

未解読文書 の解読可能性の判定

－ 世界で最も謎に満ちた写本を例に －

安形輝（亜細亜大学）

[安形麻理（慶應義塾大学）]

2008／02／26 研究談話会

本日の構成

- ・ 自己紹介、今回の研究のきっかけ
- ・ 未解読文書とは
- ・ 解読可能性判定法の提案
- ・ Voynich写本とは
- ・ 過去の解読の試み
- ・ 部分文書の類似度測定とクラスタリング
- ・ まとめ

今回の研究のきっかけ？

<http://x51.org/>より

X51.ORG : 解読不能の書 ヴォイニッチ手稿の謎 - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(I) ヘルプ(H)

[←](#) [→](#) [↺](#) [⌂](#) [51](#) <http://x51.org/x/03/10/1142.php> [RSS](#) [▶](#) [G](#) Google [🔍](#)

X51.ORG

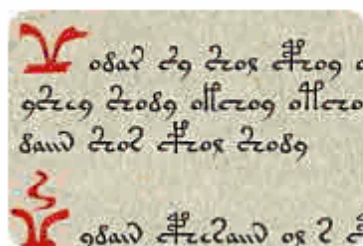
[HOME](#) | [EARTH](#) | [ANIMA](#) | [ANAL](#) | [ENEMA](#) | [FILES](#) | [CONTACT](#)

[HOME](#) > [X-FILES](#) > Article

解読不能の書 ヴォイニッチ手稿の謎

【AlchimedLab】誰も知らない場所で、誰も知らない時に、誰も知らない文字で書かれた書物を想像してみたい。ここに誰も知らない日付で、誰も知らない場所で、誰にも知られない文字で書かれた一冊の書物がある。これは決して空想の書ではない。「ヴォイニッチ手稿

(VionichManuscript)」と呼ばれる中世のいつかに何者かによって書かれたこの謎の書物は、これまで長年に渡って科学者や歴史家達を悩ませてきたのである。このヴォイニッチ手稿は1912年、イタリアのローマに程近いモンドラゴネ寺院書庫にてアメリカの古書収集家Wilfrid M. Voynich（書物の名前は彼の名前に由来する）が発見、彼はその原稿の異常さと潜在的価値を見込んですぐにその「読めない書物」を購入したのである。彼はアメリカに持ち帰るなり、すぐにコピーを取り、古文書学者、暗号学者、歴史家、言語学者、哲学者、更には天文学者から植物学者といった様々な人々にそのコピーの解読を求め、送りつけたの



X51.ORG THE ODYSSEY



- 迫真の謎解きチャレンジに脱帽する！！（荒俣宏・帯）
- 対象への飽くなき愛に支えられたオカルト世界紀行 — 素晴らしい！（香山リカ・週刊ポスト）
- 私は読みながら、胸躍ってな

らなかった。（三浦しをん・読売新聞）

■ 週刊プレイボーイ、STUDIOVOICE、文芸春秋、週刊SPA!、ダ・ヴィンチ他に書評掲載

■ 【お知らせ 01/10】現在、一時的に在庫切れの状態ですが、近日中に販売再開されますので定価以上の中古にご注意下さい。

[>>X51.ORG THE ODYSSEY特設サイトへ](#)

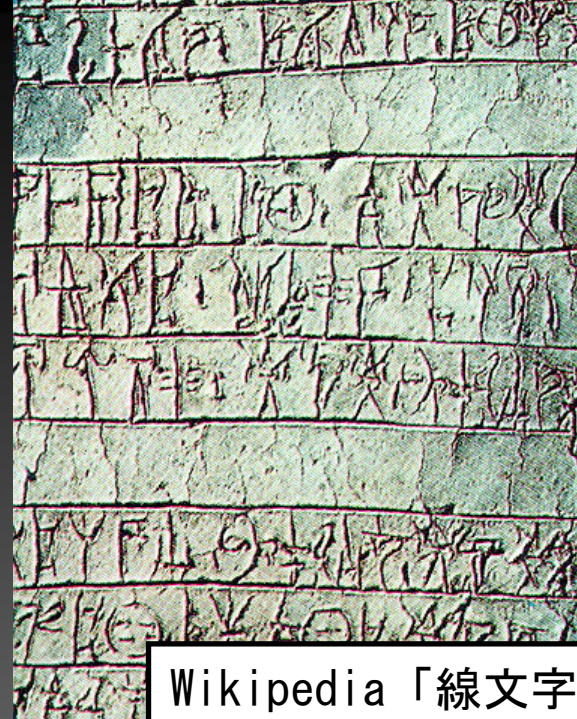
NEWS HEADLINE

[RSS](#)

