursor to the line and pressing Enter. Alternatively, the mouse can be used to click on a title.

### Searching within the texts

Pressing “F6” will initiate a full-text search for a specified character string. This command is intended to be used when reading a text and trying to find other occurrences of a term in the text that is displayed. This command will take the six characters starting from the cursor position and display them after the prompt “Search for: “. This search string can then be edited as needed. The search can then be initiated by pressing enter.

The resulting display, in this example for the term 太平經, is similar to the one on the web site as shown in Figure 3.17. In this case there are 406 matches. The location is a link that can be followed to display the text. Emacs will show up to 2000 hits in the default setting.

### A very short crash course on git: commit, push and pull

The git software is used to keep track of branches and versions of the texts on your computer. It also works behind the scenes on GitHub. It is important to follow some simple rules when using it.

リモートに利用可能なブランチすべてが表示される。カーソルキーはそれをビューするためにいずれかのブランチに移動するのに使うことができる。それはこのブランチのテキストを取得し表示する。

### 題名で検索

特別のファンクションキー「F7」を題名検索のために設定している。このキーを押すとスクリーンの下部にプロンプトが表示される:「Mandoku | Search for title containing:」。ここに文字を入力して「Enter」を押せばよい。入力した文字を含む全ての題名の検索が実行される。表示される題名のリストはウェブサイト上のものと同じで、王朝や著者の名前も同様に含まれている。

カーソルを特定の題名に移動させ「Enter」を押すことで検索結果に表示された題名リストから望むものを選択できる。あるいは、マウスを使って題名をクリックしてもよい。

### テキスト内を検索

「F6」を押すことで特定した文字列の全文検索が開始される。このコマンドは、テキストを読んでいて、ある用語をそのテキスト内で検索したい時に使うよう意図されている。このコマンドはカーソルの位置から始めて6文字を取り、プロンプト「Search for:」の後ろに表示する。必要であればこの検索列は編集できる。その後、「Enter」を押すことで検索を開始する。

用語「太平經」の例において結果として出て来る表示は、図3.17に示されるように、ウェブサイト上のそれと似ている。この場合、406の検索結果がある。「location」はリンクになっており、テキストを表示するためにアクティベートできる。デフォルトとして Emacs は1ページに2,000まで検索結果を表示する。

### gitの「コミット、プッシュ、プル」についてのとても短い速成コース

gitソフトウェアはあなたのコンピューター上のテキストのブランチやバージョンを把握するのに使われている。これはまたGitHubの舞台裏でも働いている。これを使っている時にはいくつかの簡単なルールに従うことが重要である。

### ・ Commit

Once a file has been edited and saved to the computer, `git` needs to be told about this change. This action is called “to commit this file”. As a result of this action, the current state will be saved to `git`'s internal database (the `git` term for this is “stage”) to be retrieved later. It is now possible to change to another branch as explained above. In `Magit` (activated by pressing “Control-x” then `g`” from any text downloaded from the repository), the commit action is initiated by pressing “`c`”. A pop-up window with instructions on how to proceed will then be displayed.

### ・ Push

After a change has been committed, the changed file is still only available on the local computer in its current state. Any connected folder (a “remote” as `git` calls it) will need to be informed about the changes, so that they are available there as well. This action is referred to as a “push” to the remote. In `Magit` this is done by pressing “`p`” and then following the instructions. You will need to select the `remote` and the `branch`, but usually the defaults will be fine. As a result, the change will appear on the remote repository as well, for example on the connected folder in the GitHub user account.

### ・ Pull

`Pull` is the inverse of `push`, it will check the connected remote folder to see if there are any changes that need to be incorporated into the branch that is currently on display. In `Magit`, this is initiated using the “`F`” key.

## Too complicated? Automate all of this

If this sounds a bit complicated and daunting, you can relax. `Mandoku` has a setting that can automate all this for you. You can enable or disable this setting on the `Mandoku` menu. It works by periodically committing all relevant files, pushing them to the remote if necessary, and pulling changes from the remote. If no Internet connection is available, this will be done later once a connection is available again. This function will

### ・ コミット

ファイルが編集されコンピューターに保存されると、`git`はこの変更についての報告を受ける必要がある。このアクションは「このファイルをコミットする」と呼ばれる。このアクションの結果、現在の状態が `git` の内部データベースに保存され、それが後でまた取得されるのに使われる。上記した通り、これで別のブランチに変えることが可能になる。`Magit` (レポジトリからダウンロードしたどのテキストからも「Control-x」続いて「`g`」を押すことでアクティベートする)において、これは「`c`」を押すことで開始できる。するとどのように進めるかの指示を記したポップアップウィンドウが表示される。

### ・ プッシュ

変更が行われた後、まだその現在の状態では、変更されたファイルはローカルコンピューターでのみ入手可能である。すべての関連したフォルダー (`git`用語で「リモート」)は、そこでも変更されたファイルが入手できるように、この変更について知らせを受ける必要がある。このアクションがリモートへの「プッシュ」と呼ばれる。`Magit`において、これは「`p`」を押す、その後の指示に従うことで実行される。リモートとブランチを選択する必要があるが、通常、デフォルトで大丈夫である。結果として、変更がリモートのリポジトリ、例えば GitHub ユーザーアカウントの接続されたフォルダー上にも現れる。

### ・ プル

プルはプッシュの逆である。これは、現在表示されているブランチに取り入れるべき変更があるかを確認するために、接続しているリモートフォルダーをチェックする。`Magit`においてプルを開始するには「`F`」キーを使う。

## 複雑すぎ? 全部を自動化

これがちょっと複雑で難しすぎるように聞こえても大丈夫。`Mandoku`にはあなたのためにこれらすべてを自動化する設定がある。`Mandoku`メニューでこの設定をオンにしたりオフにしたりできる。これは、定期的に、すべての関連ファイルを、それが必要であればリモートにプッシュし、またリモートから変更をプルする。インターネットに接続していなければ、後で再び接続した時に行われる。この機能はまた、GitHub ユーザーアカウントか



also automatically get all your texts from your GitHub user account if they are not already downloaded. While this will work most of the time, there could be situations when something goes wrong. In this event, it is good to know at least a little about what is going on and what could have gone wrong. If in doubt, you can go to the mailing list for Mandoku users<sup>11</sup> and ask a question there.

#### Notes

- 1 In fact, there is also a proper catalog with detailed information about each text. It is currently available only in the “stack room” at [13], but in the future it will become more integrated into the “reading room.”
- 2 Where available, the serial numbers of the project in question have been used. If no such number is available, such as in the case of the DZJY, other widely used numbers have been added; in this case the serial number used by the Daozang Jiyao Project.
- 3 Due to the importance of this for some research questions, it has been made possible to adapt this list for specific needs. This is one of the more advanced usages which is explained later.
- 4 In some cases, a portable version can also be used on removable media, such as USB memory drives.
- 5 These installation instructions are also available online at [10]. Since the instructions there will be more up-to-date, with links to required items, it is advisable to visit that page when installing the program.
- 6 This installs `cygwin` and additional supporting software, including Emacs, in a way that essentially provides an operating environment similar to Linux and Mac OS X. `gnupack` is available from [9]. Be sure to use the “development” version.
- 7 Many systems based on Debian, such as all Ubuntu systems and all Linux Mint versions come with fairly old versions of Emacs by default. If you are using such a system, it is advisable to build Emacs from the released tarball or the git repository. You can prepare for this by first installing the package `build-essential` and then issuing the command “`sudo apt-get build-dep emacs24`”. The whole process is described here, for example at [5]. This method is also valid for later versions of Ubuntu.
- 8 For an explanation of how this is done, please see below “Commit, push and pull”.
- 9 If the catalog is not visible, there are several ways to display it. First look in the “Buffers” menu for an entry called “mandoku-catalog.txt”. Or, if the “Mandoku” item is visible in the menu, use “Browse-> Show catalog”. Otherwise, the easiest way is as described above under

らそのユーザーのすべてのテキストを（それがまだダウンロードされていない）自動的に取得する。これはほとんどの場合機能するが、それでも何かが行かなくなる状況もあり得る。そのような場合、少なくとも何が起っていて何がありうる原因が少しでも知っておくことは有用である。疑問があるなら、マンドクユーザーのメーリングリスト<sup>12</sup>に行き、そこから質問できる。

#### 注

- 1 実のところ、それぞれのテキストについての詳細な情報を含んだ正当な目録も存在する。それは現在のところ「書庫」の [13] においてのみ入手可能であるが、将来的には「閲覧室」に更に統合していく予定である。
- 2 入手可能なところでは、当該プロジェクトのシリアル番号が使われている。そのような番号が入手できない場合、道藏輯要のように、広く使用されている別の番号付け方法を適用している。この場合、「道藏輯要プロジェクト」で使用されたシリアル番号である。
- 3 これが重要なことである研究課題もあるので、個別のニーズに合わせてこのリストを適合できるようにした。これはより高度な用法の一つで、これに関しては後ほど説明する。
- 4 このサイトにはサインアップの手順が日本語で説明されている：[8]
- 5 場合によっては、USB メモリなどリムーバブルメディアに入れたポータブル版が使われることもある。
- 6 これらのインストールの指示はオンラインでも提供されている（[10]）。オンラインの指示は最新で、また必要事項へのリンクも付いているので、このプログラムをインストールする時にはこのページにアクセスすることを勧める。
- 7 これは `cygwin` とその他 Emacs も含めた補助的なソフトウェアをインストールする。そうすることで本質的には Linux や Mac OS X に似たオペレーティング環境を提供する。`gnupack` は [9] で提供されている。必ず「development」版を使うこと。
- 8 すべての Ubuntu システムやすべての Linux Mint バージョンなど Debian を基盤とする多くのシステムにはデフォルトとしてかなり古いバージョンの Emacs が付いてくる。このようなシステムを使っている場合、公開されている tarball や Git レポジトリから Emacs を組み立てることを勧める。まず `build-essential` パッケージをインストールし、それから「`sudo apt-get build-dep emacs24`」コマンドを発行することでそれを準備できる。この全プロセスは、例えば [5] で説明されている。この方法は Ubuntu のより新しいバージョンにも有効である。
- 9 どのようにしてこれがなされているかの説明は以下の「コミット、プッシュ、プル」を参照。
- 10 目録が見当たらない場合、それを表示する方法がいくつ

“Starting Mondoku for the first time”, i. e., using M-x mandoku-show-catalog.

- 10 A help screen for Magit with the most important commands can be called by pressing the key “?”. Control-h i displays Emacs manuals, which includes one for Magit.
- 11 The mailing list is open to all users and all kinds of questions related to the use of Mandoku or Kanripo can be asked. The group can be found here: [14].

かある。まず「バッファ」メニューの中の「mandoku-catalog.txt」というエントリーを見る。あるいは、メニューの中に「マンドク」項目があるなら、「Browse-> Show catalog」を使う。そうでなければ、最も簡単な方法は、上記「マンドクを初めて起動する」で説明した方法、つまり M-x mandoku-show-catalog を使う。

- 11 最も重要なコマンドの入っている Magit のヘルプスクリーンは「?» キーを押すことで呼び出せる。「Control-h」「i」で Emacs のマニュアルを表示でき、そこに Magit のものも含まれている。
- 12 このメーリングリストはすべてのユーザーに開放されており、マンドクや漢リポの利用に関するあらゆる質問をすることができる。このグループの場所は次の通り：[14]

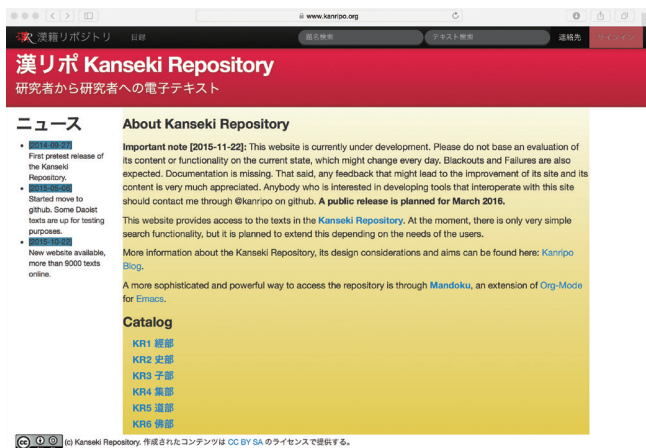


Figure 3.1: The start page of the Kanripo website, nicknamed the 'reading room'

図 3.1 「閲覧室」という愛称を持つ漢リポウェブサイトのスタートページ



Figure 3.2: The first page of the catalog

図 3.2 目録の最初のページ



Figure 3.3: Beginning of the listing for the 道藏輯要 Daozang jiyao

図 3.3 『道藏輯要』のリストの冒頭



Figure 3.4: Landing page for KR5h0001 太上中道妙法蓮花經

図 3.4 KR5h0001『太上中道妙法蓮花經』のランディングページ



Figure 3.5: Beginning of the first juan of the text KR5h0001 太上中道妙法蓮華經

図 3.5 テキスト KR5h0001『太上中道妙法蓮華經』の第一 juan の冒頭

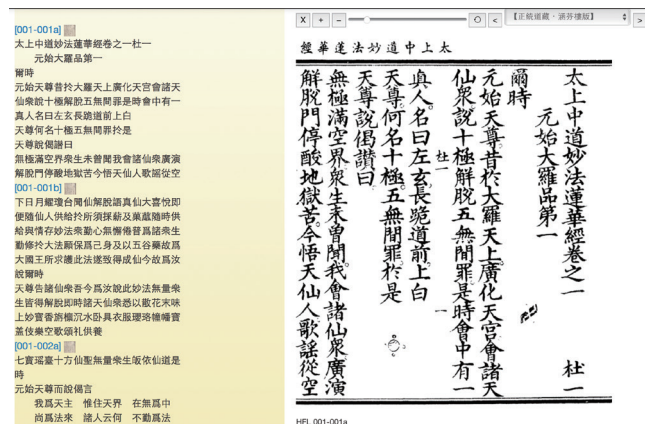


Figure 3.6: Display of the text with facsimile of the first page of the 涵芬樓 Hanfen lou edition of KR5h0001

図 3.6 KR5h0001の涵芬樓版 1 ページ目のファクシミリを付したテキストの表示

### 3 漢籍リポジトリの使い方

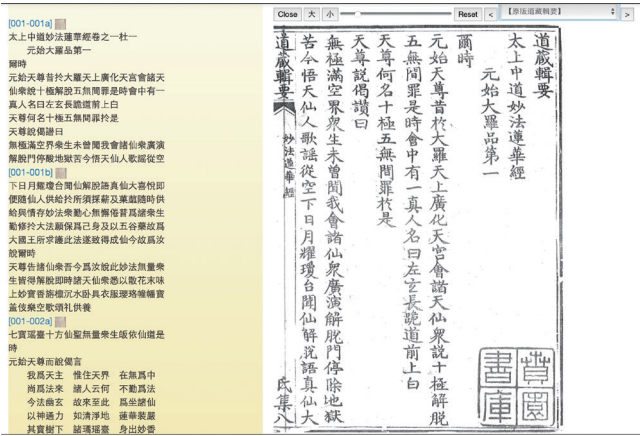


Figure 3.7: Display of the text with facsimile of the original DZJY edition of KR5h0001

