

Development of a Digital Archive for Dao-Fa Hui-Yuan and its Application to Daoism Research

Feng XiaoXiao

**Graduate School of Library,
Information and Media Studies
University of Tsukuba**

January 2014

Development of a Digital Archive for Dao-Fa Hui-Yuan and its Application to Daoism Research

Feng XiaoXiao

**Graduate School of Library,
Information and Media Studies
University of Tsukuba**

January 2014

Development of a Digital Archive for Dao-Fa Hui-Yuan and its Application to Daoism Research

Abstract

The goal of this study is to build a digital archive for a traditional Chinese compilation of Daoism resources, named Dao-Fa Hui-Yuan (abbreviated as DFHY). Also for supporting humanities research, this study has proposed to analyze the contents of DFHY using the digitized resources. The author has found some facts from an analysis using the digital archive of DFHY. The study on the digital archive of DFHY presented in this paper consists of two main parts, one is the development of a digital collection of DFHY which includes organization of the collection based on the analysis of DFHY and its contents, and the other is a set of functions built for relationship analysis of *Fu* and other components of DFHY and applied to the digital collection.

Daoism is a Chinese philosophical religious tradition and its influence on society in China precedes Buddhism and other religions. DFHY is one of the most important Daoism resources and many Daoism researchers are using it to study Daoism, e.g., the meanings of the texts and the relationships of the descriptions in DFHY to Daoism schools. DFHY is an important compilation of many graphic expressions called *Fu*, which is used in Daoism ceremonies and religious activities. Every *Fu* has multiple meanings, e.g., pray for gods, statements to devils, and so on. DFHY is composed of 268 volumes, each of which contains many *Fus* and their associated texts. .

The digital collection of DFHY developed in this study is named Digital Dao-Fa Hui-Yuan (abbreviated as Digital DFHY). Because DFHY is a large compilation of *Fus*, Digital DFHY is a digital archive of *Fus* designed for Daoism researchers not only to browse and search the *Fus* but also to analyze the contents of DFHY. It is essential for any development of a digital collection of humanities research to learn and understand the needs of the humanities scholars. The author studied the needs of Daoism researchers to improve the efficiency of their research. Those needs are used in the design of Digital DFHY.

Preservation-and-Access are, among others, important reasons to create digital archives of cultural heritage. The Digital DFHY provides useful search and browse functions for Daoism researchers. After studying the characteristics of DFHY and the needs of Daoism researchers, the author defined the organization of the database – i.e., data for pages, texts, and graphic symbols and then noted the relationships between these data. At the same time, to support the humanities researchers with a suitable user interface, the author designed a system that has multiple ways of querying and viewing the contents of the original DFHY.

As a *Fu* is a crucial expression of mystical messages by prayers at Daoism ceremonies, many Daoism researches focus on studying the meanings of the *Fus* and the schools and their relationships behind them. And other Daoism study topics are carried on related with the studies on *Fus*. A *Fu* is represented as a complex graphical symbol composed of one or more parts. The *Fus*, however, are not well studied because of the secrecy of Daoism and difficulties in understanding the meanings written in ancient Chinese and represented as symbols.

To study to understand how *Fus* are organized is the key to discovering the secrets of *Fus*, such as their semantic meaning, their school connections and the relationships between them, and so forth. To get the answer to the questions, this study proposed to use relationship information between *Fus*, their constituent parts and the volumes of DFHY in which *Fus* are included. The relationships are hard to extract manually from DFHY but can be supported by computers using the digitized contents. The author created some graphs representing relationships among the volumes of DFHY, *Fus* and their constituent parts. The graph is composed of values obtained by calculating the average values of the relationships between *Fus* contained in volumes. By using these analysis results, humanities researchers can check the relationship between constituent parts, *Fus* and volumes. The author found some important relationships between the graphics, which had not been well recognized previously by Daoism researchers.

The graph representing the relationships is useful for Daoism researchers who have been trying to investigate how parts, *Fus* and volumes are

connected overall. The analysis of the contents is developed to evaluate the usefulness of the Digital DFHY as a digital archive for Daoism research. Daoism researchers can see the relationships as a graph, chart or table and compare them with facts and knowledge already known. Some relationships between *Fus* and volumes have been proved to be accurate by Daoism researchers. Some connections they never knew are found in the relationships data and some new topics about the relationships between the *Fus* and volumes have arisen through this study. The author considers that the digital archive of DFHY is the first step to create a more complete digital archive as this study has shown the digital archive is a very useful tool for Daoism researchers to improve their research environment. She hopes that this study helps build digital archives and analysis of the contents to support humanities researchers in their studies.

In the first chapter, this paper introduces the background of the research of digital archives from the viewpoint of the use of digital resources in humanities research, i.e., Digital Humanities. It also explains the background and history of research about Digital DFHY by the author's research partners at her laboratory. This chapter also shows some related issues and facts, e.g. *Fus*, Daoism documents and DFHY. The research goal of this study is also presented in this chapter.

Chapter 2 discusses the work related to digital archives. Because the project to build Digital Dao-Fa Hui-Yuan started more than 10 years ago in Tsukuba University, several previous works are presented in this chapter. This study builds on some experiences of the previous studies such as the creation of the catalog for the shapes of parts.

Chapter 3 discusses the digitization of the printed contents of Dao-Fa Hui-Yuan. Because graphic parts of a *Fu* are presented with its texts as *San-Xing Fu* in Dao-Fa Hui-Yuan, it is important to define the regions occupied by a graphic part in a *Fu* to associate it with its corresponding text. To organize the contents in the digital archive, the digitized data was divided into pages, texts, parts and *Fus* and organized as a database. This chapter also discusses the digitization procedure.

For supporting Daoism researchers with developed information based on the original contents, the author analyzed the original contents of DFHY. Chapter 4 shows how the author has applied the digital archive to an analysis of *Fus*. Using statistical methods, the author mapped the distributions of parts and *Fus*. Then the author derived the relationships between the *Fus*, parts and volumes. This chapter also includes the user interface of the digital archives.

Chapter 5 shows the evaluation of the digital archive and analysis of Digital DFHY.

In the last chapter, the author summarizes the experiences of creating a digital archive for DFHY and analyzing the contents for supporting the Daoism researchers. She describes her efforts to make progress in building a new type of digital archive for preserving the original contents and for supporting the study of analyses of the contents. She discusses her experiences in this study and issues for the further study.

道法会元のデジタルアーカイブの開発およびその道教研究への応用に関する研究

概要

本研究は、中国の道教の経典資料の一つである「道法会元」のデジタルアーカイブの構築に関する研究であり、道法会元のコンテンツを利用する道教研究者を支援することを目的としている。この目的を実現するため、筆者は主たる二つのテーマについて研究を進めた。一つは道法会元のオリジナルの内容に基づくデジタルアーカイブの構築、もう一つは構築したデジタルアーカイブを使った道法会元に含まれる内容分析である。

コンピュータ技術の発展に伴い、様々な歴史資料が電子化されている。デジタルアーカイブの構築により、これらの歴史資料に関する内容を研究している人文学の研究者に便利な電子資料を提供して、閲覧、検索、学術交流などの研究活動を支援することが以前から進められてきている。デジタルアーカイブはアーカイブの内容の検索、閲覧機能を提供すると同時に、電子化による資料保存を進めるものである。デジタルアーカイブを構築するには、まず、コンテンツの特徴を把握し、その上で適切な電子化手法を使ってデジタルコンテンツ化と組織化、蓄積を行わねばならない。デジタル化対象のコンテンツに含まれる言葉の出現頻度や図の意味などの分析結果が利用しやすくなることで、人文学研究者は研究上のヒントを得ることができる。このようなデジタル資料を利用した人文学研究は、デジタル人文学（Digital Humanities）と呼ばれる領域で行われている。2000年代に入り、デジタル人文学はいろんな研究対象に広がりを見せてきた。デジタルアーカイブの利用性を高めることでデジタル人文学が推進され、歴史資料を保存しながら、コンテンツ分析等の面において人文学研究者に役立つことが想定できる。しかしながら、道教研究の分野ではまだ十分にデジタルアーカイブを利用ができる状況とは言うことができない。本研究ではこうした観点から、道教資料「道法会元」の電子化とデジタルアーカイブの構築を進めた。また、構築したデジタルアーカイブの有用性を確かめるために、道法会元のコンテンツ分析を行った。

本研究の研究対象は道法会元であり、道教における最も重要な資料のひとつであり、道教資料の集大成である。道教は中国で生まれた伝統的な宗教である。道教の“道”の概念は、世界と人生の根源的な真理を指している。道教は、中国のみならず中国周辺の国々にも影響を与えた。現在、中国以外でも様々な国の研究者が道教に注目している。道教資料は多数存在しており、複数の道教宗派も存在する。道教研究者は道教資料の内

容に注目した分析を行っている。多くの道教資料の中には図と文が混在しており、符と呼ばれるシンボリックな図像表現は道教資料の内容でも特に大切なものである。道士は、人間を鬼から守ったり鬼を使ったりすることなどの目的のために符を使う。符には様々なものがあり、道教宗派ごとに異なる符を作る場合もある。符はその構造によって二種類に分けられる。一つは散形符と呼ばれ、符を構成する部品とそれらの部品（パーツと呼ぶ）の説明文からできているものである。もう一つは聚形符と呼ばれ、散形符に示されたパーツからなるひとまとまりの図像である。符は道教研究に不可欠なシンボルであり、符の研究自体が道教研究における重要な課題である。

道法会元の全 268 巻は道教の呪術儀礼の集合である。道法会元は、明の初め頃に成立し、編纂されたと考えられている。「道法会元」には多数の符が納められており、大量の図と文が混在しているため、人手で内容を検索することは容易でない。特に図の一つである符の構造は複雑で、宗派による違いもあり、その研究には時間がかかる。

筑波大学では道法会元の電子化プロジェクトを 10 年余り前に開始し、様々な研究を行ってきた。ページと文章と図の電子化データを作った後、道教研究者の協力の下に、道教研究者の支援を目的として、文章に出現する言葉の頻度、図の形の分類などを研究してきた。

道法会元のデジタルアーカイブ（デジタル道法会元）を構築する為、道法会元に存在するページ、文と図などの内容の電子化方法、データベースの仕様について検討した。また、道教研究者の研究方法、研究目的を考察した後、デジタル道法会元のためのデータベースを構築した。デジタル道法会元では、道法会元の図と文章とページが互いに関連付けられているため、道教研究者は道法会元のページや図の検索、閲覧を効率的に行えるようになった。しかしながら、筆者は、デジタル道法会元には、検索や閲覧といった基本的な機能を越えて、符の分析に役立つ機能が求められると考えた。そこで、道教の分析を行う研究者のために、筆者は符の間の関連度を表し、それをデジタル道法会元に組み入れ、利用することで道法会元に含まれる符の分析に適用することを試みた。

本研究では始めに道法会元のデータ構成を分析し、モデル化を行った。道法会元は、符を中心にして考えた場合、道法会元のなかで符の位置、符を表す図、そしてそれらに関連付けられたテキストから構成される。符の位置は、具体的には巻、葉という要素に分けられ、図には符を表すものとそのパーツを表すものがあり、符には聚形符もしくは散形符の区別、そして符のパーツがある。関連テキストは「呪文」と「説明」に分けられ、「説明」には符をどのように用いるかを解説した「使用法」、符そのものが何であることを解説した「解説」、使用にあたっての留意点もしくは他の符との違いなどを指摘した「注意点」などがある。

道法会元の中にあった符の間の関連度がパーツの形と関係していると考え、デジタル道法会元を用いた符の分析を試みた。ここでは、道教研究者によるパーツの形の分類結果を利用して、同じ形のパーツを含む符同士を繋ぎ、距離と共通するパーツの数に基づく関連度を定義して、符の関係の分析を進めた。道法会元に収録されている散形符はほぼ全て含めることができた。この方法では、符の間の関連度を基礎に、符とパーツの間、符とそのパーツの関連、符と巻の関連も考察できた。こうした分析により、符やパーツ、巻といった要素間の関係について分析した。

本研究では、符の関連度を踏まえて、全体的なパーツの関連、そのパーツが存在する符の関連を分析した。これらの関連度を可視化してパーツと符の関連の全体像を作った。この全体像は符とパーツの繋がりを示し、符間の関連度も表現できた。どのような符同士が強い関連度を持っているのか、どのパーツが一番多く符につかわれるのかといったことが、符とパーツの関連度で見ることができる。さらにパーツの分布状況とパーツが存在している符と巻の状況を表現することで、道教研究者がこの状況を把握して、どの符とどのパーツが関連するかなどを見つけることができる。符がどの巻に存在しているかを示すデータを使って、巻毎の符の分布状況と符と巻の間の関連度等を知ることができる。

パーツと符の関連度に基づき、パーツを利用した巻間の関連度も本研究の分析結果の一つである。この分析結果は巻の関連度を示すものとして道教研究者に提供されている。

これらのパーツ、符、巻に対する分析結果は道法会元そのものの内容には含まれていないが、道教研究者の分析を支える有用な情報である。本研究では、この分析結果を道教研究者の協力を得て検証し、本研究の研究結果が道教研究において実用的に利用できるものであることを示した。

電子化された道法会元のデータをシステムに実装することも本研究の重要な課題の一つである。本研究で作成したシステムでは道法会元のオリジナルの内容の閲覧、検索の基本機能の他、パーツの形、符の形から関連する内容の検索、関連している符の閲覧、パーツと符の全体像の確認が可能である。

この論文の第1章では、現在のデジタルアーカイブ、デジタル人文学について述べ、道教資料をデジタルアーカイブ化することが道教研究者に期待されていることを説明する。デジタルアーカイブには多くの先行研究があるが、道教資料に適用したものはほとんどない。本章では、本研究の位置付け、道教資料の特徴、道教研究者による要求等について論じた。そして本研究の目的と実現手法を述べた。

第2章では「道法会元」に関する多くの先行研究を紹介した。

第3章は、道法会元の電子化方法、データベースの構成について述べる。道法会元に含まれている図と文章、符に含まれるパーツ画像の処理方法を説明した。

第4章は、パーツ、符、巻の分析結果のまとめであり、これらの分析結果を基に、道教研究者による実用方法も検討した。

第5章は、本研究の分析結果を使った道教研究への適用例を挙げた。

第6章では、この道教資料に対してデジタルアーカイブを構築した経験に基づき、この電子化やデータ分析を進める上での問題などを議論しながら、今後の課題と進むべき方向を検討した。最後にデジタルアーカイブと電子化分析の統合が必要であるとの視点から、本研究の実例を使って、これからの経典的な資料に対してデジタルアーカイブと分析方法をまとめた。道法会元に限定せず、種々の資料にも適用できるアーカイブ構築と分析の一般化された手法について述べた。

Table of Contents

Abstract (English).....	i
Abstract (Japanese).....	v
Table of Contents.....	ix
List of Figures and Tables.....	xi
1 Introduction.....	1
1.1 Background.....	1
1.2 Related Knowledge about Daoism, Dao-Fa Hui-Yuan and <i>Fus</i>	3
1.3 Research Goals	8
2 Related Works and Previous Works	10
2.1 Related Works for Digital Archive and Digital Humanities.....	10
2.2 Previous Works on Dao-Fa Hui-Yuan	21
3 Digitization for the content of Dao-Fa Hui-Yuan.....	23
3.1 About the content of Dao-Fa Hui-Yuan and <i>Fus</i>	23
3.2 Database for the Digital DFHY.....	27
3.3 Digitization of the Original Content	33
3.3.1 Texts: The basic elements of the digital archive.....	33
3.3.2 Digitization of <i>Fus</i>	34
3.3.3 Catalog of Parts indexed by Shape	36
4 Analysis of Dao-Fa Hui-Yuan	45
4.1 Connection between <i>Fus</i> in the original Dao-Fa Hui-Yuan	45
4.2 Analysis of Relationship Between <i>Fus</i>	48
4.3 Mapping of <i>Fus</i> and parts.....	58
4.4 Sorting Relationships of Volumes as Evaluation of the Relationships of <i>Fus</i>	63
4.5 The interface of the Digital Dao-Fa Hui-Yuan system	69
5 Evaluation of the DIGITAL DAO-FA HUI-YUAN.....	78
5.1 Evaluation for the content of Digital DFHY	78
5.2 Evaluation of the statistical analysis results of the DFHY Contents.	80

6	Conclusions and Discussion	86
6.1	About the Digital Archive for Dao-Fa Hui-Yuan.....	86
6.2	The Model to Build a Digital Archive for Classic Documents	90
6.3	Conclusion and Future work.....	92
	Acknowledgements	95
	References.....	96
	List of Publications	100
	Appendices.....	101
	Appendix A	101
	Appendix B	103
	Appendix C	104

List of Figures and Tables

Fig. 1-1. Original sample of a page of Dao-Fa Hui-Yuan.....	5
Fig. 1-2. An example case for Daoism researchers studying the meaning of <i>Fus</i>	6
Fig. 2-1. Process model of digitization of humanist resources.....	12
Fig. 2-2. Several topics in the American Memory.....	15
Fig. 2-3. Original contents and their related information in the Digital Library from Meiji Era of Japan.....	16
Fig. 2-4. Example of some original manuscripts of Archimedes Palimpsest contained in the 13rd century prayer book	18
Fig. 2-5. The Humi project collection of digitized resources.....	19
Fig. 3-1. Several graphics for <i>Fus</i> and texts contained in Page 27 of Volume 162 of DFHY.....	25
Fig. 3-2. A page in DFHY having no graphics.....	28
Fig. 3-3. A page in DFHY having several kinds of graphic symbols contained in	29
Fig. 3-4. The original schematic diagram of DFHY.....	30
Fig. 3-5. Overview of structure of the metadata for all the contents of DFHY.....	31
Fig. 3-6. Shapes of <i>San-Xing Fu</i> and <i>Ju-Xing Fu</i> for Yuan-Shuai Da Huo-Ling <i>Fu</i> in Volume 162, page27.....	35
Table 3-1. Digitized data for Yuan-Shuai Da Huo-Ling <i>Fu</i>	36
Table 3-2. Data for <i>Yuan-shuai Da Huo-Ling Fu</i> and each part contained in the <i>San-Xing Fu</i>	36
Fig. 3-7. Steps in cataloging the shapes of Parts.....	40
Fig. 3-8. Example of the connections between <i>Fus</i> having the shared parts	41
Fig. 3-9. The framework of the Digital DFHY.....	43
Fig. 4-1. Example of the shared parts shape indicating relationships between <i>Fus</i>	46
Fig. 4-2. Three pages of Dao Fa Hui Yuan containing Yuan-Shuai Da Huo-Ling <i>Fu</i> , Kui-Huo <i>Fu</i> and Jiu-Tian Zhuo-Long Jiu-Tian Zhuo-Long Fei-Dan <i>Fu</i>	48
Fig. 4-3. An example of a social network.....	49
Fig. 4-4. Shared parts in the three <i>Fus</i>	51

Fig. 4-5. Process to find all shortest paths from part A to all Other connected <i>Fus</i>	52
Fig. 4-6. Network of the <i>Fus</i> and their parts in DFHY.....	53
Table 4-1. Values of relationships for the Yuan-Shuai Da Huo-Ling <i>Fu</i> , Kui-Huo <i>Fu</i> and Jiu-Tian Zhuo-Long Fei-Dan <i>Fu</i>	54
Table 4-2. Examples of values of the relationships between <i>Fus</i>	56
Table. 4-3. Data for shared parts in the digital DFHY.....	58
Fig. 4-7. The distribution of shared-parts (part).....	59
Fig. 4-8. Distribution map for the three parts - sf030208, sf010103, sc021302.....	60
Fig. 4-9. Example of the ranking of <i>Fus</i> relationships.....	61
Fig. 4-10. Examples for shapes of Ju-Xing <i>Fu</i> contained in volume 118, 260 and 247.....	64
Table 4-4. Relation values for <i>Fus</i> contained in volume 118 and Volume 260....	64
Table4-5 Example of Inter-Fu and Inter-volume relationship values.....	66
Fig. 4-11. Relationship between volumes (part).....	67
Fig. 4-12. Flow chart for the building of the Digital DFHY system.....	70
Fig. 4-13. Flow chart of the Digital DFHY for presenting the original contents of the Digital DFHY system.....	71
Fig. 4-14. Process chart of the system of Digital Dao-Fa Hui-Yuan for showing the analysis results.....	72
Fig. 4-15. Page view in the Digital DFHY.....	73
Fig. 4-16. Interface of The digital DFHY for supporting Users to select the title of <i>Fus</i> and the number of shared parts to study the contents.....	74
Fig. 4-17. All values of the parts relationships.....	75
Fig. 4-18. The options for the shapes of parts.....	75
Fig. 4-19. The list of names of <i>Fus</i>	76
Fig. 4-20. The relationship value chart for all of the parts.....	77
Fig. 5-1. Interface for Daoism researchers to search and browse related <i>Fus</i> having similar meanings and parts in the system.....	79
Table 5-1. Visualization of the distribution of parts.....	81
Fig. 5-2. Distribution map for parts.....	82
Fig. 5-3. Relationships between <i>Fus</i> and the sorting of the relationships.....	83

Table 5-2. The relationship value for inter-Volumes.....	84
Fig. 6-1. Example for the gap between the location by the natural reading order and the order of instruction text in San-Xing <i>Fu</i> (散形符).....	92

1 Introduction

1.1 Background

With the development of the digital information environment, more and more humanities scholars are working with computers and Internet. The World Wide Web (WWW) provides humanities scholars with a variety of digitized cultural resources and their information. Digital Archive is a widely recognized term meaning a network-oriented service whose fundamental function is long-term service of a large collection of digital resources. Since the early 1990s, there have been many programs to build digital archives of cultural resources. Those developments accelerated the use of digital resources for humanities research. Thus, both humanities scholars and information technology researchers started cross-disciplinary activities, which have grown as a new scholarly area called Digital Humanities. Digital archives are one of the central pieces in the development process of the new scholarly area. The main topic of this paper is a digital archive of a well-known Daoism resource, Dao-Fa Hui-Yuan. The study was initiated as a digitization project and has grown as a study to clarify the advantages of digital archives in Daoism research.

In academic circles, there are various definitions for “Archive”. It is generally recognized that the primary service of archives is to keep records and information resources over time. A digital archive is grounded in the archives and information technologies. The primary purpose of a digital archive is to preserve and provide access to the archived resources [1]. The key functionalities of digital archives are to create digital resources from physical resources (i.e. digitization), to select and collect digital resources which may be born-digital or digitized, to organize digital resources, to provide access to the resources and to maintain the resources over time (i.e., preservation). There have been many great developments in computer technologies over the last decades, and the current “age of information” has brought innovation in every aspect of our society [2]. Digital humanities is a new scholarly area born as an inter-disciplinary area, and applies

information technologies to humanities research. Digitized contents can be supplied to a variety of users, including the general public, students and professional researchers. Humanities scholars use digital tools, e.g. WWW browsers and word-processors, as well as conventional tools like pens, paper and ink. In general, however, many scholars are not familiar with advanced information technologies, even though advanced technologies have been used in the humanities, e.g. archeology research using CT scanners and the Text Encoding Initiative. The digital information environment is still not well-integrated for humanities scholars [3]. How to create good practices for combining the humanities resources with digital technologies has become a hot topic.

Daoism researchers have to investigate a large amount of resources in Daoism and related areas, e.g. books, letters, manuscripts, pictures, inscriptions, and so forth. They also need to collect incantations, folk proverbs, myths, legends and so on. When they verify and prove the authenticity of resources, they have to read and understand them in the historical context of Daoism. It is straight-forward to use computers to collect and analyze Daoism resources.

1.2 Related Knowledge about Daoism, Dao-Fa Hui-Yuan and *Fus*

The main goal of this study is to build a digital archive of Dao-Fa Hui-Yuan and tools to use the digital archive to help Daoism researchers use the archive. The main contribution of the author was technological but some knowledge of Daoism and Dao-Fa Hui-Yuan was indispensable for this research and collaboration with Daoism researchers was also critical for building the digital archive. In this section, some related knowledge about Daoism and Dao-Fa Hui-Yuan is introduced briefly.

Daoism, (also referred to as Taoism) is a Chinese philosophical religious tradition. It was born in China nearly 2000 years ago and its influence on society in China precedes Buddhism and other religious [4]. “The Dao is the ultimate creative principle of the universe” [5]. Daoism includes many deities, worshipped in Daoism temples (道觀).

For Chinese people, Daoism refers not only to a type of faith but also to work rules and life principles [6]. Several fields of study cannot avoid Daoism in China, such as political culture, education, medicine, health care and so on [7]. Daoists (Taoists) have various schools, and each school has its own set of behavioral guidelines and texts defining ethical considerations.

These texts and documents are collected and compiled as the collections for Daoism. Many Daoism document collections are available for the humanities studies of Chinese history, religions, cultures and literature [8].

Beyond the ancient philosophers, who are well-known for the moral dimension of their teachings, religious Daoist rules cover both ethics, i.e., the personal values of the individual, and morality, i.e., the communal norms and social values of the organizations [9]. Even for the surrounding areas such as Korea and Japan, the influence of Daoism cannot be overlooked. Also, the nationalities of Daoism researchers are not limited to Chinese but Korean, Japanese, German and so on. In this way, Daoism research is a scholarly subject studied by international humanities researchers.

Until now, a large number of Daoism documents have been preserved as important Daoism, historic, cultural and spiritual heritages [8]. Dao-Fa

Hui-Yuan (abbreviated as DFHY) is one part of the most important compilations. It contains 268 volumes such as the records of Daoism schools from the Song (宋) Dynasty to the Ming (明) Dynasty, spanning over 200 years. DFHY contains different types of contents, e.g. original documents used by Daoist priests, documents created by Daoism researchers, and so on. The original content of DFHY is already compiled and arranged by the Daoism schools living during the Song Dynasty to the Ming dynasty (A.D.1100- A.D.1400) by several Taoist priests.

DFHY is a very rich resource for Daoism scholars, not only because of its great volumes but also because it contains important and representative resources [6]. Many Daoism researchers are focused on the contents of DFHY and try to analyze the meanings of the texts and the rules of *Fu* (符). At the same time, the distinction between the Daoism schools is one of the important issues. The *Fus* are the key for each of the Daoism schools, so *Fus* are the primary entities used in this study to help study features in DFHY and relationships between *Fus* and the Daoism schools.

A *Fu* is an important expression for communication with gods in Daoism. In many humanities research subjects, *Fus* and some scholarly works about *Fus* are collected and arranged to study the rules and/or meanings of *Fus*. Jiang-Sheng [10], Li Yuan Guo [11], Liu Ye [6], Matsumoto Koichi [12] and other humanities researchers have been studying to determine the rules, origins and the functions of *Fus* in their papers and other academic publications. Some institutes and colleges are also devoted to the study of Daoism. Thus, Daoism research is meaningful for discovering the spiritual world [9].