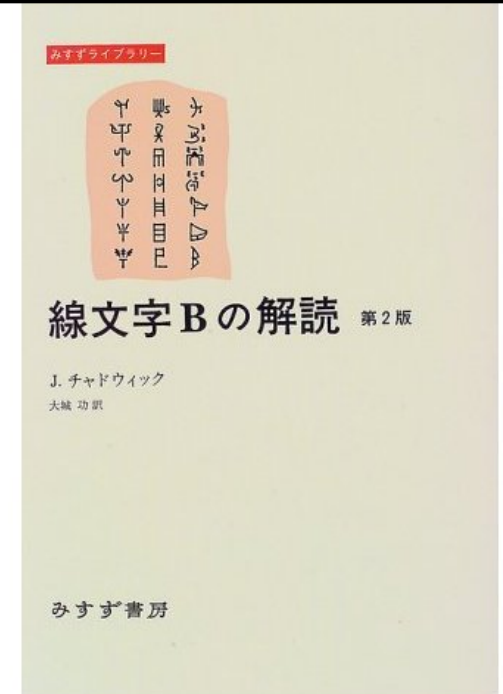
- ▶ 男性の顔色が真っ青になる 米 **NEW!**
- ▶ 「X51.EARTH」リリースのお知らせ **NEW!**
- ▶ 1933年、自殺の聖歌『暗い日曜日』の謎 **NEW!**
- ▶ 自動車性交 - 車とセックスする男 英