図 3.7 KR5h0001のオリジナル『道藏輯要版』のファクシミリを付したテキストの表示



Figure 3.8: Full-text search for 七經, first page  
図 3.8 「七經」の全文検索, 1 ページ目

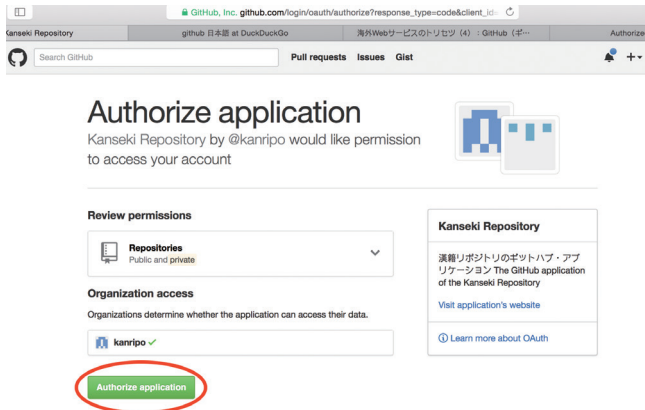


Figure 3.9: Request for authorization from GitHub  
図 3.9 GitHub からの認証リクエスト



Figure 3.10: The workspace of user krptest  
図 3.10 krptest ユーザーのワークスペース

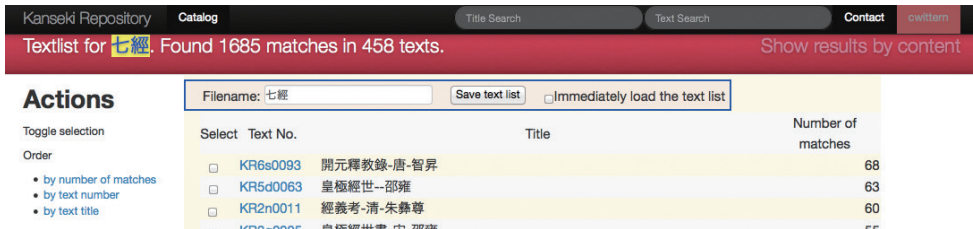


Figure 3.11: Saving a text list now becomes possible  
図 3.11 テキストリストの保存が可能になった





漢籍リポジトリ 目録 題名検索 テキスト検索 連絡先 rewritten

「全唐」を検索した。フィルター：\$元始 合計2434内1から20の表示。 テキスト毎まとめて表示  
フィルターを外す

Sort

- テキスト番号順
- 日付順
- 検索語順
- 検索語の前の字で

特別 (Filter)

- 元始

朝代

- 清 2462
- 唐 69
- 明 20
- 宋 18

部

- 集部 2280
- 子部 159
- 史部 139
- 道部 8
- 佛部 2

KR4f0005	御製詩集-清-高宗弘曆, 514, 4a	什幾與全唐一代詩人篇什相輝可謂
KR2o0020	御批資治通鑑綱目-宋-朱..., 52, 62b	必能保全唐主為之愴然由是無易太子 (夾注)
KR2b0037	御批歷代通鑑輯覽-清-高..., 67, 11b	冀得自全唐主魏范延光日卿今遠去事
KR2o0020	御批資治通鑑綱目-宋-朱..., 79, 43b	冀得自全唐主魏范延光日卿今遠去事 (夾注)
KR2m0014	欽定續文獻通考-清-嵇璜, 229, 33b	為斷自全唐之地分為遼宋幾與南北朝
KR2m0014	欽定續文獻通考-清-嵇璜, 229, 1b	及也且全唐之地宋與遼分馬氏之世遼
KR2m0014	欽定續文獻通考-清-嵇璜, 229, 2b	遼對而全唐之舊地沿革始完至於各代
KR2m0037	國朝宮史-清-高宗弘曆, 33, 19a	習者由全唐之詩沿波討瀾以上溯乎汾
KR2m0015	皇朝文獻通考-清-高宗弘..., 238, 7a	習者由全唐之詩沿波討瀾以上溯夫汾
KR4f0001	聖祖仁皇帝御製文集-清..., 237, 10a	習者由全唐之詩沿波討瀾以上溯夫汾
KR4h0161	御定全唐詩錄-清-徐伸, 2b	習者由全唐之詩沿波討瀾以上溯夫汾
KR4f0025	曝書亭集-清-朱彝尊, 37, 4b	目不覩全唐人之詩輒隨響附影未知正
KR4h0152	明文海-清-黃宗羲, 216, 11b	而全唐人之由質以成文故體備而並
KR4f0062	肇經室一集-清-, 5, 15a	于繙讀全唐人詩之人也然則亦各窮至
KR4f0062	肇經室一集-清-, 5, 15a	傳夥矣全唐人詩亦夥矣予未見繙讀唐
KR4h0152	明文海-清-黃宗羲, 273, 2a	履宋代全唐以公安代思下仲登必不為
KR4f0055	唐音彙編-明-胡震亨, 31, 5b	瑤補成全唐備成初盛而不克終客散去 (夾注)
KR2k0036	欽定四庫全書-清-于敏..., 69, 8b	此篇得全唐元度十體書補篇文其跡有 (夾注)
KR3b0012	武經總要-宋-曾公亮, 25, 19a	大軍得全唐社谷澤寇邊詔右僕射李靖
KR3k0013	册府元龜-宋-王欽若, 485, 6b	二鏡獲全唐宋王成器為韓州刺史獻馬

Figure 3.12: Searched for 全唐, but showing only texts that are in the list called 元始.  
図 3.12 全唐を検索し、その内の「元始」というリストにあるテキストのみを表示

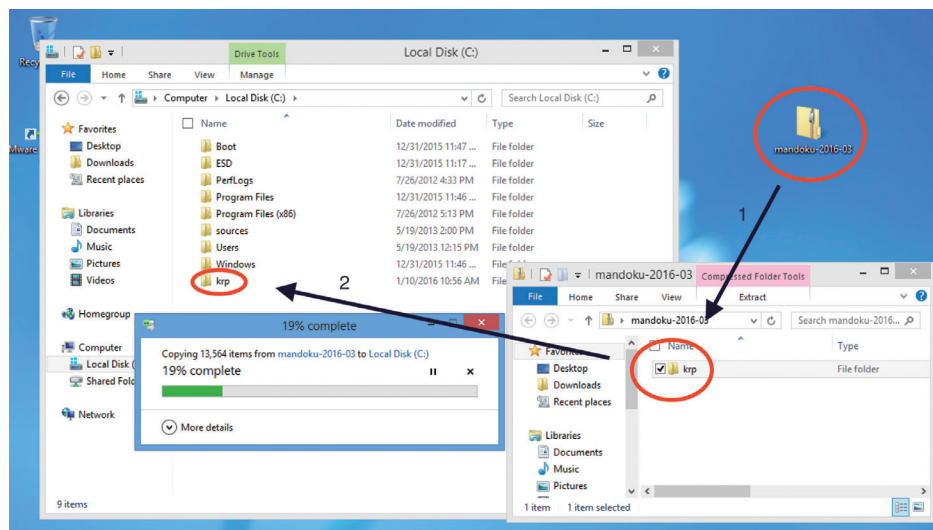


Figure 3.13: Dragging the folder krp to the root of drive C:  
図 3.13 krp フォルダを C: ドライブへ移動する

```

KR5h0001 HFL, - 目次 -- p00 BR:
# *_mode: mandoku_view; -*
# Don't edit this file. If you want to edit, press C-c d to download it first.
#+TITLE: 太上中道妙法蓮花經 / HFL
#+DATE: 2015-08-29 03:00:47.367493
* 版本
| CK-KZ| 【重刊道藏輯要】 |
| HFL| 【正統道藏・涵芬樓版】 |
| YP-CI| 【原版道藏輯要】 |
| ZTDZ| 【正統道藏・三家本】 |
* 目次
- 太上中道妙法蓮花經 0
- 太上中道妙法蓮花經 1
- 太上中道妙法蓮花經 2
- 太上中道妙法蓮花經 3
- 太上中道妙法蓮花經 4

```

Figure 3.14: A text opened for browsing in a temporary location  
図 3.14 一時的な場所でブラウズするために開けられたテキスト

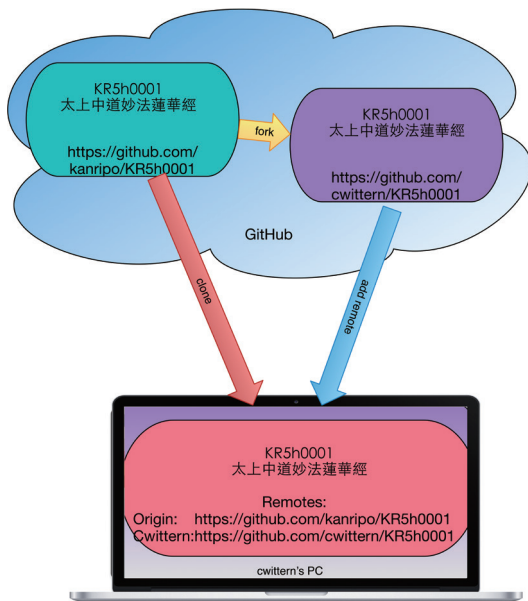


Figure 3.15: Situation after cloning, forking and adding remote of text KR5h0001

図 3.15 テキスト KR5h0001 のクローン、フォーク、リモート追加後の状況

```

KR5h0001 太上中道妙法蓮華經, 1 - HFL p001a04 BR:
# *- mode: mandoku-view; -*
#+DATE: 2015-08-29 03:00:47.178337
#+TITLE: 太上中道妙法蓮華經
#+PROPERTY: DZID DZ1432
#+PROPERTY: BASEEDITION HFL
#+PROPERTY: WITNESS HFL
#+PROPERTY: ID KR5h0001
<001-001a>
太上中道妙法蓮華經卷之一杜一
元始大羅品第一
爾時
元始天尊昔於大羅天上廣化天宮會諸天
仙衆說十極解脫五無間罪是時會中有一
真人名曰左玄長跪道前上白
天尊何名十極五無間罪於是
天尊說偈讚曰
無極湛空界衆生未萌明我命該仙聖廣濟
U:** - (Mandoku Index) Top (12,0) Git-master (mandoku-vie

Branches (1)
@ master [origin/master] Initial commit.

cwitern (https://github.com/cwitern/KR5h0001.git) (7)
cwitern/CK-KZ Initial commit.
cwitern/CK-KZ-jye Initial commit.
cwitern/HFL Initial commit.
cwitern/YP-C Initial commit.
cwitern/ZTDZ Initial commit.
cwitern/_data Initial commit.
cwitern/master Initial commit.

Origin (https://github.com/kanripo/KR5h0001.git) (7)
origin/CK-KZ Initial commit.
origin/CK-KZ-jye Initial commit.
origin/HFL Initial commit.
origin/YP-C Initial commit.

U:%* - magit-refs: KR5h0001 Top (4,0) (Magit Refs) Mon J
    
```

Figure 3.16: Text (upper half) and references in Magit branch manager (lower half) for text KR5h0001

図 3.16 KR5h0001 のテキスト (上半分) と Magit ブランチマネージャーでのリファレンス (下半分)

Location	Match	Source
* 太平經 (406)		
** 4,17b9	人于吉太平經一百七十卷其要曰且人	《KR5c0170 猶龍傳--賈善翔》...
** 1,11a4	授于吉太平經一百七十卷後漢明帝佛	《KR5c0169 太上混元老子史略--謝守灝》...
** 225,1a6	十二子太平經一百七十卷後漢書襄楷	《KR2m0009 文獻通考-元-馬端臨》...
** 0,2b2	人又傳太平經一百七十卷甲乙十部後	《KR5c0183 太上老君經律--》...
** 39,1a7	人又傳太平經一百七十卷甲子十部後	《KR5d0055 雲笈七籤-唐-張君房》...
** 9,10b2	始故作太平經一百七十卷象天地為數	《KR5e0001 太平經鈔-唐-》...
** 13,176c22	十戒并太平經一百七十篇為二時至漢	《KR6r0138 廣弘明集-唐-道宣》...
** 6,525c18	十戒并太平經一百七十篇為二時至漢	《KR6r0143 辯正論-唐-法琳》...
** 0,3b5	上篇并太平經一百五十卷桓帝時老君	《KR5a0092 太上洞玄靈寶無量度人上品妙經註解-
** 0,5a3	修又有太平經一百八十卷是蜀人于吉	《KR5a0177 玄風慶會錄--移刺楚才》...
** 3,569c5	君所出太平經三部村人驚異廣於經所	《KR6r0145 甄正論-唐-玄嶷》...
** 1,17a5	君所出太平經三部村人驚異於經所起	《KR5e0041 三洞珠囊--王懸河》...
** 666,4b2	今道家太平經也其經以甲乙丙丁戊己	《KR3k0012 太平御覽-宋-李昉》...
** 225,1a8	今道家太平經也其經以甲乙丙丁戊己	《KR2m0009 文獻通考-元-馬端臨》...
** 28,46b4	今道家太平經也其經以甲乙丙丁戊己	《KR3k0054 廣博物志-明-董斯張》...
** 21,1a9	今道家太平經也其經以甲乙丙丁戊己	《KR5c0308 歷世真仙體道通鑑--稽道一》...
U:** - Mandoku Index* Top (3,0) (mandoku-index-mode Wrap) Mon Jan 11 9:38PM 2.02		

Figure 3.17: The Mandoku Index for 太平經

図 3.17 太平經のマンドクインデックス

## 4 Some use cases for the Kanseki Repository

### 漢籍リポジトリの使用例

#### Overview

In this section four examples of how to make use of the data in the Kanseki Repository will be presented:

- (I) Correcting and reporting problems in the data
- (II) Translating
- (III) Collaborative work
- (IV) Analytic processing

These processes are increasingly complex, but each of these sections stands alone and can be read and understood separately.