DFHY is a compilation of many *Fus* collected from many Daoism schools. These *Fus* are important resources for Daoism research. Yet the creator of *Fus*, Daoism schools tried to keep their secrets from people outside their Daoism schools. They deliberately made their documents hard to understand and established strict inheritance systems by using codes or abbreviations not known by others in order to protect their religious documents from outsiders. To uncover the secrets of the Daoism documents is one of the

important topics for Daoism researchers. Fig. 1-2 shows a page in DFHY. This page includes some graphics mixed with the texts.

Daoism researchers have concluded the rules for the constitution of *Fus* and meanings of some texts in the manuscripts. They also admit that it is difficult to discover the meanings of all of the *Fus* because most *Fus* are not directly explained in Daoism documents [11]. Daoism researchers are faced with many difficulties in their studies.

There are several kinds of graphic Daoism symbols in DFHY in addition to *Fu*, such as *Tan*(壇), and *Gang*(罡) and so on. Among them, *Fu* is the most important symbol [12]. *Fu* is rooted in Chinese ancient characters and used as an expression of prayer by Daoism priests for communication with gods. *Fus* are generally considered to be an expression or sentence compiled by Daoism priests using basic terms, each of which is given a meaning. Among the *Fus*, some *Fus* contain only text, some contain graphic symbols, and others contain both texts and

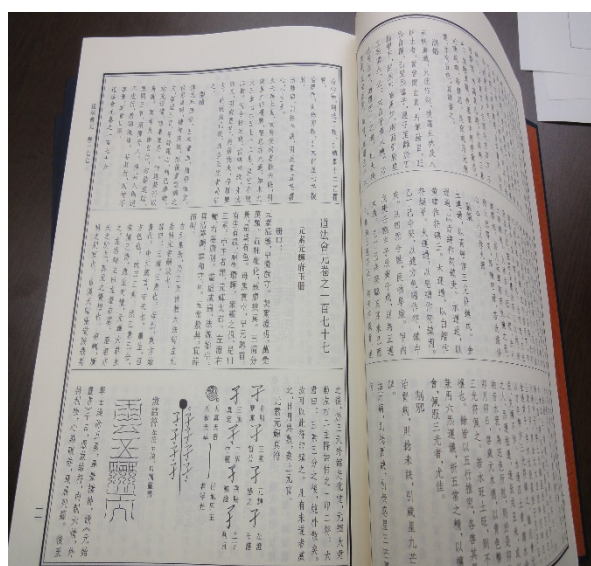


Fig. 1-1. Original sample for a page of Dao-Fa Hui-Yuan

graphic symbols. Each of Daoism schools has its own interpretation of their *Fus* so that different schools may use different *Fus* to express the same prayer. Thus, *Fus* can be used as symbols to distinguish Daoism schools. This is an important topic for Daoism researchers [4]. In general, *Fus* are mysterious and hard-to-interpret expressions but very important in Daoism.

A graphic symbol of a *Fu* is usually composed of several parts. In DFHY, a *Fu* may or may not include descriptions of the graphic parts, each of which explains the meaning and usage of a graphic part. These descriptions are written in ancient Chinese, and the meanings are ambiguous and hard to understand. Li reported that there are more than 2000 *Fus* contained in DFHY and not all of them are recognized and understood [11].

To study *Fus*, Daoism researchers have to compare the shape, and identify and recognize the meaning of each *Fu* by checking many documents, which is a very time consuming task if it is done manually.



Fig. 1-2. An example case for Daoism researchers studying the meaning of *Fus*

Fig.1-2 shows a picture of a rubbing containing incantations. *Fus* are hard to read and understand even for Daoism researchers because of the damage to the surface of the unearthed item and many unknown symbols. History researchers assumed that these *Fus* are shaped similarly to the radicals of ancient Chinese. Coincidentally, after several years, another set of inscriptions having Daoism *Fus* was excavated in 1990 in Chengdu, China. And the rubbings made from these inscriptions are shown in the left of Fig. 1-2. It was found that in the inscriptions, corresponding Chinese characters are written to explain each *Fu*. Furthermore, some *Fus* are the same as the *Fus* on the inscriptions unearthed before. Daoism researchers are excited to find

that many meanings for *Fus* are now uncovered and explained. They can now till progress their studies [4].

Unfortunately, cases like Fig. 1-2 are rare, and more *Fus* are expressed as complicated graphic symbols than expressed as Chinese characters. It is hard to find reliable explanations of the meanings of *Fus*. Daoism researchers have to refer to other resources to uncover the meanings of *Fus* and they used several steps to study them. First, the meaning of a *Fu* is hypothesized if there are no authority descriptions of it. Then, Daoism researchers check the supposed meanings of the *Fu* and compare them with others. Finally, they reach a conclusion about the meanings and roles of the *Fus*. The relationships between *Fus* are important to help the researchers obtain useful information to understand the semantics, history, school association and so on. Many of these information resources, i.e. the original content of *Fus*, bibliographic information of and relationships among the *Fus*, can be digitized. The digitized information resources will greatly help the Daoism researchers.

1.3 Research Goals

The goals of this study are to build a digital archive for DFHY and to analyze the contents of DFHY using the digitized archive. Therefore, this study consists of two main parts, one is the development of the Digital DFHY which is a digital collection of DFHY with a set of functions to use and analyze the contents of DFHY, particularly *Fu* (符), its constituent parts and the volumes of DFHY, and the other is an experimental study of the analysis of the relationships of those components contained in DFHY using the Digital DFHY.

In this study, the author collaborated with Daoism researchers, in particular Professor Koichi Matsumoto, to learn about important research topics for Daoism researchers and their research methodology, in addition to the scholarly knowledge about the relevant contents. Another important thing the author learned from the collaboration is the user's need to work with DFHY in the digital and networked environment. The collaborators' assessment of the Digital DFHY is also a very important in this study. Daoism researchers contributed to the work in its validation by evaluating the results of analysis of the DFHY contents and so helped improve the analysis methods of this study.

The rest of this paper is organized as follows. Chapter 2 discusses the related works of this study, in particular, digital archives. The project to build the digital DFHY at Tsukuba University started more than 10 years ago. Some of the previous works are discussed in this chapter. The lessons learned in the previous studies such as the creation of the catalog for the shapes of parts, are the important base of this study.

Chapter 3 is about the digitization of the original contents of DFHY. The key object in the organization of the Digital DFHY is *Fu*. *Fu* is presented as a graphic symbol, which is explained in DFHY in several different styles. As the digitization of DFHY is page-based, the raw digital data is image files. This chapter explains the objects presented in the pages, e.g. *Fu* and their descriptions, and then describes the organization of the objects into a

database. The author divided the data into pages, texts, parts and *Fus* and added them to the database.

Chapter 4 shows the application of the Digital DFHY in an analysis of the relationships between *Fus*, the parts contained in *Fus* and the volumes of DFHY. The analysis is based on statistical analysis of the graphs composed of those constituent parts. The author found some unknown relationships in DFHY from the analysis. This chapter also includes the user interface of the digital archives.

Chapter 5 discusses the evaluation of the digital archive and analysis. Chapter 6 concludes this study. In these chapters, experiences in creating a digital archive for DFHY and an analysis of the contents for supporting the Daoism researchers are discussed. Issues found but not solved in the study are discussed as the future work.

2 Related Works and Previous Works

2.1 Related Works for Digital Archive and Digital Humanities

As discussed in the previous chapter, with the development of digital technologies, digital archives have been built in many different subject domains. LTDP (Long-Term Digital Preservation Reference Model) says “*a digital archive system has three essential components: the digital records being retained, the organization that administers those records, and the infrastructure that supports the retention, protection, and accessibility of those records over their designated retention periods* [1].” In this definition, a digital archive system is used to preserve the information about the original content as well as the content.

There are some standard models that serve as guides to creating digital archives, e.g. OAIS (Open Archival Information System) [13]. OAIS defines a reference model of digital archives, which are responsible to preserve digital resources, consisting of an organization of people and systems as its components [13]. Digital preservation is an important topic for digital archives, but, in this study, digital preservation is not included in the main research topics. This is because the main goal of this study is to provide a useful digital archive of DFHY for Daoism researchers.

A number of digital archives have been and are being created for different purposes and by different organizations. SAADA, the South Asia American Digital Archive is built to reflect the unique interests, needs, and priorities of diverse South Asian American communities [14]. For achieving a range of purposes for digital archives, the model for the different types of digital archives is useful. NESTOR, which is one of the newest digital archive models, is developed as a formal model of a digital archive. NESTOR is composed of two set-based data models: the Nested Sets Model (NS-M) and the Inverse Nested Sets Model (INS-M) that express the hierarchical relationships between objects through the inclusion property between the sets [15].

Many aspects of digital archives have been discussed, e.g. building digital archives, features of the original resources, user support methodologies and technologies, and so forth.

It should be pointed out that a “digital archive” is a different concept from the original concept of “archive”. The difference is not limited to the materials stored but exists in the many other features, e.g. access method, preservation technologies, and so forth. The primary purpose of digital archives is not only to collect digital resources but also to make archives used by users for a long period. There are many interesting features in very advanced digital archives - an interesting example is that advanced computer technologies such as three dimensional (3D) technologies help archives virtually restore the contents, e.g. overcoming scratches and stains in the original resources. Some libraries and museums have been converting their precious resources into very high quality 3D images and providing them on the Internet. The advanced technologies are useful to provide convenient access to the resources for the general public [16].

The important functions of digital archives are to save the physical resources in a digital form and to organize the digitized resources for users and maintain usability over time. There are many different methods for building digital archives, satisfying the various demands of libraries, museums and their projects.

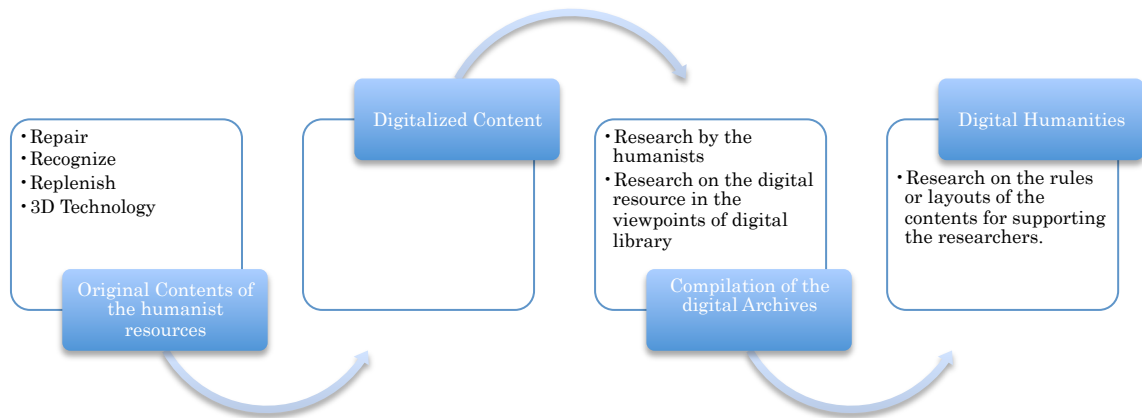


Fig.2-1. Process model of digitization of humanist resources

As shown in Fig. 2-1, there are several steps in the process of the digitization. To convert the original resources into a digital form, digital archive builders have to deal with many different types of resources such as texts, photos, videos and so forth. In the case of damaged resources automated contents recognition tasks, e.g. OCR (Optical Character Recognition). By using those technologies, content in the original documents is digitized and used to support humanities research. The compilation of digitized contents is usually reduced to the digital archives. Digital Humanities as a scholarly area was born to use digital technologies to analyze humanities resources [16].

Digital humanities is defined in several ways as follows [17].

“A diverse and still emerging field that includes the practice of humanities research in and through information technology. It also includes the development of digital educational/research/teaching/archival/publishing resources for the specific use and study of the humanities and interconnected disciplines. The digital humanities is also concerned with an exploration of how humanities may evolve through their engagement with technology.” (Ernesto Priego)

“Digital Humanities is the critical study of how the technologies and techniques associated with the digital medium intersect with and alter humanities scholarship and scholarly communication.” (Julia Flanders)

“Digital Humanities as a whole is something I am extremely reluctant to define, because any definition will inevitably cause controversy. However, I have not heard any dispute over my work being part of digital humanities, and that work involves computationally analyzing data sets from the field of Slavic linguistics, and de-siloing scholarship in that same field through a collaboratively authored academic wiki.” (Quinn Dombrowski)

Digital archives are built from the original content of documents to help various users [18]. In this sense, the author considers that there is no difference between digital archives and digital humanities. Data creation from original content is a very important step in order to build a digital archive. During the process of designing the digital archive system, users' needs should be considered, such as analyzing the contents of the digital archive. Digital humanities uses computers and other developed technologies to analyze the rules hidden in the content and should not overlook the facts in the original documents. Including analysis in a digital archive is considered reasonable.

Generally, humanities scholars focus on the original resources and their contents. On the other hand, developers of digital archives focus on the technologies to build digital resources for users and organization of the data. It is important to design a digital archive for different types of users, e.g. the humanities scholars, librarians and information scientists. In this study, Daoism researchers, who are the primary archive users, may want to use the digital archive to find new facts from the digitized resources.

Collaboration between subject specialists and archive specialists is crucial to build a digital archive for the humanities research – collaboration between Daoism scholars and digital archive developers, in the case of this study. Daoism researchers are, in general, not familiar with new digital archive technologies. They may be reluctant to use new technologies for their research although they use PCs. On the other hand, digital archive developers, e.g. librarians and library technology specialists, have

technologies applicable to many domains but not tuned for Daoism resources. Thus, collaboration is essential.

There are several typical and important digital archive projects - *American Memory*, a large-scale digital collection managed by the US Library of Congress [19], *Europeana* [20] in Europe, the National Digital Archive Program in Taiwan [21] and the Digital Library from the Meiji Era of the National Diet Library of Japan [22]. These are digital archives of cultural resources and generally designed for researchers, educators and learners who have basic knowledge of the collections. The resources are primarily digitized collections of primary materials, books, manuscripts, maps, physical objects, pictures, movies, sound recordings, etc. The arrangement and layout of the digital resources is done by librarians.

Digital archives have to have appropriate organization of the contents and access functions including user interfaces, which are designed according to the results of an analysis of the contents in the collections. The digital archive covers a very wide range of digital resources, which means any single classification system or metadata schema should work for all of them. It is important for the developers, e.g. librarians, to know the important technologies and appropriate resource organization schemes for their digital archives and users. The projects (noted above) are leading projects, which have been influencing humanities studies and guiding them to new frontiers where information technologies are applied effectively.

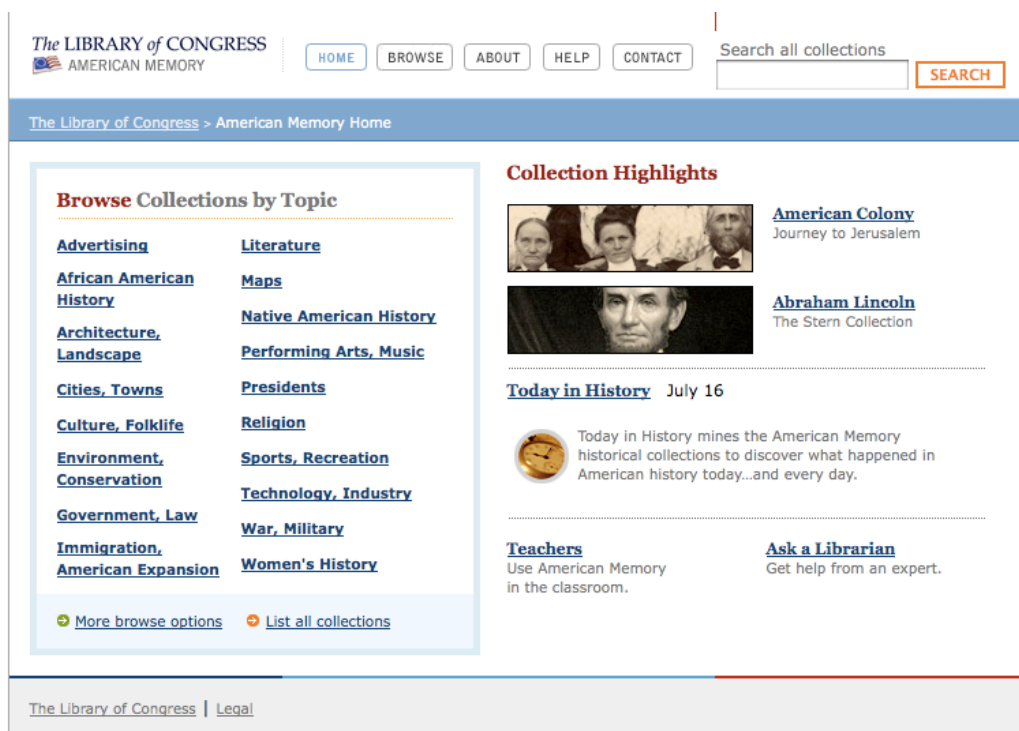


Fig. 2-2. Several topics are created in the American Memory

(1) American Memory by Library of Congress, USA

As shown in Fig. 2-2, many topics are displayed in the American Memory. The topic terms included in the bibliographic data for the digital collections are created by librarians who have the domain and cataloging knowledge. The resources are collected and listed in this digital archive, and the work includes analysis of their contents.

(2) Digital Library of Books published in Meiji-Era (Kindai Digital Library) by National Diet Library, Japan

The Digital Library of books published in the Meiji Era at the National Diet Library of Japan is built for preservation and access to cultural heritage. This digital library helps humanities scholars access rare materials online. Related resources, such as criticism articles and books, are also displayed in the digital library and arranged as a collection. For example, a famous novel “*I am a Cat*” by Natsume Soseki¹ is collected, and many of the published versions are listed and some critical articles can be checked as shown in Fig.

¹ Natsume Soseki (1867-1916) was the foremost Japanese novelist of the twentieth century, known for such highly acclaimed works as *Kokoro*, *Sanshiro*, and *I Am a Cat*.

2-3. The arranged collections support users who are browsing the contents. Citations to some other related resources are also displayed, such as books and articles written by the same author and comments and critiques written by others about the work. In this way, multiple resources are shown as related information resources.



Fig. 2-3. Original contents and related information are displayed in the Digital Library from Meiji Era of Japan

Thus, researchers can browse many other versions, related resources and related information about a particular resource. The interface of the digital

library has functions to support users who want to view related resources as well as the primary resources. However, the related information is limited to the arranged resource. As shown in Fig. 2-3, for the related contents, some commended essays or articles about the book, *I am a Cat* created by Natsume Soseki have no titles. To mark their relationships with the book, to build the digital archive and for users to find related content easily, the library, which created the digital library marked those contents and included them in the digital library system. The work of marking contents and connecting related contents in the digital library for users is a result of analysis of the contents. The analysis of the original contents should be considered as a part of developing the digitized content. Now in the digital library, more and more works are now available for the various kinds of users. For example, even the function to display a list of all connected contents is useful to support their research.

These digital libraries are focused on the on-line reading service but digital libraries are different from digital archives in some respects. The primary purpose of digital archives is to protect the content of the original resources and provide users with access to those preserved resources. The digitized resources may be edited for users; e.g. cleaning of stains on the surface, image processing to repair damaged resources and to find contents otherwise invisible from the surface of the original resources.

In the Digital Library of Meiji Era, all items of the collection are processed by their contents. But, in general, each of the digital libraries has its own digitization policies. The National Diet Library has decided their policies for the digitization [22]. In their policy, recording of collections is the primary purpose. On the other hand, to support users studying the collections in the digital library, some related contents and collections are organized in the same catalogue. The Kindai Digital Library of the National Diet Library is not focused on the work of research, however.

(3) Archimedes Palimpsest Project, USA



Fig. 2-4. Example of some original manuscripts of Archimedes Palimpsest contained in the 13rd century prayer book [23]

Fig. 2-4 shows a prayer book created in the thirteenth century. This prayer book contains erased texts that were written several centuries earlier. These erased texts include two treatises by Archimedes that can be found nowhere else. The Archimedes Palimpsest Digital Project [23] converted the Archimedes Palimpsest manuscript into digital form and organized the contents as a dataset to provide authoritative data, and derived information such as transcriptions to offer a standardized product for future users. In general, the most important aspects of digitization projects for humanities research are how to deal with the context of the manuscript, how to preserve the digitized contents over time, and how to use the information for content analysis [24]. Content analysis using digitized materials is a common goal. The users would want to use the digitized contents in order to find some unknown facts and information. This is similar to this study by the author, as she is focusing not only on the organization of the digitized contents but also on analysis of the contents via a user-friendly interface. She also aims to

provide users with an open platform based on statistic data to help them find more information hidden in the classic resources, such as the relationships between graphics, volumes and so on.

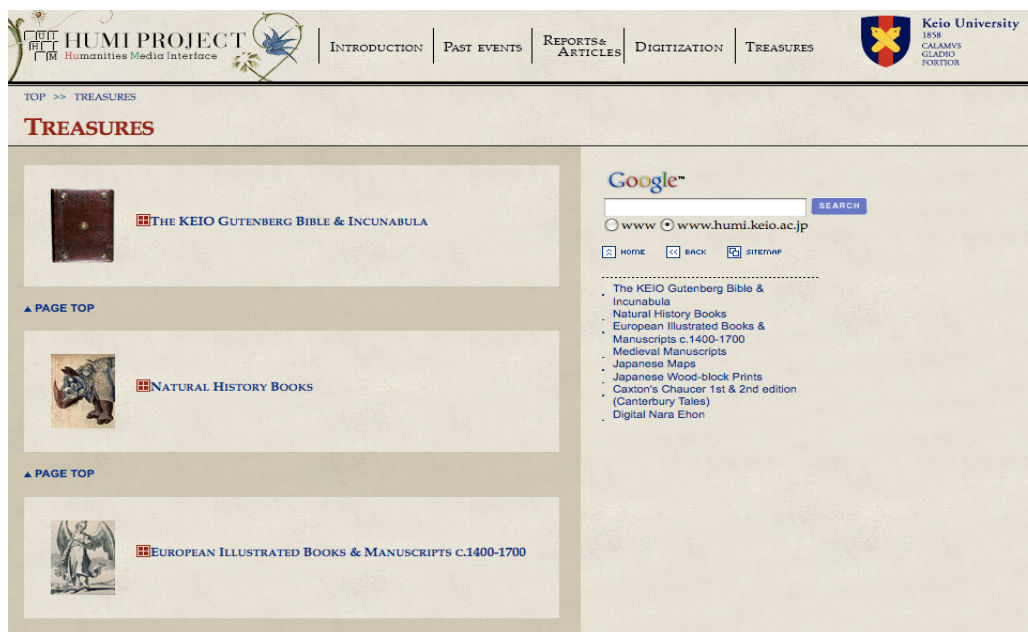


Fig. 2-5. The Humi project collection of digitized resources

(4) Humi Project by Keio University, Japan

The HUMI (Humanities Media Interface) project at Keio University created a very high quality image of Gutenberg's Bible [25] as a digital archive for supporting humanities studies aided by computers as shown in Fig. 2-5. The HUMI Project was launched by Keio University in order to pursue research in the field of digital bibliography and to construct a research environment for that field. This project used a set of advanced technologies to construct high quality digital data for the humanities research, e.g. non-destructive analysis technologies, and a 3-CCD digital camera. The aim of the HUMI Project is not simply to accumulate digital facsimiles. By compiling the images, texts and other information into databases, this project also made them accessible to researchers over an intranet and the Internet. For the participants of the project, the digital research library is not so much a digital archive as a virtual research environment [26].

The HUMI project is a good example of a digital archive of a precious cultural heritage resource used by humanities researchers like the digital archive pursued in this study. It is very well known and the data produced in the projects are an important resource for researchers and learners of the Gutenberg Bible. However, until now, the significant outcome of the HUMI project is its digital exhibition. It is not clear what functions and services the project provides for the researchers. Thus, the HUMI project aims to support humanities researchers using information technologies to create high-quality images and to preserve the original content. The HUMI project provides a platform for the researchers to share their academic endeavours.

In this section, the author showed and discussed some major digital archive and digital library projects. American Memory and Kindai Digital Library are large-scale digital collections of library holdings, which contain regular quality images and are designed for off-the-shelf browsers. On the other hand, Archimedes Palimpsest and Humi projects concentrate on a specific resource and providing high quality digital images for researchers. The Digital Dao-Fa Hui-Yuan presented in this paper is a collection of regular quality digital images taken from Dao-Fa Hui-Yuan, which is organized for use by Daoism researchers.

2.2 Previous Works on Dao-Fa Hui-Yuan

As discussed in Chapter 1, DFHY is an important collection of Daoism documents. Many humanities researchers refer to DFHY in their studies on Daoism, and ancient Chinese cultures, traditional Chinese medicine, philosophy and so on. The technological progress in digitization and computer technologies helps those researchers pursue computer-aided studies of DFHY.

A group of researchers took the first step to work on the digitized DFHY. In early 2000 Koichi Matsumoto, a Daoism researcher, proposed to digitize the original contents of DFHY and collaborated with Norihiko Uda and his research group of information scientists [28]. To build the Digital DFHY, a number of steps had to be taken. Information scientists focused on organizing the original contents of DFHY as a set of databases for humanities researchers to analyze the contents of DFHY graphics and texts contained in DFHY. Analysis of the frequency of the same words was discovered using computers and the distribution analyses of co-existing words in the contents were useful for Daoism researchers [28]. Some of the research results of Daoism researchers were used to improve the contents of the Digital DFHY. Meanwhile, studies on the *Fus*, which are a crucial component of DFHY for Daoism research, gave hints to find facts hidden in DFHY and what topics should be discussed to meet the needs of humanities researchers.

Yasoda started digitization of DFHY. She proposed an early digital archive for DFHY and tried to analyze the parts of *Fus*. Yasoda's research showed how the pages of DFHY should be arranged in the database [27]. It discussed the analysis of *Fus*, compared the shape of parts and catalogued them according to the types of shapes. Comparing the parts having the same shapes and then creating the catalogs for the shapes of parts contained in *Fus* is an important step for connecting all the parts and *Fus* in DFHY. By using the network of *Fus* and parts, the studies on the relationships between parts and *Fus* can be carried on. In the previous work, Yasoda organised the digitized pages and graphic symbols of *Fus*. Also, the meanings of a part of

Fus contained in DFHY, which were developed by cooperating Daoism researchers, were also put into the system. The system enabled viewing of every *Fu* as an image. And it also enabled users to search for parts of *Fus* by keywords. *Fus* having parts with the same shape are also analyzed. However, not all of the *Fus* were included in the database, because the analysis of *Fus* was not yet completed.

