E-Processing of Historical Manuscripts Collections:

A Project of the French Research Organization CNRS

Nikolaos E. Myridis

Assistant Professor, Aristotle University of Thessaloniki

University Campus, GR-54124, Thessaloniki, Greece

Tel: 30-2310-996970   E-mail: nmyridis@theo.auth.gr

Abstract

We focus in this paper on a CNRS (Centre National de la Recherche Scientifique, France) project that involves a collection of texts more than 60 years old. This project is the Sources Chrétiennes collection, which assembles in 520 volumes various series of most valuable texts: early Christian, Greek Byzantine, medieval Latin, etc. The content of the texts is theological/historical or even philosophical. The collection is the only bilingual one among the various similar collections. That is, each manuscript is translated in French, no matter what the original language was. We can state, without reservation, that this project is unique —not only amongst French studies— but from a global perspective. Among the various innovative results of our work, we should note in particular the structural analysis of the collection, the hierarchical permutation of the collection's content, the reordered representation of related information, etc. The information contained in this series is thus displayed in a way that demonstrates its great worth and impact on French and European/international studies.

Keywords: CNRS, Sources Chrétiennes, E-keyword, Databases

1. Introduction

Sources Chrétiennes is a multilingual french editorial project under the auspices of CNRS (Centre National de la Recherche Scientifique) (Note 1). It consists of more than 500 volumes. The Sources Chrétiennes collection is one of the most important and respected worldwide collections regarding the critical edition of manuscripts of theological sources. Consequently it incorporates text-milestones of the Christian literature as well as of the human spirit in *French* (Mondesert, 1988). That is, all projects are always translated in French. The basic languages of prototype manuscripts are Greek, Latin, Armenian, Syrian etc.

The collection ranks among other collections (such as the Patrology of Migne (Migne, 1857-1866), *Philokalia* (Kalliakmanis, 2009) of St. Nikodimos from the Holy Mountain). It is being gradually completed and its development so far covered 65-years running (Table 1-a)

The rate of progress of the series is shown in Fig.1. This is the picture of the pace of the work per year. Each period has its indicative average progress. The overall average is about 10 (volumes / year) [straight line (Fig.1)].

2. Manuscripts and feature elements

Prior to proceed to enumerate particular collection's features, it is necessary to set out the key elements of the manuscripts in the Sources Chretiennes anthology. These are also the collection's coordinates, namely:

Author name - Text Title - Volume Number

which are the fundamental attributes of a database too.

We now focus on texts of the Sources Chrétiennes collection written in Greek language.

It is therefore appropriate to develop a kind of statistical analysis which reveals the inherent structure of the entire collection. Initially we should note that there are five fundamental areas of interest (*The Sources Chrétiennes Collection,* 1943-). The key feature of each region is briefly identified respectively as:

1. [Hebraic] This family of texts pertains to the early period of Christianity. [10 volumes, 2%]

2. [Greek] The common characteristic of these texts is the Greek language. [242 volumes, 48.4%]

3. [Latin] Manuscripts written in Latin language. [165 volumes, 33%]

4. [Medieval] Texts following the Schism and written during the Medieval period (roughly up to fourteenth-century). [70 volumes, 14%]

5. [Eastern] Manuscripts written by eastern authors in Armenian, Syrian etc. language (e.g. works of St.Efraim the Syrian). [13 volumes, 2.6%]

Analysis of the statistical distribution of the collection's volumes among the aforementioned periods is indicated in the Tables 1-a and 1-b, regarding the first 500 volumes.

3. Topology of the project

We now present the modular backbone of the collection which is equivalent to the citation of the concise project's components. We characterize as *concise* those elements which describe the complete work in the shape of anchors or 'topographical signs'. Such elements (by volume) are given in column 'Field' in Table 2.

Thus the fundamental elements of the project may construct a DataBase (DB), whose distinct fields are these elements. A scheme of this DB (Myridis, 2009) can be defined.