#### (I) Correcting and reporting problems in the data

The editors of the Kanseki Repository rely on contributions from the user community to improve the accuracy of the texts. For reasons of efficiency and transparency we recommend that the reporting mechanism built into GitHub and git is used for this purpose. Here we present an example of how this can be done. The procedure is still somewhat tedious and overly complicated, since it follows the general protocol for reporting problems in the source code of the software. Hopefully, the procedure can be streamlined to reduce the steps needed by the user, but the more important steps will still be necessary. Here is an outline:

- (1) Spot an error or problem in the text. Make sure that the transcribed version and the digital facsimile where the problem has been noted represent the same edition of the text<sup>1</sup>.
- (2) Fork the text in question to your user account.
- (3) Open the text for editing in GitHub to make the

#### 概要

この章では漢籍リポジトリのデータを活用する四つの例を紹介する：

- (I) データ中の問題を訂正し報告する
- (II) 翻訳する
- (III) 共同作業
- (IV) 分析的な処理

これらのプロセスは順を追って複雑になるが、それぞれの節は独立しており、別々に読んでも理解できる。

#### (I) データ中の問題を訂正し報告する

漢籍リポジトリの編集者は、ユーザーコミュニティからの貢献に頼ってテキストの正確さを向上している。効率と透明性を高めるためには、その目的のために GitHub と git に組み込まれている報告メカニズムを使うことを勧める。それをどのようにして行うかの例をここに紹介する。その手順はソフトウェアのソースコードの問題を報告する一般的なプロトコールに準じているので、いくぶん面倒であり、ひどく複雑である。ユーザーが踏むべきステップを減らすためにこの手順が簡素化されることを期待するが、そうなったとしても主要なステップはまだ必要である。ここにそれらの概要を述べる：

- (1) テキスト中の誤りや問題を見つける。その問題が見つかった翻刻バージョンとデジタルファクシミリがテキストの同じ版を表していることを確認する<sup>1</sup>。
- (2) 当該テキストを自分のアカウントにフォークする。
- (3) GitHub で編集するためにテキストを開き、必要な

necessary changes.

- (4) Commit the changes. These should now become visible if you look at the page while signed into your account on kanripo.org.
- (5) If everything looks fine you can now report the problem to @kanripo by creating a pull request.
- (6) The editors will respond to your request and if there is no problem merge it into the branch on @kanripo.
- (7) The corrected text will now be available to all users.

The following example illustrates these steps in more detail.

### (1) The problem

In this case, while looking at a text discussing the 太極圖 *Taiji* diagram in Figure 4.1, we notice a problem: The characters for the five elements 五行 *wu xing* have not been entered, as seen by the fact that the text contains circled characters.

### (2) Forking the text

In Figure 4.1, the user `krptest` is currently logged in to kanripo.org. Clicking on the link labeled GitHub takes the user to the corresponding page on the GitHub account of @kanripo, as seen in Figure 4.2.

Clicking on the fork button will initiate the forking process and after a short time the text should appear in the user's account, as shown in Figure 4.3. When the user now views this text on the kanripo.org website, the copy in @krptest's account will be used in preference to the one in @kanripo's account<sup>2</sup>.

### (3) Open the text on GitHub for editing

With the forked version of the text ready, a click on the GitHub link in Figure 4.1 will now open the text directly in @krptest's account for editing. This is always the fastest and safest method to open the file for editing, since it will also be automatically located on the correct branch, WYG in this case.

Figure 4.4 shows the text in editing mode on GitHub. The location we want to change in line 145 is

変更を行う。

- (4) 変更をコミットする。これで、kanripo.org の自分のアカウントにログインしてこのページを見ると、これらの変更が反映されているはずである。
- (5) すべて問題なければ、プルリクエストをすることでこの問題を @kanripo に報告できる。
- (6) 編集者がこのリクエストに対応し、何の問題もなければ、その変更を @kanripo 上のブランチにマージする。
- (7) これで、訂正されたテキストがすべてのユーザーに提供される。

以下の例がこれらのステップをより詳細に説明している。

### (1) 問題

ここでの場合、図 4.1 にある太極圖を論じるテキストを見ている時に、私たちはある問題に気づいた：テキストには円で囲まれた文字が記載されているのに、その五行の文字が入力されていないことである。

### (2) テキストをフォークする

図 4.1 において、ユーザー `krptest` はすでに kanripo.org にログインしている。それで GitHub リンクをクリックすると、図 4.2 に示されているように、@kanripo の GitHub アカウント上の対応ページが表示される。

フォークボタンをクリックするとフォークするプロセスが開始される。ほんの少しの間を置いて、図 4.3 に示されているように、テキストがユーザーのアカウントに表示されるはずである。これでユーザーが kanripo.org ウェブサイトのテキストを見る時、@krptest のアカウントにあるコピーが @kanripo のアカウントのそれよりも優先して使われる<sup>2</sup>。

### (3) 編集のためにテキストを GitHub 上で開ける

テキストのフォークされたバージョンが準備されたので、図 4.1 にある GitHub リンクをクリックすると、今度はテキストが直接、編集するために @krptest のアカウントで開かれる。これが編集するためにファイルを開けるのに常に最も早く安全な方法である。自動的に正しいブランチ、この場合には WYG に置かれるからである。

図 4.4 は GitHub 上の編集モードにあるテキストを示している。145行目の私たちが変更したい箇所が高い



highlighted. Since the page on kanripo.org that showed the transcribed text and the digital facsimile side by side is no longer visible, it is a good idea to open the GitHub link in a new browser window. This will make it possible to view both at the same time, as shown in Figure 4.5. This figure also shows the character viewer used to input the desired characters. This will of course not always be necessary or even available, since it is a feature of the underlying operating system.

#### (4) Commit the changes

Figure 4.6 shows the situation where the characters have been added and the work is ready to be committed. It is good practice to include a commit message that explains what has been done. Figure 4.6 below shows the page reloaded for the user @krptest. The five elements are now readable.

#### (5) Make a pull request

If everything looks good, a pull request can be made. This is done from @krptest's GitHub page for the text. A pull request can be initiated by pressing the green "Create pull request" button shown in Figure 4.2. Care is needed to ensure that the correct branch is designated, as shown in Figure 4.8. Here, the "base fork" (the object of the pull request) is set to branch "WYG" of "kanripo/KR3a0023", and the "head fork" (the origin of the pull request) has been set to the same branch of "krptest/KR3a0023. If the branches are not set correctly, the pull request will be of no use. To make it easy to confirm that everything is correct, a comparison of the two versions is shown at the bottom of the screen. A system message "Able to merge" indicates that the branches can be automatically merged. This is a necessary condition for a successful pull request. Since everything is fine, @krptest now just needs to click on the green button "Create pull request" to actually issue the request. As the text next to the button suggests, this will initiate a discussion about the changes with the editors and possibly other users.

ライトされている。ここでは、翻刻テキストとデジタルファクシミリが並んで表示されている kanripo.org 上のページをもう見るができないので、新しいブラウザウィンドウに GitHub リンクを開けるとよい。図 4.5 に示されているように、両方を同時に見ることが可能になる。この図はまた、望みの文字を入力するために使われる文字ビューアーも表示している。もちろんこれは常に必要ではなく、またパソコンの環境によっては利用できないかもしれない。

#### (4) 変更をコミットする

文字が追加されて変更をコミットする準備ができた状況が図 4.6 に示されている。何がなされたかを説明するコミットメッセージを付けることを習慣にしておくといよい。その下の図 4.6 はユーザー @krptest のリロードしたページを示している。今では五行がきちんと載っている。

#### (5) プルリクエストをする

すべて問題がなければプルリクエストをすることができる。これはそのテキストの @krptest の GitHub ページから行う。図 4.2 にある緑のボタン「Create pull request」をクリックすることでプルリクエストを開始できる。図 4.8 に示されているように、正しいブランチを指定するよう注意しなければならない。ここでは「ベースフォーク」（プルリクエストの対象）が「kanripo/KR3a0023」の「WYG」ブランチに設定されており、「ヘッドフォーク」（プルリクエストの出所）が「krptest/KR3a0023」の同じブランチに設定されている。ブランチが正しく設定されていなければ、プルリクエストは何の役にも立たない。すべてが正しいことを簡単に確認できるように、スクリーンの下部に二つのバージョンの比較が表示されている。システムメッセージ「Able to merge」はブランチが自動的にマージされることを示している。これはプルリクエストの成功に必要な条件である。すべて問題がないので、@krptest は緑のボタン「Create pull request」をクリックしてこのリクエストを実際に実行する。このボタンの隣りに書かれているように、これによって編集者と、また事によれば他のユーザーとも、変更についての話し合いが開始される。

### (6+7) response from the editors and merge

Figure 4.9 shows how the pull request will look to users that have editing permission on @kanripo. Any discussion concerning this change will take place here, until the editors finally decide to press the button “Merge pull request”. From this point on, the change will be available to all users of @kanripo. However, users who forked or cloned their version before this change will have to update their copies to take advantage of the change<sup>3</sup>.

### Conclusion

The steps described here show how changes can be made to the text in the Kanseki Repository. The record of who initiated this change, when and why will continue to be available and can be verified by all users.

At the moment, the series of steps to initiate the change is still somewhat cumbersome. Future updates to the website will likely make this process easier, but the general sequence of forking, editing, committing, requesting a pull, discussing the change with the editors and the final merging by the editors will not change.

If this all sounds too complicated, or if the problem in the text not so easily resolved, it is always possible to open an “issue” on the text. This is done by visiting the text in the @kanripo account on GitHub. When you view a particular text, the “GitHub” link on that page will lead to this location. (If you have already forked the text, however, it will lead to your own copy, but from there it is easy to go to the “origin” via the link on the page of the forked text.) A small button with a circled “?” and the word “Issues” will bring you to the “Issues” page. Here, a new issue can be opened by clicking the green “New Issue” button. Give it a short title and describe the problem as clearly as possible, including the edition, juan number and page of the problem.

### (6+7) 編集者からの応答と統合

図 4.9 はプルリクエストが @kanripo 上で編集する許可を持つユーザーにどのように見えているかを示している。この変更についての話し合いはすべて、編集者が最終的に「Merge pull request」ボタンをクリックすると決定するまで、ここで行われる。ボタンをクリックした時点から、この変更は @kanripo のすべてのユーザーに提供される。しかしながら、この変更の前に自分のバージョンをフォークしたりクローンしたユーザーは、この変更を利用するためには、それらのコピーを更新しなければならない<sup>3</sup>。

### 結論

ここで説明したステップは、漢籍リポジトリのテキストにどのように変更を加えるかを示している。誰が、いつ、なぜこの変更を提起したかの記録は継続して入手でき、すべてのユーザーが確認できる。

現在のところ、変更を提起する一連のステップはまだいくらか複雑である。ウェブサイトの今後の更新でこのプロセスはより簡単になるだろうが、フォークし、編集し、コミットし、ユーザーがプルをリクエストし、編集者と変更を話し合い、そして最終的に編集者がマージするという全般的な順序に変わりはないであろう。

このプロセスが複雑すぎるようなら、あるいはテキスト中の問題を簡単に解決できそうにないなら、テキストの「issue」を開くという方法をいつでも使うことができる。これは GitHub 上の @kanripo アカウントにあるテキストにアクセスすることで行う。特定のテキストを見ている時、そのページの「GitHub」リンクからその場所に行くことができる。(しかし、すでにテキストをフォークしてある場合、このリンクはあなたのコピーを表示する。しかしそこからフォークされたテキストのページ上のリンクを通り、簡単に「origin」に行くことができる)。円で囲まれた「？」と「Issues」の小さなボタンから「Issues」ページに行くことができる。そこでは緑のボタン「New Issue」をクリックして新しい問題を開くことができる。それに短い題を付け、その問題の版、巻番号、ページ番号も含め、それをできるだけ明確に説明すること。



## (II) Text and translation

A frequent task that requires close reading is the preparation of a translation of a text. This section outlines the process for preparing a branch for translation and reformatting the text, and describes ways to facilitate this work.

The procedure can be broken down into several steps.

- ・ creating a new branch to contain the translation.
- ・ reformatting the text, adding punctuation where necessary.
- ・ adding the translation to the text.
- ・ pushing the translation to the GitHub account (optional).

### Creating a new branch

The translation will be located on a new branch, together with the text. Since a translation requires long and continued effort, web browsers are not well suited to this task, which is why **Mandoku** is used. It is assumed that the program has been installed and set up as described in Chapter 3. The text of interest can be looked up, for example using the “Title search” function, called by pressing F7. As explained, the text can be displayed (enter the text name on the line) and downloaded (cloned) using the command “C-c d”. In this case, a fork is created and the remote is added.

In most cases, it is convenient to start the translation from a **master** branch. As shown in Figure 4.10, with the text displayed, “C-x g” has been pressed and in the **Magit** overview screen at the bottom of the window, “y” has been pressed to display the branch manager. This branch manager is divided into three parts. The topmost one is labelled “Branches”; this displays the local branches. The next part, labelled “Krp-test”, displays the branches connected to the forked text of @krptest. The third part shows the branches in the original location of the text, @kanripo. The cursor selection is set to the local **master** branch. We want to create a new branch based on this one, so we press “b” (for branching) followed by “c” for “creating

## (II) テキストと翻訳

よく行われる作業でクローズリーディングが必要とされるのはテキストの翻訳である。この節では、翻訳のためのブランチの準備やテキストのリフォーマットのプロセスを概説し、またこの作業を行う方法について紹介する。

この手順はいくつかのステップに分けられる。

- ・ 翻訳を保持するための新しいブランチを作る。
- ・ テキストをリフォーマットし、必要な場所に句読点を付ける。
- ・ テキストに翻訳を追加する。
- ・ 翻訳を GitHub アカウントにプッシュする (任意)。

### 新しいブランチを作る

翻訳はテキストと一緒に新しいブランチに置かれる。翻訳には長く継続した努力が必要とされるので、ウェブブラウザはこのような作業に向いておらず、それで**Mandoku**が使われる。この**Mandoku**プログラムは、第三章で説明されているようなインストールと設定がなされているとみなす。興味のあるテキストは、例えば、F7を押して呼び出す「題名で検索」機能を使って調べることができる。すでに説明したように、テキストは「C-c d」を使って表示し（その行にテキスト名を入力する）、ダウンロードする（クローンする）。ここでの場合、フォークが作成され、リモートが追加される。

ほとんどの場合、マスター（**master**）ブランチから翻訳を始めると都合がよい。表示テキストと共に図 4.10に示されているように、「C-x g」が押されており、ウィンドウの下部に表示された**Magit** オーバービュースクリーンにはブランチマネージャーを表示するために「y」が押されている。このブランチマネージャーは三つの部分に分けられている。一番上の部分は「Branches」で、これはローカルのブランチを表示している。次の部分は「Krp-test」で、@krptestのフォークされたテキストにつながっているブランチを表示している。また三つ目の部分は、テキストの元の位置である@kanripoにあるブランチを示している。カーソルはローカルのマスターブランチに置かれている。ここで私たちは、このブランチを基にした新しいブランチを作ろうとしている。

and checkout”. (As mentioned, a help screen with available commands can be displayed by pressing “?”.) On the line at the very bottom of the window, you are now prompted for the name of the branch to start from. “master” is suggested because that is the branch where the cursor is positioned. Pressing enter will confirm this and generate the next prompt, asking for the name of the new branch. We choose a different name, “trans-en” to indicate that this is going to be the branch for a translation into English. The name is of course entirely up to the user, although it is advisable to avoid using capital letters, as recommended. After again pressing enter, the action is executed, and a new branch appears under local branches. “@” precedes the name to indicate that the branch is active. The header line of the text displayed at the top of the window changes to show that the active branch is “trans-en”. We can now close the branch manager and the Magit status window by pressing the key “q” twice. This will restore the window displaying the text at full size.