Hayakawa, one of the research fellows of Matsumoto, created a catalog of constituent parts contained in *Fus* in DFHY where she used shapes to classify the parts [28]. All of the parts contained in *Fus* are classified by shape. Some of the parts are composed of graphics, some of them are composed of radicals of Chinese characters, standing for the meaning of the Chinese character, and some parts are composed of complicated symbols mixed with graphics and Chinese characters. Recognition and classification of the shape of parts are part of the Daoism studies. *Fus* can be compared by their constituent parts more efficiently than before because the comparison function of the previous system was limited to comparison by pages. Using this catalog, Daoism researchers can now study the structures of *Fus* more systematically using the parts.

This current study has re-examined the organization of the parts catalog, i.e., index, and extended the catalog database of the *Fus*. The author has also added a new system for the digital archive of DFHY for supporting Daoism researchers.

3 Digitization for the content of Dao-Fa

Hui-Yuan

3.1 About the content of Dao-Fa Hui-Yuan and *Fus*

As discussed in Chapter 2, the author proposed to develop a digital archive not just into a collection of digitized documents but a newly organized resource collection for current and future users. The author needs to take future developments into account when developing a digital archive, e.g. the management and analysis of the contents.

DFHY is a well-used Daoism document and is a compilation of *Fus*. In this study, the author not only converts DFHY into a digital form but also designs a data set for a digital archive of DFHY, which is called Digital DFHY. Digital DFHY should be usable for researchers as a primary Daoism resource. In this chapter, the author shows the process for the digitization of the contents and the steps of building a digital archive of DFHY.

As generally known, each digital archive has its own features, and is designed in accordance with the source resources, the audience and the main goals of the digital archive. For example, in the case of the digitization of books of Meiji Era of the National Diet Library, the main goal is preservation of the collections, and then providing users with easy access and a new reading environment for the digitized books. User-friendliness and keeping costs low are two of the most important factors in the development. In the Archimedes Palimpsest Project, the crucial issue is analyzing the usability of the digital collection of the original resources for supporting humanities research. In the case of Digital DFHY, the primary purpose is to build a database for the digital archive to help researchers search for and study the contents of DFHY.

A *Fu* is the key to understanding any of the contents of DFHY, according to Daoism researchers. Because a *Fu* is an expression of prayer or phrase for Daoism priests to communicate with their gods. *Fus* may be drawn on any medium, but paper is the most common. In this study, to simplify our

understanding of *Fu*, we assume that a single *Fu* is a piece of paper on which a graphic symbol and associated texts are drawn. In other words, in DFHY, a *Fu* is a page (or continuous pages, or a part of a page) included in a volume of DFHY and the contents of the page as well. In Daoism ceremonies, the priests use *Fus* drawn on paper. Thus, *Fu* means both a logical entity and a physical entity, i.e. logical entity as a symbolic expression of a prayer or phrase, and a physical entity as a piece of paper on which a symbolic expression is drawn.

Each *Fu* has a graphic symbol, which expresses a religious statement. Each graphic symbol is composed of more than one constituent part. There are two types of *Fus* – those with and those without descriptions of their constituent components. The former are named Sanxing-*Fus* and the latter named Juxing-*Fus*. Fig. 3-1 includes a Juxing-*Fu* and a Sanxing-*Fu*, respectively. The meaning of some *Fus* is still unclear because not all of the *Fus* are accompanied by a description of their constituent parts.

The primary users of Digital DFHY are people having deep knowledge of Daoism but not familiar with information technologies. In DFHY, as defined above, a single *Fu* can be explained in one or more pages and a single page can include one or more *Fus*. These irregularities made the digitization process quite complicated.

As mentioned before, this study was done following previous studies of digitization of DFHY. The author re-organized the database building on the lessons learned from the previous studies and her own experiences. This re-organization was done in collaboration with Daoism researchers in order to define appropriate database schemas benefitting from their deep knowledge of Daoism. The task of recognizing the characters and graphics cannot be carried out reliably by computer technologies such as OCR because the original pages are written by hand and the kind of graphics in a *Fu* cannot be divided into their constituent parts automatically by a computer.

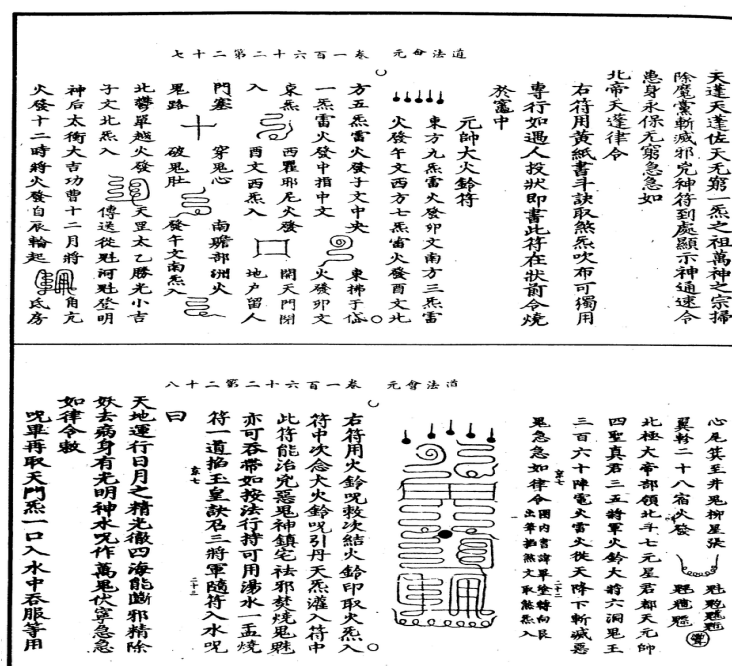


Fig. 3-1. Several graphics for *Fus* and texts contained in Page 27 of Volume 162 in DFHY

As shown in Fig. 3-1, in the page all the texts and graphics are written by hand and some of the texts are similar to the graphics. It is hard to identify the end of a graphic and/or the end of text. Thus, for creating the Digital DFHY, the author revised the database schemas created in earlier studies several times in accordance with the requirements identified by the collaborating Daoism researchers.

The needs of Daoism researchers were used to determine the functions of the Digital DFHY. Moreover, in academic circles related to Daoism, there is still considerable controversy regarding the content of DFHY.

The primary purpose of the Digital DFHY is to provide Daoism researchers with a digitized collection of *Fus* with tools to search and browse the collection. The data extracted from the digitized pages of the DFHY is the core of the Digital DFHY, which includes graphic symbols and other textual contents. The digitization process was mostly carried out manually because there are no usable software tools to parse the page images of DFHY into segments fit for the database schema. Having carried out this

digitization, the database was designed to help Daoism researchers use DFHY efficiently and discover new findings for the research community, e.g., the structures of the contents and the rules not immediately visible on the surface. The analysis of the contents of DFHY, which is an important application of Digital DFHY, is discussed in more detail in the next chapter.

3.2 Database for the Digital DFHY

As introduced in Chapter 1, DFHY is an important resource for Daoism research. Many researchers are using DFHY in their studies. In the information study field, there are several analyses and database building studies for DFHY from an information science perspective. For example, researchers have tried to analyze frequencies of the words in the texts, to create catalogs of the shapes of *Fus* and so on. These previous works provided many hints for this study.

The correct and precise digitization of the original contents of the DFHY is important. In the original DFHY, some pages contain several graphics and some pages contain only texts. Different from Fig. 3-1 shown above, the pages shown in Fig. 3-2 have no graphics but only texts. From previous experience and discussions with Daoism researchers, the author found that some users or researchers wish to view individual pages and sometimes they wish to search the related graphics and texts from the pages. How to arrange the metadata for the items of DFHY is the first issue in the database design. Once done, the author could organize the items in the database of DFHY.

In the Digital DFHY, the information about the graphic associated with information about the page is helpful to Daoism research. So the author investigated the needs of Daoism researchers and then formulated policies for designing the Digital DFHY:

- 1 The data should contain the original information and save the original resources in the digital archive. As discussed in Chapter 2, preserving DFHY in a digital form is a part of its goals.
- 2 The data should include information about the pages and their contents, e.g. *Fus*, parts and associated texts. There are many kinds of graphics in DFHY, and those graphics are important for the studies of Daoism researchers, so it is necessary to show the location of those graphics.
- 3 In order to use the Digital DFHY for analysis of *Fus* the digitization process has to be based on the user requirements, e.g. items being prepared as a digital image or text.

Fig. 3-3. A page in DFHY that containing several kinds of graphic symbols

In Fig. 3-2 there is a page with no graphics, while there are several kinds of graphic symbols presented in the page in Fig. 3-3. Basing on whether graphics are contained in pages, the author divided the data in the Digital DFHY into two types as shown in Fig. 3-4. First, pages are saved as .jpg files so Daoism researchers can view complete pages. Secondly, the graphic entities such as *Fus* are saved in the database as individual entities. The component *Fus*, as named parts are saved into .jpg files, as well as *Fus*. Finally, as *Fus* are an important research subject for Daoism researches, the graphics of *Fus* are saved separately as .jpg files. In some case, some *Fus* are explained in more than one page and all pages are arranged with adjacent page numbers in the original contents in DFHY. Thirdly, the author decided that the data to mark the location of each *Fu* should be presented as a combination of a page number where the *Fu*'s title is contained and the sequence number of the *Fu*. This data is used as a unique identifier of a *Fu*. For keeping the original shapes of *Fus*, the author saved the two pages together as one in the database and created the file name according to the number sequence of each file.

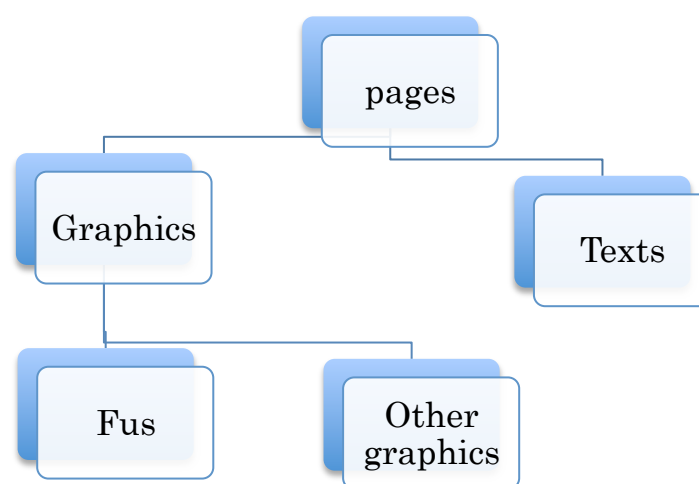


Fig. 3-4. The original schematic diagram of DFHY

The texts in pages are picked out and saved as .txt files. The graphic symbols are extracted from the pages separately. First, the scales of the

graphics are defined. In DFHY, the start of *Fus* and other graphics is complex and hard to understand, Daoism researchers helped the author to recognize *Fus* and mark them out for the analysis in the next steps.

As shown in Fig. 3-1 the *Fus* often include image symbols, their constituent parts, their accompanying texts, explanations of the meaning of the related references and so on. These accompanying texts and component parts should be a part of the main subject of the research using the Digital DFHY. The author also defined the scale of the graphics and their type with the assistance of the humanities scholars and then saved the component parts in a .jpg file and the text accompanying the graphics in a .txt file. And then, the identifiers for the *Fus* and each parts contained in it are set to mark the relationships between *Fus* and its parts. The metadata are used to describe the locations of pages, texts and *Fus*, the relations between *Fus* and parts and other texts and graphics.

The author also hoped to process the other graphic instance referred to in Chapter 1 – Tan (壇) and Gang (罡) – but they are not included and left for the future study. As introduced in Chapter 1, many Daoism researchers are focused on the *Fus*. In this study, *Fus* and their analysis are a key for supporting Daoism researchers. Therefore, metadata for *Fus* is necessary. In this study, the metadata is designed for *Fus* based on the structure of the content of DFHY.

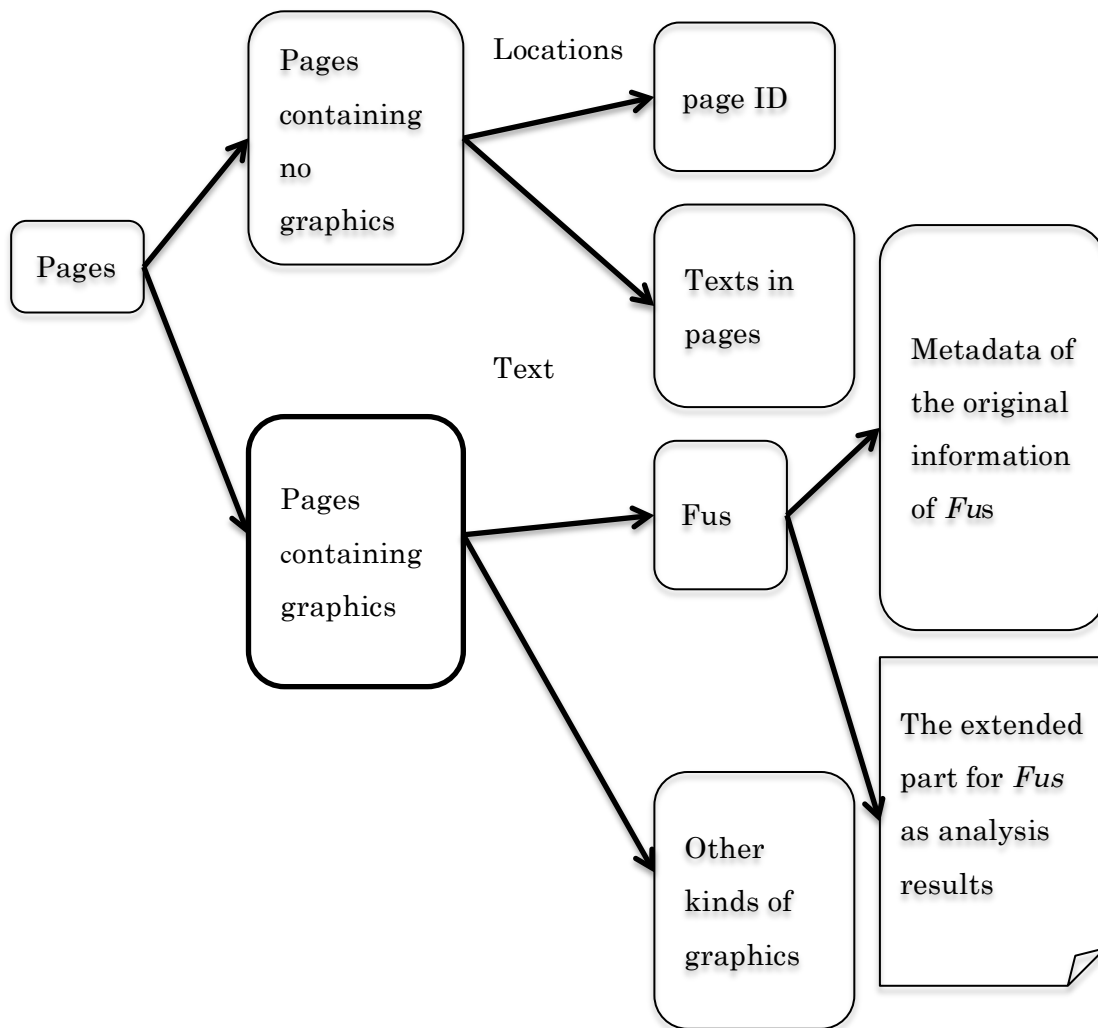


Fig. 3-5. Overview of structure of the metadata for all the contents of DFHY

The author defined the relationship between the different kinds of content in DFHY. Fig. 3-5 shows the metadata entities describing the data contained in the Digital DFHY. First, the pages are divided into two types – pages with/without graphic symbols, i.e., a *Fu* and its parts. Whether the graphic instance is a *Fu* is defined in the second step. In the third step, some structural data is manually extracted from each page and added to the database. Every page containing a *Fu* is given a unique identifier. Furthermore, the metadata model presented in Fig. 3-5 defines the foundation of the Digital DFHY. The relationships between pages and graphics play an important role in the analysis of DFHY. The relationships are not found in the original texts, so the relationships information is stored

in the Digital DFHY separately from the information directly created from DFHY. The relationship information is used in the analysis of DFHY, which is discussed in detail in the next chapter.

3.3 Digitization of the Original Content

Dao-Fa Hui-Yuan (DFHY) contains both texts and graphic symbols. The original texts in DFHY are written in ancient Chinese characters. Digitized contents of the ancient Chinese texts vary according to the complication of the version of DFHY [11].

The digitization of the contents was done in several steps. The first step is to deal with texts. In this study, the author used the texts of DFHY with sentence segmentation and punctuation already done by Daoism researchers. And then she determined graphic symbols and the texts explaining the symbols. The second step is digitization of the *Fus*. To create catalogs for the shapes of parts contained in *Fus* is an important task for the analysis of the content of DFHY.

3.3.1 Texts: The basic elements of the digital archive

The first issue to be solved in building the Digital DFHY was to split the graphics and texts and organize them into a database, without losing the connections between them. The 2651 pages are digitized into image files using scanners. The author used the scanned pages to create the database for the Digital DFHY. As shown in Fig. 3-1, the pages of the DFHY contain complicated graphics and texts. All of the texts are written in ancient Chinese and the graphics are all hand-drawn. It is important to accurately read the texts and digitize the contents correctly. The OCR technologies are tested to create texts from the page images but the level of recognition was not high enough. So in this study, the digitization of the texts was done manually and then proofread by Daoism researchers.

In some cases, blank pages are included in the original DFHY and in the database, the blank pages are saved as .jpg files in order to maintain the page sequence of the original resource.

3.3.2 Digitization of *Fu*s

To make the Digital DFHY really usable for users in this study, determining how to deal with the *Fu*s for building the digital DFHY and analyzing the content must come first. . The following paragraphs show some examples of *Fu*s.

Fig. 3-6 shows a San-Xing *Fu* (散形符) from the right to center and a Ju-Xing *Fu* (聚形符) on the left. Both of the San-Xing *Fu* and Ju-Xing *Fu* are expressions of Yuan-Shuai Da Huo-Lin *Fu* (元帥大火鈴符). The graphical components included in the San-Xing *Fu* can be found in the Ju-Xing *Fu*. Therefore, we call these images constituent parts or simply parts. The meaning of each constituent part of the *Fu* is explained by the text in the San-Xing *Fu*. The Ju-Xing *Fu* is a graphical composition of these parts and is given its meanings as Yuan-Shuai Da Huo-Lin *Fu*. Thus, the meaning of a *Fu* may be based on the meanings of the parts. However, some parts may not be clearly identified in the Ju-Xing *Fu*. In some cases, not all parts in a Ju-Xing *Fu* are expressed in its San-Xing *Fu*. In DFHY all *Fu*s are drawn by hand and the texts following the graphics are in classic Chinese, which causes problems for Daoism researchers trying to understand the structure of a Ju-Xing *Fu* and its meanings. Because applying OCR to the DFHY texts is not feasible, all of the original scripts were converted manually into text files under the guidance of Daoism researchers. Fig. 3-6 shows page 27, Volume 162 of DFHY. There is a *Fu* titled Yuan-Shuai Da Huo-Ling *Fu* (元帥大火鈴符) on the page. Following the title, there are the San-Xing *Fu* (散形符) and Ju-Xing *Fu* (聚形符) of Yuan-Shuai Da Huo-Ling *Fu* (元帥大火鈴符).

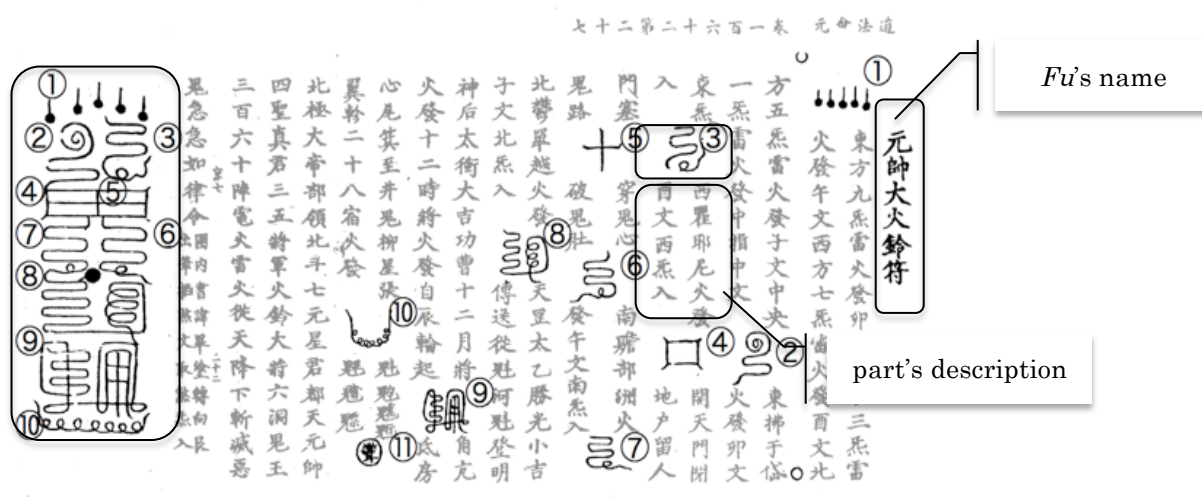


Fig. 3-6. Shapes of San-Xing *Fu* and Ju-Xing *Fu* for Yuan-Shuai Da Huo-Ling *Fu* in Volume 162, page27

In the database, the title of the *Fu* is saved as Yuan-Shuai Da Huo-Ling *Fu*, and the image for the *Fu* shown in this figure is the 28th page of volume 162, so the image file for the *Fu* is named 16228.jpg. Yuan-Shuai Da Huo-Ling *Fu* is the No.1 *Fu* appearing on this page, so the data for the *Fu* is saved in a file named after the identifier, e.g. 1622801.jpg. In other words, every *Fu* in a page is identified by its location on the page and page number. The author gave each part of this *Fu* an identifier, which is a combination of the *Fu* identifier and its sequence number. The texts following a part are saved in a plain text file. So the identifier for the first part in Yuan-Shuai Da Huo-Ling *Fu* is 162280101. In this way, we can connect the *Fus*, parts and the texts by these identification numbers in the database. In the Table 3-1, there are examples of *Fus* and parts and texts for Yuan-Shuai Da Huo-Ling *Fu*. In Table 1, the identifier for the Yuan-Shuai Da Huo-Ling *Fu* is 1622801 because it is in page No. 162, volume number 26 and the sequence number of the *Fu* is No. 1 on this page.

To build the Digital DFHY database, individual graphic symbols were extracted and saved as a JPEG file. Each graphic symbol was associated with the accompanying texts and stored in the database. Every constituent part of a *Fu* was extracted manually and stored in the database.

Table 3-1. Digitized data for *Yuan-Shuai Da Huo-Ling Fu*

<i>Fu</i>	<i>Fu_Id</i>	The mark for having San-Xing <i>Fu</i> or not
<i>Yuan-Shuai Da Huo-Ling Fu</i> (元帥大火鈴符)	1622801	1

Table 3-2. Data for *Yuan-Shuai Da Huo-Ling Fu*(元帥大火鈴符) and each parts contained in the *San-Xing Fu*

<i>Fu_ID</i>	Parts_id	Parts_shape Id[28]	Texts for parts
1622801	162280101	Sf010506	東方九炁雷火發卯文南方三炁雷火發 午文西方七炁雷火發酉文北方五炁雷 火發子文中央一炁雷火發中指中文
1622801	162280102	Sf011203	東拂于岱火發卯文東炁入
1622801	162280103	Sf011204	西瞿取尼火發酉文西炁入
1622801	162280104	Sf030301	開天門閉地戶留人門塞鬼路

Table 3-2 shows the data for the parts contained in *Yuan-Shuai Da Huo-Ling Fu*. The identifier of each part is composed of the identifier of the *Fu* in which they are contained and the sequence number in the *Fu*. The first part contained in the *Fu* is given the parts-Id 162280101. This identifier is also composed of numbers of volume, page and relative order in the page. The parts shape identifier is discussed in a later chapter. Texts associated with each part are copied from the original texts in a *San-Xing Fu*.

3.3.3 Catalog of Parts indexed by Shape

After getting the text data from a *San-Xing Fu*, the images for *Ju-Xing Fu* and its parts, Daoism researchers can find and read the *San-Xing Fu* and *Ju-Xing Fu* stored in the database. However, Daoism researchers are not satisfied by only functions to find and browse texts and to manually compare the shapes of *Fus* on the display. Those functions are fundamental as a

digital archive but not sufficient. For example, a *Fu* may have the same shape as another *Fu* but different explanations. On the other hand, some *Fus* may belong to the same school or have been used by the same Daoism priests. However, how *Fus* are used by a Daoism school or in a Daoism ceremony is not clear in DFHY. Daoism researchers want to study the relationships between *Fus* and those entities.

A description scheme for *Fus* is crucial for this study. Each graphic symbol and component is given a unique identifier. Yasoda has suggested cataloguing the parts of *Fus* by their shape [29]. Hayakawa followed the definition of the parts given by Yasoda and classified them into 4 types, which are “a part of Chinese Characters”, “a part of graphics”, “combination of Chinese Characters and graphics” and “Unknown” [28].

Creating the catalogs and classifying the shapes of parts contained in *Fus* are crucial because some components look similar but have different meanings. Similarity of the two parts is judged by their shape and description, i.e. the shape for visual similarity and the description for semantic similarity.

After digitizing all pages, we found that there were no San-Xing *Fus* in volumes 1 to 55. And from Volume 56, there are 2143 Ju-Xing *Fus* and 500 San Xing *Fus*. In all, 4458 parts were identified in DFHY.

From detailed investigation of the parts of the *Fus*, the author found that some of the explanation texts attached to a part do not express any meanings about any *Fus*. Some of the explanation texts may be curses used by Daoism priests. However, most of the texts are reliable resources for *Fu* studies.

Generally, in Daoism documents, *Fus* that look similar have similar meanings [11]. As a *Fu* is composed of its constituent parts and every part is given its meaning, *Fus* that shared many parts might have similar meanings. To investigate this, this study hypothesizes that a pair of *Fus* which share one or more parts, share some meanings. In this study, a database of the parts is created to find the relationship between *Fus*, using the occurrence frequency, locations and other information about *Fus*.