The DB is formed by implementations of the Cartesian product of its fields. Thus the nine remaining fields (omitting the date of writing) form the graph (Gross & Yellen, 2004) shown in Fig. 2(a) (cubic graph in Fig.2(b)).

Implementations of graph's subsets in Fig.2(b) construct the DB of Sources Chrétiennes. The generalized ranges of values are cited in Table 2, where L stands for the length of strings (names, titles etc.). In the fields '*author'*, *'area'* and *'project category'* an indirect redirection (*lookup table*) between numbers and strings shall be performed. The field value of 'project category' is infinite (∞) as we assume an infinite number of possible thematic subjects.

We thoroughly identify the values of fields in Table 2. We also incorporate a second column in Table 2, wherein the specific fields' values are depicted, regarding Sources Chrétiennes collection.

The '*project category*' value for the Sources Chrétiennes case is 65, equal to the number of distinct thematic categories in the collection (Table 4).

4. E-Keyword

The work performed during this study leads to the construction of an electronic 'keyword' (eKeyword) for the Sources Chrétiennes collection (Note 2). An excerpt from the printed version of this e-Keyword is shown in Table A [Appendix] (in alphabetical authors ranking). The formation of such a (digital) table gives the best resource for the undertaking of contemporary research relating to the statistical and structural organization of the series. This is particularly important. Indeed the content and functionality of the collection can be better evaluated. Moreover the complete operation of this multi-annual critical editorial work may be understood in depth. We indicatively present some of the main results which are the effects of hierarchical classification in the e-Keyword.

It should be noticed that, both the analyses referred to hereinafter, as well as those not listed (e.g. for reasons of space), are significantly accelerated and particularly facilitated, by using the e-Keyword and, generally speaking, by using Information Technology (IT) resources.

4.1 Greek authors in Sources Chrétiennes

The table of the published volumes until now (1-519) indicates that there are 61 authors of Greek texts in this collection. We now cite the final list of authors in Greek Literature (Table 3) (Note 3).

4.2 Analysis of Subjects

It is certainly difficult to define precisely all the various subjects dealt with by the Greek library of *Sources Chrétiennes*. However, in a general effort to analyse and determine the subject matter of Greek Literature we find that there are, at least, three hundred (300) different subjects in the Greek texts of this collection.

The reader could find more on e-processing of texts collections in Bibliography (Myridis, 2006).

5. Subjects & Conclusion

We have previously reported about 300 subjects in Greek texts in the *Sources Chretiennes* collection. Given the excellent facility provided by the e-Keyword, we construct a general thematic classification of subjects with which the Greek manuscripts of *Sources Chretiennes* deals. This classification is given in Table 4. We observe that about 65 different topics can be identified. Further analysis and consequent formulation of new categories certainly results in considering a larger table of thematic units.

References

[Online] Available: www.sources-chretiennes.mom.fr (October 18, 2010).

Gross J. & Yellen J. (2004). *Handbook of Graph Theory.* BocaRaton: CRC Press.

Kalliakmanis V. (2009). *From fear to love. Study in the texts of Filokalia.* Thessaloniki: Pournaras Pbls. [in Greek].

Migne J.-P. (1857-1866). *The Patrologiae Cursus Completus, Series Graeca.* Paris (161 volumes).

Mondesert Cl. (1988). *Pour lire les Peres de l' Eglise dans la collection Sources Chrétiennes.* Paris: Foi Vivante-Les Editions du Cerf. [in French].

Myridis N.E. (2006). *The Information of Observation: From Analog to Digital.* Thessaloniki: Kyriakides Pbls [in Greek].

Myridis N.E. (2009). *Processing Modules of Digital Information*, (3rd ed.) p.90. Thessaloniki: Kyriakides Pbls [in Greek].

The Sources Chrétiennes Collection. (1943-). Lyon: Les Editions du Cerf.

Notes

Note 1. The government-funded research organization of France

Note 2. For a printed keyword to the Patrologia Graeca collection see F. Cavallera, *Indices: Patrologiae Cursus Completus, Series Graeca.* (Paris: Apud Fratres Garnier Editores, 1912).

