

テキストデータベース「ゲート・ファイル」書簡集 の完成

樋口, 忠治
九州大学言語文化部

<https://doi.org/10.15017/1470255>

出版情報：九州大学大型計算機センター広報. 27 (4), pp.364-367, 1994-09. 九州大学大型計算機センター
バージョン：
権利関係：



テキストデータベース「ゲーテ・ファイル」書簡集の完成

樋口 忠治*

まえがき

ゲーテ・ファイルは当初ハンプルク版ゲーテ全集全14巻のテキストデータベースとして構想され、データの入力が終わったものから順次公開提供されてきた。その構築が予想以上に早く終わったこともあって、次の段階として、ゲーテの最大の全集であるワイマル版（全133巻）のうち書簡集（全50巻）のデータベース化を計画するに至った。この作業は1988年9月10日に開始し、1994年5月19日に終了した。このことを本日ここに報告できることは私にとって大きな喜びである。

1 ワイマル版ゲーテ全集について

数多いゲーテ全集の中で最も権威ある全集として知られている。別名ゾフィア版ともいう。全体は133巻143冊からなる非常に大きな全集で、その構成は次のようになっている。

- 第1部 55巻 文芸著作物など
- 第2部 13巻 自然科学論文など
- 第3部 15巻 日記
- 第4部 50巻 書簡

このうち10巻はそれぞれ2冊に分冊されているので、合計143冊となっている。第1部および第2部の中のポピュラーなものは、その多くがハンプルク版全集に収められている。

2 書簡集について

若干の例外を除いて、それぞれの書簡には宛先がある。その宛先のみをリストを取り出したい場合を考えて、宛先の前にはすべて“%”記号を付けた。従ってこれを検索すれば、宛先一覧表を入手することができる。書簡の先頭や末尾には日付がある。末尾の日付、発信地、署名などは行数が厳密とはいえないので注意して欲しい。書簡の本文を検索の主たる対象と考えていることから、それ以外については、ピリオドがなければ、原文のように行毎に分けるという方法はとっていない。

それぞれの文の先頭に出典を示す数字が付けられていることは、他のテキストデータベースの場合と同じであるから、1から50までの巻数を示す数字が付いているわけであるが、50ものファイルに分かれているというのは長所もあるが、やはり多すぎて不便を感じる。そこで、5巻ずつを一つのファイルにまとめ、全体を10個のファイルとした。検索は全体を対象として行うことが多いであろうということもあって、各巻毎の検索を犠牲にしたということになる。

*九州大学言語文化部

ゲータ・ファイル全体の一覧表(資料1)を下に示す。上の部分の GA01 から GC14 まではハンブルク版全集に相当し、下の部分の WBR01 から WBR10 までがワイマル版書簡集に相当する。また最後に、格納されているデータの一部をサンプルとして示した。

資料1

```

READY SIGMA
SIGMA> DDIR S
PASSNUMBER:=
OPTION:= A
FILENAME                ID ALIAS      SIZE    DATE TIME
BIBEL                    437K    0    1380033 92:11:02 14:39
    Text-Database von dem neuen Testament
GA01                      1869E    0    687411 92:06:19 13:27
    Bd. 1 ( 7-391) Bd. 2 ( 7-125)
GA02                      1899I    0    805583 92:06:19 13:29
    Bd. 2 (126-270;285-536)
GA03                      1934S    0    689649 92:06:19 13:32
    Bd. 3 ( 9-364;367-420)
GA04                      1964Z    0    732842 92:06:19 13:34
    Bd. 4 ( 7- 72;176-202;455-462) Bd. 5 ( 7-167;215-308;332-454)
GA05                      19963    0    812846 92:06:19 13:35
    Bd. 4 ( 74-175;203-454) Bd. 5 (168-214;309-331)
GB06                      1407X    0    1198907 92:05:14 13:16
    Bd. 6 ( 7-513)
GB07                      1460I    0    2580270 92:05:14 13:17
    Bd. 7 ( 9-610) Bd. 8 ( 7-516)
GB09                      1573I    0    1849524 92:05:14 13:18
    Bd. 9 ( 7-598) Bd.10 ( 7-187)
GB10                      1654I    0    833500 92:05:14 13:19
    Bd.10 (188-547)
GB11                      1690Z    0    1265685 92:05:14 13:35
    Bd.11 ( 9-556)
GB12                      1746H    0    1121377 92:05:14 13:36
    Bd.12 ( 9-364;365-547)
GC13                      1795L    0    1096918 92:05:14 13:37
    Bd.13 ( 7-523)
GC14                      1843N    0    586775 92:05:14 13:49
    Bd.14 ( 7-269)
*****
NUM                        792Z    0    360 92:06:23 18:05
README                    1278N    0    2474 92:12:15 14:52
SMALL                     668W    0    919 91:06:03 17:28
*****
WBR01                     1047I    0    1964693 91:06:07 18:36
    Weimarer Ausgabe , Briefe Bd. 1- 5
WBR02                     851C    0    2337330 91:06:03 17:32
    Weimarer Ausgabe , Briefe Bd. 6-10
WBR03                     9530    0    2141324 91:06:03 17:33
    Weimarer Ausgabe , Briefe Bd. 11-15
WBR04                     668Z    0    2111046 91:06:03 17:29