未解読文書とは

- ・ 例えば「線文字B」
によるクレタ島文書
 - 当初はクレタ文明独自の言語？
 - 1952年 ヴェントリスとチャドウィックがギリシャ語として解読



Wikipedia「線文字B」より





解読できない文書



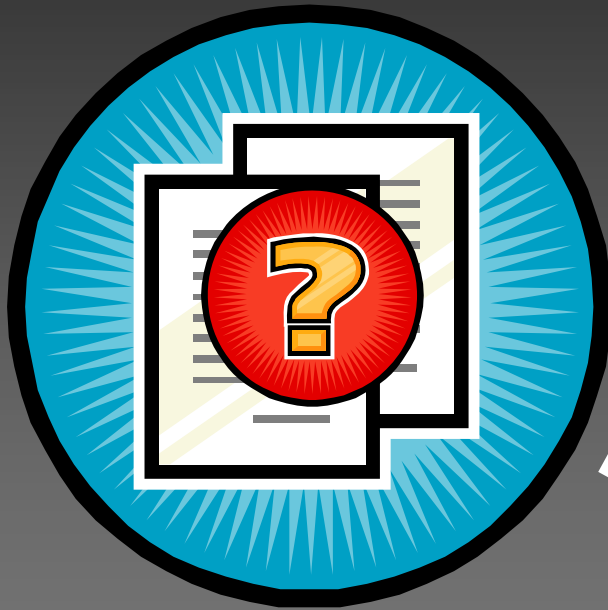
暗号解読、翻訳



言語学的分析
→言語的特徴があるか

仮
説

多くの言語に適用可能な
テキスト処理技術は未知
の言語にも応用可能



解読できない文書

価値判断、解
読に傾注すべ
き労力の判断

テキスト処理技術



図表やページ構造な
どの他の手がかりと
対応をみる
＝捏造文書の判定

The Vinland Map



他の要素からの構造

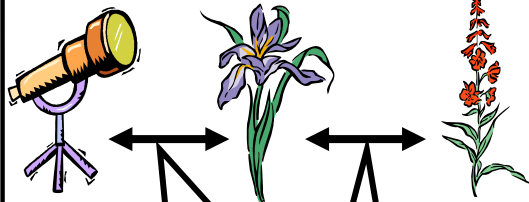
ページ順

1, 2, 3, ..., $n-2$, $n-1$, n

内容的
に近い

内容的に
近くない

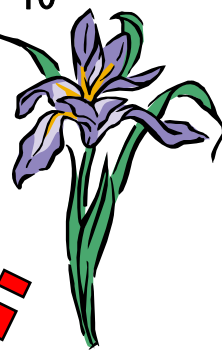
図の類似性



似ていない

似ている

Page 10



未解読文書

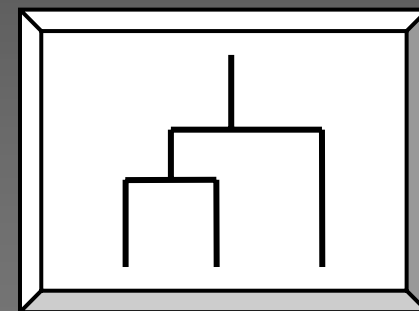
<比較>

捏造／真正？

テキストからの構造

読めない
テキスト

テキスト
クラスタリング



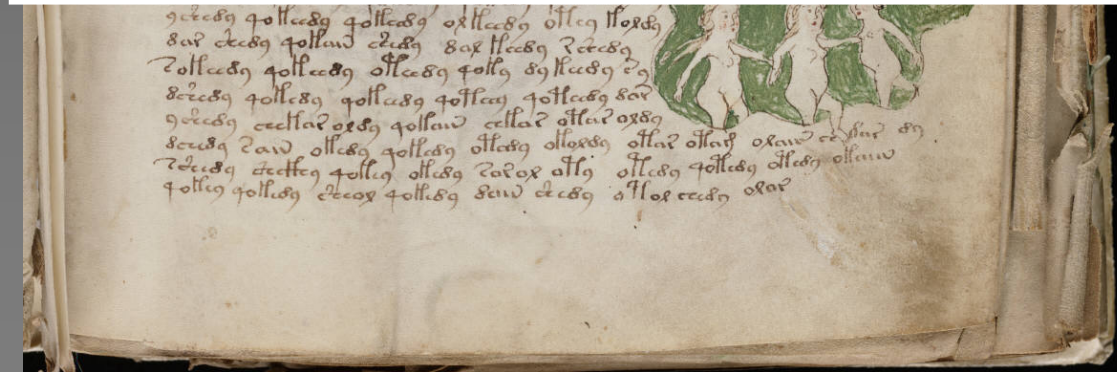
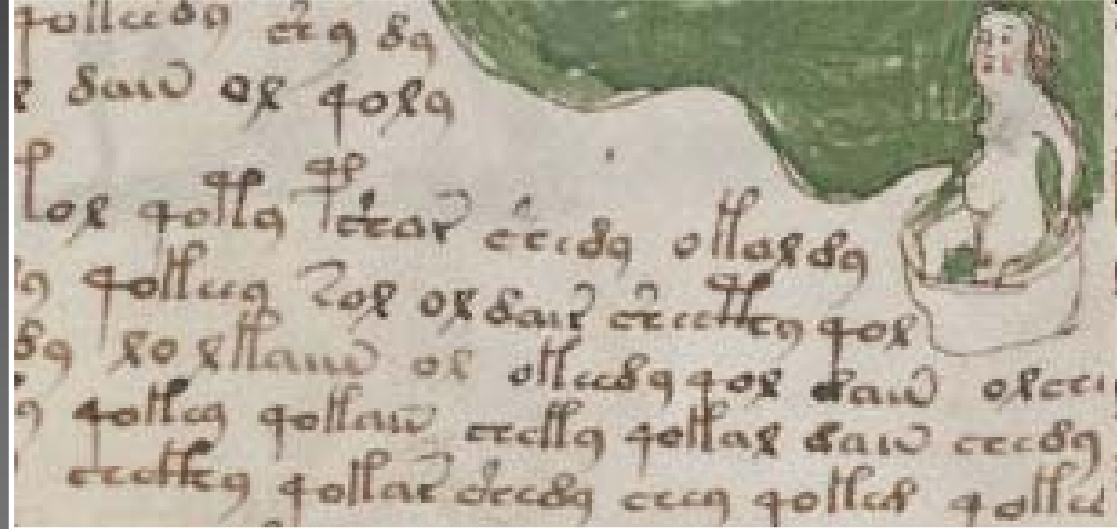
Voynich写本

- ・ 1912年Wilfred M. Voynichがイタリアで発見
- ・ 未知の言語／暗号、いままで全ての解読の試みは失敗
- ・ 現在はイエール大学のバイネッケ図書館の蔵書 (MS 408)

- ・ 制作地や年代は、目録によれば15末～16世紀中欧と推定

- ・ 102葉の羊皮紙
(14葉欠落あり)