### Reformatting the text

Before actually starting the translation, we will reformat the text a little to make it possible to write the translation of the text on the same line as the text itself. This is useful for a raw translation, since it makes it easy to look for the translations of terms that have already been encountered. We use a “one phrase, one line” 一逗一行 format here. While editing in the file, it is a good idea to show the hidden markers that indicate page and line breaks. This can be done from the menu using “Mandoku > Display > Show markers” (Figure 4.11). As shown in Figure 4.12, a “¶” character appears at the end of every line. In the Mandoku text format, this indicates the end of a line in the base edition. Furthermore, page numbers are now displayed in full, to show not only the current page, but also the text number (KR6q0332) and the edition (X in this case, which stands for the 新纂大日本續藏經 Shin-san dainippon zokuzōkyō). This information is important for the correct functioning of the system, so it should not be deleted. However, the file can be rearranged as needed without causing problems, as long

それで、「[branching]」の「b」を押し、続いて「creating and checkout」の「c」を押し。（前にも述べた通り、「?」を押すことで利用可能なコマンドの記されたヘルプスクリーンを表示できる）。ウィンドウの一番下の行で、プログラムが使用するブランチの名前を聞いてくる。その名前として、カーソルがそこに置かれているので、「master」を提案してくる。「enter」を押すとこれが確定され、次の質問、新しいブランチの名前を聞いてくる。私たちは、これが英語翻訳のためのブランチになることを示すために「trans-en」という別の名前を選んだ。大文字だけを使わないという勧めに従うことは推奨されているものの、もちろん名前はまったくユーザー次第である。もう一度「enter」を押すとアクションが実行され、ローカルブランチの下に新しいブランチが現れる。名前の前の「@」はブランチがアクティブであることを示している。テキストを表示しているウィンドウ上部のヘッダ行が変わり、アクティブブランチが「trans-en」であることを示している。これで「q」キーを二度押すことでブランチマネージャーと Magit ステータスウィンドウを閉じることができる。これでテキストを表示しているウィンドウがフルサイズに戻る。

### 翻訳を開始する

実際に翻訳を始める前に、テキストと同じ行にその翻訳が書けるようテキストを少しリフォーマットする。これはすでに出てきた用語の翻訳を探しやすくするので、下訳の役に立つ。ここでは「一句一行」の形式を使う。このようにファイルを編集している間、改ページや改行を示す不可視文字を表示するとよい。これはメニューから「Mandoku-> Display-> Show markers」を使って行う（図 4.11）。図 4.12 に示されているように、各行の終わりに「¶」文字が見られるが、これはマンドクテキストのフォーマットではベース本における一つの行の終わりを意味している。更に、ページ番号がそのフルフォーマットで表示され、現在のページだけでなく、テキスト番号（KR6q0332）と版（この場合は「新纂大日本續藏經」を意味する X）も示している。この情報はシステムが正しく機能するために重要であり、削除してはならない。しかしながら、必要であれば、そして文字の順序さえ変更しなければ、ファイルは問題を起こすことなく整理し直すことができる。





as the sequence of characters is not altered.

### Starting the translation

Figure 4.13 shows the file with lines reformatted, ready for starting the translation. This can of course also be done line by line if the translator prefers. The markers can also be hidden again (using same menu command as above) to reduce distraction while working with the text. Figure 4.14 shows the result of translating the first few lines. Translation and text have been separated by a “tab” character. (To enter a “tab” use “C-q C-i”. This complicated key combination is necessary because the tab key is already assigned for other purposes.) The tab-stop has been set to 30 (M-x set-variable tab-width <enter>30).

### Pushing to the GitHub account.

If the translation is pushed to the GitHub account, it will also be available to @krptest when using the kanripo.org, and it can be used on other computers. It can also be used as a backup in case of emergency.

To initiate the push, we again press C-x g to call up Magit. Now the display looks like Figure 4.16. There are two things to note: (1) The push destination is absent (because the branch was created locally, not pulled from a remote) and (2) there are “unstaged changes”, which means there are changes to the file that have not yet been committed to the internal registry of changes. Before pushing, we therefore need to stage and commit the changes. The easiest way to do this is to press “c” two times. This will call up the buffer for editing commit messages, as shown in Figure 4.17. The cursor is at the lower part of the window, where the commit message is entered. All the other parts are for information only. The top part of the window shows the changes to be committed, while the bottom part below the cursor shows the files affected. After writing a short message to explain what has been done, pressing “C-c” twice will conclude the action. The part of the window about “Unstaged changes” now disappears.

We can now initiate the actual push. Pressing “P” (i. e., capital letter “P” or shift-“p”) brings up another popup buffer as shown in Figure 4.18. This buffer in-

### 翻訳を開始する

図 4.13は、必要に応じて行がリフォーマットされ、翻訳開始の準備が整ったファイルを示している。これはもちろん、翻訳者がそちらを好むのであれば、一行ずつ行うこともできる。より快適にテキストで作業できるように、再び不可視文字を非表示にしてもよい（上記と同じメニューコマンドを使う）。図 4.14は最初の数行を翻訳した結果を示している。翻訳とテキストは「tab」文字で分けられている。（「tab」を入力するには「C-q C-i」を使う。このいくぶん複雑な入力方法が必要なのは、tab キーがすでに別の目的に割当てられているからである）。tab-stop は30に設定されている（M-x set-variable tab-width < enter >30）。

### GitHub アカウントにプッシュする

翻訳が GitHub アカウントにプッシュされたなら、kanripo.org を使っている時に @krptest でも利用可能になり、他のコンピューターでも使えるようになる。緊急事態の場合のバックアップとしても使える。

プッシュを開始するには、Magit を呼び出すために再び「C-x g」を押す。この時、表示は図 4.16 のようになる。ここで留意点が二つある：(1) プッシュの宛先が欠けている（ブランチがローカルで作られており、リモートからプルされていないため）、及び(2) 「unstaged changes」がある、つまり内部レジストリーにまだコミットされていないファイルの変更がある。それゆえ、プッシュする前に、この変更をステージし、コミットする必要がある。これを行う最も簡単な方法は、「c」を二度押すことである。図 4.17 に示されているように、これがコミットメッセージを編集するバッファを呼び出す。カーソルがウィンドウの下部、コミットメッセージを入力すべき位置に置かれている。その他の部分はすべて情報のみである。ウィンドウの上部はコミットすべき変更を表示しており、下部のカーソルの下は変更を反映したファイルを表示している。何かなされたかを示す短いメッセージを書いた後、「C-c」を二度押すことでアクションが行使される。今では「Unstaged changes」についてのウィンドウの部分が消えている。

これで実際のプッシュを開始できる。「P」（つまり大文字の「P」あるいは shift-「p」）を押すと図 4.18 に示されているようなもう一つのポップアップバッファが

dicates what keys are available to conclude the action. Here the second section, labelled “Push trans-en to” is most relevant. The first item in this section is the one we need. This will push the commit we just made to the user account on GitHub. After pressing p to initiate this, the program might prompt for the GitHub user name and password if these have not been saved before, but usually it will do its magic and then display the commit message just entered for both the local and the remote branch.

The change is now available on GitHub and it is also visible on kanripo.org. However, we first need to tell the website that we now want to see the translation branch and not the master branch of this text. This is done by adding the line “KR6q0332=krptest/KR6q0332/trans-en” (“user account/text number/branch”) to one of the configuration files, as shown in Figure 4.26. Once this is set and loaded, the result should be visible for the user @krptest as in Figure 4.15. It should also be noted here, that the change on the GitHub repository will be **visible to everybody** who visits the page there, not just the owner. Therefore, if you do not yet want to share unfinished work, it’s better to use a private repository, as explained in the next part, or simply avoid pushing to the repository on GitHub.

### (Ⅲ) Workflow for accessing and sharing texts

Sometimes the reading and translating of a text is done by multiple members of a research group. In this case, a separate account should be set up for the research group, designated as an “organization” account, not a user account. The members of the group can then be added to the organization. This enables the group to work with a common, shared copy of the text in addition to, or instead of, maintaining a copy in every individual members’ account.

Here, we will assume that the account of the research group is called “krp-zinbun” and has the permissions necessary to create private repositories. Private repositories can not be directly forked from

表示される。このバッファはアクションを完了するのにどのキーが利用できるかを示している。ここでは二番目のエントリー「Push trans-en to」が最も適切である。私たちが必要としているのはその一つ目の項目である。これが、今作ったばかりのコミットを GitHub 上のユーザーアカウントへプッシュする。これを開始するために「p」を押した後、GitHub のユーザーアカウントとパスワードが以前に保存されていなければ、プログラムがそれを聞いてくるかもしれない。しかし通常はうまく行き、今入力したばかりのローカルとリモート両方のブランチのコミットメッセージを表示する。

これで変更が GitHub 上で入手可能になり、kanripo.org にアクセスするとそれを見ることができる。しかしながら、これからはテキストのマスターブランチではなく翻訳ブランチが見たいということを、まずウェブサイトに伝える必要がある。これは「KR6q0332=krptest/KR6q0332/trans-en」(user account/text number/branch) を図 4.26 に示されている設定ファイルの一つに追加することで実行できる。これがひとたび設定されロードされれば、図 4.15 にあるように、その結果がユーザー @krptest に見えるようになるはずである。GitHub リポジトリ上の変更はオーナーだけでなく、そのページにアクセスする誰にも見えるようになることにも留意すること。それゆえに、未完成の仕事でまだシェアしたくなければ、次の項で説明するプライベートリポジトリを使うか、あるいはただ単に GitHub 上のリポジトリをプッシュしなければよい。

### (Ⅲ) テキストにアクセスし共有するためのワークフロー

時として、研究班の複数のメンバーでテキストを読んだり翻訳したりすることがある。このような場合、その研究班用の別のアカウントを設定し、ユーザーアカウントとしてではなく「organization」(組織) アカウントとして指定しなければならない。そして研究班のメンバーはこの組織アカウントに追加される。そうすることで、各メンバーは、それぞれのアカウントに保持しているコピーに加えて、あるいはその代わりに、共通で共有されたテキストコピーを持つことができる。

ここでは研究班のアカウントを「krp-zinbun」と呼ぶことにし、プライベートリポジトリを作るのに必要な許可を持っているものと仮定する。プライベートリポジトリはパブリックリポジトリから直接フォークすること



public repositories; they have to go through some local clone. An easy way to create such a local clone is to use Mandoku. Simply display the text as explained above and then press “C-c d” to download the text. We do not want a fork in this case, so simply answer “no” when prompted to. This clone can then be pushed to a repository created as private on GitHub. A group administrator will have to set up the repository. (The details of this are beyond the scope of this introduction.)

We will also assume that the **master** branch of this text in the account @krp-zinbun will have the text all users see on kanripo.org, while individual users will have their own branches on which they do their work. Only when the work is ready, will this branch be merged into the common master. We will further assume that users will each be responsible for certain sections of the text, which they will prepare in private, and present the results of their work to the group later.

The steps that need to be completed for this task depend on whether the work is done on Mandoku or on the Kanripo.org website:

- For work on Mandoku:
  - Clone (=download) the text from @krp-zinbun
  - Create a branch
  - Prepare the text
  - Add translation and notes
  - Merge and push
- For work on Kanripo.org
  - Fork text to the user account
  - Create a branch
  - Prepare the text, add translation (this can be done outside the browser)
  - Merge and push
- Confirm visibility to all (both methods)

### Mandoku

The workflow on Mandoku is very similar to the one already described for translations prepared individually. The description offered now focuses on the differences.

ができず、ローカルクローンを経由しなければならない。そのようなローカルクローンを作る簡単な方法はマンドクを使うことである。上記で説明したようにテキストを表示し、そのテキストをダウンロードするために「C-c d」を押せばよい。この場合、フォークは必要ないので、それを聞かれた時には「no」と答えること。このクローンが GitHub 上でプライベートとして作られたりポジトリへとプッシュされる。このグループの管理者はリポジトリの設定をしなければならない。(その詳細はここでの説明の範囲外である)。

更に、アカウント @krp-zinbun にあるこのテキストのマスターブランチが、kanripo.org 上ですべてのユーザーが見ているテキストを持っているものと仮定する。一方、各ユーザーはそれぞれ自分のブランチを持ち、そこで自分の仕事をし、その仕事が完了した時にだけそのブランチを共通の master にマージするものとする。更に、ユーザーはそれぞれがテキストの異なる部分を担当し、それをそれぞれ個人で準備し、後でその結果をグループに提供するものとする。

この作業のために完了すべきステップは、仕事がマンドク上で行われるか kanripo.org ウェブサイトで行われるかによって異なる：

- マンドクでの作業の場合：
  - @krp-zinbun からテキストをクローン (=ダウンロード) する
  - ブランチを作る
  - テキストを準備する
  - 翻訳と注釈を追加する
  - マージとプッシュ
- kanripo.org での作業の場合：
  - テキストをユーザーアカウントにフォークする
  - ブランチを作る
  - テキストを準備し、翻訳を追加する (これはブラウザの外で行なってもよい)
  - マージとプッシュ
- 誰にも見えることを確認する (両方の方法)

### マンドクでの作業の場合

マンドクのワークフローは、すでに説明した個別で準備する翻訳の場合ととても似ている。ここでは異なっている点に焦点を当てて説明する。

## Preparation

With the setup above, user @krptest is a member of the research group and is starting to work on a part of the text assigned to her. She will first need to get the files from the private repository, where contributions from others might already be recorded. First, she will look for the text needed in the catalog and display the file that contains the line “Don’t edit this file. If you want to edit, press C-c d to download it first.” However, instead of using “C-c d” to get the file from the @kanripo account, this time she uses the command “Mandoku > Maintenance > Download this text from other account” available on the menu<sup>4</sup>. Selecting this function now prompts for the name of the account to use, to which she enters “krp-zinbun”. This will of course work only if she indeed has access to this repository. If everything goes well, the text will be cloned and will be available for editing, as shown in Figure 4.19.

## Starting work

Before she actually starts editing the text, @krptest will create a branch where she can work without being affected by others. This branch is named “work”, but any name will do. As shown in Figure 4.19, the user now creates a branch using the procedure explained above.

As soon as the line “@ work” appears under “Branches”, the branch “work” is active, setup is finished and the actual work can start. While working, a digital facsimile of this page can be displayed at any time by pressing C-c i, as shown in Figure 4.24.

When the work is finished, it can be committed as usual (C-x g, then c and c). The work is still only on the local computer and not available to anybody else.