Prior to the development of the database, the catalog for the shapes of *Fus* is important. In this study, the author referenced the research results of Hayakawa presented in the examples as followings [29]:

1. Some *Fu* are a single image. The following graphic is an example of a *Fu* with a ghost shape.



2. Some *Fus* are similar to Chinese characters or parts of Chinese characters. The symbol below resembles the Chinese character “田” and it is sorted as ‘similar to a Chinese character’.



3. Some *Fus* consist of images and Chinese characters. In the image below, the top part resembles a Chinese character “山” (shān), but the part below is not similar to any Chinese character, so the *Fu* is sorted as a ‘mixed shape’.



4. Some *Fus* consist of more than one Chinese Character. In the image below there are two Chinese characters, “漸” (jiàn) and “耳” (ěr) so the graphic is sorted as a ‘graphic complied with two characters’.



5. Part structures that are hard to define are initially marked as 'Unknown'. In the table below, images not included in the catalogs are defined as 'unknown'.

The steps for sorting the shapes of *Fus* are presented in Fig. 3-7. The first step is to define the structure of parts. There are two types of part structures. One is a simple structure composed of a single elementary part, and the other is a complex structure composed of more than one elementary part. Fig. 3-7 shows the decision process. First, whether a *Fu* has San-Xing *Fu* or not is noted, and then each part contained in the San-Xing *Fu* is cataloged and the shape ID for the parts is saved in the database.

Using the results of the cataloging of parts, every *Fu* is connected and also, there is an overall mapping for the *Fus* Based on the theory of Daoism that the more shared parts two *Fus* have, the closer semantic connection they have. This catalog of the parts is also useful for recognizing the shapes of *Fus* and the symbols of various religious schools [10].

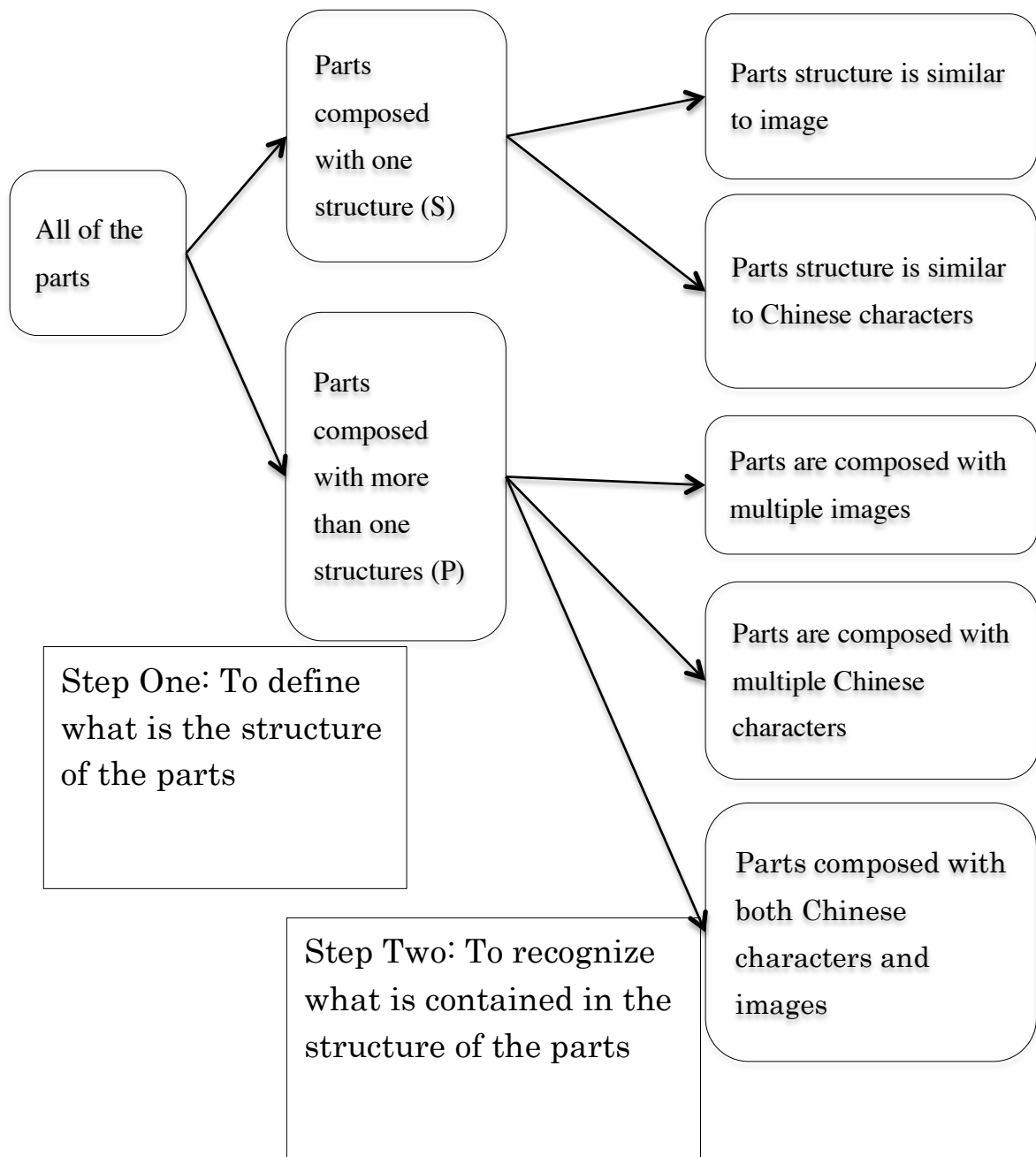


Fig. 3-7. Steps of cataloging for the shapes of Parts

The shapes of the parts contained in the *Fus* are sorted and the shapes of parts are identified and saved into the database. Table 3-2 also presents the identification of the shapes of parts named “shapeID” for each part contained in the Yuan-Shuai Da Huo-Ling *Fu* (元帥大火鈴符). Keywords for annotating the shapes of parts are useful for Daoism researchers to easily perform look-ups to refine shapes. Also, this study on the content of DFHY could not

be done without this catalog of part shapes containing *Fus*. The catalogs for the shapes of *Fus* referenced the precious works based on the research results of Daoism researchers [28].

Thus, with the assistance of Daoism researchers, the author created a catalog of the part shapes contained in *Fus* by using the identification scheme of *Fus* in order to correctly organize the information about the alignment of graphic symbols and texts. Because the technology for image recognition is not applicable to these *Fus*, the cataloging process was done manually. In this sense, sorting the shapes of parts contained in *Fus* and applying the sorting into building the Digital DFHY is analysis of the *Fus*. Based on the sorting of shapes of parts, more analysis work is in process.

Using this cataloging method, the author saved every part expressed in San-Xing *Fus* in a single JPEG file. She assigned an identifier to all the indexed shapes of the parts contained in a *Fu*. Thus, the parts are given identifiers and the *Fus* are connected to each other using the part identifiers. At the same time, by using the part identifiers (parts ID) and the part shape identifiers (part shape ID), parts having the same shapes can be found and then in what *Fus* parts are contained can also be checked.

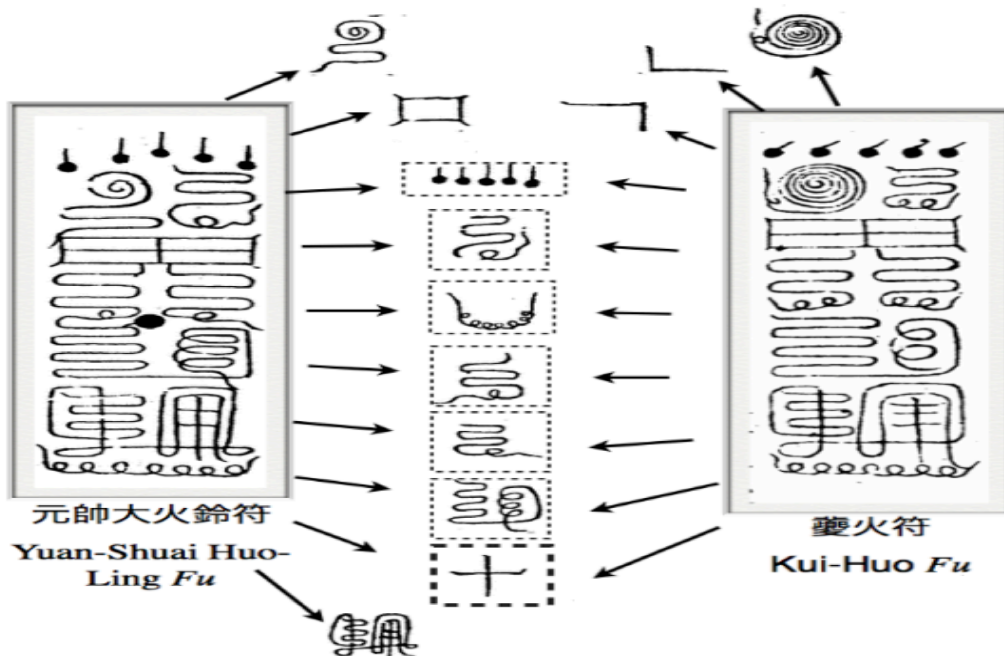


Fig. 3-8. Examples of the connections between *Fus* that share the same part shapes

In summary, the data saved into the database in this study are:

1. The title of the *Fu*
2. The image of the *Fu*
3. The page number where the *Fu* is contained
4. Texts attached to the *Fu*
5. Whether San-Xing *Fu* and Ju-Xing *Fu* are contained or not
6. Number of parts contained in the *Fu*
7. List of shapes of the parts
8. Description texts of each part.

The original information for each *Fu* is saved to connect the *Fus* to each other. The connections between *Fus* make it possible to study the relationship between *Fus* and deepen the analysis of the contents of DFHY. As shown in Fig. 3-8, the *Fus* that have the same shapes can be connected.

By using these connections, the relationships network for all of the *Fus* includes those that have no connection to other *Fus*.

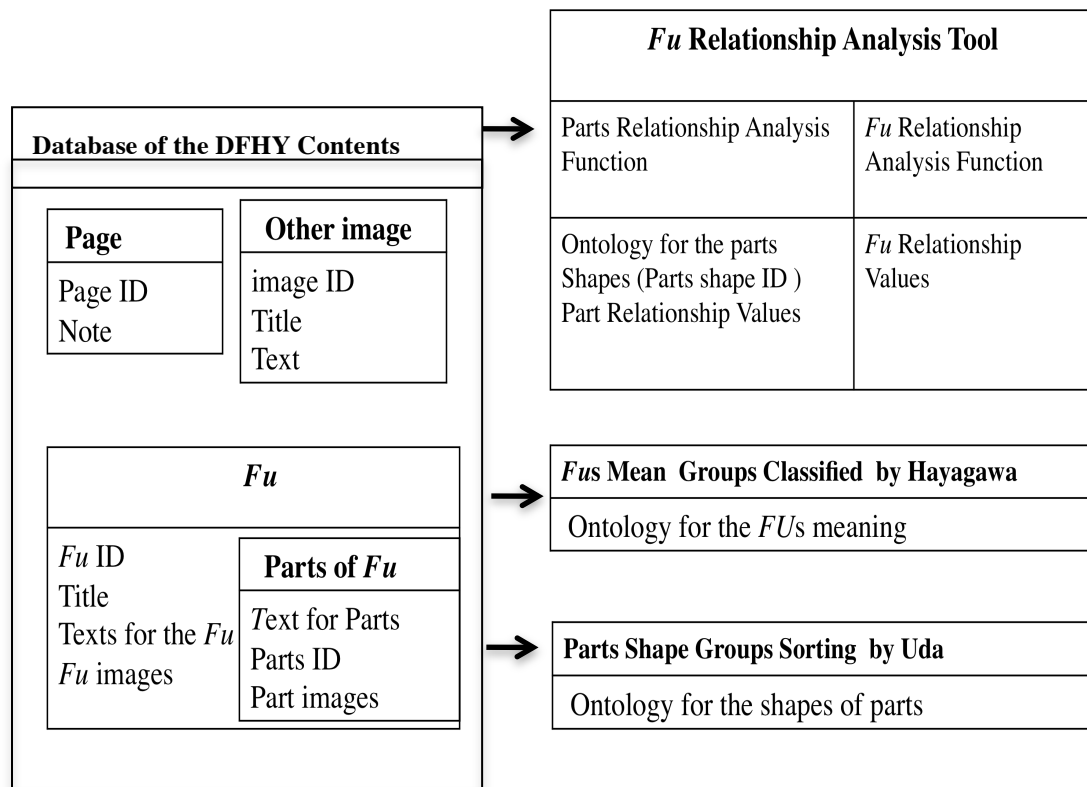


Fig. 3-9. The framework of the Digital DFHY

Fig. 3-9 shows the proposed framework for DFHY. In this framework, there are two major components. The left part shows the core database of the Digital DFHY – a set of databases of images and texts of the *Fus* and constituent parts. Images are saved as JPEG files. Texts are taken from the titles of *Fus*, scripts attached to *Fus*, and other texts are included in the pages. This set of data is used to search *Fus* and the constituent parts. The database is used in several different applications as shown on the right hand side of this figure. The author used the data set obtained from the relationship analysis of *Fus* and their constituent parts. The output of these studies has been included in the Digital DFHY as a database and a tool to use the database. Thus, this study augments the Digital DFHY by adding new functions and databases from the application research. From another viewpoint, this augmentation is a joint collaborative process between Daoism and information researchers - Daoism researchers help information

researchers create new functions and databases from the core databases, for example to decide and create the ontology for *Fus*, and, in turn, the new functions and databases are used to enrich the core databases by the information researchers and then the developed digital archives based on the database can help the studies of Daoism researchers. The analysis study using the Digital DFHY is discussed in later chapters.

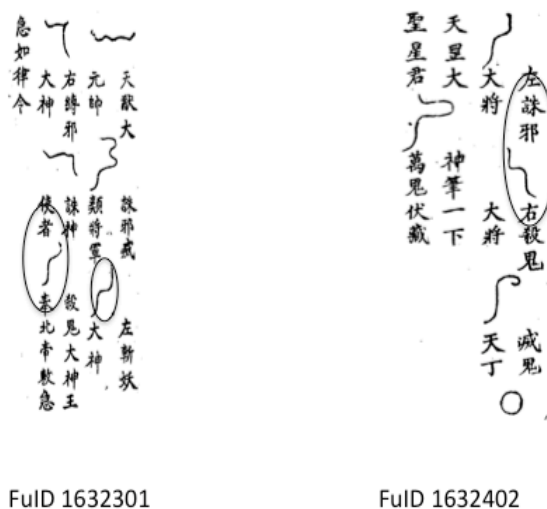
4 Analysis of Dao-Fa Hui-Yuan

4.1 Connection between *Fus* in the original Dao-Fa Hui-Yuan

A well-organized archive of digitized contents of DFHY will help Daoism researchers find new facts and knowledge. So the Digital DFHY should not only be a convenient tool for searching the contents but it should help researchers analyze the contents of DFHY to find new facts and knowledge.

As discussed in former chapters, *Fus* are important symbols for all Daoism schools and humanities scholars are trying to study their meanings, usages, directions, formation age etc. Daoism researchers study *Fus* by reading the texts and symbols. Other history documents are also important for the study. However, the *Fus*' texts are not easy to understand for those who have no deep knowledge about classic Chinese and Daoism. *Fus* with the same shape occasionally have different meanings. Daoism researchers have an assumption that the shapes of *Fus* may be affected by the context of the *Fus*, e.g., Daoism schools and history. For example, the same shapes of *Fus* indicate belonging to the same religious school or having similar function. The example can be seen in Fig. 4-1.

In Fig. 4-1, two *Fus* are contained in Volume 163. One is in the 23rd page and the other is in the 24th page. Those *Fus* are given identifiers 1632301 and 1632402, respectively. In the San-Xing *Fu* of this pair of *Fus*, there are several parts named as shared-parts with the same shape. The shared parts are marked out in Fig. 4-1. The illustration texts for the shared parts are different. For *Fu* 1632301, one of the shared parts has the illustration texts as “大神王大神殺鬼”, and the other has the texts as “奉北帝敕急”.



The texts for the shared parts of *Fu*-1631402 scripted as “左誅邪” and “右殺鬼” appear in both *Fus*. These texts have similar meanings, namely ‘Kill devils’. This means, in some cases a *Fu* can be a sentence and the parts can be considered as a word in the sentence. On the other hand, *Fus* of the same shape do not always have the same meaning. Yet many Daoism researchers consider that the similarity of shapes means somehow a relationship between the *Fus*, e.g., the usages of *Fus*, the Daoism schools, which use the *Fus*. The more parts the *Fus* share, the closer the relationship between the *Fus* would be. It could be assumed that the more parts a pair of *Fus* share, the closer the *Fus* are.

The author studied the positions of each of the shared parts, and then created the network of shared parts. By using the network for the shared parts, Daoism researchers can study what shared parts are likely to be included in what *Fus* and what volumes. As a part ID is defined by what

volume and what *Fu* it is in, where the part is can be checked by the part ID. Then, in what volumes is a part is most common, and in what volume of DFHY, and where is this part mostly used, can be checked. Also, the distribution of parts in *Fus* and volumes can be used in Daoism research and help Daoism researchers find something new from the content of DFHY.

In a word, parts are contained in *Fus* and shared parts are the key for studying the content of DFHY.

4.2 Analysis of Relationship Between *Fus*

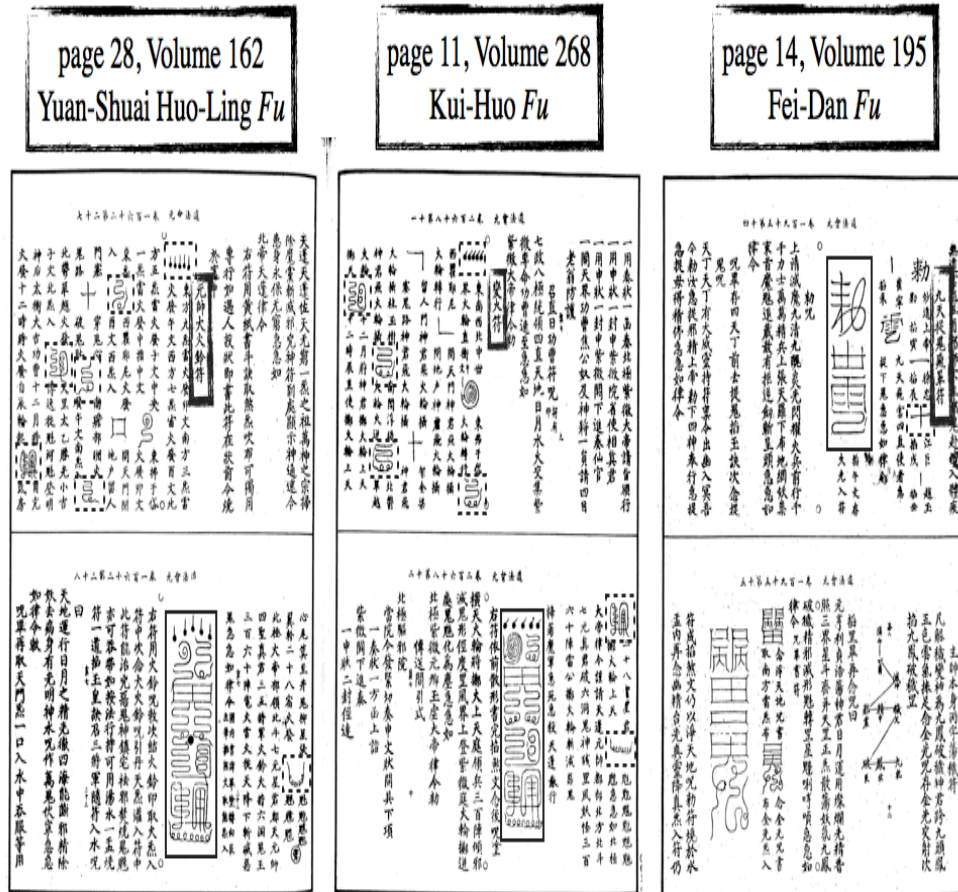


Fig. 4-2. Three pages of Dao Fa Hui Yuan containing Yuan-Shuai Da Huo-Ling *Fu*, Kui-Huo *Fu* and Jiu-Tian Zhuo-Long Jiu-Tian Zhuo-Long Fei-Dan *Fu*

Fig. 4-2 shows three pages of DFHY, and each page shows a *Fu*. The left-most page shows a *Fu* named Yuan-Shuai Da Huo-Ling *Fu* (元帥大火鈴符) in the 27th page of Volume 162 of Fig.4-2. The middle page shows Kui-Huo *Fu* (夔火符). Jiu-Tian Zhuo-Long Fei-Dan *Fu* (九天捉龍飛丹符) is shown on the right. The Ju-Xing *Fu* of these three *Fus* are sub-divided into parts and explained as Sang-Xin *Fus*. Yuan-Shuai Da Huo-Ling *Fu* has seven shared-parts with Kui-Huo *Fu*. On the other hand, Kui-Huo *Fu* has only one shared part with Jiu-Tian Zhuo-Long Fei-Dan *Fu*.

As discussed in Chapter 4, the number of shared parts affects the relationship between *Fus*. For the three *Fus* shown in Fig. 4-2, it can be seen that there are 7 parts shared between Yuan-shuai Da Huo-Ling *Fu* and Kui-Huo *Fu*, while Kui-Huo *Fu* and Jiu-Tian Zhuo-Long Jiu-Tian Zhuo-Long Fei-Dan Fu *Fu* have only one shared part.

How to calculate the relationships between every pair of *Fus* contained in the DFHY is the important issue in this study. The author proposed the network built of *Fus* and parts contained in them and then attempted to analyze the relationships between parts and *Fus* in this network.

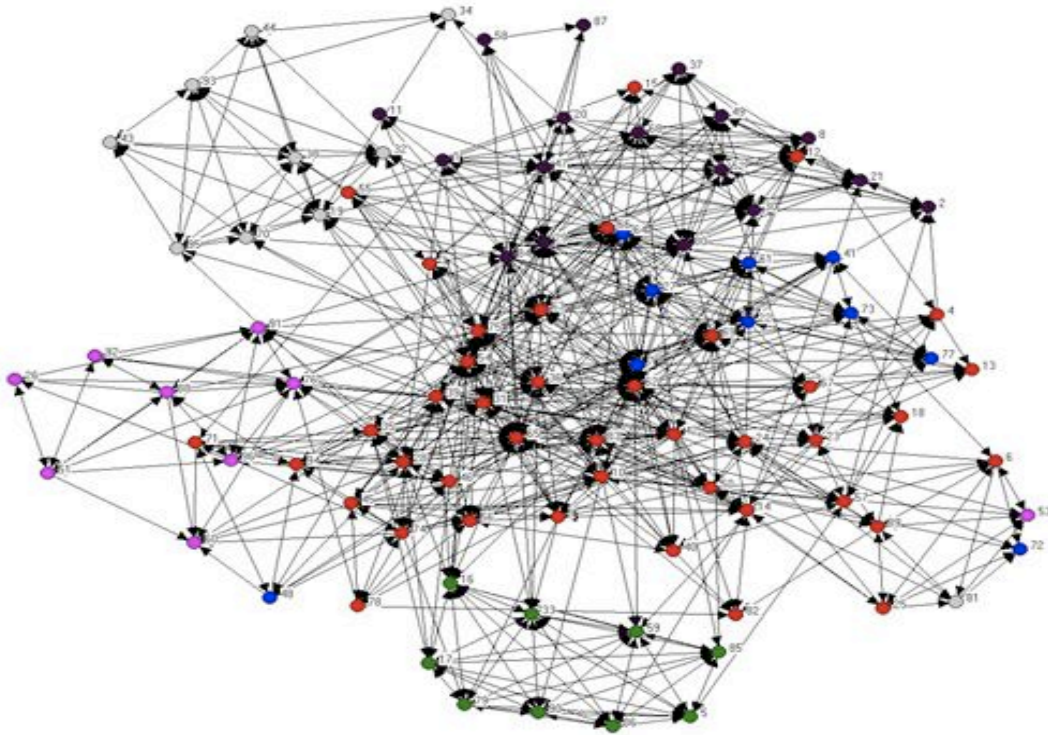


Fig. 4-3. An example of a social network

Network analysis has been applied to many fields, such as social, behavioral sciences, economics and computer science. Network analysis is also called network theory. In network theory, the node, the curial component in the network, relates with other nodes in the network [30]. The general social network is studied as the characteristics of the social network [31]. As shown in Fig. 4-3, nodes in social network science represent a person.

The relationships between persons can be calculated by social behaviors and then persons in the social network can be catalogued into groups.

By comparing the shapes of the parts contained in each *Fu*, and getting the numbers of shared parts from previous works, *Fus* and parts can be connected and the network of *Fus* can be created. The author referenced the social network analysis methods and tried to analyze the relationship between *Fus* to apply the analysis results to Daoism research [36].

The network of *Fus* is analogous to the organizational network shown in Fig. 4-3. In this study of the Digital DFHY, a part stands for each individual and the *Fus* stand for the organizations in the network.

Comparing the Digital DFHY with general social networks, it has some special features as follows.

- 1) The relationship between parts and *Fus* is not exactly the same as between individuals and the organizations in a social network. In most cases, a part is given a single meaning but there are some parts given more than one meaning.
- 2) The meaning of a *Fu* cannot be determined entirely by the meaning of parts contained in the *Fu*. As we have mentioned, the meaning of a *Fu* is complex and hard to understand. And the mistakes that occurred during the writing process by the original schools are especially hard to find.
- 3) Understanding the relationships between *Fus* is an important issue for Daoism researchers. They also want to explain the relationships between volumes to understand the history of DFHY. The network of *Fus* developed in this study cannot express the relationship between volumes and other instances contained in DFHY.

Because each of the *Fus* is connected to other *Fus* by one or more parts and these parts are named shared parts, the number of shared parts is used to indicate the relationship between each pair of *Fus*. However, in some particular cases, even though some parts looked very similar to each other, they do not have the same meaning. In this study, the author first considered the number of shared parts contained in *Fus* as the general values and the exceptions for the shared parts are left for future work.

In the example shown in Fig. 4-4, there are 10 parts in *Yuan-Shuai Da Huo-Ling Fu* (元帥大火鈴符), and *Kui-Huo Fu* (夔火符) and 9 parts in *Jiu-Tian Zhuo-Long Fei-Dan Fu* (九天捉龍飛丹符). All of these parts are picked up from the *San-Xing Fus*. Shared parts are drawn between the *Fus* using pointed arrows.