```

Weimarer Ausgabe , Briefe Bd. 16-20					
WBR05	11464	0	2828963	91:08:27	15:12
Weimarer Ausgabe , Briefe Bd. 21-25					
WBR06	7921	0	2104583	92:08:25	17:58
Weimarer Ausgabe , Briefe Bd. 26-30					
WBR07	160B	0	1915414	93:03:06	11:00
Weimarer Ausgabe , Briefe Bd. 31-35					
WBR08	3480	0	1976714	93:08:25	12:23
Weimarer Ausgabe , Briefe Bd. 36-40					
WBR09	611P	0	1945675	93:12:25	11:20
Weimarer Ausgabe , Briefe Bd. 41-45					
WBR10	1345D	0	1682911	94:05:19	13:36
Weimarer Ausgabe , Briefe Bd. 46-50					

PREFIX = A70152C					

資料 2

@#

#0100100 %1 AN <LUDWIG <YSENBURG VON <BURI .
 #0100101 WOHLGEBOHRNER , INSONDERS <HOCHZUEHRENDER <HERR , <EW. <WOHLG
 EBOHREN WERDEN <SICH WUNDERN , WENN EIN UNBEKANNTER SICH UNTERSTEHET ,
 BEY <IHNNEN EINE <BITTE VORZUBRINGEN .
 #0100105 DOCH BILLIG FOLTEN <SIE MIT ALLEN <DENJENIGEN , DIE IHRE <VER
 DIENSTE KENNEN , NICHT ERSTAUNEN .
 #0100107 DA <SIE WOHL WISSEN K=ONNEN , DA\$ IHRE <EIGENSCHAFTEN SELBST
 AUCH NOCH IN FERNERN <L=ANDERN ALS WO ICH WOHNE DIE <GEM=UTHER <IHNNEN
 EIGEN ZU MACHEN VERM=OGEND SIND .
 #0100110 <SIE SEHEN AUS MEINER <VORREDE , DA\$ ICH ZUR <ZEIT , UM NICHT
 S ALS IHRE <BEKANNTSCHAFT ANHALTE , BI\$ <SIE ERFAHREN , OB ICH WERTH
 BIN , IHR <FREUND ZU SEYN , UND IN IHRE <GESELLSCHAFT EINZUGEHEN .
 #0100114 WERDEN <SIE =UBER MEINE <K=UHNHEIT NICHT UNWILLIG , UND VERZE
 IHEN <SIE IHR .
 #0100115 ICH KANN NICHT ANDERS , DENN WENN ICH AUCH L=ANGER SCHWEIGEN
 UND IHRE GROSE <EIGENSCHAFTEN INSGEHEIM VEREHREN WOLTE , WIE ICH BI\$H
 ER GETHAN HABE , SO W=URDE MIR DIESES DIE GR=0\$TE <BETR=UBN=U\$ VON DER
 <WELT ERWECKEN .
 #0100119 KEINER VON MEINEN <FREUNDEN DIE <SIE KENNEN , G=ONNT MIR DIES
 ES UNSCH=ATZBARE <GL=UCK .
 #0100121 VIELLEICHT IST AUCH EIN KLEINER <NEID <SCHULD DARAN .
 #0100122 ABER EBEN F=ALLT MIR DIE BESTE <URSACHE EIN , <SIE WOLLEN KEI
 NEN <MENSCHEN , DER MEINEN <FEHLER HAT IN IHRE <BEKANNTSCHAFT BRINGEN
 , DAMIT <SIE DESWEGEN NICHT ZUR <VERANTWORTUNG GEZOGEN WERDEN .
 #

3 検索に際しての注意

ワイマル版ゲートル全集は正書法以前のものであるから、現代のドイツ語に馴れた者にとつては検索は必ずしも簡単ではない。すなわち、語の綴り方が一貫していないことがある。そ

ればかりではなく、植字工のミスと考えるには余りにもたびたび出くわすところの besuchen(besuchen)、hergegestellt(hergestellt)、あるいは Hofnung(Hoffnung) のような綴りなどがあり、このような事実を知った上で検索をしなければ誤った結果を得ることになりかねない。検索をする際にはある程度このテキストの癖に馴れていることが必要である。

追記

今回、ゲーテの書簡集がゲーテ・ファイルに完全な形で加わったことにより、九州大学大型計算機センターの公用データベースとしてのゲーテ・ファイルはハンプルク版全集全14巻とワイマル版全集の書簡集全50巻とからなるかなり大きなテキストデータベースとなった。

参考文献

- [1] テキスト・データベース「ゲーテ・ファイル」の公開, 広報 Vol. 21, No. 3-4, 1988.
- [2] ゲーテ・ファイルの完成について, 広報 Vol. 22, No. 1, 1989.
- [3] テキストデータベース管理システムSIGMA第2版について, 広報 Vol. 20, No. 6, 1987.