Note 3. Details in *The Oxford Dictionary of Christian Church* (1974), London: Oxford Publishing, etc.

Note 4. The symbol > is used in order to declare a very large number, while >> declares a huge number.

APPENDIX

Table A. An excerpt of the printed version of the e-keyword of Sources Chrétiennes

	Author	Title (in Greek)	Volume
1	*Aristeas*	Ἐπιστολή πρός Φιλοκράτιν	89
2	*Aristides of Athens*	Ἀπολογία	470
3	*Athanasius St.*	Ἐπιστολές πρός Σεραπίωνα	15
		Ἀπολογία περί τῶν διαβαλλόντων τήν ἐν τῷ διωγμῷ φυγήν αὐτοῦ	56bis
		Ἀπολογία πρός τόν βασιλέαν Κωνστάντιον	
		Λόγος περί ἐνανθρωπίσεως τοῦ Λόγου	18bis (1st Edt.), 199
		Βίος Ἁγίου Ἀντωνίου	400
4	*Athenagoras the Athenian the Philosopher*	Πρεσβεία περί Χριστιανῶν	3 (1st Edt.), 379
		Περί ἀναστάσεως νεκρῶν	379
5	*Barsanuphius & John*	Ἐπιστολαί & Ἀποκρίσεις I[1]	426, 427
		II	450, 451
		III	468

6	*Basil the Great*	Κατά Εὐνομίου 1	299
		Κατά Εὐνομίου 2-3	305
		Ὁμιλίαι εἰς τήν Ἑξαήμερον, 1-9	26bis
		Περί Βαπτίσματος	357
		Περί τῆς τοῦ ἀνθρώπου κατασκευῆς	160
		Περί τοῦ Ἁγίου Πνεύματος πρός Ἀμφιλόχιον	17
7	*Basil of Seleucia*	Λόγος εἰς τό Ἅγιον Πάσχα	187
8	*Gregory Thaumaturgus*	Εἰς Ὠριγένην Προσφωνητικός	148
9	*Gregory of Nazianzus*	Λόγοι 1-12	247, 309, 405
		20-43	270, 284, 250, 318, 358, 384
		Πρός Κλυδώνιον ἐπιστολαί I-II	208
		Πρός Νεκτάριον ἐπίσκοπον Κωνσταντινουπόλεως	
		Χριστός Πάσχων	149
		Ἐπιστολή πρός Θεόδωρον ἐπίσκοπον Τυάνων	302
10	*Gregory of Nyssa*	Περί τῆς τοῦ ἀνθρώπου κατασκευῆς	6
		Λόγοι κατηχητικοί	453
		[Ἐξήγησις] εἰς τόν Ἐκκλησιαστήν [τοῦ Σολομῶντος]	416
		Ἐπιστολαί	363
		Εἰς τάς ἐπιγραφάς τῶν Ψαλμῶν	466
		Περί παρθενίας	119
		Εἰς τόν βίον τοῦ Μωϋσέως	1bis
		Ἐπιστολή εἰς τόν Βίον τῆς Ὁσίας Μακρίνης	178

Table 1-a. The publication progress of the *Sources Chrétiennes* Collection

Volume number	1	50	100	150	200	300	400	519
Year of publication	1943	1957	1964	1969	1973	1982	1994	2008

Table 1-b. The chronological distribution of the Collection's content

Period	First until third century	Fourth until eight century	Ninth until fourteenth century	Total
Number of volumes	129	288	83	500
%	25,8%	57,6%	16,6%	100%

Table 2. Generalized ranges of values for the fields of *Sources Chrétiennes* database (Note 4)