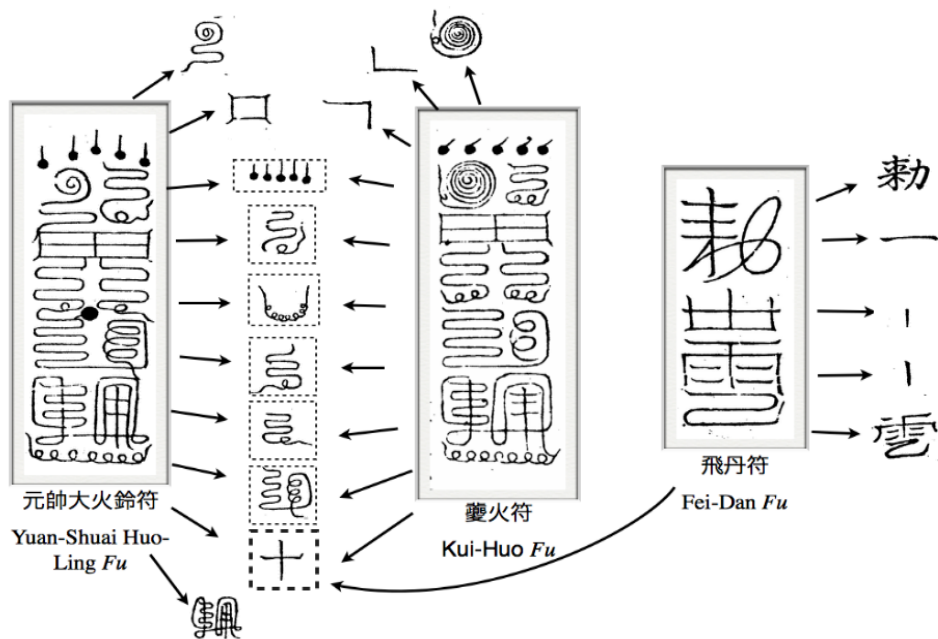


Fig. 4-4. The shared parts in the three *Fus*

Structurally, *Yuan-Shuai Da Huo-Ling Fu* is more similar to *Kui-Huo Fu* than *Jiu-Tian Zhuo-Long Fei-Dan Fu*, because there are 7 parts shared between *Yuan-Shuai Da Huo-Ling Fu* and *Kui-Huo Fu*, which is more than between *Yuan-Shuai Da Huo-Ling Fu* and *Fei-Dan Fu*. In other words, the distance between the *Fus*, i.e., inter-*Fu* distance, determined by the number of shared-parts indicates the closeness in the meaning of or usage between each pair of *Fus*.

For the shortest path between *Fus*, the author referenced the Dijkstra's Algorithm [32]. The author defined the parts and *Fus* as nodes in the network of *Fus*. In Fig. 4-5, a process of finding the shortest path between *Fus* is shown.

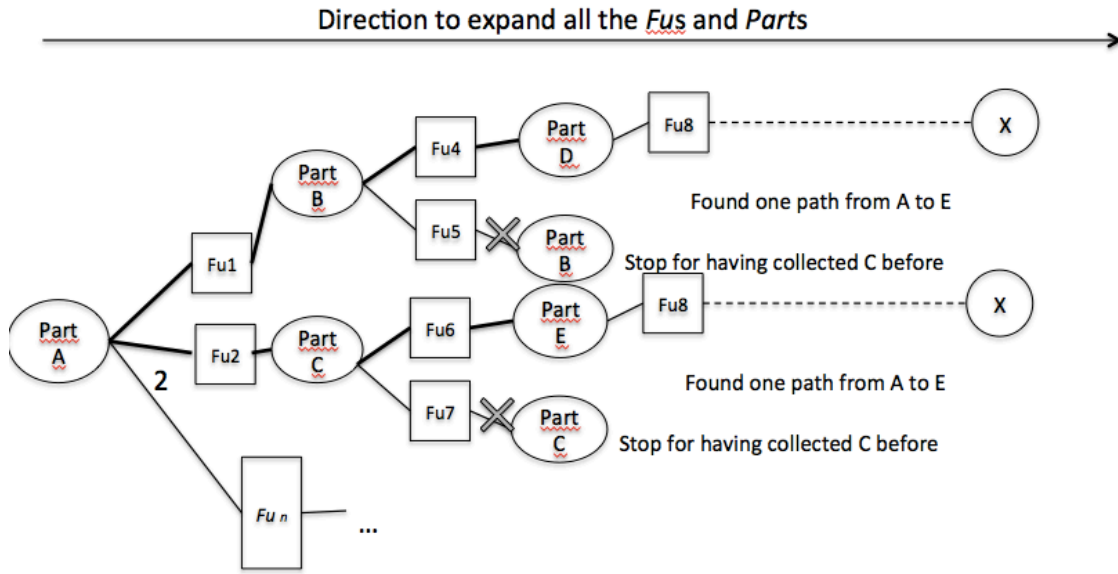


Fig. 4-5. Process to find all shortest paths from part A to all other connected *Fus*

In Fig. 4-5, nodes A, B, C, D, E each represent a part. For finding all connections having the same shapes with part-A contained in other *Fus*, all of the *Fus* containing part A are listed in the second step. Then all the parts except part A are listed after getting all *Fus* contained the part-A. In this way, all of the *Fus* containing these parts are listed. The process of listing *Fus* stops if the same parts are found in the list of *Fus*. All of the connections are saved into the database. Thus, using the inter-*Fu* distance, the network of *Fus* shown in Fig. 4-5 is created. In the network, the parts contained in the *Fus* are shown. The mapping shows that most of the *Fus* are connected by shared-parts. *Yuan-Shuai Da Huo-Ling Fu* and *Jiu-Tian Zhuo-Long Fei-Dan Fu* are also included in the network. As the *Fu ID* indicates in which volume a *Fu* is contained, the relationships between volumes and *Fus* and their parts can also be checked. In this way, all of parts, *Fus*, and volumes are included in the network. All of the *Fus* are connected with each other by the shared-parts. For example, those *Fus* located at the center of the network

have the closest connections. In other words, they are the *Fus* that have the most frequently used parts.

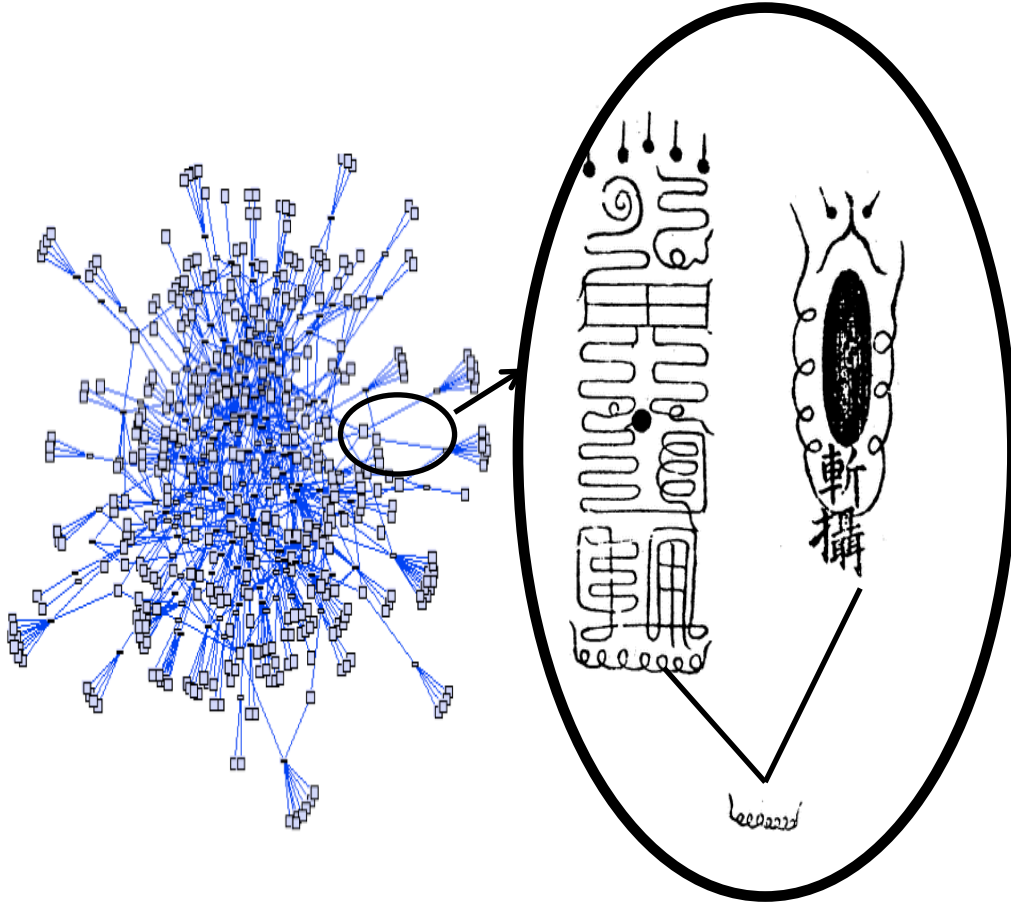


Fig. 4-6. Network of the *Fus* and their parts in DFHY

In order to calculate the relationship values for *Fus* in the network, we defined the algorithm below:

d---Distance between two parts i, j (d_{ij}): the number of the *Fus* on the path between two parts on the graph.

t---The shortest path length between two parts i, j (t_{ij}): the minimum distance between the two parts. i.e., $\min(d_{ij})$

w---Number of shortest paths between two parts i, j (w_{ij}): the number of paths whose lengths are equal to t_{ij} .

r---Relationship between parts r (P_i, P_j) is determined by the expression below. Here, a is a heuristic value and we have set it as 1.2. (Subscripts i, j are omitted.)

$$(r) : r = a^{w-1} / t \quad (1)$$

The relationship values are in inverse proportion to the distance of each pair of parts. The relationship value between two *Fus* (FA and FB) is defined by the expression below:

$$r(F_A, F_B) = \frac{\sum_{1 \leq i \leq m, 1 \leq j \leq n} r(P_{Ai}, P_{Bj})}{m \times n}$$

In this way, we get the relationship between *Yuan-Shuai Da Huo-Ling Fu*, *Kui-Huo Fu* and *Jiu-Tian Zhuo-Long Fei-Dan Fu* in Table 4-1.

Table 4-1. The values of relationships for the Yuan-Shuai Da Huo-Ling Fu, Kui-Huo Fu and J Jiu-Tian Zhuo-Long Fei-Dan Fu.

<i>Fu</i> pairs (<i>Fu</i> ID)	Relationship value
1180201, 2600301	2.23
1180601, 2600801	6.25
1180701, 2470101	6.84
1180802, 2470101	3.75

The number of shared parts is the fundamental factor to determine the relationship between the *Fus*. In the algorithm, the shortest path between two parts is used to calculate the relationship values. To evaluate the rationality of the shortest path, the author used the mapping for the *Yuan-Shuai Da Huo-Ling Fu*, *Kui-Huo Fu* and *Jiu-Tian Zhuo-Long Fei-Dan Fu* as shown in Fig. 4-4.

Some examples of the values of the relationships between *Fus* are shown in Table 4-1. To evaluate the method of calculating the shortest path, the author proposed to use the average value for the paths between two *Fus*. The value of the shortest path is shown in Table 4-2. The left part of the table shows the pair of *Fus* that have a close relationship. The right part of the table shows the pairs of *Fus* that have less close relationships. The relationship values between *Fus* suggest that the number of shared parts directly affects the value of relationships. However, *Fus* having no shared parts are also evaluated by the relationships between parts. In this way, the mapping of connections between *Fus* and their parts is built for expressing the relationships between all *Fus* in DFHY.

Table 4-2. Examples of the values of the relationship between *Fus*

The pair of <i>Fus</i>	Relationship value	The pair of <i>Fus</i>	Relationship value
煎鬼籙符 火繩萬丈符	5.71	神化火鷄符 造化	0.027
煎鬼籙符 北斗然骨符	5.196	追魂令符 造化	0.037
洗手符 太極六府符 <4>	4.28	通天關符 造化	0.041
煎鬼籙符 辟屍千里符	3.81	五方陽雷符 (四季主令十 五員即毒陽雷 符) 下書龜符	0.042
洗手符 太極六府符	3.57	火伯符 追魂令符	0.046
太極六府符 斗罡煞符	2.85	驅役萬靈符 通天關符	0.05
火繩萬丈符 明通報應符	2.85	五方陽雷符 (四季主令十 五員即毒陽雷 符) 追魂令符	0.05

For Daoism researchers, the relationship values of the *Fu* pairs provide useful indications about the connection and the semantic relationship between the *Fus*. The verification of the results is an important topic for the discussion in this study. The value of relationships between *Fus* are based on

the numbers of shared parts, however, there may be other factors affecting the relationships between *Fus*, which are not investigated in this study. The distances between *Fus* are set by the value of relationships between them. The author used two methods to evaluate the methods to calculate the relationship between *Fus*. One is to analyze relationships between volumes in DFHY by using the values from the relationship between *Fus*. The other is evaluation made by Daoism researchers.

The evaluation methods still have some limits as follows.

1. The values for the relationships between *Fus* are only a part of all the relationships between all *Fus*. As we explained in Chapter 3, *Fus* having no San-Xing *Fu* are not divided into parts in the original contents, so the methods in our study cannot be applied. How to manage and analyze data for *Fus* that have only Ju-Xing *Fu* still needs further investigation to extract parts efficiently.
2. Some volumes, which contain no *Fus* have not been included in the *Fu*-Volume relationship analysis.
3. Other kinds of graphic instances are left for future analysis. As discussed before, the *Fu* ID presents what volume and what page a *Fu* is on. So in the network of *Fus*, it can be checked where *Fus* having close relationships are set in DFHY and in this way, analysis of the network of the *Fus* and relationships between them can give the researchers some hints to uncover the secrets that are difficult to find from the original contents. Distribution of the relationship between each pair of *Fus* can be checked. Further analysis on the contents of DFHY can be performed.

Through the process of digitization of DFHY, we found that by analyzing the contents of DFHY (such as the shapes of parts, the relationships on the *Fus*, etc.) we can supply researchers with analysis results to help them find some of the secrets hidden in the contents.

To deepen our study of the contents, we studied the needs of Daoism researchers and tried to develop analysis methods to support them with useful results. In the next chapter, some analysis methods will be introduced.

4.3 Mapping of *Fus* and parts

During the analysis process, we noticed that the location of every *Fu* and part is important to understand both the logical and chronological sequences of the original contents. Different from the calculation of the relationship between *Fus*, to create the network connected by *Fus* and parts aims to help Daoism researchers find hidden rules for parts as hints for their study.

Some examples of the relationships between parts are listed in Fig. 4-6. The figure contains attributes and values for the relationships of shared-parts. The author counted how many times each shared-part appears in a single volume.

The distribution of the relationships of parts and *Fus* are also mapped. Table 4-3 shows that the shared parts mostly are found in Volume 162 to Volume 268. It is found from the analysis results that one shared-part in Volume 164 was contained in more than 30 *Fus*.

Table 4-3. Data for shared parts

Texts of Parts	Parts Shape	Shape ID	Contained in <i>Fu</i>	In Volume
穿鬼心破鬼肚門 (Chuan Gui Xin Po Gui Du Men)		sf030208	<i>Yuan-Shuai Da Huo-Ling Fu</i>	162
東弗手岱火輪轉行 (Dong Fu Shou Dai Huo Lun Zhuan Xing)		sf010103	<i>Kui-Huo Fu</i>	268
妙道上帝勅招寅 (Miao Dao Shang Di Ci Qia Yin)		sc021302	<i>Jiu-Tian Zhuo-Long Fei-Dan Fu</i>	195

Fig. 4-7 shows distribution graphs for shared-parts in the volumes. In this chart, the x-axis and y-axis show the volumes sorted by the volume number given in DFHY, and the number of occurrences of a part in each volume, respectively. In this chart, the author uses different symbols to express the different shapes of shared parts. As mentioned in Chapter 3, the parts are cataloged where parts are indexed by their shapes and the shape IDs are defined. The maximum value in the chart means the most frequently used shared parts. Volume 262 contains a part whose shape ID is sf030102 as it is used more than 30 times.

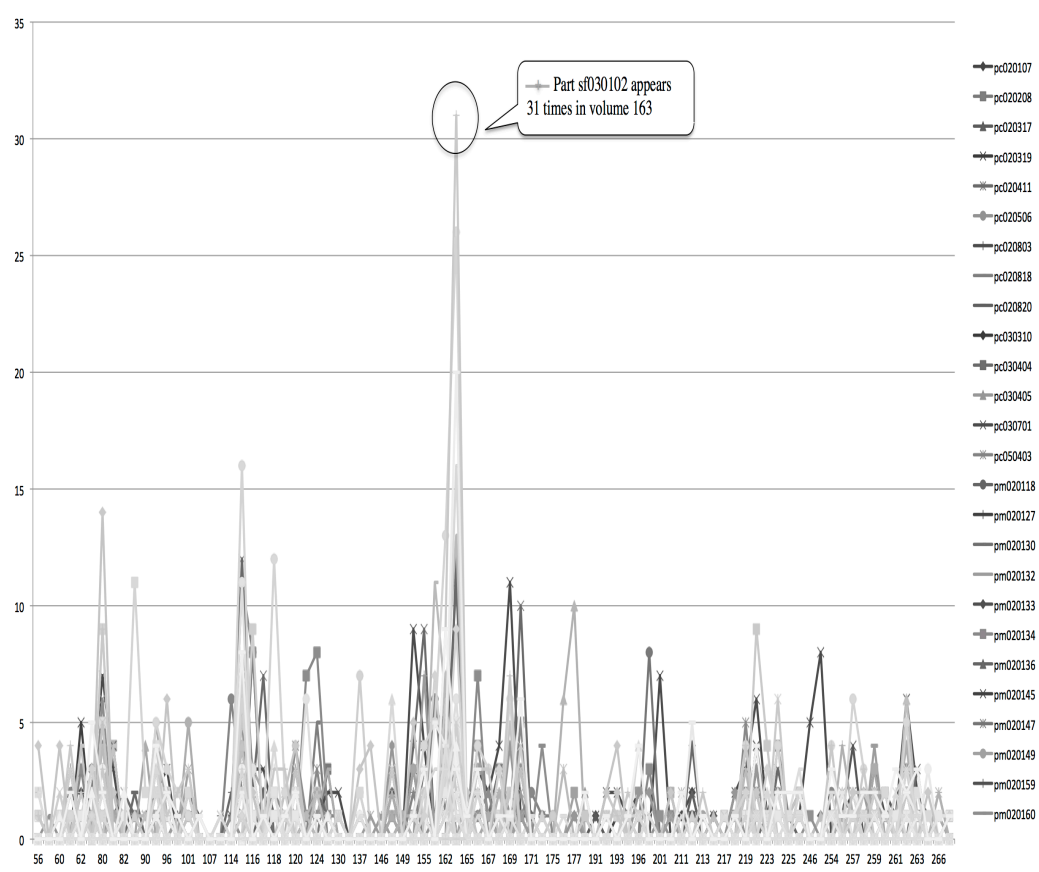


Fig. 4-7. The distribution map of shared-parts (part)

Fig. 4-8 shows a few shared parts extracted from Fig. 4-7. In this chart, three parts are selected and their distribution among volumes can be found

in this chart. It shows for these three parts, which one is used most among the *Fus* and in which *Fus* and volumes they are contained.

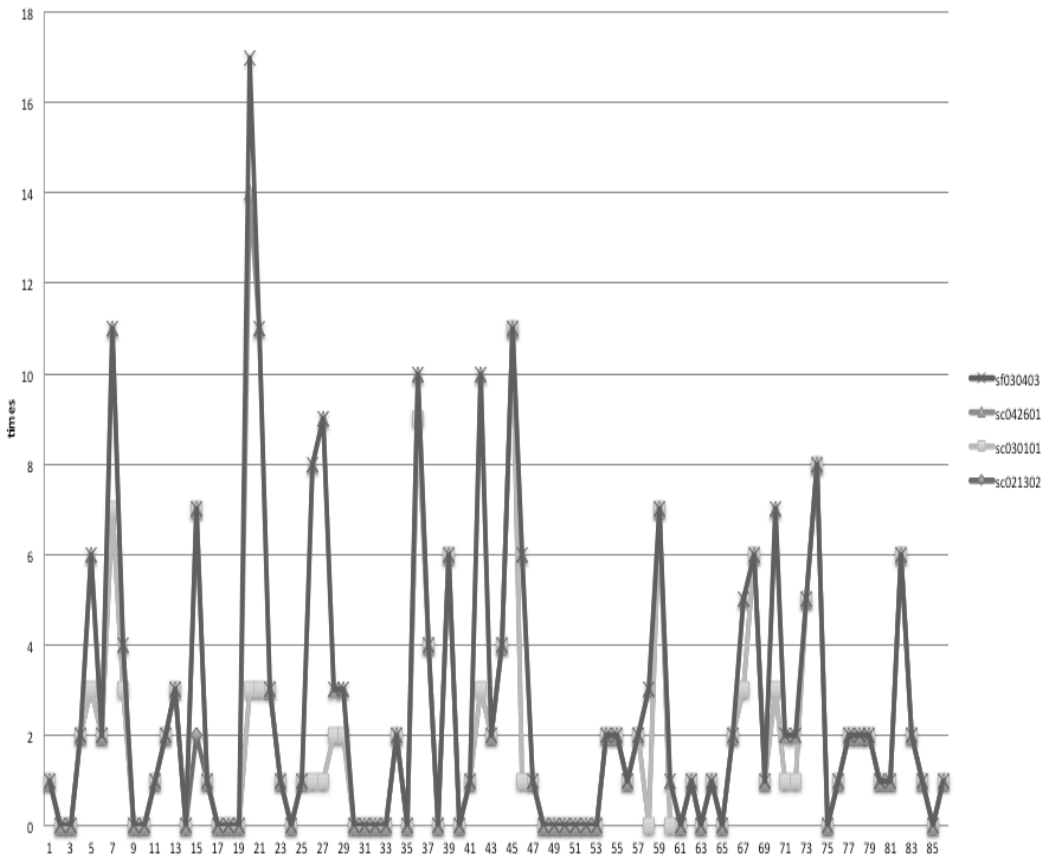


Fig. 4-8. Distribution of the three parts - sf030208, sf010103, sc021302

By studying the relationship values of all pairs of *Fus*, Daoism researchers can see that a pair of *Fus* shares the same meaning if the pair contains one or more shared parts.


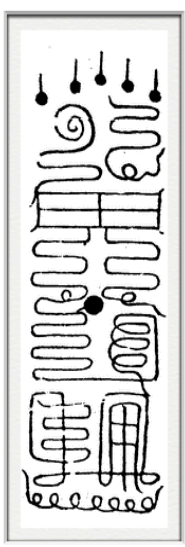

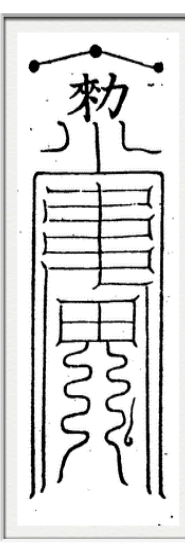
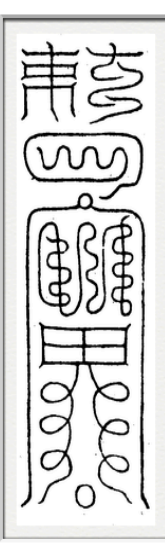
Fu	Related Fus			
				
page 11, Volume 268 Kui-Huo <i>Fu</i>	page 28, Volume 162 Yuan-Shuai Huo-Ling <i>Fu</i>	page 10, Volume 168 San-Kui Lu <i>Fu</i>	page 10, Volume 166 Kui-Zhi-Ran <i>Fu</i>	page 6, Volume 246 Qi-Tu <i>Fu</i>

Fig. 4-9. Example of the ranking of *Fus* relationships

The values for the relationships between Kui-Huo *Fu*(夔火符) and related *Fus* are presented in Fig. 4-9. *Fus* related to Kui-Huo *Fu* are sorted in descending order by their relationship value.

By checking the structure of each of the related *Fus* for Kui-Huo *Fu*, Yuan-Shuai Da Huo-Ling *Fu* not only has the most shared-parts with Kui – Huo *Fu* but also a similar structure. A *Fu* is an important symbol to identify Daoism schools. The structure of a *Fu* is the key factor in the school identification process. In Fig. 4-9, not only the relationship values of *Fu* pairs but also the relationships between *Fus* and volumes can be investigated. By comparing the structure of each *Fu* in this list, we can see the features of the structures of related-*Fus*. Comparing the related *Fus*' meanings with the structure, it is possible to find some hidden relationships, which are not explicitly written in DFHY.

In the analysis of this study, the author found that Kui-Huo *Fu* and Yuan-Shuai Da Huo-Ling *Fu* have a close relationship with each other. The

close relationship between Kui-Huo *Fu* and Kui-Zhi-Ran *Fu* (the third *Fu* in the list) is a new finding made from our research.

More analysis results about the relationships between *Fus* and parts contained in *Fus* are presented in the Appendix A and Appendix C.

4.4 Sorting Relationships of Volumes as Evaluation of the Relationships of *Fus*

One of the most important issues in this study is to find the relationships among volumes using the Digital DFHY. DFHY states that some volumes are compiled in accordance with Daoism schools. However, the relationships between volumes and schools are not yet clear in the Daoism research community. Assuming that *Fus* would represent styles of charms and secret terms owned by a Daoism school, the author proposed to use the *Fus*-Volumes relationships as a hint to find Volume-School relationships.

Because about 20% of the volumes of DFHY contain no *Fus*, the author analyzed the inter-volume relationships, i.e., the relationships between volumes connected by *Fus*. The average values of all the relationships between *Fus* contained in volumes are used to stand for the relationships between volumes.

The expression below is the formula to calculate the inter-volume relationship value for a pair of volumes:

$$r(V_A, V_B) = \frac{\sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} r(f_{Ai}, f_{Bj})}{m \times n}$$

Here, f_{Ai} is a *Fu* of volume V_A , f_{Bj} is a *Fu* of volume V_B , and $r(f_{Ai}, f_{Bj})$ is the relationship value between f_{Ai} and f_{Bj} . Topologically, all volume pairs were sorted using the volume-volume relationship values. Table 4-4 shows some of the results of relationship values between volumes.

The result shows the volume pair (118, 260) has the maximum inter-volume relationship value in the volumes pairs. And the volume pair (118, 247) has a close relationship, while the volume pair (260, 247) has a looser relationship.