## Merging and pushing

When @krptest is satisfied with her work, she will want to make it visible to the other members of the group. This is done by pushing it back to GitHub. Since she is on a different branch, however, she will

## 準備

ユーザー @krptest はこの研究班のメンバーであり、上記の設定の元で担当箇所の仕事を始めるところである。ここでまず、このユーザーは他からの貢献がすでに記録されているかもしれないプライベートリポジトリからファイル入手する必要がある。まず目録から必要なテキストを探し、「Don’t edit this file. If you want to edit, press C-c d to download it first.」を含むファイルを表示する。しかし、@kanripo アカウントからファイル入手するための「C-c d」を使う代わりに、今回は「Mandoku-> Maintenance-> Download this text from other account」エントリーを使うことにする<sup>4</sup>。この機能を選択することで、使用するアカウント名を入力するよう促され、そこに「krp-zinbun」と入力する。これはもちろん、このリポジトリへのアクセス権を持っていてこそ可能である。すべてがうまく行ったなら、図 4.19示されているように、テキストがクローンされ、編集に利用できる。

## 仕事を開始する

実際にこのテキストを編集し始める前に、@krptest はこのユーザーが他から影響されることなく働けるブランチを作る。このブランチは「work」と名付けられるが、他のどんな名前でもよい。図 4.19のスクリーンに示されているように、このユーザーは上記で説明された方法でブランチを作る。

リンク「@work」が「Branches」の下に現れたなら、この「work」ブランチは使用でき、設定は完了し、実際の仕事が始められるということである。仕事をしている間、図 4.24に示されているように、「C-c i」を押すことで、いつでもそのページのデジタルファクシミリを表示できる。

仕事を終了したなら、いつもの方法でそれをコミットできる (C-x g, そして c と c)。その仕事はまだローカルコンピューター上にあるだけで、他の誰にも提供されていない。

## マージし、プッシュする

@krptest がその仕事に満足した場合、このユーザーは研究班の他のメンバーにも見えるようにしたいと考える。これは仕事を GitHub にプッシュして戻すことで実行できる。しかしながら、このユーザーは別のブランチ上に



first have to merge her work back into the master. Before doing so, she will check if anybody else has pushed to master in the meantime. To make sure the master is up-to-date, she will go to the branch manager (C-x g and then “y”), and change to the master branch. (This is done by pressing enter on the line with “master”. “@” will then indicate the active branch. This requires that all changes have been committed.)

@krptest will now press “F” (shift-“f”) to pull changes from GitHub into master. On the popup screen in (Figure 4.21) this is done by pressing “u”. The master branch is now up-to-date and @krptest can merge her changes in and push. Pressing “m” to merge will display a screen with additional options, as in Figure 4.22. To see what will be merged, “p” can be pressed to generate a preview. This confirms the branches from which the merge (“work”) will take place and also shows what is going to be merged. The preview is displayed at the lower part of the screen, as in Figure 4.23. The lines of text in red are the lines that will be deleted; the green ones (only partially visible) are the lines that will replace them. Pressing “q” closes the preview, and if everything seems fine, “m” can be pressed to do execute the merge.

Figure 4.25 shows the new situation. After master the description now reads “ [origin/master: ahead 1]”. This means that we are now one commit ahead of the master branch from the remote “origin”. The bottom of the window shows that the origin/master is still at the previous commit. Pressing “P” (push) followed by “u” (destination “origin/master”) executes the push action. After this action, the master branch of KR5e0001 in the group’s account @kzp-zinbun will have the same content as the version on @krptest’s computer.

## Kanripo.org

### Forking the text/Creating a new branch

Depending on the policy of the team, a branch can either be created on a separate fork, or directly in the group’s repository. In this case, as can be seen in Figure 4.20, a branch is created without prior forking.

いるので、まずこの仕事をマスターにマージする必要がある。それをする前に、この仕事をしている間にも他の誰かがマスターにプッシュしていないかチェックする。マスターが最新のものであることを確認するために、ブランチマネージャーに行き (C-x g, そして「y」), マスターブランチに変更する (「master」の行で「enter」を押す。「@」は使用できるブランチを示している。これをするにはすべての変更がコミットされていなければならない)。

ここで @krptest は、変更を GitHub からマスターへとプルするために「F」(shift「f」)を押す。ポップアップスクリーン (図 4.21) では、これは「u」を押すことで行われる。これでマスターブランチは最新のものになり、@krptest はこのユーザーの変更をマージしプッシュできる。マージは「m」を押すことで行われる。これは図 4.22 に示されているように、追加オプションのスクリーンを表示する。何がマージされるかを見るためには「p」を押してプレビューを表示する。これがどのブランチからマージされるか (「work」)を確認し、何がマージされるかを表示する。図 4.23 において、下部がプレビューを表示している。赤い線が削除されるものを、緑 (部分的に表示) が置き換えられるものを示している。「q」を押すことでプレビューを閉じ、すべてが良さそうであれば「m」を押して実際のマージを行う。

図 4.25 が新しい状態を示している。マスターの後に「[origin/master: ahead 1]」とある。これは、今では私たちがリモート「origin」からのマスターブランチより 1 コミット分先であることを意味している。このウィンドウの下部には origin/master がこれより前のコミットの状態であることを示している。「P」(プッシュ)を押し、続いて「u」(宛先「origin/master」)を押すことでプッシュを実行する。このアクションの後、グループのアカウント @kzp-zinbun にある KR5e0001 のマスターブランチは、@krptest のコンピューターのバージョンと同じ内容を持つことになる。

## kanripo.org での作業の場合

### テキストをフォークする / 新しいブランチを作る

チームの方針次第で、ブランチを別のフォークに、あるいは直接グループのリポジトリに作成できる。ここでの場合、図 4.20 に示されているように、前もってフォークすることなくブランチが作られている。

## Preparing the text and translating

The text can then be directly edited in the browser on the GitHub site, or the text can be copied to a different place and then pasted back in later.

## Merging and pushing

The merge will again be initiated through a “Pull request” in the same way as above in “(I), (5) Make a pull request”. In this case, however, a member of the group will need to have the necessary permissions to complete the pull request by merging into master.

## Visible to all members

Regardless of how the change is initiated, it should now be visible on the website kanripo.org for all users who are members of the group @krp-zinbun, provided that they have requested this in their settings, as shown in Figure 4.26. (The settings are accessed most conveniently from the link to the file `global.cfg` on the profile page.) This means that all members, even those who are not using Mandoku are able to see the results.

## Conclusion

The combination of Mandoku and the Kanseki Repository provide new tools for collaboration. While these tools might initially seem complicated and cumbersome, they are easier to use than the explanation might suggest. In time, the tools will also improve. It should become possible for example to directly prepare printouts for research meetings without the need to use a word processor.

## (IV) Analytical processing

In this last section, some advanced ways of working with the Kanseki Repository are presented. The example research challenge considered here is: “For a given list of terms, what are the texts most relevant to these terms?”

To answer this question, every single one of these terms will have to be searched in the database. For each term, the texts that contain the term are noted

## テキストを準備し翻訳する

これで、テキストを GitHub サイト上のブラウザで直接編集することができる。あるいは、テキストを別の場所にコピーして、後でそれをペーストしてもよい。

## マージし、プッシュする

マージは上記「(I)(5)プルリクエストをする」と同じ方法で、「プルリクエスト」によって開始する。しかしここでの場合、グループのメンバーが、master にマージすることでプルリクエストを完了するのに必要な権限を持つ必要がある。

## すべてのメンバーに見えるようにする

変更がどのように開始されたとしても、それがグループ @krp-zinbun のメンバーであるすべてのユーザーに、彼らが図 4.26 に示されているようにそれを設定中でリクエストしたならば（設定にはプロファイルページの `global.cfg` ファイルへのリンクから最も簡単に行き着くことができる）、kanripo.org ウェブサイト上で見えるようにしなければならない。これは、すべてのメンバーが、マンドクを使っていない者でさえ、この結果を見られることを意味する。

## 結論

マンドクと漢籍リポジトリの組合せは共同作業のための新しいツールを提供する。最初は複雑で面倒に思えるかもしれないが、実は説明の方が実際に使うよりもずっと複雑である。そのうちにツールも改良されるであろう。例えば、それは研究会議のために、ワードプロセッサを使う必要なく、直接プリントアウトを準備するようになるはずである。

## (IV) 分析的な処理

この最後の節では、漢籍リポジトリを使って作業する、より高度な方法を紹介する。ここで検討する研究課題の例は以下の通りである：

「与えられた用語のリストについて、これらの用語に最も関連のあるテキストは何か？」

この質問に答えるためには、これらの用語の一つ一つをデータベースで検索しなければならない。そしてその一つ一つの用語について、その用語を含むテキストを記



and the number of matches per text is calculated. Obviously, this is a task that cries out for automation, but the question is how can it be done?

With Mandoku it is actually quite straightforward to do this. Org-mode, on which Mandoku is based, enables executable code to be embedded within documents, in the manner of so-called 'literate programming'. Such documents can also be exported easily to a word processor format, or even to PDF. This makes it an ideal tool for **reproducible research**, which is a hot topic in some fields, because it allows the description of the research, the data that serves as the basis of research, and the programs used to analyze the data to be bundled together in one single package.

To demonstrate this, a document has been prepared that can be used to produce reports on data.

### Downloading the document and preparing the data

A document that can be used to generate a report on the texts that contain the listed terms is available at [16]. Download and save the document (to your Documents folder, for example), then open it in Emacs.

The input to this document has to be in a folder called "input", located at the same level as the file. In this case, it would be Documents/input. For this example, a text file containing the terms in the following list, named `bing.txt`, will be used.

二種病	任病	五臓病相	為病聽酒
八萬四千病	身病	六神病相	結病
九病	咒時氣病經	病起六緣	業病
三病	治禪病祕要經	鬼病	滅病
三病人	治禪病祕要法	除一切疾病陀羅尼經	煩惱病
女病	看病	為病聽酒	熱病
五種惡病	重病堂	能淨一切眼疾病陀羅尼經	瞋病
瞻病五德	耆婆治病	貪病	樹提伽往昔給病道人緣
心病	病子	淫欲病	薄拘羅無病
四百四病	病行	得病十緣	應病與藥
四病	病苦	眼疾病陀羅尼經	療痔病經
玄沙三種病人	病暇	救療小兒疾病經	禪病
生老病死	病導師	雲門藥病相治	瞻病
生老病死苦	病魔	無明病	難治三病
百一病惱	四大病相	眾病悉除	魔病

録し、テキストごとに一致した用語数を計算しなければならない。この作業が自動化を必要としていることは明らかであるが、問題はそれをどのように行うかだ。

これは実のところマンドクを使えばわりと簡単にできる。マンドクの基盤である org-mode は、実行可能コードを「文芸的プログラミング」と呼ばれる形でドキュメントに埋め込むことを可能にする。そのようなドキュメントはまた、ワードプロセッサ形式に、あるいは PDF にさえ簡単にエクスポートできる。これがいくつかの分野で話題になっている再現可能な研究 (reproducible research) の理想的なツールにしている。なぜなら、これが研究の記述、研究の基礎となるデータ、そしてデータを分析するのに使われるプログラムを単一パッケージの中にひとまとめにするからである。

このことを実証するために、データのレポートを作成するのに使われるドキュメントを準備した。

### ドキュメントをダウンロードしデータを準備する

用語リストの入ったテキストについてのレポートを作成するのに使われるドキュメントは [16] から入手できる。このドキュメントをダウンロードし、(例えば自分の Documents フォルダーに) 保存し、そしてこれを Emacs で開ける。

このドキュメントへのインプットはこのファイルと同じレベルに位置する「input」と呼ばれるフォルダーの中に置く。ここでの場合は「Documents/input」である。この例では、以下の用語リストが入っている `bing.txt` という名前のテキストファイルが使われる。

## Executing the analysis

Figure 4.27 shows a so-called “source block”, which contains the code to be executed. To execute the code and start the analysis, move the cursor to the line starting with `#+BEGIN_SRC`, then press “C-c C-c” (that is, press “control-C” two times in succession). Emacs will ask you to confirm your intention of running the code with the prompt “Evaluate this emacs-lisp source block on your system? (yes or no)” in the minibuffer at the bottom of the screen. Type “yes” then press enter. This also needs to be done twice.

Emacs now performs searches in the Kanseki Repository for each of the terms in the files located in the “input” folder. The results of these search are saved for later analysis. Note that for long lists, searching can take considerable time. Upon completion of this phase, another part of the program will look at the results and compile some reports.

## Results

The intermediate results of the search are placed in a folder called `data`. A `data` folder is created for every input file and each one contains two further folders, `raw` and `index`. The `raw` folder records the results of the search as they are returned from the Kanseki Repository index server. These results are used to generate the report. The second folder, `index` contains the formatted index as it would be displayed in Mandoku. The reports can link to these files to allow further inspection of the details of the results.

Figure 4.28 shows the first few lines of the generated report, with the texts listed in order of the highest number of matching terms. The blue text code is a hyperlink to a more detailed display, as shown at the top of Figure 4.30. The terms are again hyperlinks; the lower part shows the details of search results for the term 鬼病 *gui bing*. This is the same information displayed when a live search is conducted for this term and the links can similarly be used to jump to the location in the text. (In fact, it is possible to change settings so that the link performs a live search instead of calling up the results saved in the file.) Since this is simply a file in the same format as the texts of the

## 分析を実行する

図 4.27 は実行されるべきコードの入ったいわゆる「source block」を示している。コードを実行して分析を始めるためには、カーソルを「`#+BEGIN_SRC`」で始まる行に移動させ「C-c C-c」を押す（つまり、「control-C」を続けて二度押す）。Emacs がスクリーン下のミニバッファの中に「Evaluate this emacs-lisp source block on your system? (yes or no)」という問いでコードを実行する意図を確認してくる。「yes」をタイプし「enter」キーを押す。これもまた二度行う必要がある。

これで Emacs は「input」フォルダーに置かれたファイルの中の用語一つ一つを漢籍リポジトリで検索する。この結果は後で分析を行うために保存される。長いリストは検索にだいぶ時間がかかることに留意すること。この段階が終了するとこのプログラムの別の部分が結果を見てレポートを作成する。

## 結果

この検索の中間結果は「data」と呼ばれるフォルダーに置かれる。各インプットファイルごとに「data」フォルダーが作られ、各フォルダーには更に二つのフォルダー、`raw` と `index` が入っている。`raw` フォルダーは漢籍リポジトリのインデックスサーバーから戻されたこの検索の結果を記録する。これらの結果はレポートを作成するのに使われる。二つ目のフォルダー、`index` はマンドクで表示されるようにフォーマットされたインデックスを含んでいる。このレポートは結果についてより詳細に検査するために、これらのファイルとリンクできる。

図 4.28 は作成されたレポートの最初の部分で、用語の一致結果の多い順にテキストを示している。青いテキストコードはハイパーリンクになっており、これが図 4.30 の上部に示されているように、より詳細なディスプレイを表示する。これらの用語もまたハイパーリンクになっている。ディスプレイの下の部分は用語「鬼病」の詳細を表示している。これは、この用語をライブサーチした時に表示される情報と同じものであり、これらのリンクもまた同様にしてテキストのその場所にジャンプするために使うことができる。(実のところ、ファイルに保存された結果を呼び出す代わりにライブサーチを行うようにリンクの設定を変えることもできる)。これは漢籍リポジトリのテキストと同じ形式のファイルなので、例





Kanseki Repository, it can be edited, saved, and copied, for example to eliminate matches that are not relevant to the question at hand, or to annotate it with comments from the researcher.