Field	Field values of arbitrary Collection	Field values of Sources Chrétiennes
Title	L>> *number of (possible) titles on the collection*	N_τ *the number of individual titles in the specific collection*
author	N >> *the overall population of humans in History*	N=60 *the number of greek authors in Sources Chrétiennes*
number of individual texts (per volume)	$M_1 \in [1, N_s]$ *the total number of pages per volume, assuming that each volume can at most include one text per page*	$M_1 \in [1, N_s^{\max}]$ the maximum number of volume pages in the Collection
number of introduction pages	$M_2 \in [1, N_s]$	$M_2 \in [1, N_s]$
index	[0,1]	[0,1]
number of index words	$N_\lambda >$	$N_\lambda^{\max}$ *Sources Chrétiennes* case
date of publication	$N_\chi >$ *the number of human beings in History*	[1943,today] the period of Sources Chrétiennes publications
area	$N_\pi >$ *the dynamic number of geographical regions allover the world (by defining a minimum unit of surface fragmentation)*	N_π^{SC} the set of geographical regions where the scripts of this Collection were written
project category	$N_c = \infty$ *the infinite number of possible project's categories*	$N_c = 65$

Table 3. List of Greek Literature authors in *Sources Chrétiennes*

Authors	
St.Athanasius [Patriarch]	Joannes Moschus [Spiritual]
Athenagoras the Athenian [the Philosopher]	John Chrysostom [Antiochene 4th century]
Aristeas [Historian]	(Pseudo)Chrysostom [Theology]
Aristides of Athens [Apologist]	Callinicus monk [Biographer]
Barsanuphius & John [Spiritual]	Clement of Alexandria [Alexandrian Theologian]
Basil the Great [Cappadocian]	Clement of Rome [Apostolic]
Basil of Seleucia [Antiochene 5th century]	Cosmas Indicopleustes [Geographer]
Gregory Thaumaturgus [Theologian]	Cyril of Alexandria [Theologian]
Gregory of Nazianzus [Theologian]	Cyril of Jerusalem [Speeches]
Gregory of Nyssa [Theologian]	Leontius of Byzantium [Theologian]
Diadochus Bishop of Photike [Spiritual]	(Pseudo) Macarius of Egypt [Speeches]
Didymus the Blind [Alexandrian]	Manuel II Palaeologus [Apologist]
Dionysius the Areopagite [Apologist]	Maximus the Confessor [7th century]
St. Dorotheus of Gaza [Ascetical]	Marcus Eremita [Ascetical]
Irenaeus of Lyon [Contra Heresies]	Methodius of Olympia [Spiritual]
Hermas [Apostolic]	Melito of Sardis [Orthodox Christology]
Hermias Philosophus [Philosophical writer]	Nilus of Ancyra [Ascetical]
Evagrius Ponticus [Cappadocian]	Nicetas Stethatos [Spiritual]
Eudocia Augusta [Rhetoric]	Nicephorus Blemmida [Theologian]
Eunomius [Arian]	Nicolaus Cabasilas [Theologian reformer]
Eusebius of Caesarea [Arian Historian]	Palladius of Elenopolis [Historian]
Hesychius of Jerusalem [Hermeneutic]	St.Pamphilus [Theologian]
Theodotus of Ancyra [Minor Asian 5th century]	St. Polycarp Bishop of Smyrna [Apostolic]
Theodoretus of Cyrrhus [Antiochene 5th century]	St. Romanos 'Melodus' [Poet]
Theophilus of Antioch [Apologist]	St. Simeon the new Theologian [Spiritual]
Ignatius of Antioch [Apostolic]	Simeon Studita [Spiritual]
Justin martyr [Philosopher]	Sozomenus [Historian]
Hippolytus of Rome [Traditional theologian]	Socrates of Constantinople, 'Scholasticus' [Historian]
Isidorus Pelusiota [Spiritual]	Firmus of Caesaria [Cappadocian]
John of Beirut [Middle East]	Photius of Constantinople [Canon law]
John of Damascus [Poet]	Origen [Alexandrian writer]

In brackets we cite the characterization of each writer.