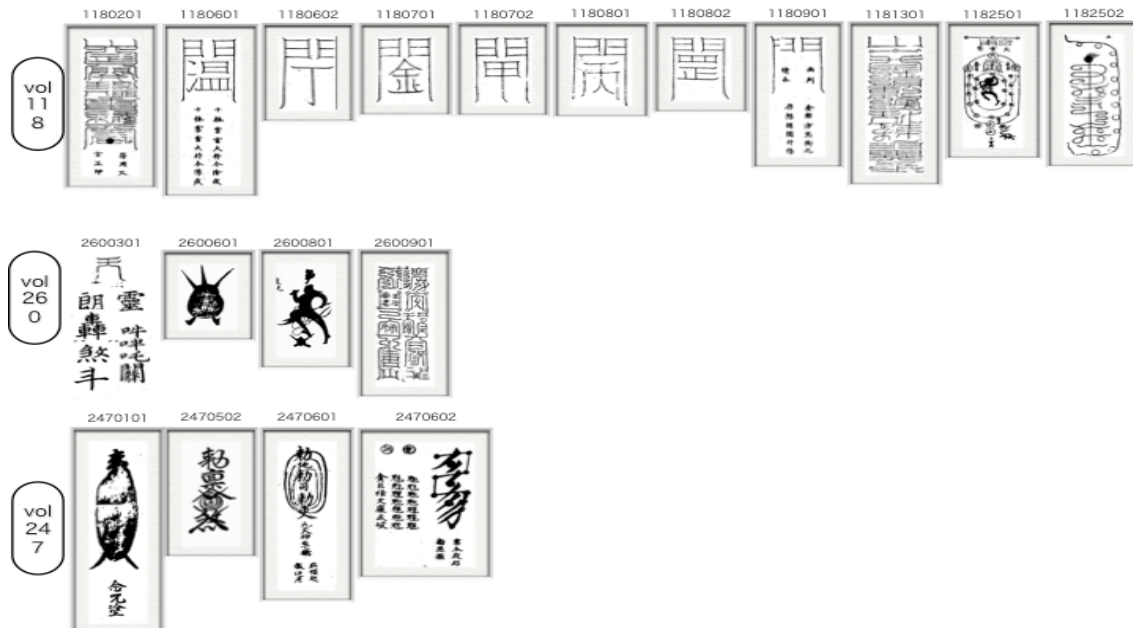


Fig. 4-10. Example of shapes of Ju-Xing *Fu* contained in volumes 118, 260 and 247

As shown in Table 4-4, there are all the *Fus* contained in the volume 118 and Volume 260. The formula to calculate the inter-volume relationship value the author proposed shows that the relationships between *Fus* contained in volumes effect the inter-volumes value directly. The author tried many methods and found that the *Fus* pairs having high relationship effect the inter- value relationships value more than *Fus* pairs having less relationships. So the author set the value of relationships between *Fu* pairs less than the average value among all the *Fus*- pair as 0.

Table 4-4. Relation values for *Fus* contained in volume 118 and volume 260

Fu ID in Vol. 118	Fu ID in Vol. 260	Relation values
1180201	2600301	0
1180201	2600601	0
1180201	2600801	0
1180601	2600301	8.16526E+13

1180601	2600601	0
1180601	2600801	4.40671E+13
1180602	2600301	1.36088E+14
1180602	2600601	0
1180602	2600801	7.3402E+13
1180701	2600301	2.04131E+14
1180701	2600601	0
1180701	2600801	1.10103E+14
1180702	2600301	2.44958E+14
1180702	2600601	0
1180702	2600801	1.32124E+14
1180801	2600301	1.36088E+14
1180801	2600601	0
1180801	2600801	7.3402E+13
1180802	2600301	2.04131E+14
1180802	2600601	0
1180802	2600801	1.10103E+14
1181301	2600301	0
1181301	2600601	0
1181301	2600801	0
1181501	2600301	0
1181501	2600601	0
1181501	2600801	4.92742E+11
1181901	2600301	0
1181901	2600601	0
1181901	2600801	0
1182502	2600301	0
1182502	2600601	0
1182502	2600801	0
	Summary	1.55074E+15

By using the formula to calculate the inter-volume relationship value the author proposed, the relationships between volume 118 and volume 260 are calculated as following:

$$\frac{1.55074 \times 10^{15}}{13 \times 4} = 2.98 \times 10^{13}$$

Then all the relationships between Volume 118, 260, and 247 are calculated as followings:

Table 4-5. Example of inter-volume relationship values.

Volume pairs	Relationship value (E+13 is omitted)
118, 260	2.98
118, 247	2.59
260, 247	0

For in the volumes having *Fus* contained in are all calculated and the values for the relationships between inter volumes are presented that only about 12 pair of volumes has the high relationships between each other. The author presented the relationship values omitted the E+13.

In Fig. 4-10 the shape of Ju-Xing *Fu* for each *Fus* contained in the volumes 118, 260 and 247 are presented. Table 4-5 shows the inter-volume relationships. It can be seen that the number of *Fus* affects the values of relationships between volumes.

In Fig. 4-12, all the *Fus* having San-Xing *Fu* and Ju-Xing *Fu* contained in volume 118, 260 and 247 are listed in the graphics. The author used the Ju-Xing *Fu* to present each *Fus*. As shown in Fig 4-10, there are 11 *Fus* contained in volume 118 and 260, while volume 247 contained 4 *Fus*. It can be proposed that Volume 118 containing more *Fus* has a higher potential for having closer relationships with another volumes than Volume 260 and Volume 247. Even though the relationships between them still need to be tested, the results are an attempt for analysis on the inter-volume

relationship, which is an important issue for Daoism researchers. It can be proposed that the shapes of *Fus* are the important factor for the relationships between volumes.

Fig.4-11 shows the analysis results of the volumes having *Fus* contained in them. This figure shows that volumes 260 and 247 have a close relationship, which was not known previously by Daoism researchers.

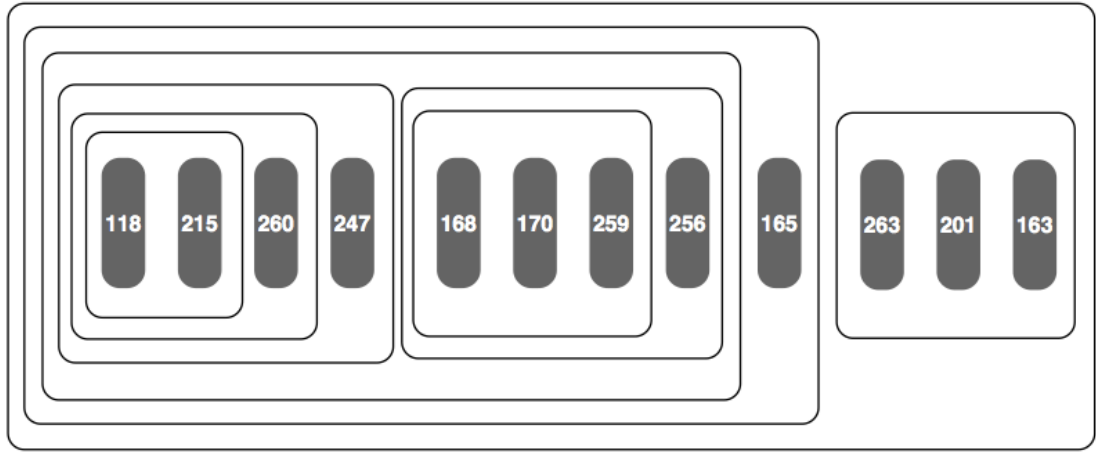


Fig. 4-11. Relationship between volumes (part)

All the relationships between volumes are expressed in Table 4-4, however, for Daoism researchers, to find the important volumes pair among DFHY is helpful instead of the number of the relationship values. According to the analysis results of the values of relationships between volumes, the volume pair (118, 215) have the maximum value. And the author selected the volume 118 as a standard item of the first class. Then, by the values of the relationships, Volume 260 has the close relationship with Volume 118, so Volume 260 is arranged as the second class. On the other hand, volumes 168, 170, 259 have close relationships with each other, and Volume 168 also has a close relationship with Volume 118, so volumes 168, 170, 259 are arranged as the 4th class.

In this way, the relationships between volumes are expressed in classes by the values of the relationships of each pair as shown in Fig. 4-13. Now the

inter-volumes relationships are easier to check. More details of the relationships between volumes are shown in the Appendix.

In this study, not all of the *Fus* included in DFHY were used for the relationship analysis. The main reason is the difficulty of extracting parts from *Fus* , which are not expressed as San-Xing *Fu*. It is left as a future study to extend the coverage of the relationship analysis to those *Fus* not examined in this study.

4.5 The user interface of the Digital Dao-Fa Hui-Yuan system

Building a user-friendly interface for the digital DFHY is an important aspect for this study.

It is commonly recognized that the digitized cultural heritage contents need to be managed in a contents management system and the management system should be designed so as to present the contents completely [33]. The Digital DFHY should supply researchers with quick and convenient functions.

Digital DFHY has functions to help users browse and search the digitized *Fus* and other aspects of DFHY. Using the Digital DFHY, researchers cannot only search the contents, such as the images for pages, *Fus* and parts, but they can also analyze the relationships between *Fus*, parts and volumes. The first step in building the whole system is to decide how to arrange the digitized contents. The author first created a category to identify the shape of parts [34] and by using the category of the shape of parts, *Fus* having shared parts can also be presented in the system.

Fig. 4-12 shows how the Digital DFHY is built step-by-step. First, input the digitized contents such as *Fus*, parts, pages into the database, and make the contents in the system of the Digital DFHY available for searching the information about the original contents. The analysis of the *Fus* relationships is extended to analysis of the relationships between volumes. Then the network of all *Fus* helps search for the related *Fus* in volumes. This also gave researchers new hints to find the hidden rules in DFHY. Then the relationship values for each pair of volumes were calculated and these results are shown in the Digital DFHY.

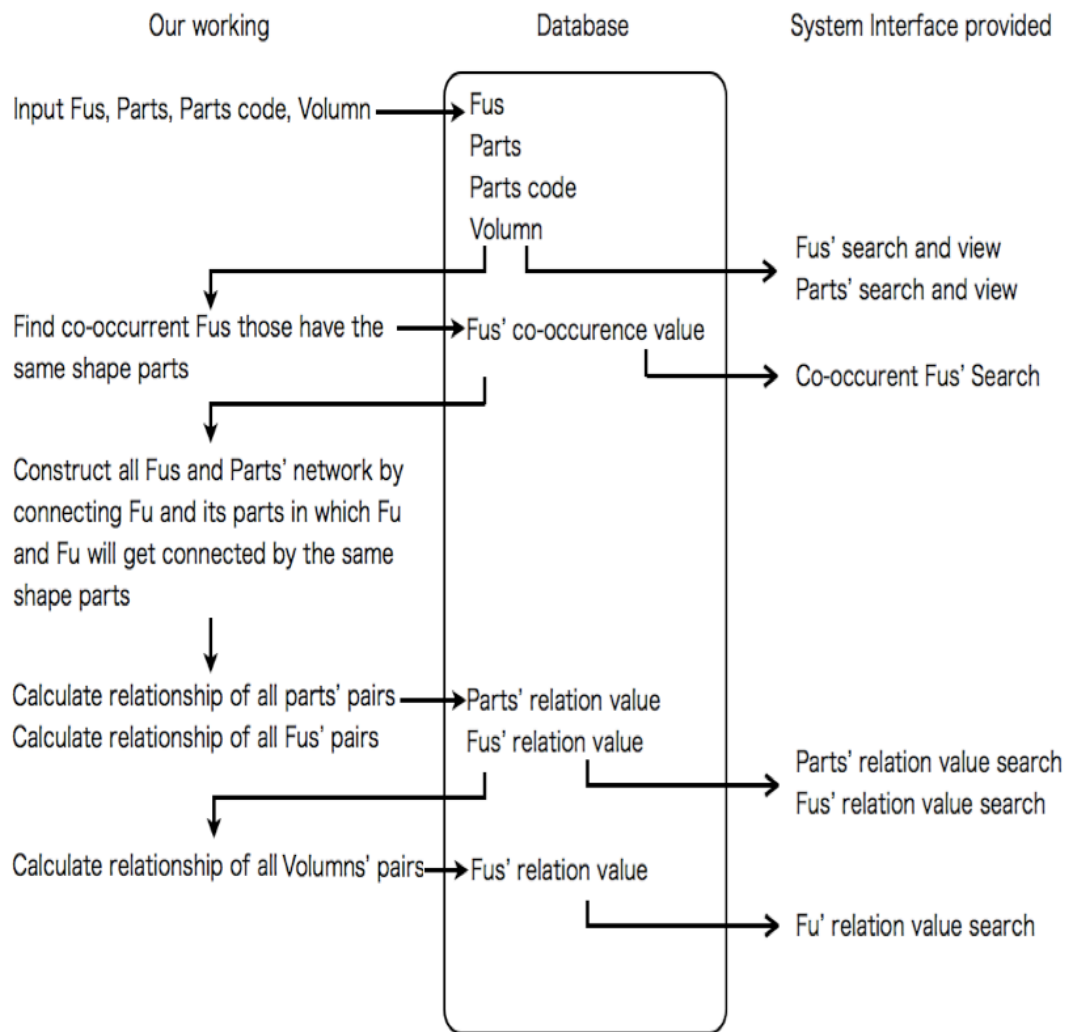


Fig. 4-12. Flow chart for the building of the Digital DFHY system

The Digital DFHY is designed for preservation and access to the original content of DFHY and, in addition, it shows the analysis results of the content for supporting Daoism researchers. The system flow chart of the Digital DFHY contains two parts. One part is for presenting the original content of DFHY and the other is for presenting the analysis results.

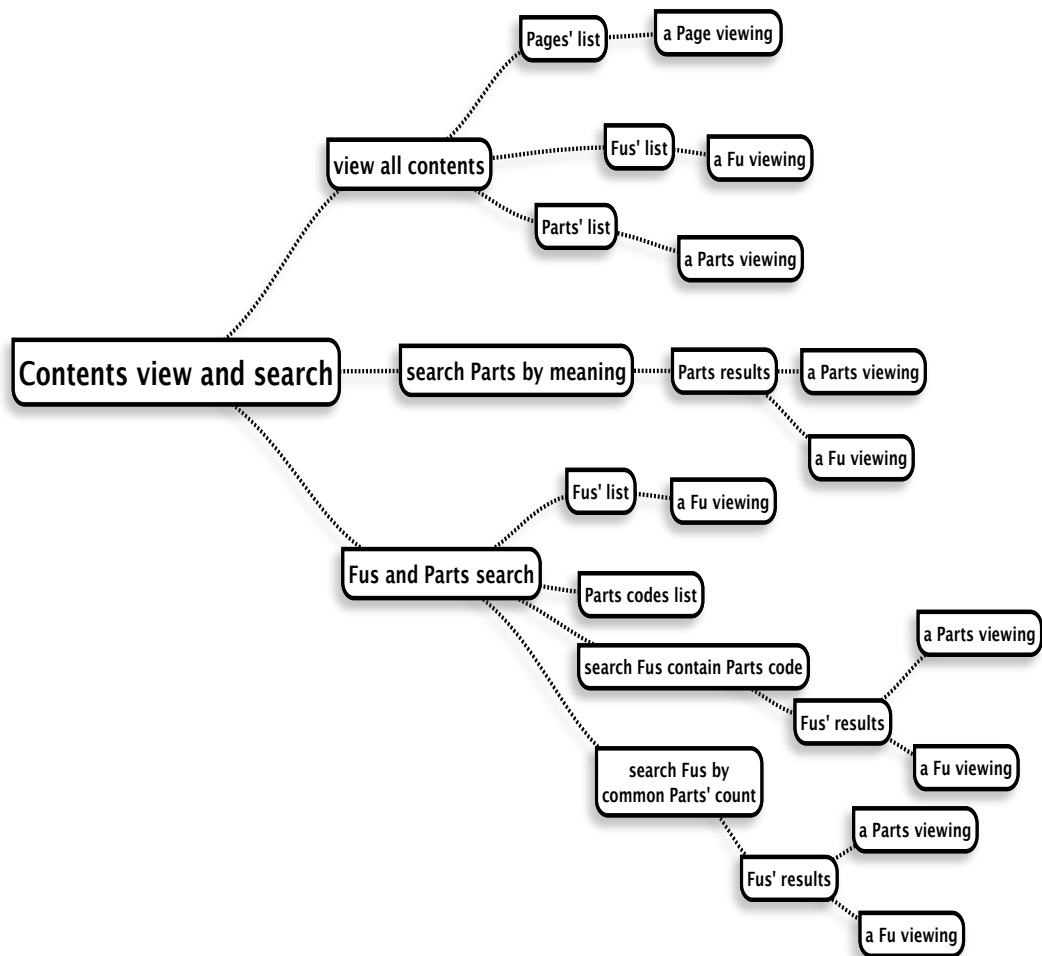


Fig. 4-13. The process chart of the Digital DFHY for presenting the original contents of the Digital DFHY system

In Fig. 4-13, the system process chart for Digital DFHY for presenting the original contents of DFHY is shown. The functions for viewing and searching for the content include pages, *Fus* and parts. For searching the original content, the data of graphics and texts should be in the system and connected to page and graphic symbols. Arranging, sorting, and cataloging the original data is set to be used for analyzing the content.

Fig. 4-1 shows some analysis results of the content of DFHY.

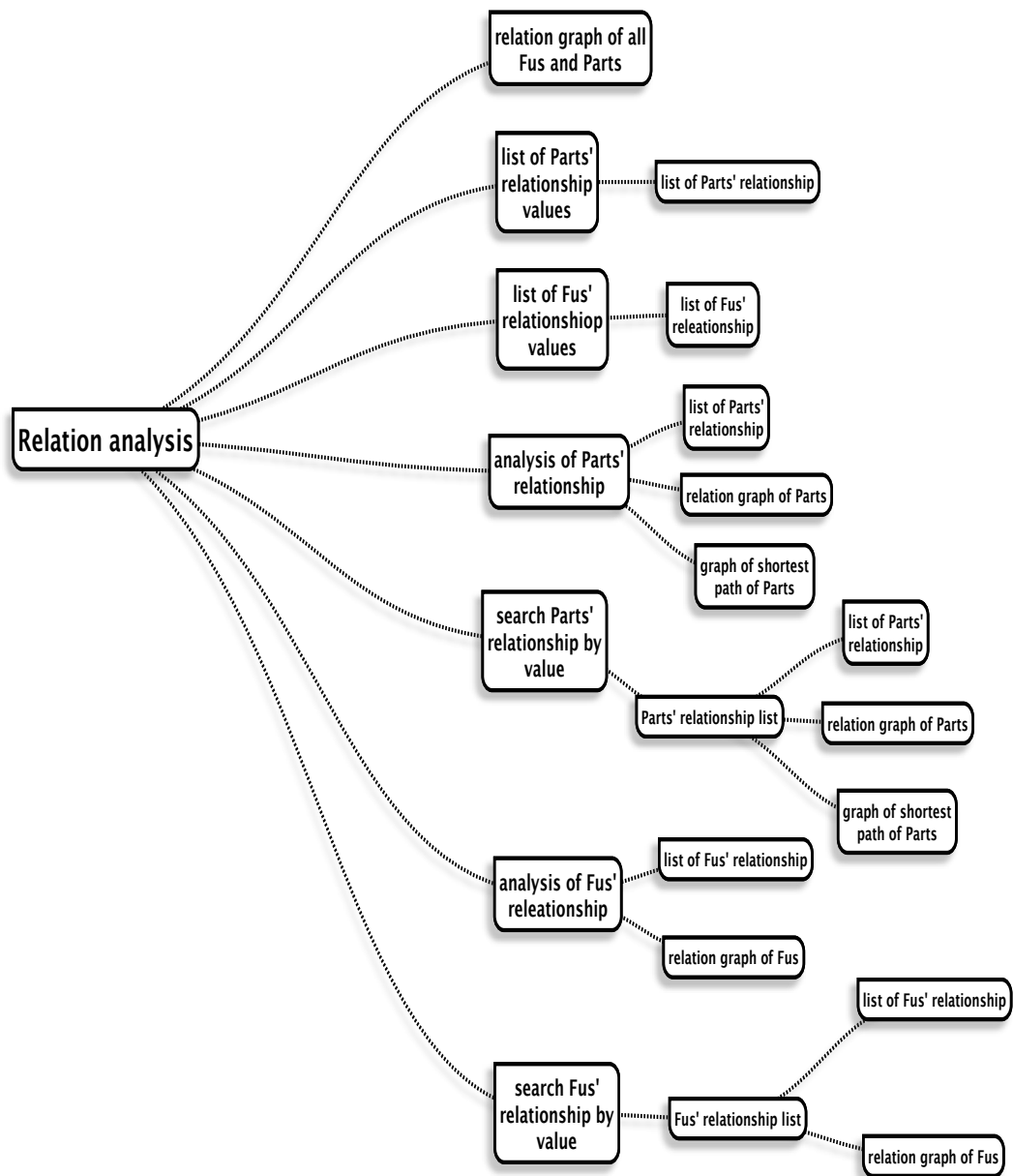


Fig. 4-14. The process chart of the system of Digital Dao-Fa Hui-Yuan for showing the analysis results

In Fig. 4-14, the relationship analysis in the digital DFHY includes the relationship graph of all *Fus* and parts, such as list of *Fu*'s relationship values, parts relationship values and so on. The relationships between

volumes and parts are part of analysis results. Even the distributions for *Fus* in volumes, parts in *Fus* and volumes exist in DFHY, as trying to find this distribution is helpful for the work of Daoism study.

User interface is an important topic for the overall design of the Digital DFHY.

In some cases, Daoism researchers need to search for the parts by the shape of parts instead of searching by title or illustration text. To design a user interface for users to view the relationships between *Fus* and parts as well as the pages, the following parameters need to be investigated - shape of the parts, name of the *Fus*, list order of the *Fu* pairs (in ascending value), pairs of parts contained in *Fus*. Fig.4-18 shows parts prepared as a parameter. Obviously, the list of the shapes of parts is limited to the images stored in the database. In Fig. 4-18, 90 shapes are available in the option part of the interface. However, they do not include all of the shapes of parts in DFHY. In future work, the *Fu* having only Ju-Xing *Fu* (聚形符) should be included.



The screenshot shows the '道法會元' (Daofa Huiyuan) search interface. The top navigation bar includes links for '1. 閲覧検索' (Browse Search), '2. パーツ/符関連度分析' (Parts/Symbols Relatedness Analysis), and '3. パーツ/符関連度分析 (2)' (Parts/Symbols Relatedness Analysis (2)). The main content area displays a list of parts (Fus) and their corresponding page numbers. The interface includes a search bar, a list of parts, and a table of results.

目録	ページ	目録	ページ
2	479p	\N	目録第01
3	480p	目録第02	目録第03
4	481p	目録第04	目録第05
5	482p	目録第06	目録第07
6	483p	目録第08	目録第09
7	484p	目録第01	目録第01
8	485p	目録第01	目録第01
9	486p	目録第01	巻1第01

Fig. 4-15. Page view in the Digital DFHY

Fig. 4-15 is one example of the interface for viewing the original content. A DFHY page is shown on the right when the page is chosen in the menu on the left.

The screenshot shows the '道法會元分析' (Do-Ho-Kai-Gen Analysis) interface. The left sidebar contains navigation options like 'ページ閲覧' (Page View) and '符閲覧' (Character View). The top header shows the user 'guest' and search filters. The main table lists search results with columns for item number, title, volume/page, and shared parts (Fus). A callout box highlights the '黒煞大将符' entry, with an arrow pointing to the 'Select the title of *Fus* or and the numbers of shared-parts.' instruction.

Item	Title	Volume	Page	Shared Parts (Fus)
1	黒煞大将符	1651301	巻165第13	muw102 muw015 muw050 muw005 muw006 muw038 muw018 muw019 muw074 muw106
2	北帝真鏡符	1663002	巻166第30	muw155 muw102 muw260
3	二變錄符	1680801	巻168第08	muw023 muw131 muw059 muw008 muw009 muw061 muw009 muw130 muw106 muw007
4	煎鬼藤符	1681201	巻168第12	muw136 muw106 muw022
5	太一救災符	1681701	巻168第17	muw266 muw107 muw124 muw102 muw012 muw010
6	又催災符	1691501	巻169第15	muw107 muw107 muw107 muw107 muw102 muw120 muw114
7	天蓬刀圍符			muw112 muw133 muw133 muw278 muw140 muw279 muw106 muw007
8	熾火符			

Fig. 4-16. Interface of the digital DFHY for supporting users to select the title of *Fus* and the number of shared parts to study the contents

Fig. 4-16 shows a function of searching for the related *Fus* selected in the Digital DFHY. Users select the number of shared-parts for a *Fu* and then find out all the *Fus* having shared parts with it. For Daoism researchers, this function of viewing must work conveniently so a *Fu*'s title is listed in the selection options. In Fig. 4-16, all *Fus* having more than one part shared with Hei-Sha Da-Jiang *Fu* is requested, and the list of titles of those *Fus* and parts contained in those *Fus* is shown.

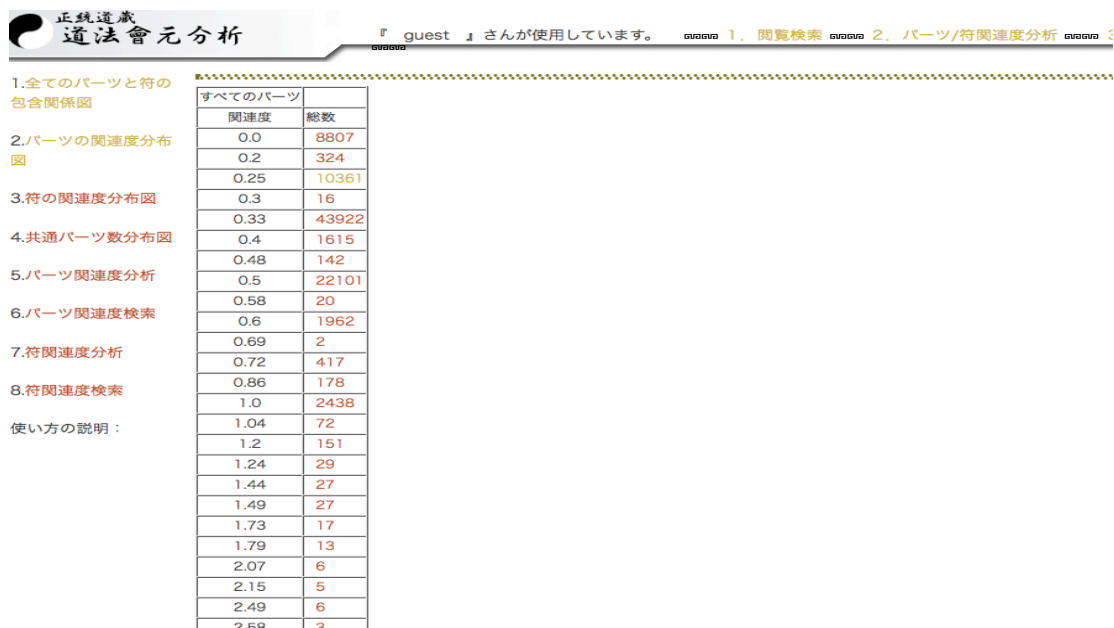


Fig. 4-17. All values of the parts relationships

In Fig. 4-17, the relationship between parts is presented. By using the list of inter-part relationships, what parts are frequently connected with each other, and what parts are never connected can be checked. Daoism researchers can check the details by clicking the number displayed in red.