## Conclusion

The report is generated by the function `mdplus-print-results` contained in the file. This particular function is also written in Emacs Lisp, but it could be written in any of various other languages (such as Python, Ruby, R, and Perl<sup>5</sup>) and executed in the same way from within Emacs. If a different analysis is required, it is usually sufficient to replace this function by another one that yields the desired results. Having the source code for the analysis bundled in this way with analysis instructions, results, and collected and analyzed data makes it easier to share the source code and research findings with other researchers and enables other researcher to reproduce the results easily.

The functionality demonstrated in this last section depends mostly on `org-mode`, which is a part of Emacs and also the basis for the extension Mandoku. `Org-mode` has many more features that are relevant to researchers. For example, it allows research publications to be written in `org-mode` with the ability to export to Libre Office, HTML and PDF formats. This functionality was used to generate this report.

## Notes

- 1 The master edition is of course an exception, since it does not represent one single edition. However, since that is an interpretative edition, any discrepancy might have been intended by the editors, so the request for a change has to be argued differently.
- 2 This is a default setting, but the setting can be changed if necessary, as demonstrated later (see Figure 4.25).
- 3 For a fork, this involves “rebasing” the fork; for a clone, it is simply a pull from the remote, where it was originally cloned from.
- 4 This command will be available only if the text has not yet been downloaded. If the text has been previously downloaded from `@kanripo`, the text at `@krp-zinbun` can be simply added as an additional remote.
- 5 More information about this feature can be found in the Org Manual, Chapter 14. [27]

えば現在の質問には関係のない一致結果を削除したり、研究者からのコメントをアノテートするために、編集し、保存し、コピーすることができる。

## 結論

レポートはファイルに含まれている `mdplus-print-results` 機能で作成される。この特定の機能は Emacs Lisp で書かれているが、さまざまな他の言語 (Python, Ruby, R, Perl など<sup>5</sup>) のどれで書くこともでき、同様にして Emacs 内からの実行が可能である。異なった分析が必要な場合、この機能を望みの結果を生み出す別の機能に変えるだけで通常は事が足りる。

分析のためのソースコードを、分析の指示や結果、集められ分析されたデータと一緒に、このような形でまとめて持つことは、他の研究者とソースコードや研究成果を共有することを容易にし、また他の研究者が結果を簡単に再現することを可能にする。

最後の節で実証したこの機能性は、そのほとんどを Emacs の拡張でありマンドクの基盤でもある `org-mode` に依存している。`org-mode` は研究者が活用できる更に多くの機能を持っている。例えば、研究出版物を `org-mode` で書き、それを Libre Office や HTML, PDF 形式にエクスポートすることを可能にする。実のところ、今回のこのレポートを作成するためにもこの機能性は使われた。

## 注

- 1 マスター版は一つの版だけを表しているのではないので、もちろん例外である。しかしながら、それは解釈的版なので、どのような不一致も、それは編集者が意図したものかもしれない。それで変更のリクエストも異なった形で議論されるべきである。
- 2 これはデフォルト設定であり、後で実証するように (図 26を参照)、必要であればこの設定を変更できる。
- 3 フォークの場合、フォークを「リベース」すること。クローンの場合、それが元々クローンされた場所であるリモートからプルを行うこと。
- 4 このコマンドはテキストをまだダウンロードしていない場合にのみ利用できる。テキストをすでに `@kanripo` からダウンロードしているならば、そのテキスト `@krp-zinbun` を追加リモートとして加えればよいだけである。
- 5 この機能についての更なる情報は Org Manual の第14章にある [27]。

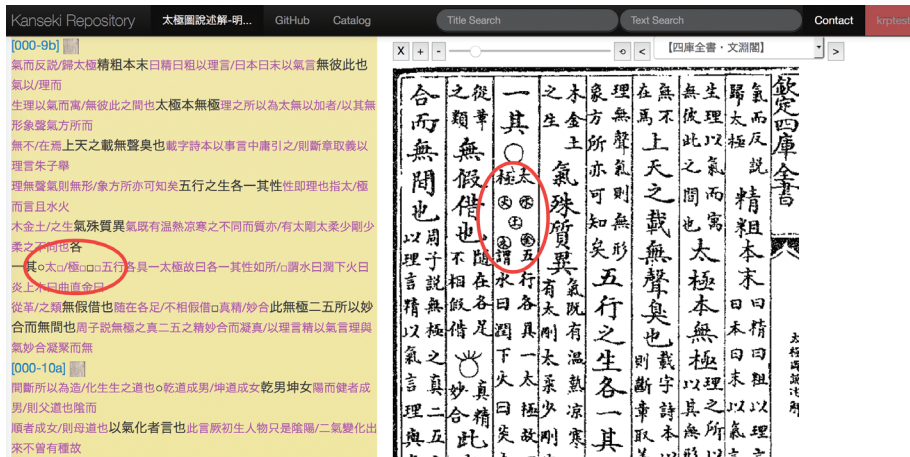


Figure 4.1: The problem we want to correct  
 図 4.1 私たちが訂正したい問題

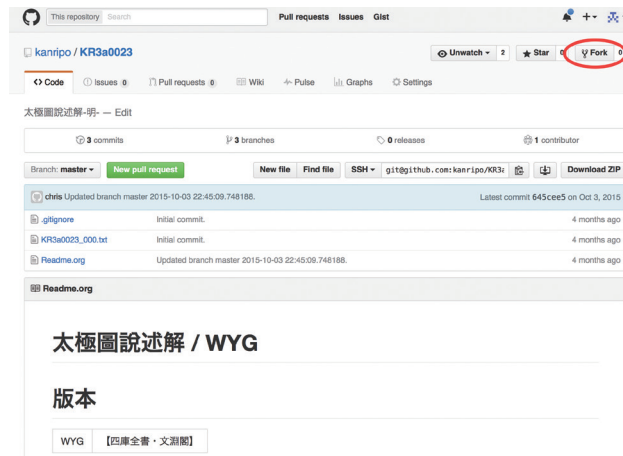


Figure 4.2: The GitHub page for text KR3a0023 太極圖說述解-明-曹端  
 図 4.2 テキスト KR3a0023 太極圖說述解-明-曹端の GitHub ページ

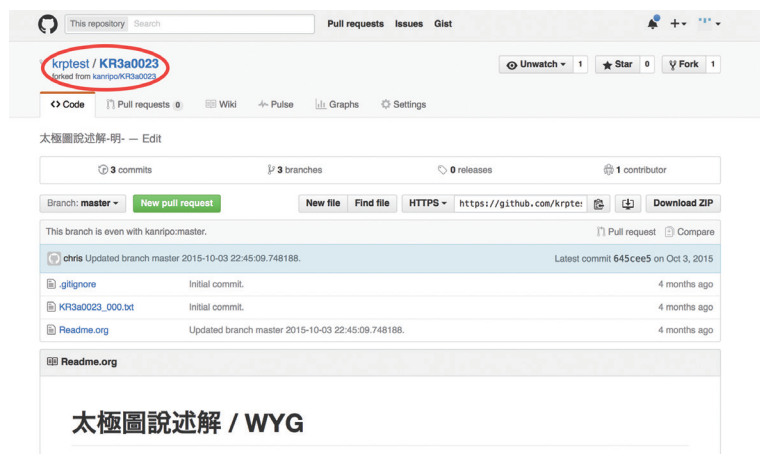


Figure 4.3: The text is now forked in @krptest's account  
 図 4.3 テキストが@krptestのアカウントにフォークされた

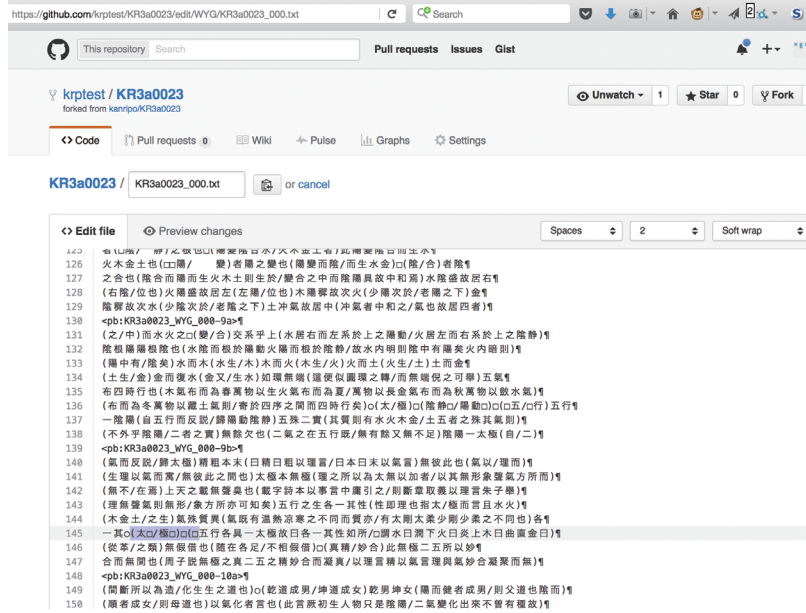


Figure 4.4: Text KR3a0023 in editing mode  
 図 4.4 編集モードにあるテキスト KR3a0023

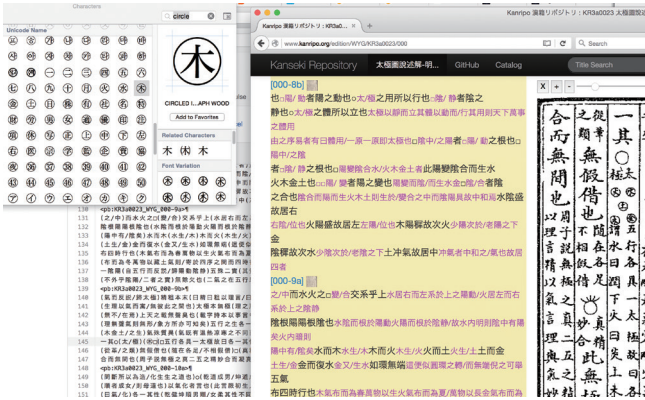


Figure 4.5: Two browser windows and the character viewer  
 図 4.5 二つのブラウザウィンドウと文字ビューアー

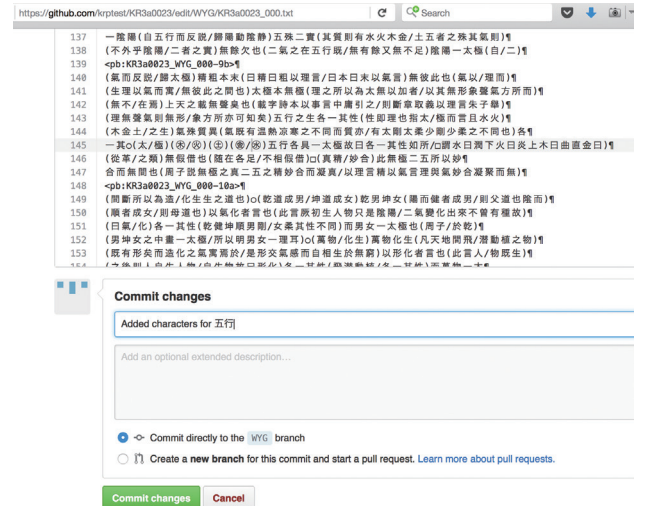


Figure 4.6: Changes completed, ready to commit  
 図 4.6 変更が完了し、コミットする準備ができた

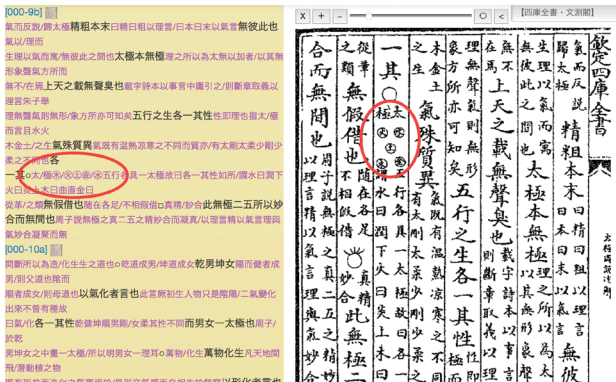


Figure 4.7: Confirm the changes in kanripo.org (visible only to @krptest)  
 図 4.7 kanripo.org で変更を確認する (@krptest にのみ見える)

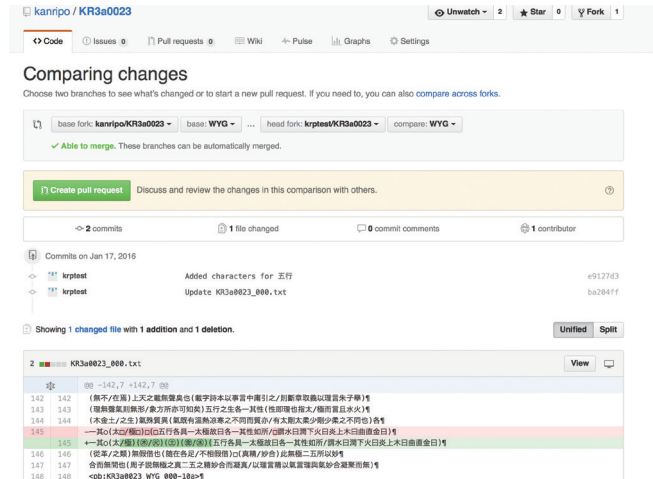


Figure 4.8: Confirm changes and create pull request  
 図 4.8 変更を確認し、プルリクエストをする

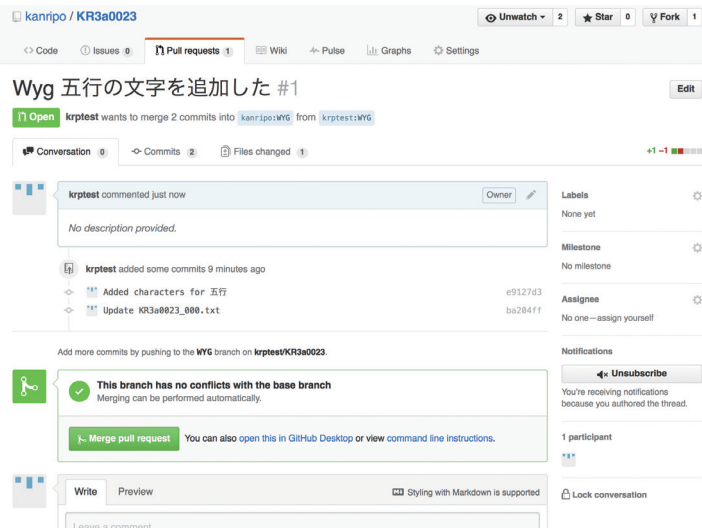


Figure 4.9: Pull request ready  
 図 4.9 プルリクエストの準備完了

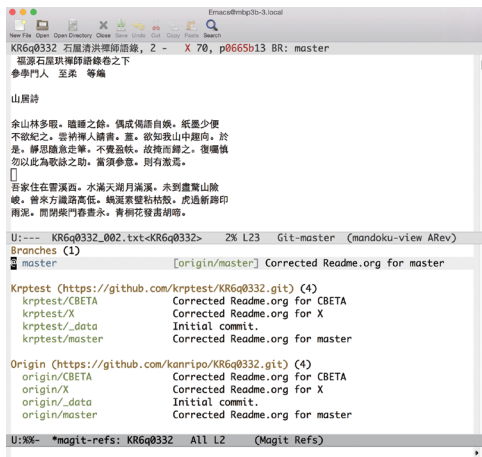


Figure 4.10: Ready to create a new branch  
 図 4.10 新しいブランチを作る準備ができた

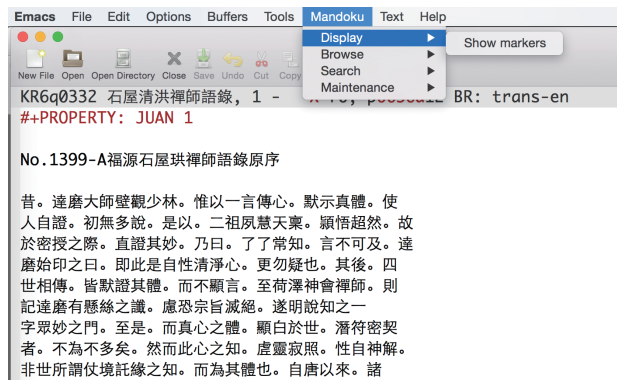


Figure 4.11: Mandoku menu Display-> Show markers  
 図 4.11 マンドクメニュー "Display-> Show markers"