Table 4. Thematic catalogue of Greek Literature in Sources Chrétiennes

Thematic catalogue (categories)	32.History of Church
1.Annotations	33.Hortative orations
2.Anthropology	34.Hymns
3.Apocrypha	35.Instructive
4.Apologetic	36.Le pre spirituel
5.Apothegms of Fathers	37.Legendaries
6.Ascetics	38.Love
7.Baptesmatic	39.Martyrdoms
8.Canons	40.Mystic
9.Catecheses	41.Objection apologetics
10.Codices	42.On Celestial Hierarchy
11.Comments	43.On Holy Easter
12.Compendia	44.On Incarnation
13.Contra Heresies	45.On Principles
14.Dialogues	46.On Resurrection
15.Dogmatic	47.On the Holy Passion
16.Encomia	48.On the Holy Spirit
17.Encouraging speeches	49.On the Holy Trinity
18.Epistles	50.On thoughts
19.Ethics	51.On Wedding
20.Evangelic	52.Pastoral
21.Exegetic	53.Pedagogic
22.Festive	54.Practical
23.Fests of Lord	55.Questions & responses
24.Fests of Theotokos	56.Speeches
25.Filokalia	57.Teaching
26.Gnostic	58.Theology
27.Hermeneutics	59.Therapeutics
28.Hermeneutics of Divine Liturgy	60.To kings
29.Hermeneutics of New Testament	61.Topography
30.Hermeneutics of Old Testament	62.Virginity
31.Historic Theology	

Progress rate per period
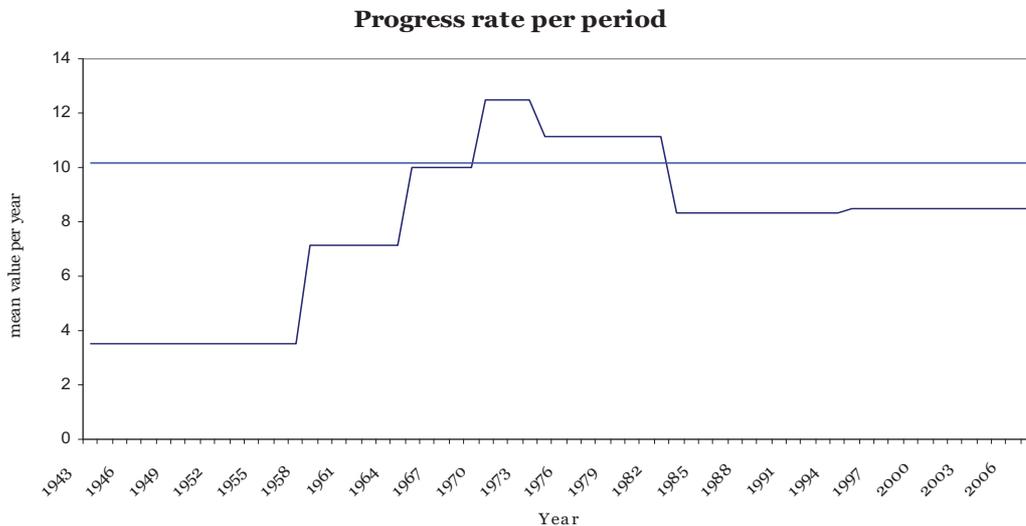

Figure 1. [small graph] The rate of progress for the Sources Chrétiennes collection [large graph] The rate of progress per time period

(a)   (b)
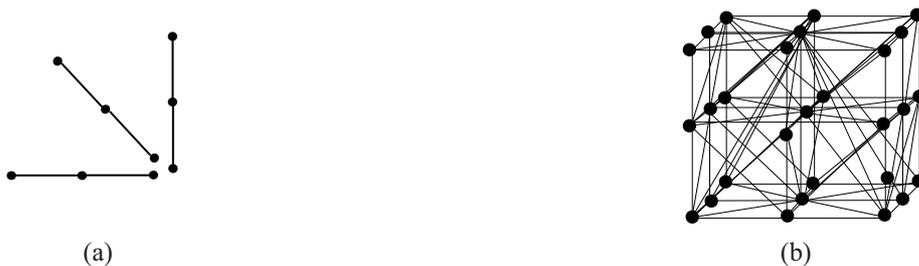
Figure 2. (a) Graph of the *Sources Chrétiennes* Collection. (b) The cubic graph