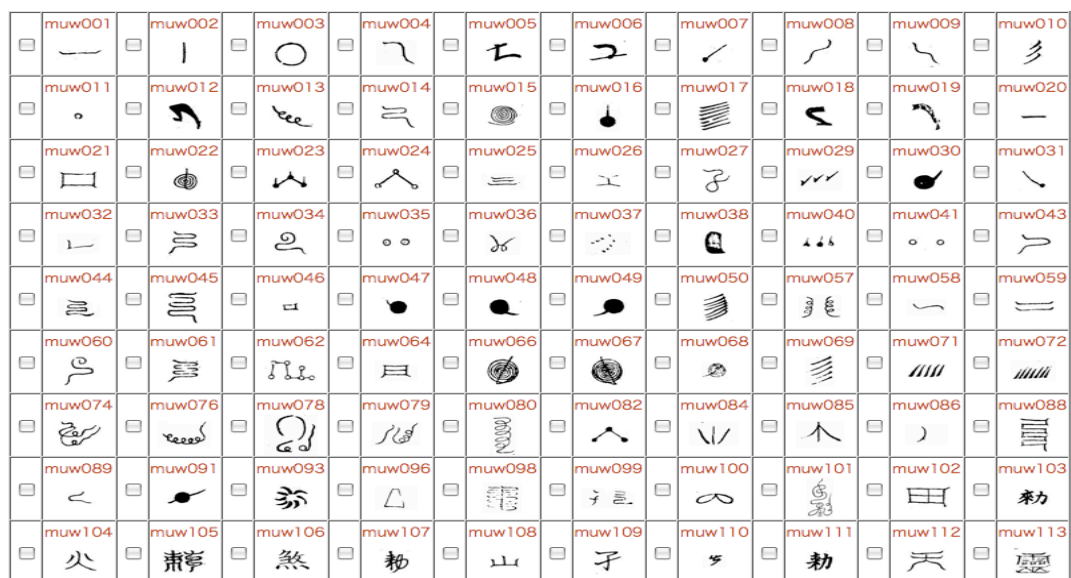


Fig. 4-18. The options for the shapes of parts

As mentioned in Chapter 1, parts are contained in *Fus* and graphics are complex and difficult for viewing and analyzing. In the Digital DFHY, how to list the graphic symbols is an issue. The author first proposed to list all of the shapes of parts as shown in Fig. 4-18. Parts contained in *Fus* are listed visually for Daoism researchers. After selecting the shapes of parts, the relationships between them and the *Fus* are shown.

Fus can be listed by their titles as shown in Fig. 4-19. All *Fu* titles are listed in the Digital DFHY and, after selecting several *Fus*, relationships between them can also be checked.

<input type="checkbox"/> 黑煞大符	<input type="checkbox"/> 北帝圓鏡符	<input type="checkbox"/> 煎鬼錄符	<input type="checkbox"/> 二變錄符	<input type="checkbox"/> 三變錄符	<input type="checkbox"/> 煎鬼錄前符	<input type="checkbox"/> 趙帥救法符	<input type="checkbox"/> 太一救捉符	<input type="checkbox"/> 變龍頭符
<input type="checkbox"/> 催捉妖魔符	<input type="checkbox"/> 又催捉符	<input type="checkbox"/> 收捕精靈諸龍 兇惡大崇寶錄	<input type="checkbox"/> 收捉鬼神并選 捕寶錄	<input type="checkbox"/> 縛符	<input type="checkbox"/> 翊聖擎山鞠邪 符法	<input type="checkbox"/> 天蓬火山燒鬼 折魔符	<input type="checkbox"/> 天蓬刀圍符	<input type="checkbox"/> 神運符
<input type="checkbox"/> 皇主定元符	<input type="checkbox"/> 丹天太皇真符	<input type="checkbox"/> 玉曆延康符	<input type="checkbox"/> 三光符	<input type="checkbox"/> 元素元輝真符	<input type="checkbox"/> 敵火符	<input type="checkbox"/> 役召萬神符	<input type="checkbox"/> 敵火真形符	<input type="checkbox"/> 辛天君符
<input type="checkbox"/> 主帥本身丙字 遍爐聖	<input type="checkbox"/> 副帥本身天中 大神遍爐符	<input type="checkbox"/> 總召飛靈符	<input type="checkbox"/> 金字捉鬼符	<input type="checkbox"/> 斗煞符	<input type="checkbox"/> 捉三煞符	<input type="checkbox"/> 捉九良星符	<input type="checkbox"/> 二帥總符	<input type="checkbox"/> 陶將軍符
<input type="checkbox"/> 氣字符	<input type="checkbox"/> 氣字符	<input type="checkbox"/> 金梁倒掛破魂 斗聖八卦攝祟 符	<input type="checkbox"/> 天丁本身玉符	<input type="checkbox"/> 斬滅三尸符	<input type="checkbox"/> 五雷混天霹靂 真符	<input type="checkbox"/> 天罡符法	<input type="checkbox"/> 靈符	<input type="checkbox"/> 天罡七火符
<input type="checkbox"/> 太乙攝捉邪巫 符	<input type="checkbox"/> 玉策符	<input type="checkbox"/> 洗手符	<input type="checkbox"/> 安魂定魂符	<input type="checkbox"/> 烈血符	<input type="checkbox"/> 煎鬼符	<input type="checkbox"/> 斬瘟符	<input type="checkbox"/> 六畜瘟符	<input type="checkbox"/> 天敵符
<input type="checkbox"/> 金磚符	<input type="checkbox"/> 火驅符	<input type="checkbox"/> 火鴉符	<input type="checkbox"/> 封山破洞符	<input type="checkbox"/> 火&MO78327 飛雷霹靂符	<input type="checkbox"/> 追魂令符	<input type="checkbox"/> 趙大符	<input type="checkbox"/> 馬帥親征符	<input type="checkbox"/> 井字符
<input type="checkbox"/> 捉魂符	<input type="checkbox"/> 靈官急降現形 符	<input type="checkbox"/> 大力符	<input type="checkbox"/> 火字符	<input type="checkbox"/> 火驅符	<input type="checkbox"/> 通天關符	<input type="checkbox"/> 都天大雷火獄 符	<input type="checkbox"/> 太一飛符	<input type="checkbox"/> 白蛇符
<input type="checkbox"/> 神化火攝符	<input type="checkbox"/> 飛捉符	<input type="checkbox"/> 速現符	<input type="checkbox"/> 封山破洞符	<input type="checkbox"/> 靈官白蛇考召 符	<input type="checkbox"/> 靈官飛符	<input type="checkbox"/> 魁字符	<input type="checkbox"/> 昨字符	<input type="checkbox"/> 白蛇大符
<input type="checkbox"/> 黑罩符	<input type="checkbox"/> 太一合同符	<input type="checkbox"/> 總召符	<input type="checkbox"/> 召正一靈官馬 元帥符	<input type="checkbox"/> 遣捉符	<input type="checkbox"/> 擲山捉祟面符	<input type="checkbox"/> 背符	<input type="checkbox"/> 馬元帥急捉符	<input type="checkbox"/> 遣捉符
<input type="checkbox"/> 魁神符	<input type="checkbox"/> 萬雷動變三火 符	<input type="checkbox"/> 玄壇和合符	<input type="checkbox"/> 治邪師符	<input type="checkbox"/> 催水衣符	<input type="checkbox"/> 和合如意符	<input type="checkbox"/> 離散符	<input type="checkbox"/> 又和合符	<input type="checkbox"/> 召遣如意符
<input type="checkbox"/> 舞袖符	<input type="checkbox"/> 九泉號令符	<input type="checkbox"/> 合景符	<input type="checkbox"/> 符法	<input type="checkbox"/> 起土符	<input type="checkbox"/> 符法	<input type="checkbox"/> 裏令符	<input type="checkbox"/> 三勒符	<input type="checkbox"/> 本身符

Fig. 4-19. The list for name of *Fus*

For Daoism researchers to search for the analysis results, all of the relationship values between *Fus* and parts are counted. The number of pairs of *Fus* and parts and the relationship values between all the pairs are provided. Daoism Researchers can perform statistical analysis using the data. The data should give Daoism researchers some hints to progress their studies.


















	muw102	muw015	muw050	muw005	muw006	muw038	muw018	muw019	muw074	muw106
										
muw024 	1.2	1.0	0.5	0.72	0.72	0.6	0.6	0.6	0.5	0.86
muw062 	1.0	1.0	0.6	0.6	0.6	0.6	0.6	0.6	0.5	0.5
muw142 	0.5	0.5	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
muw142 	0.5	0.5	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
muw151 	0.6	0.5	0.4	0.5	0.5	0.4	0.4	0.4	0.4	0.4
muw088 	0.5	0.5	0.33	0.4	0.4	0.33	0.4	0.4	0.33	0.4
muw120 	1.0	0.6	0.5	0.6	0.6	0.5	0.6	0.6	0.5	0.6

Fig. 4-20. The relationship value chart for all of the parts

In Fig. 4-20, the relationship values for all of the parts are listed in the chart. Users can check which part has the closest connection between parts.

Design of the system interface of the Digital DFHY is a crucial issue in this study. Fig. 4-18 shows the shapes of parts for Daoism researchers to select from for getting the results of relationships between each pair of parts. In the original DFHY, there are no names for separate parts and in the catalog for the shapes of parts, there are no descriptions for the shape of parts, so the author arranged the shapes for each part as selection options in the digital DFHY system. Users can select a *Fu* by its position page number, name of the *Fu* and names of the parts. All of the *Fus* related to this *Fu* are shown on the right in the figure. All of the parts that these related *Fus* contain are also shown. Daoism researchers can use these analysis results as hints to help them in their studies.

5 Evaluation of the DIGITAL DAO-FA

HUI-YUAN

So far, the contents of DFHY have been digitized. Pages, *Fus* and the parts contained in DFHY are all digitized and stored as image data. The analysis works on the contents are carried out using the data. Evaluation based on the lessons is an important measure of the success of the Digital DFHY – how the researchers apply the analysis results into their studies.

The evaluation discussion below is divided into two parts – evaluation of the digitized contents of DFHY and then of the analysis results.

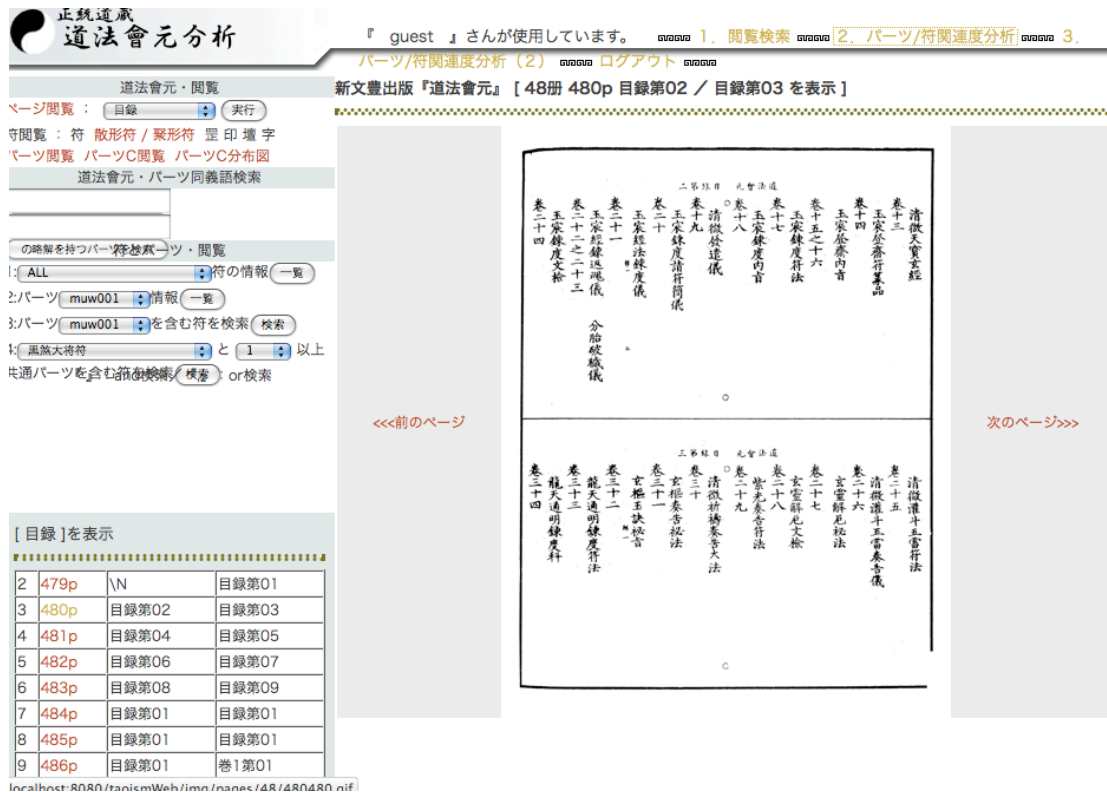
5.1 Evaluation for the content of Digital DFHY

The processing of digitization of contents was done as discussed in Chapter 3. For humanities scholars, searching and viewing of the contents is not merely to display page images on a computer but to use them in their research activities. The primary propose of Digital DFHY is to provide augmented search functions using the digital archive. Organizing parts, *Fus* and volumes and making them searchable was carried out in collaboration between humanities scholars and information scientists. In this process, the author proposed to create catalogs indexed by the shapes of parts to connect all of the *Fus* and other components with each other.

The illustrations of the *Fu* texts are primary sources for the studies of Daoism. Matsumoto and other Daoism researchers have studied their meaning. Fig. 5-1 shows a search and browsing interface for the parts and *Fus* based on the findings from DFHY. In this system, users can search and view the parts by their shapes and by meanings, and they can look up specific *Fus*.

It is important for the humanities scholars to be able to search and browse those contents related to an instance displayed on the screen or to data found

in the search-and-browse process in order for them to efficiently use the results in their scholarly activities.



正統道藏 道法會元分析

『 guest 』さんが使用しています。 1. 閲覧検索 2. パーツ/符関連度分析 3. パーツ/符関連度分析 (2) ログアウト

新文豊出版『道法會元』 [48冊 480p 目録第02 / 目録第03 を表示]

道法會元・閲覧

ページ閲覧 : 目録 実行

守閲覧 : 符 散形符 / 聚形符 聖印 壇字

パーツ閲覧 パーツC閲覧 パーツC分布図

道法會元・パーツ同義語検索

の略解を持つパーツ符を閲覧

1: ALL 符の情報 一覧

2: パーツ muw001 情報 一覧

3: パーツ muw001 を含む符を検索 検索

4: 黒旗大将符 と 1 以上

共通パーツを含む符を検索 検索 or 検索

<<<前のページ

次のページ>>>

[目録] を表示			
2	479p	\N	目録第01
3	480p	目録第02	目録第03
4	481p	目録第04	目録第05
5	482p	目録第06	目録第07
6	483p	目録第08	目録第09
7	484p	目録第01	目録第01
8	485p	目録第01	目録第01
9	486p	目録第01	目録第01

localhost:8080/tanemWeb/ima/nanac/48/480480.nif

Fig. 5-1. Interface for Daoism researchers to search and browse related *Fus* having similar meanings and parts in the system

The user interface of the Digital DFHY is designed primarily for Daoism researchers. For example, researchers can login in using their user name and add nodes to the contents in the Digital DFHY. Humanities researchers are debating the use of computers in their research, as more and more digital archives aimed to support humanities studies are available. User-interfaces friendly to them are crucial to promote the use of digital resources in humanities research.

5.2 Evaluation of the statistical analysis results of the DFHY Contents

The author initially proposed to create the system for the Digital DFHY to support the studies of humanities researchers, and then processed the DFHY data and built the database. The system for displaying the original contents is ready for searching and viewing DFHY contents, however, the use of the Digital DFHY is not limited to search and browse. As introduced in Chapter 1, to read and understand the meaning of documents is important for Daoism research. Humanities scholars have found it useful to use computers to do statistical analysis to get objective data. The experimental studies on the frequency and co-existence of words and the relationship analysis were done from this viewpoint. The author has consulted Li Yuan-Guo, a professor at Sichuan University, China and other Chinese Daoism researchers for the evaluation of the Digital DFHY. They considered the Digital DFHY is a new method for studying the content of Daoism documents and that the research analysis will have to be revised and include more effects besides the same shapes of parts.

Koichi Matsumoto, a professor at University of Tsukuba [36] helped the author present the analysis results at the 62nd Japan Daoism Research Conference [37]. Many Daoism researchers talked about the new type of digital archive for Daoism documents and Matsumoto has reported and evaluated the analysis results. The author summarized the evaluations and comments as follows.

1) Distribution of the parts contained in *Fus* and the shared parts in *Fus*:

How many times every part appears in a volume is counted and sorted in one table. Using this table, the most frequently used parts and the volumes containing the maximum numbers of parts can be checked. Daoism researchers assumed some parts frequently used were modal parts and some of them determined meaning in Daoism studies. By this analysis, some volumes containing different contents for Daoism schools are found connected with each other even though they are not closely located in the original contents.

Table 5-1. Visualization of the distribution of parts

vol_no 巻 NO、整数	fu_count 巻に記載され ている符の数	count_of_fu_owning_parts 散形符を 持つ符の数	parts_count 巻に含まれているパーツ の数（重複パーツ含む）
59	1	1	2
107	1	1	6
108	24	1	6
138	10	1	7
176	1	1	10
182	24	1	11
191	28	1	4
198	3	1	17
201	54	1	11
213	3	1	10
215	2	1	7
217	10	1	13
264	1	1	12
266	2	1	15

Moreover, co-existing parts may show the rules of parts in DFHY. Learning what parts always exist with each other can provide evidence to say that a pair of parts belonging to the same school or the same Daoism cursing system.

Daoism researchers identified the parts that have been confirmed as belonging to a cursing system with shared parts [10] if some of the parts are located in different volumes.

As shown in Fig. 5-2, as a part of our analysis of the shapes of shared parts, the most frequently used parts are contained in the region between Volume 155 and Volume 165. Based on the distribution, Daoism researchers may deduce that some parts in this region should be study subjects and that the reason of the unusual location should be studied and verified.

It is important to be able to guide the humanities researchers to recognize these problems and provide support even though the Digital DFHY is not able to analyze these reasons in detail by itself.

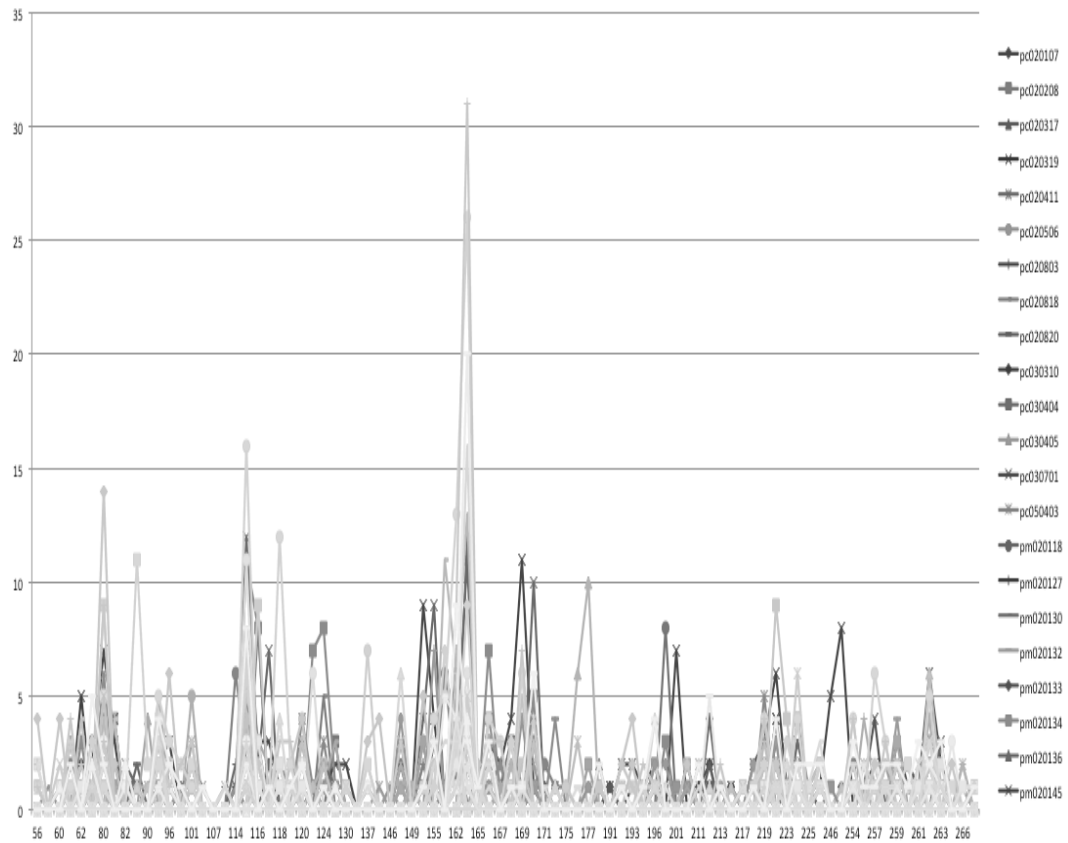


Fig. 5-2. Distribution map for parts

2) The sorting of the relationship between Fus and the distribution of the relationship between Fus .

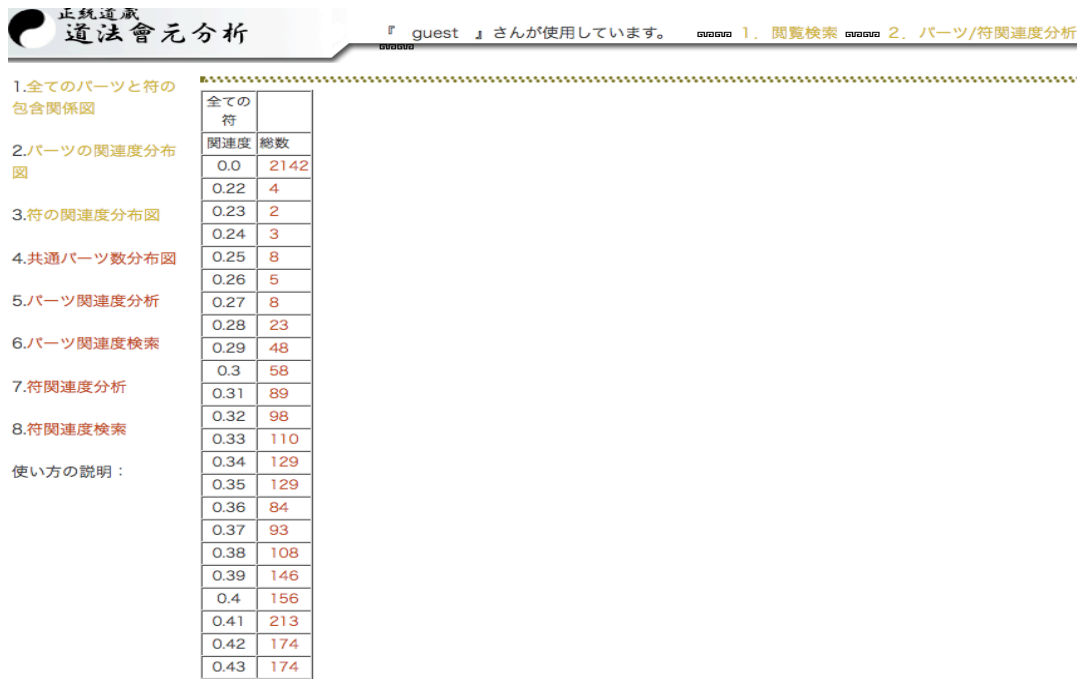


Fig. 5-3. Relationships between *Fus* and the sorting of the relationships

The relationship between *Fus* is key to this study. By mapping the relationship between *Fus*, all *Fus* can be connected and a network of *Fus* can be created. As we described before, the mapping of the relationships between each pair of *Fus* are shown in Digital DFHY. In the network, some *Fus* and their constituent parts have close relationships, and some *Fus* have less shared parts and are shown further apart from the center of the mapping. The distribution of the relationships between *Fus* provides hints to humanities researchers. Combined with the functions, uses, meanings and other information of *Fus*, finding all *Fus* having close relationships can help study of the *Fus* in the network. To study the cursing, the relationships between *Fus* are analyzed from the number of shared parts and the pages they are in. The relationships between *Fus* can be used in several ways for Daoism researchers. The closest *Fu* can be used for verification of the Daoism research. On the other hand, the *Fus* having no connection with each other can prove some hypothesis about the relationships between these *Fus*. Also the relationships between *Fus* can help Daoism researchers find the

Daoism allegiance of *Fus*, and the shapes of shared *Fus* are created in the journal for the Daoism studies [28].

The ranking of related *Fus* sorted by relationship could help researchers study the relationships between *Fus* in detail. In Fig. 5-3, some *Fus* connected with the Kui-Huo *Fu* are listed. It indicates that the two *Fus* - Kui-Huo *Fu* in Volume 268 and Yuan-Shuai Da Huo-Ling *Fu* in Volume 162 have the closest relationship in our analysis results although the distance between the volumes is large. It was found that the two *Fus* belong to the same cursing system [35].

By using the average relationship values between *Fus*, the value for the relationship between volumes can be found. Because some volumes contain no *Fus*, the research analysis results are limited to volumes that contain *Fus* as shown in Table 5-2.

Table 5-2. The relationship value for inter-Volumes

Volume1	Volume2	Value of Relationship between Volumes (E+13 is omitted)
118	215	3.87
118	260	2.98
118	247	2.25
118	262	1.70
215	258	1.70
163	215	1.66
215	224	1.34
258	260	1.31
163	260	1.28
118	168	1.27
115	215	1.17
118	257	1.08
169	215	1.07

Daoism researchers have studied the results and listed the different possible uses of the cursing system in DFHY, for each volume presented. And by using the results, Daoism researchers have found that there is a close relationship between volumes 260 and 262, 256 and 259, 118 and 262, 155 and 227. However, some parts, *Fus* and volumes are not included in this analysis. It is crucial to augment the Digital DFHY from the viewpoint of information science and also from the viewpoint of more accurate and rich data to support Daoism studies.