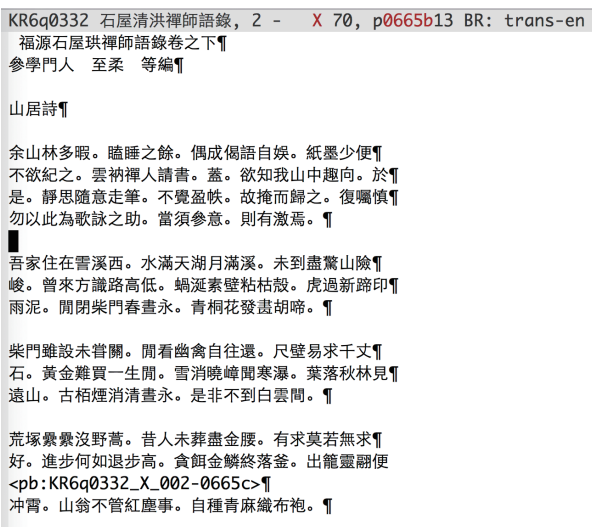


Figure 4.12: Text with visible markers  
 図 4.12 可視文字 [visible markers] の付いたテキスト

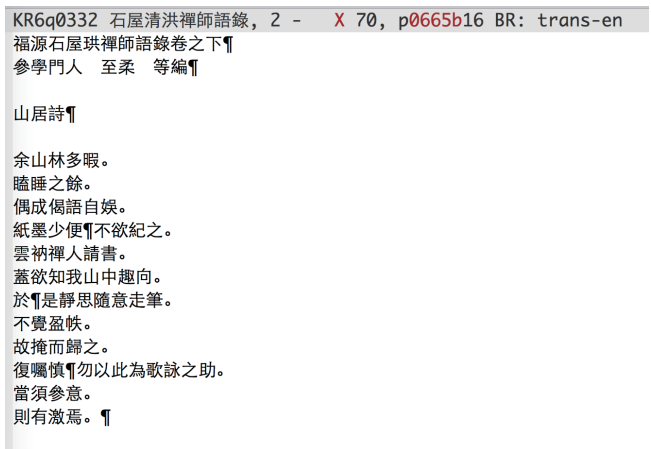


Figure 4.13: Text reformatted for translation  
 図 4.13 翻訳のためにリフォーマットされたテキスト



KR6q0332 石屋清洪禪師語錄, 2 - X 70, p0665b15 BR: trans-en	
福源石屋瑛禪師語錄卷之下 參學門人 至柔 等編	Recorded sayings of Chan master Fuyuan Shiwu Hong, Volume 2 Compiled by disciple Zhi Mao and others
山居詩	Mountain poems
余山林多暇。 瞌睡之餘。 偶成偈語自娛。 紙墨少便不欲紀之。 雲衲禪人請書。 蓋欲知我山中趣向。 於是靜思隨意走筆。 不覺盈帙。 故掩而歸之。 復囑慎勿以此為歌詠之助。 當須參意。 則有激焉。	In the mountains and woods, I am always free. The time I don't spend sleeping, I entertain myself writing down verses I encounter. Because paper and ink are rare, I did not wish to write them down But now some Chan monks asked me to write, since they wish to now what I find interesting in the mountain. Therefore I have sat quiet and let my mind follow my brush, suddenly the volume is full. So I close it and send it back, with the admonition not to try to sing these poems. You will need to consider there meaning, only then will they be able to move you.

Figure 4.14: Translation done

図 4.14 翻訳が完了

The screenshot shows the Kanripo.org website interface. At the top, there are navigation links for '漢籍リポジトリ', '石屋清洪禪師語錄...', 'ギットハブ', '目録', '題名検索', 'テキスト検索', '連絡先', and 'krptest'. Below this is a red header with the title 'KR6q0332 石屋清洪禪師語錄--(參學門人)至柔 (trans-en)'. The main content area is divided into two columns. The left column contains a '版本' (Version) section with links for '[master]', '【電子佛典集成】', '[社藏]', and '【使用者版】', and a '目次' (Table of Contents) section with links for 'No.1399-A福源石屋瑛禪師語錄原序1' and '山居詩1'. The right column contains a table with two columns: the original Chinese text and its English translation. The text and translation are identical to those shown in Figure 4.14.

Figure 4.15: Translation on the Kanripo.org website

図 4.15 kanripo.org ウェブサイト上の翻訳

```

KR6q0332 石屋清洪禪師語錄, 2 - X 70, p0665b13 BR: trans-en
於是靜思隨意走筆。           Therefore I have sat quiet and let my mind follow my brush,
不覺盈帙。                   suddenly the volume is full.
故掩而歸之。                 So I close it and send it back,
復囑慎勿以此為歌詠之助。   with the admonition not to try to sing these poems.
當須參意。                   You will need to consider there meaning,
則有激焉。                   only then will they be able to move you.

吾家住雲溪西。水滿天湖月滿溪。未到盡驚山險峻。
曾來方識路高低。蝸涎素壁粘枯殼。虎過新蹄印雨泥。
聞閉柴門春晝永。青桐花發畫胡啼。
U:--- KR6q0332_002.txt<KR6q0332> 5% (30,25) Git:master (mandoku-view ARev)
#lead: trans-en Corrected Readme.org for master
Push: krptest/trans-en is missing

Unstaged changes (1)
modified KR6q0332_002.txt

U:%%- *magit: KR6q0332 All (1,0) (Magit)
(No files need saving)

```

Figure 4.16: Magit display: Unstaged changes, no remote to push branch trans-en for @krptest

図 4.16 Magit ディスプレイ : 「Unstaged changes」 @krptest の trans-en ブランチをプッシュするリモートがない

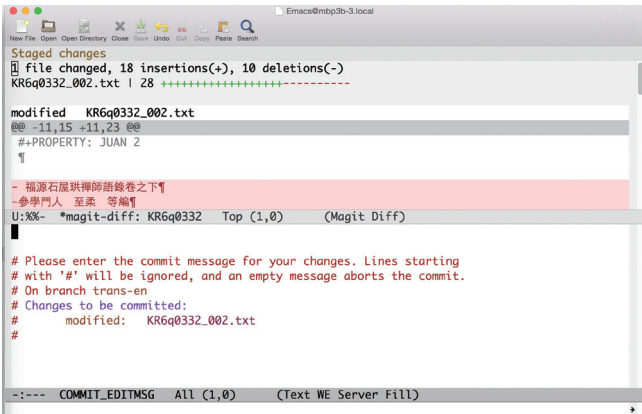


Figure 4.17: Commit message editor and display of changes  
 図 4.17 メッセージエディターと変更ディスプレイをコミットする

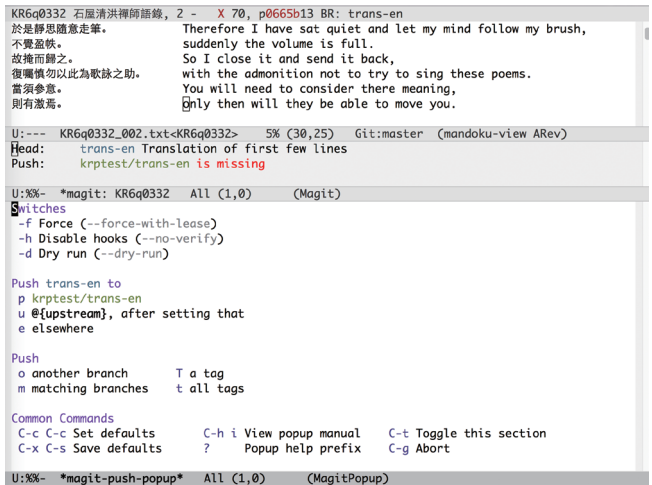


Figure 4.18: Magit popup menu: "p" will push to krptest/trans-en  
 図 4.18 Magit ポップアップメニュー: 「p」が krptest/trans-en にプッシュする

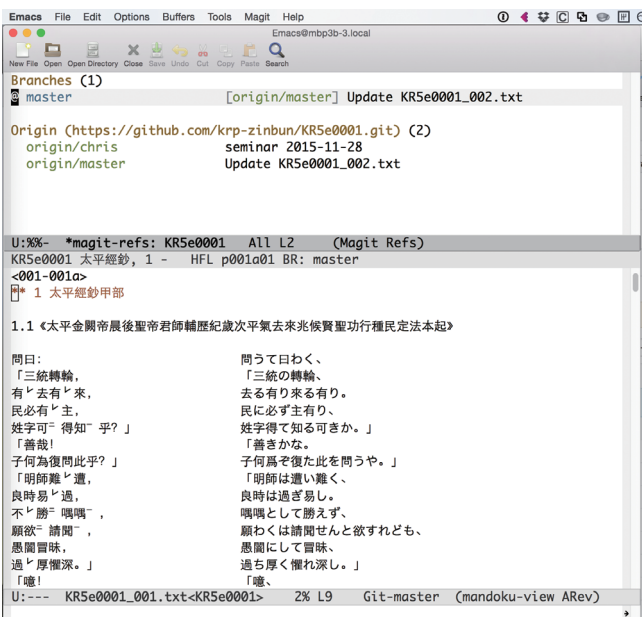


Figure 4.19: Magit display: On local branch master, ready to create a new branch starting here  
 図 4.19 Magit ディスプレイ: ローカル・マスターブランチ上で、ここから始まる新しいブランチを作る準備ができた

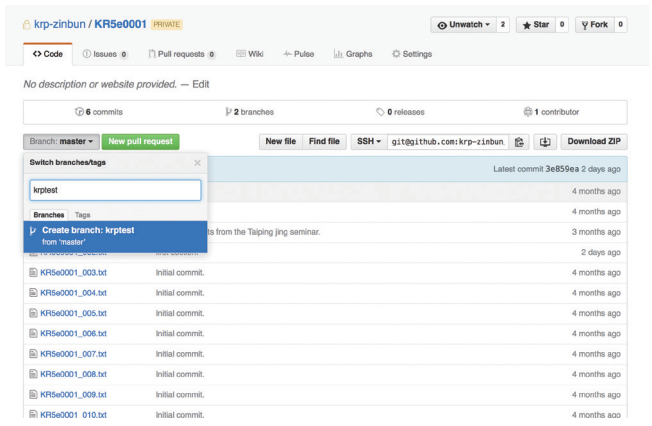


Figure 4.20: Create a branch from krp-zinbun/KR5e0001  
 図 4.20 krp-zinbun/KR5e0001からブランチを作る

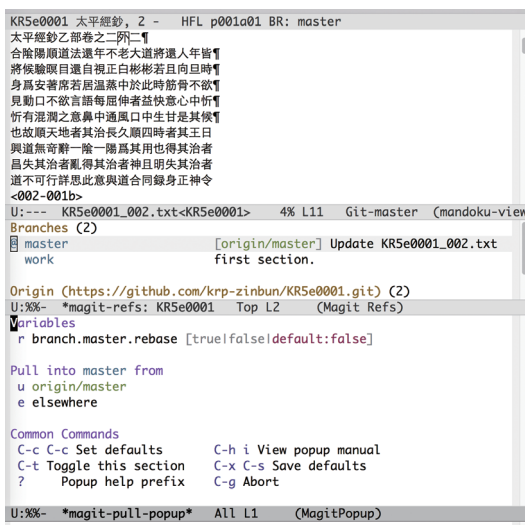


Figure 4.21: Magit display: press "u" to pull into master  
 図 4.21 Magit ディスプレイ: 「u」を押して master にプルする

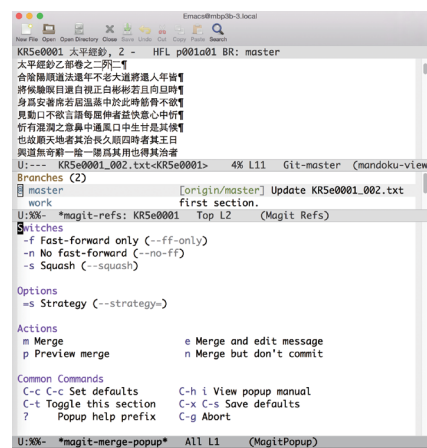


Figure 4.22: Magit display: press "p" to preview the merge process  
 図 4.22 Magit ポップアップメニュー: 「p」が krptest/trans-en にプッシュする

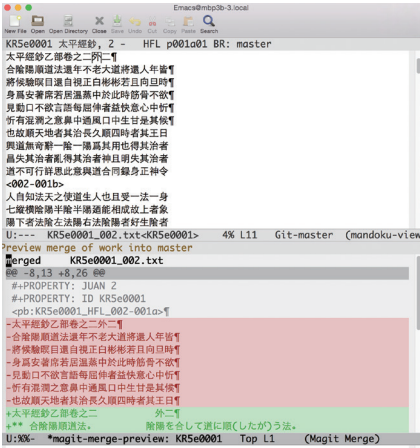


Figure 4.23: Magit display: changes that will be merged to the local branch master