6 Conclusions and Discussion

6.1 About the Digital Archive for Dao-Fa Hui-Yuan

The original purpose of building the Digital DFHY was to make a tool to support Daoism researchers. First, a digital archive is crucial to provide access to the contents in it. For building the Digital DFHY, the author analyzed *Fus* as the primary purpose of this study. A *Fu* is expressed as a complex graphical entity composed of more than one constituent part. There are still many facts and rules in DFHY unknown among Daoism researchers. To support the researchers in uncovering these unknown facts and rules, the author proposed to build an environment to help the Daoism research community analyze DFHY using its digitized contents and data derived from the contents, for example, analysis of the structure of *Fus* and relationships among *Fus*. Up till now, in the Digital DFHY, all the *Fus* containing San-Xing *Fu*(散形符) defined in Chapter 3 have been digitized. The graphic instances called Tan(壇) and Gang(罡) are not included in the database and are not used for the analysis of DFHY in this study. Archiving and analysis of these instances are left for future study.

Wider and more advanced use of digital technology is an important topic for humanities scholars. In this study, the author has focused on the digital archive for humanities scholars, in particular Daoism researchers. Digital archives should not be defined merely as a set of digitized data but as a well-organized set of data that enhances the intellectual activities of the users. The Digital DFHY presented in this paper is built as an experimental system for the study of the effective use of digitized Daoism contents for research by humanities scholars. The Digital DFHY has proved its usefulness for the humanities scholars. On the other hand, the author has learned many lessons from the development. She hopes to improve the methods for building digital archives and analysis of them for supporting professional users.

It is important for any digital archive to meet the users' needs. In other words, digital archives in the humanities must be built not only on the

information technology but also on the subject knowledge. The digital archive is based on the digitized contents and can be used to supply more analysis results. However, it should be pointed out that the analysis results are neither equivalent to the results done by humanities scholars who study the contents in this digital archive, nor just statistical data created or arranged by computers.

Resources stored in a digital archive should be organized to help humanities researchers. The author hopes the Digital DFHY will help researchers make new discoveries. Furthermore, the analysis results can be used to improve the user interface of the digital archive, arrange the original contents in a more logical order and even to find more hidden relationships with other digital archives. Indeed, the author started the study on DFHY after studying its various texts, pages and graphics. The Digital DFHY was created to supply users with both digitization of data for the original content and some analysis results of the DFHY.

The digital archive provides users with convenient tools not only to find and browse the contents but also to help analyze them. The definition of digital archives can be different according to the individual researcher. The definition is generally established as “A specific storage system, whether a tier of storage, a cloud-based storage repository, or a specialized storage device”[38].

- *A service providing extensions to an organization's storage for the purposes of offloading whether onsite, off-site, or at end-of-life as storage expansion.*
- *Any device or practice vendors or IT calls "archiving", "database archive", or "archival storage" - from a "digital archive systems" perspective, these are not "Level 1" systems or services by themselves* [38].

This means that a digital archive is not created as a digitized collection of resources, but as an integration of the digitized resources and services which add more value to the contents.

Digital archives have been well developed and there are many existing archives. As discussed in Chapter 2, the *Europeana* Project [20] and other

projects provide many types of users with information arranged in accordance with user types – amateurs and scholars [41]. The author has discussed in Chapter 2 that the compilation of digital archives is different from Digital Humanities. Digital Humanities was born several years ago. Mapping for the historic events [39], visualization of relationships among historical persons from historical documents are all included as events for Digital Humanities. Other theory and technology about Digital Humanities is also developed. However, still some new arguments about the definitions of Digital Humanities are emerging [40]. The definition of Digital Humanities seems not fixed well but it is obvious that a digital archive is the key component of infrastructure to support Digital Humanities.

In this study, the author tried to compile the digital archive of DFHY from research on the content. After digitization of the original contents of DFHY, the author focused on analysis of the relationships between *Fus* and volumes. A network of *Fus* was built, where all of the *Fus* are connected by shared-parts. The author created a network of volumes, which is used to calculate relationship values between volumes. This study of the relationships among the volumes helps uncover the secrets of *Fus* and volumes. *Fus*, volumes and their relationships are difficult and complex subjects even for Daoism researchers because of the large number of *Fus*, and that the classic texts are difficult to understand [10]. The relationships between parts, *Fus*, and volumes are presented in various formats in this study. By using the distribution of the relationships of parts, usage of the parts in the volumes is investigated in this study. The co-location of parts shown in our research suggests the possibility of the co-location of parts in other places. The relationship between *Fus* can also help researchers discover semantic connections even where this connection is not written in DFHY explicitly. In addition, the relationships can be used to verify new hypotheses, which have been proved in some cases.

The relationship analysis presented here has revealed a few new facts about the DFHY, which has shown the value of the Digital DFHY and its potential to help Daoism researchers. Moreover, this result encourages us to enhance the Digital DFHY. The database of *Fus* does not cover all of the *Fus*

yet, so the first priority in the future work is to extend this coverage. To help humanities research of DFHY, it is crucial to refine the functions to determine the relationship values between volumes. Nevertheless, this study has clarified the advantages of the Digital DFHY to help research in Daoism and more generally in the humanities. The Digital DFHY has shown that we can add extra value to a digital archive of historical documents and humanity resources instead of just having documents in a digital format. The author hopes to deepen this research and the methods of analysis of the graphics in classic documents to assist humanities scholars more effectively and efficiently. In any attempt to build a new digital archive for humanities resources, investigating what the users of these resources need is necessary. The author tried to build a new analysis method for images contained in Daoism resources and also to develop a new method for building digital archives to support Daoism research.

6.2 The Model to Build a Digital Archive for Classic Documents

In chapter 3, the framework of the Digital DFHY is discussed. It shows two crucial points for the digital DFHY - creation of the *Fu* database and metadata design.

The *Fu* database consists of images and texts of the *Fus* and constituent parts. The database is valuable not only to support Daoism researchers but also to preserve the original contents of DFHY. Images of pages, *Fus* and other components are saved as JPEG files in the database. Titles and scripts attached to the *Fus*, and other texts included in the pages, are extracted from DFHY and stored in the database. This set of data is used to search *Fus* and the constituent parts. The database organization is based on experiences in the preceding studies, for example, the classification and catalog of the constituent parts by Yasoda [27], which was examined by Daoism researchers.

The design of the metadata for all resources is a key in the development of digital archives. DFHY contains many different types of graphics and texts written in ancient Chinese. The metadata schemas are determined by the structure of the tables in the database. Metadata design in this study was done in collaboration with Daoism researchers because the subject knowledge in Daoism was essential.

The analysis of the original contents of DFHY is an important part of the research to organize the context and supply these findings to the humanities researchers. Distributions of the *Fus* were used as an important analysis of the features of DFHY.


Based on the experience of building the digital DFHY, the author learned that to create a digital archive of classic documents for humanities research, the first step is to study the characteristics of the documents, such as how the texts and the graphics are arranged, what humanities scholars are focusing on and how the digital archive can support the humanities studies. Digital archives should be able to provide users with the contents as close as possible to the original resource but within the limitation of the cost and

restrictions in accordance with user requirements. The following issues need to be considered: How to supply the functionality for searching and viewing the contents, how to digitize the contents into the database and how to set up the interface of the digital archives.

6.3 Conclusion and Future work

This paper presents the Digital DFHY, a digital archive created from Dao-Fa Hui-Yuan (DFHY) – a very important resource for Daoism researchers. The goal of this study is to build the Digital DFHY and experimentally apply it to the content analysis of DFHY.

The relationship analysis presented here has revealed a few new facts about the DFHY. The database of *Fus* does not yet cover all of the *Fus*, so the first priority for our future work is to extend the coverage.

As shown in Fig. 6-1, the texts following the parts are not arranged by the natural order for parts contained in the San-Xing *Fu*. The part  is located on the top of the *Fu*, but the explanation of this part does not appear at the first position in the San-Xing *Fu*. For this reason, to define the location for each part was done manually using the natural reading order.

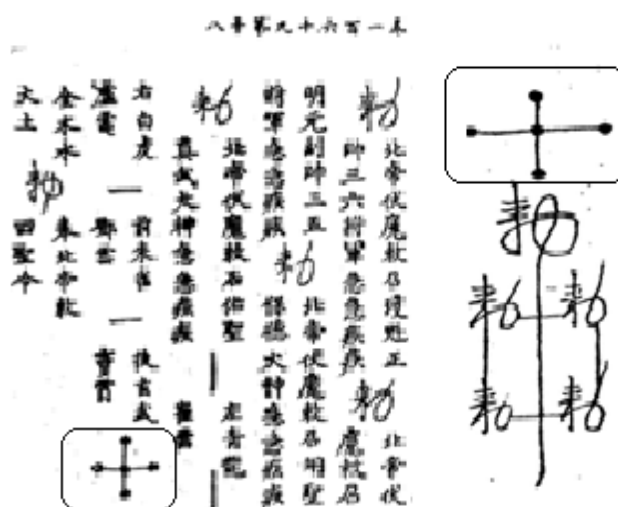


Fig. 6-1. Example for the gap between the location by the natural reading order and the order of instruction text in San-Xing *Fu*

In the catalog of parts shapes, it may also be possible to divide Ju-Xing *Fu* into parts. The data for all the *Fus* contained in DFHY can be collected to make the analysis of *Fus* more comprehensive. Moreover, we propose to

improve the calculation of the relationships between parts and *Fus* as future work.

This study has clarified the advantages of the Digital DFHY to help research in Daoism and more generally in the humanities. It showed the values added to the collection of the humanities resources. Furthermore, the research is proposed to be the methods of analysis of the classic documents. In the attempt to build a new digital archive for humanities resources, we need to investigate what users of the resources need.

It is necessary to point out that the digital archive should be designed using good archival practices. It is our goal to save all the contents in the original documents and other related information such as the versions, history, creator and so on. Furthermore, analysis of the contents is different from the study of the contents, which is the work of humanities researchers. To conclude, the experiences of building the digital archive and achieving the analysis results it are proposed as a model for creating digital archives for classic documents.

The author tried to produce an interface for digital archives and publish the analysis results in a suitable format. To combine humanities research with digital technology has been a popular topic since the 1990s. Digital Humanities as a scholarly area was born from the cross-disciplinary community based on the humanities and information science [42]. Yet, there are many problems that should be considered. Besides the gap between humanities research and information technology research, the differences between the study methods in those areas are also challenge for us.

The digital humanities resources need not be limited to digitized content but any resources useful to support humanities researchers. The content analysis shown in this paper is a good example of the support functions attached to the Digital DFHY. Because classic cultural resources are complex and we need to have background knowledge to understand the resources, it is necessary for us to study the characteristics of the resources and use them to organize digital archives. How to create digital archives for users and the contents analysis methods is also crucial.

In this study, the collaborating Daoism researchers have helped evaluation of the analysis results. However, some questions about the content of DFHY still remain. The quality of the analysis has to be improved. And, at the same time, the evaluation method has to be investigated because the goal of the Digital DFHY is to help Daoism researchers discover new facts.

The author thinks the reasons for this result are the following:

1. Ju-Xing *Fu* is still not covered comprehensively in the analysis of this study. It is necessary to enrich the data of the Digital DFHY and cover all of the *Fus* to get accurate analysis results.

2. The relationship value is calculated by using the number of shared parts. However, it is left for future study to include the position of the shared parts, the number of total parts and other elements in the calculation of the relationship values.

In order to avoid any bias for the analysis data in this study, the author analyzed the data of all the *Fus* contained in DFHY. The relationships between volumes revealed some hidden connections between volumes, not written in the original texts.

The analysis data supports Daoism researchers in making new discoveries using the data about the relationships between parts, *Fus*, and volumes. Building the digital archive for DFHY, a collection of Chinese Daoism documents, how to digitize the contents, how to manage the data, how to analyze the contents in DFHY, and then how to supply the analysis results to researchers are the objectives of our research. Finally, the author also tried to sum up the experiences of creating the digital archive for DFHY.

Acknowledgements

I would never imagine without my supervisor Professor Sugimoto and Professor Matsumoto how I accomplish my doctoral program. So I would like to express my deep gratitude to Professor Sugimoto and Professor Matsumoto for their patience and advising on my research. During my doctor program, I experienced not only the guidance from Professor Sugimoto and Professor Matsumoto but also their personalities. After giving birth to my daughter, I attempted to suspend my doctoral program because nearly all my time was spent in taking care of my family. Professor Sugimoto encouraged me and gave me a plenty of time to recover. Slowly but relentlessly make our study better. Professor Matsumoto supported me a plenty of study methods and life lessons.

I would like to express my deep sincere gratitude to Professor Uda, and Professor Hasegawa. At the first time I entered Tsukuba University knowing nothing about Japan, Professor Uda and Professor Hasegawa gave me warm guidance about life and study.

Professor Sakaguchi, Professor Nagamori and Professor Morishima also gave me suggestions and comments about my studies. I wish to express my deep gratitude for many kinds of helping from them.

Jan Askhoj, my friend and research colleague in Sugimoto Lab spent much time to help me check my PhD thesis. I would like to express my gratitude to him and other friends for their kindness.

The JASSO Scholarships supported me during my doctoral program. I would like to be grateful for the financial support.

I would like to show my heartfelt gratefulness to all the people who supported me and devoted to my study.

Finally, I would devote this PhD thesis to my family, especially my husband and my daughter. Even sometimes I have to take care of them, but they supported me with more understanding, patience and encouragements.

References

- [1] LTDP: "What is a Digital Archive?"; Long-Term Digital Retention and Preservation Reference Model site. 2013; pp.26-11.
- [2] Landow George, P: "The Digital Word: Text-Based Computing in the Humanities"; MA.USA: MIT Press Cambridge, 1993.
- [3] Katz Stanley, N: "Why technology matters: the humanities in the twenty-first century"; Interdisciplinary Science Reviews (Institute of Materials, Minerals and Mining); No. 2 (2005): 105; pp30.
- [4] Liu, XiaoMing: "Zhong Guo Fu Zhou Da Guan"; Bai Hua Zhou Wen Yi Press; Nanchang;1995.[in Chinese]
- [5] Nikaido, Yoshihiro: "The Chane of the Marshal Gad in the Daoism and Folk Beliefs"; The Institution of Oriental and Occidental Studies;Kansai University;Osaka ;2006 [in Japanese].
- [6] Liu ,Ye: "Lun Dao: Dao Jiao Ru Men 600 Jiang"; Chinese women books Press, Beijing ,2012 [in Chinese].
- [7] Zhong HeJi: "The book of Balance and Harmony A Taoist Handbook"; Translated by Thomas Cleary; Boston: Shambhala Press; 2003 [in Chinese].
- [8] Ren, Zong Quan. "Dao Jiao Zheng Biao Fu Yin Wen Hua Yan Jiu ". Zong Jiao Wen Hua Press;Beijing;2012 [in Chinese].
- [9] Livia, Kohn: "Cosmos And Community: The Ethical Dimension Of Daoism. Hawaii": Three Pines Press, 2004.
- [10] Jiang, Sheng:" Zong Jiao Yu Ren Lei Zi Wo Kong Zhi: Zhong Guo Dao Jiao Lun Li Yan Jiu"; Bashu Press; Chengdu 1996 [in Chinese].
- [11] Li, Yuan Guo:"Shen Xiao Lei Fa : The History and Thought of Daoism Shen Xiao Lei Fa"; Si Chuan Ren Min Press;Chengdu; 2003 [in Chinese].
- [12] Matsumoto, Koichi:"Dao Fa Hui Yuan- A Great compile of Magic, Reading the Classic Daoism Book".; Taishukan Publishing; Tokyo ,2001 [in Japanese].
- [13] CCSDS: "Reference Model For an Open Archival Information System "; Magenta Book;; Washiton, DC; 2012.

- [14] The Inclusion for SAADA:
<http://www.saadigitalarchive.org/tides/article/20120418-704> (accessed Nov. 2013).
- [15] Nicola Ferro, Gianmaria Silvello. "NESTOR : A formal model for digital archives." *Information Processing & Management* (Elsevier) 49, no. 6 (2013): 1206–1240.
- [16] David Koller, Bernard Frischer, and Greg Humphreys: "Research Challenges for Digital Archives of 3D Cultural Heritage Models." *ACM Journal on Computing and Cultural Heritage* 2, no. 3 (2009): pp.7
- [17] Susan, Schreibman, and Siemens: ". A Companion to Digital Humanities". Malden: Black Well Publishing, 2004.
- [18] DCMI. About DublinCore. 2013:Purpose of the digital archive:
<http://dublincore.org/about-us/> (accessed Nov. 2013).
- [19] The library of Congress. "About the American Memory";
<http://memory.loc.gov/ammem/about/index.html> (accessed Nov. 2013).
- [20] Europeana project group: "Europeana Professional. National Library of the Netherlands"; <http://pro.europeana.eu/web/guest/home> (accessed Nov. 2013).
- [21] TELDAP:"The overview of TELDAP".
<http://teldap.tw/en/Overview.html> (accessed Nov. 2013).
- [22] National Diet Library of Japan: "Digital Library From the Meji Era".
http://kindai.ndl.go.jp/?__lang=en (accessed Nov. 2013).
- [23]Toth, Michael B: "The Archimedes Palimpsest";
<http://www.archimedespalimpsest.org/about/> (accessed Nov. 2013).
- [24] Berry, David M. "Understanding Digital Humanities. Hampshire"; Palgrave Macmillan, 2012.
- [25] KeioUniversity: "HUMI Project Humanities Media Interface";1999;
<http://www.humi.keio.ac.jp/en/introduction/index.html> (accessed Nov. 2013).
- [26] L.Borgman, Chistine: "Introduction of the host of the HUMI project". Los Angels: University of California, 2010.
<http://www.humi.keio.ac.jp/en/introduction/index.html> (accessed Nov. 2013).

- [27] Yasoda, Y: "The Digitizing of History Documents for Supporting the Researching"; Graduate School of Information Science, Tsukuba University, 2001[in Japanese].
- [28] Hayagawa, M, and Matsumoto,K and Uda,N: "Classifying terms used in Taoism charms appearing in "Dao-Fa Hui-Yuan""; Tsukuba: Tsukuba University;Digital Libraries study, 2010, pp57-59 [in Japanese].
- [29] Yasoda, Y; and Matsumoto,M and Uda,N: "Construction of Charm Analysis Support System for "Dao-fa Hui-yuan" Research."; Joho Shori Gakkai Shinpojiumu Ronbunshu;2004; pp143-150 [in Japanese].
- [30] Scott, John P: "Social Network Analysis: A Handbook".: SAGE Publications; London 2000.
- [31] Jun Karamon, Yutaka Matsuo, and Mitsuru Ishizuka:"Generation Social Network Features for Link-based Classification."; Information Processing Society of Japan; Tokyo: 2008; pp.45.
- [32] Dijkstra, E.W. "A node on two problems in connexion with graphs." ; Numerische Mathematik ; 1 (1959); pp269 ~ 271.
- [33] Norio, Togiya and Akira, Baba: "The Digitization of Cultural Resources: An Exploration of Current Issues and the Future Outlook -A Comparison of Japan and Other Countries-." World Library and Information Congress; 75 th IFLA General Conference and Council; Milan: IFLA; 2009; pp1 [in Japanese].
- [34] Feng XiaoXiao, Matsumoto Koichi, and Sugimoto Shigeo. "Digital Archive "Dao Fa Hui Yuan" for Daoism Research." ICADL 2011 Proceeding of the 13th international conference on Asia-pacific digital libraries. Beijing: Springer, 2011. pp26-34.
- [35] Feng XiaoXiao, Matsumoto Koichi, and Sugimoto Shigeo: "Creating the Database and analysis the contents using statistical methods of Dao-Fa Hui- Yuan ."; Nippon Dokyo Gakkai: The journal of Eastern Religions, 2012: pp.119.[in Japanese]
- [36] Feng XiaoXiao, Matsumoto Koichi, and Sugimoto Shigeo: "Uncovering the Secrets of Daoism Fus using Dao Fa HuiYuan."; ICADL 2012 Proceeding of the 14th international conference on Asia-pacific digital libraries; Taipei: Springer, 2012.

- [37] The program for the 62th Japanese Daoism Conference :
<http://www.taoistic-research.jp/event/062> ; Dec. 2012
- [38] Peterson, Michael. LTDP: long term preservation Reference Model. 2013
<http://www.ltdprm.org/reference-model/what-is-an-archive> (accessed Nov. 2013).
- [39] Baba Yuko, Nagamori Mitsuharu, and Sugimoto Shigeo. “V-HAWKS: Development of a Historical Information Browser Based on RDF/OWL.”;JPSJ ;SIG-CH-73 (2008): 49-56 [in Japanese].
- [40] Fuminori Kimura, Takahiko Osaki, Taro Tezuka, and Akira Maeda: “Visualization of relationships among historical persons from Japanese historical documents.” (Literary and Linguistic Computing) 28, no.2 (2013) pp.271-278 [in Japanese].
- [41] Erway, Ricky L: “Digital Initiatives of the Research Libraries Group.” D-Lib Magazine; 996; pp 6.
- [42] Berry, David M.: “The Computational Turan: Thinking About The Digital Humanities .” CULTURE MACHINE , 12 2011: pp1-22.

List of Publications

1. Xiao-Xiao Feng, Shigeo Sugimoto, and Koichi Matsumoto. Digital Dao-Fa Hui-Yuan: Developing a Digital Archive of Daoism Documents, *Journal of Library and Information Science*, 38(2); Oct, 2012; pp.24-38.
2. 馮曉曉, 松本浩一, 杉本重雄. 『道法會元』データベースの構築とその計量的分析, *東方宗教*, 第 120 號, 2012, pp. 63-80.
3. XiaoXiao Feng, Koichi Matsumoto, and Shigeo Sugimoto. Digital Archive “Dao Fa Hui Yuan” for Daoism Research, *Proceedings of 13th International Conference on Asia-Pacific Digital Libraries, ICADL 2011 Beijing, China*, October 24-27, 2011, pp. 26-34.
4. XiaoXiao Feng, Koichi Matsumoto, and Shigeo Sugimoto: Uncovering the Secrets of Daoism Fus Using Digital Dao-Fa Hui-Yuan, *Proceedings of 14th International Conference on Asia-Pacific Digital Libraries, ICADL 2012, Taipei, Taiwan*, Nov. 12-15, 2012, pp.1-10.
5. 馮曉曉, 松本浩一, 杉本重雄. 符の電子化構造分析: 「道法會元」における護符を対象として, *情報処理学会研究報告*, vol. 2008-CH-79, 2008; pp. 7-16.
6. 馮曉曉, 松本浩一, 杉本重雄. 道法會元における符関連度の分析, *情報処理学会研究報告*, Vol. 2010-CH-88, 2010, pp. 1-3.

Appendices

Appendix A

Relationship between *Fus*

In the Table of Relationship between *Fus*, the *Fu* ID of a pair of *Fus* are expressed in each column as *Fu1* and *Fu2*. The number of parts contained in each *Fus* is also shown. The number of shared parts shows how many parts in the pair of *Fus* have the same shape. The Values for the relationship between *Fus* can help Daoism researchers to study the hidden relationship of *Fus*.

<i>Fu1</i>	<i>Fu2</i>	Number of Parts contained in <i>Fu1</i>	Number of Parts contained in <i>Fu2</i>	Number of Shared parts	Values for the Relationship between <i>Fus</i> (E+14 is omitted)
1180702	1681201	5	3	0	5.71
1621602	1681201	5	3	0	5.71
1632102	1681201	11	3	0	5.19
1681201	2242301	3	6	0	4.76
1681201	1692002	3	6	0	4.76

1180701	1681201	6	3	0	4.76
1180802	1681201	4	3	0	4.76
1180702	2182203	5	4	0	4.28
1621602	2182203	5	4	0	4.28
1632102	2182203	11	4	0	3.89
1631701	1681201	5	3	0	3.81
1180701	2182203	6	4	0	3.57
1180802	2182203	4	4	0	3.57
1692002	2182203	6	4	0	3.57
2182203	2242301	4	6	0	3.57
1180801	1681201	3	3	0	3.17
1180602	1681201	3	3	0	3.17
1632701	1681201	6	3	0	3.17
1681201	2580402	3	9	0	3.17
1681201	2621502	3	6	0	3.17

Appendix B

Relationships between Parts

The values of relationship between parts contained in *Fus* having different shapes are shown in this table. *Fu*'s number on the shortest path stands for how the *Fus* containing the pair of parts are connected with each other. The number of the shortest path expresses how many times *Fus* contained in this pair of parts are connected.

Parts Code	Parts Code	<i>Fu</i> 's number on the shortest path	Number of the shortest path	Value of the Relationship between parts (E+12 is omitted)
sf030101	sc042602	2	200	2860
sf020402	sf020401	2	181	89.5
sc100801	sf030103	2	165	4.84
sf030101	sf020201	2	165	4.84
sf020202	sf030103	2	163	3.36
sf010101	sc100801	2	161	2.33
sf020202	sf020201	2	152	0.45
sf030101	sc042601	2	152	0.0045
sf030101	sf012905	2	140	0.00051
sf030104	sf013104	2	136	0.00025
sf030101	sf014101	2	134	0.00017
sf030104	pm020155	3	136	0.00016
sf030104	pm020167	3	136	0.00016
sf030104	pm020168	3	136	0.00016
sf030104	pm020169	3	136	0.00016
sf030104	pm020170	3	136	0.00016
sf030104	pm020171	3	136	0.00016
sf030104	pm020172	3	136	0.00016
sf030104	pm020173	3	136	0.00016
sf030104	pm020174	3	136	0.00016

Appendix C

Relationship of inter-volumes

The inter-volume relationships are calculated by the relationships between *Fus* in volumes. In Dao Fa Hui Yuan, some volumes contained no *Fus*, so the relationships between volumes are limited to each pair of volumes having the *Fus*.

Volume1	Volume2	Value of Relationship between Volume (E+13 is omitted)
118	215	3.87
118	260	2.98
118	247	2.25
118	262	1.70
215	258	1.70
163	215	1.66
215	224	1.34
258	260	1.31
163	260	1.28
118	168	1.27
115	215	1.17
118	257	1.08
169	215	1.07
118	170	1.07
118	259	1.03
224	260	1.03
247	258	0.99
163	247	0.97
115	260	0.90
215	257	0.87