図 4.23 Magit デisplay: ローカルマスタブランチにマージする変更

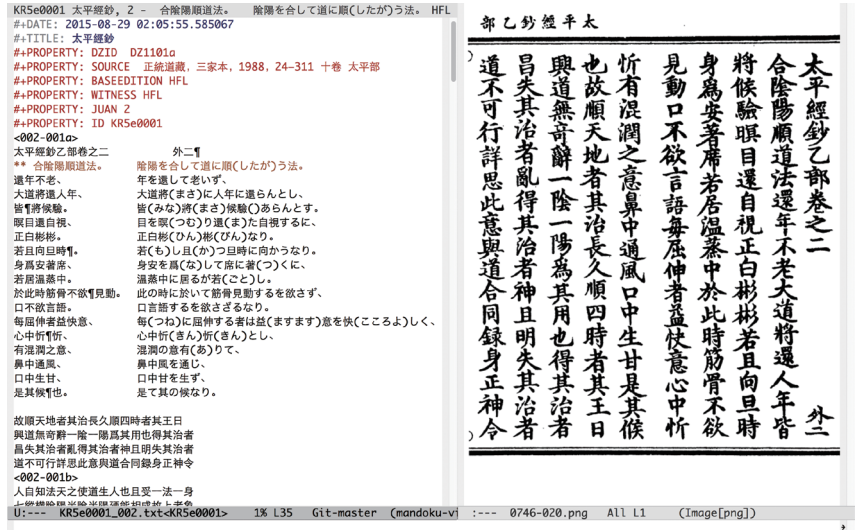


Figure 4.24: The text after completing the merge, displaying a digital facsimile to the right

図 4.24 右にデジタルファクシミリを表示しているマージ完了後のテキスト

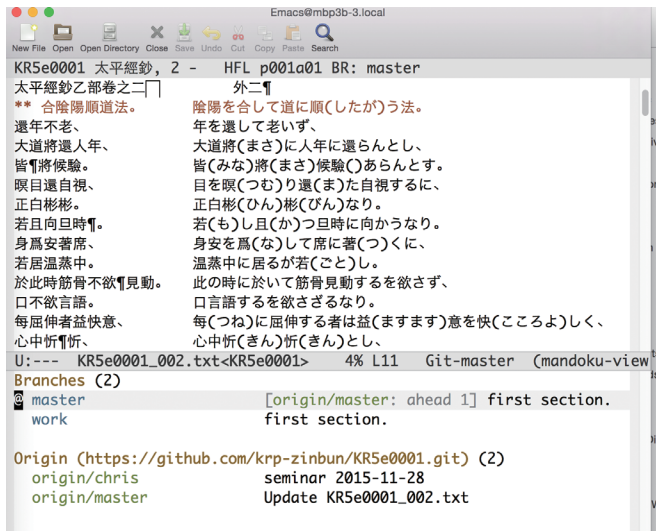


Figure 4.25: After merge: 1 commit ahead of origin

図 4.25 マージ後: origin より 1 コミット分先

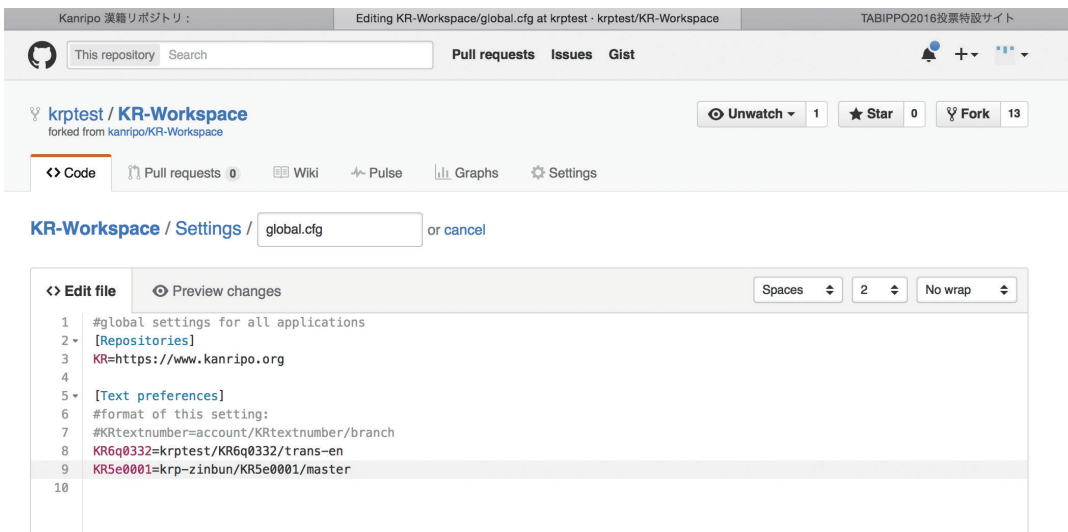


Figure 4.26: Display in kanriPO.org for members of @kzp-zinbun

図 4.26 @kzp-zinbun メンバーのための kanriPO.org での Display

```
#+BEGIN_SRC emacs-lisp
(unless
 (featurep 'mdplus-print-results)
 (save-restriction
 (save-excursion
 (re-search-forward "^* Source code")
 (org-narrow-to-subtree)
 (org-babel-execute-buffer)
 (widen))))
 (mdplus-print-results)
#+END_SRC
```

Figure 4.27: The code that will be executed with "C-c C-c"

図 4.27 「C-c C-c」で実行されるコード

TextNumber	Title	Number of terms	matched
KR6k0206	法界聖凡水陸大齋法輪寶懺-清-咫觀	0021	
KR6d0131	止觀輔行傳弘決-唐-湛然	0020	
KR6c0005	大智度論-後秦-龍樹菩薩	0018	
KR6q0022	五燈全書-清-超永	0018	
KR6s0002	法苑珠林-唐-道世	0018	
KR6g0016	涅槃經會疏-隋-灌頂	0017	
KR6s0104	閱藏知津-明-智旭	0016	
KR6f0001	大寶積經-唐-菩提流志	0016	
KR6e0021	大方廣佛華嚴經疏鈔會本-唐-實叉難陀	0016	0016
KR6k0173	四分律名義標釋-明-弘贊	0016	
KR6d0030	法華經玄贊要集-唐-栖復	0016	
KR6i0557	圓覺經大疏釋義鈔-唐-宗密	0016	
KR3e0067	普濟方-明-朱橚	0016	
KR6e0012	大方廣佛華嚴經隨疏演義鈔-唐-澄觀	0015	0015
KR6g0003	大般涅槃經-宋-慧嚴	0015	
KR3k0045	喻林-明-徐元太	0014	
KR6d0144	釋禪波羅蜜次第法門-隋-智顛	0014	0014
KR4h0117	石倉歷代詩選-明-曹學佺	0014	
KR6q0092	宗鏡錄-宋-延壽	0014	
KR6i0564	圓覺經夾頌集講義-宋-周琪	0014	0014
KR6d0130	摩訶止觀-隋-智顛	0014	

Figure 4.28: Start of the report

図 4.28 レポートの冒頭

Section	Count
KR3e 醫家類	3673
KR6k 律部類	2458
KR6i 經集部類	1785
KR6q 禪宗部類	1684
KR6d 法華部類	1194
KR6g 涅槃部類	0937
KR6j 密教部類	0925
KR6e 華嚴部類	0681
KR6s 事彙部類	0676
KR6c 般若部類	0546
KR4h 總集類	0505
KR4d 別集類-宋	0486
KR6a 阿含部類	0389
KR6b 本緣部類	0385
KR6p 淨土宗部類	0383
KR3k 類書類	0358
KR1a 易類	0329
KR6r 史傳部類	0329
KR4e 別集類-明	0313
KR6c 論衡部類	0277

Figure 4.29: Sections with the most matches

図 4.29 検索結果の多い順でのセクション

```

* Text and terms Sun Jan 24 18:11:42 2016
** Total: 22018
** Sections...
* Texts
** KR6k0206 法界聖凡水陸大齋法輪寶懺-清-咫觀 (21/34)
:PROPERTIES:
  Location      Match  Source
  鬼病 (336)
  ** 12.185b5  祝治惡鬼病*千轉陀羅尼*觀世音隨
  ** 1.401c15  行一切鬼病一切疫病不能為害若有小
  ** 7.844a12  治一切鬼病一切諸佛菩薩等眾皆大歡
  ** 12.600a6  呪治惡鬼病一千轉陀羅尼一觀世音隨
  ** 22.930c7  腫呪一鬼病一千轉陀羅尼一觀世音隨
  ** 6.35c9    語謂呼鬼病下問譴察無所不至無能知
  ** 9.856c18  眾生為鬼病下於餘時中能救病苦第四
  ** 2.150b16  上一切鬼病不入若加持頂上少髮作一
  ** 7.70a2    為凶并鬼病不可療五藥香金丑木火
  ** 2.797b8  知惡若鬼病不得語者取白芥子呪二十
  ** 161.53a2  倍信理鬼病不服藥允焚至祠開藥局卒
  ** 3.60b    殺人祭鬼病不求醫借妻孥深宜化導
  ** 30.8b1   殺人祭鬼病不求醫借妻孥等事深宜
  
```

Figure 4.30: Detail for text KR6k0206 with index display for 鬼病

図 4.30 テキスト KR6k0206の詳細と鬼病のインデックスディスプレイ



## References 参考文献

- [ 1 ] *Activate Mandoku*. <http://www.mandoku.org/data/activate-mandoku.el>.
- [ 2 ] *CBETA P5 XML on GitHub*. <https://github.com/cbeta-org/xml-p5>.
- [ 3 ] *Chinese Buddhist Electronic Text Association* 中華電子佛典協會。 <http://www.cbeta.org>.
- [ 4 ] *EBTI meeting 1996, Tripitaka Koreana*. <http://www.buddhism-dict.net/ehti/textinput/koreana.html>.
- [ 5 ] *Emacs 24.4 Released, How to Install it in Ubuntu 14.04*. <http://ubuntuhandbook.org/index.php/2014/10/emacs-244-released-install-in-ubuntu-1404/>.
- [ 6 ] IFLA Study Group on the Functional Requirements for Bibliographic Records. *FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS*. 1997.
- [ 7 ] *Git documentation*. <https://git-scm.com/documentation>.
- [ 8 ] *GitHub* サインアップの手順。 <https://git-scm.com/book/ja/v2/GitHub-アカウントの準備と設定>。
- [ 9 ] *Gnupack*. <https://osdn.jp/projects/gnupack/>.
- [10] *Installing and using Mandoku*. <http://www.mandoku.org/mandoku-install-en.html>.
- [11] *Internet Archive, Wayback Machine*. <http://archive.org>.
- [12] *Kanseki Repository at GitHub*. <https://github.com/kanripo>.
- [13] *Kanseki Repository Catalog*. <http://github.com/kanripo/KR-Catalog>.
- [14] *Kanseki Repository User Forum*. <https://groups.google.com/d/forum/kanripo>.
- [15] *Mac OS X Emacs patch*. <https://github.com/railwaycat/homebrew-emacsmacport/releases>.
- [16] *Mandoku analytic processing*. <http://www.mandoku.org/data/mandoku-plus.org>.
- [17] Peter Shillingsburg. *Global Textual Scholarship : An American View*. [http://sunrisetc.blogspot.jp/2015\\_04\\_01\\_archive.html](http://sunrisetc.blogspot.jp/2015_04_01_archive.html). 2015.
- [18] *Tang Knowledgebase Project*. <http://tkb.zinbun.kyoto-u.ac.jp/index.html>.
- [19] *TEI Consortium*. <http://www.tei-c.org>.
- [20] *The Mandoku text format*. <http://www.mandoku.org/mandoku-format-en.html>.
- [21] *The NTEmacs project*. <https://raw.githubusercontent.com/chuntaro/NTEmacs64/master/emacs-24.5-IME-patched.zip>.
- [22] *WWW Database of Buddhist Texts*. <http://kanji.zinbun.kyoto-u.ac.jp/~wittern/can>.
- [23] Hans Zeller. "Contemporary German Editorial Theory (Editorial Theory and Literary Criticism)". In: Univ of Michigan Pr, 1995. Chap. Record and Interpretation: Analysis and Documentation as Goal and Method of Editing, pp. 17-58.
- [24] マンドクのテキスト形式。 <http://www.mandoku.org/mandoku-format-ja.html>.
- [25] 人文情報学の基礎研究— *Fundamental topics in Digital Humanities*. <http://kanji.zinbun.kyoto-u.ac.jp/~wittern/kkh/dhbasic/>.
- [26] 道教學術資訊網站 *Daojiao xueshu zixun wang zhan*. <http://www.ctcwri.idv.tw>.
- [27] *The Org Manual* <http://orgmode.org/manual>.

## Some commands for Mandoku, Magit, Emacs

### Mandoku

Key	Command	Explanation	備考
F6	mandoku-search-text	Search for a string of characters in the Kanseki Repository	漢籍リポジトリで全文検索
F7	mandoku-search-titles	Search within the title list	目録中の検索
Shift-F7	mandoku-show-catalog	Display catalog top page	目録のトップを表示
C-c d	mandoku-get-remote-text-now	Download text (only from temporary text)	臨時的表示されたテキストからダウンロード
C-c i	mandoku-open-image-at-page	Display facsimile for current page	現ページのファクシミリを表示

### Magit

All Magit commands except C-x g (magit-status) are valid only within the **Magit** display buffers.

C-x g を除くすべての Magit コマンドは **Magit** バッファ内のみ有効となっていある。

C-x g	Display state of the (text) repository	テキストの現状を表示する
y	Branch manager	ブランチマネージャー
P	Push to remote repository	リモートへのプッシュ
F	Pull from remote repository	リモートからのプル
m	merge two branches	ブランチのマージ
s	stage change of the file at point	今のファイルの変更内部へ記録する
c	commit staged changes	記録された変更をコミットする

### Emacs

C-h k (key)	describe command bound to (key)	キーに連携されたコマンドを説明する
M-x (command) <enter>	execute command without key binding	キーに連携されていないコマンドを実行する
C-x C-s	save file	ファイルを保存する
C-x 1	One window on current buffer	今のバッファのみ表示する
C-x 2	Split window vertically	上下分けて二つのバッファを表示する
C-x 3	Split window horizontally	左右分けて二つのバッファを表示する

既刊：

センター研究年報2010 特集 中国旅行雑誌（井波陵一編）

センター研究年報2013 特集 科学史研究室所蔵資料データベース（武田時昌編）

センター研究年報2015

特集 漢籍リポジトリ（ウイッテルン・クリスティアン編）

---

発行日2016年3月18日

発行所 京都大学人文科学研究所附属  
東アジア人文情報学研究センター

〒606-8265 京都市左京区北白川東小倉町47

電話 075-753-6997 FAX 075-753-6999

<http://www.kita.zinbun.kyoto-u.ac.jp/>

印刷所 共同印刷工業株式会社

---