- ・ アラビア数字やアルファベットに似たヴォイニッチ文字
- ・ 同じ文字の資料は存在しない



・ 大部分のページ
に挿図

・ 図の内容は、植
物、天文学、小
さな裸の女性、
十二宮図、薬草
の調合用壺など



・ 真正性の根拠は200
ページを超える
＝捏造にしては作成
の労力が大きい
⇔強い動機を持つ捏
造者

・ 外的証拠として、こ
れと思われる写本に
言及した17世紀の複
数の書簡が現存



書簡 MS 408A

- ・ バイネツケ図書館に収蔵
- ・ 1665年にMarcus Marciが暗号専門家Atanasius Kircher宛に解読を依頼する内容

Reuerendo et Eximie Viri
in Christo Pater

Librum hunc ab Amico singulari mihi testamento relictum, mox
eundem tibi amicissime Athanasi uti primum possidem capi, animo des-
tinavi. Siquidem persuasum habui à nullo nisi attulsi legi posse. Sed
aliquando per litteras ejusdem libri tum possessor judicium tuum parte
aliqua libri à se descripta et tibi transmissa, ex qua reliqua à te legi
posse persuasum habuit, verum librum ipsum transmittere tum reu-
cabat, in quo disciffrando posuit indefessum laborem, uti manifestum ex
conatibus ejusdem hic una tibi transmissis, neg prius huius spei quam
vile hoc finem fecit. Verum labor hic frustraneus fuit, Siquidem non
nisi suo Kircheri obediunt ejusmodi ptinges. Accipe ergo modo quod
nunc tibi debebatur hoc quaecumque mei erga te affectus indicium,
huiusq; seras si qua sunt, consueta tibi felicitate per rumpe. redulit mihi
D. Doctor Raphael Terdinandi tertij Regis tum Boemia in lingua boemica
instructor dictum librum fuisse Rudolphi Imperatoris, pro quo ipse latoni
qui librum attulisset 600 ducatos protendarit, Rethorem vero ipsum pu-
tabat esse Rogerium Baconem Anglum. ego judicium meum hic sus-
pendo. tu vero quid nobis hic sentiendum detine, cupis fauori et gra-
tis me totum commento manco.

Atia vestre

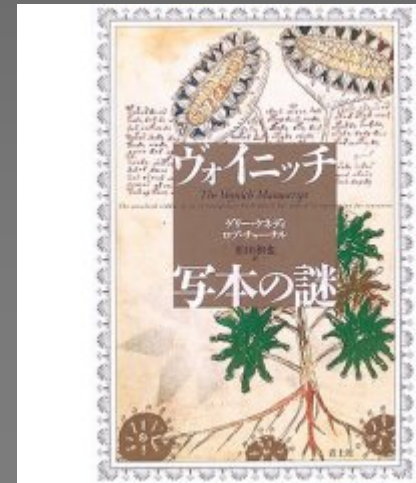
Praga 19 Augusti.
1665

Ad hunc uia

Joannes Marcus Marci
in Cracovia.

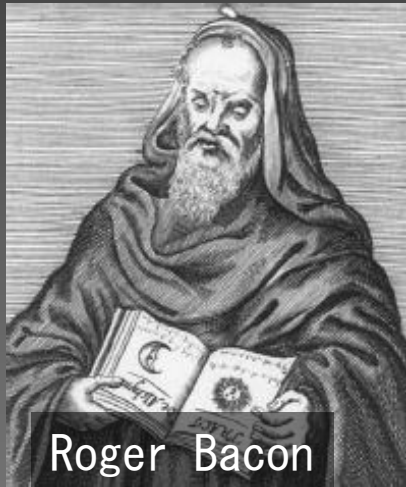
「ヴォイニッチ写本の謎」

- ・ イギリスのBBCで放送されたドキュメンタリー「ヴォイニッチの謎」の書籍版
 - 著者の一人がWilfred M. Voynichの子孫
- ・ ISBN:4-7917-6248-7



解読の初期の試み

- ・ 1920年代の研究
- ・ 13世紀 Roger Baconによる暗号書？



Roger Bacon

- ・ 英語学者J. M. Manlyは、Newboldの研究が著しく主観的かつ不完全であると批判



WILLIAM ROMAINE NEWBOLD

1865-1926

The Portrait by Joseph Sacks

Friedman夫妻

- ・ 第二次大戦日本軍暗号解読で有名な夫妻
- ・ Voynich写本は「ア・プリオリなタイプの人工的もしくは普遍的言語を作成しようとする初期の試みである」とアナグラムメッセージを遺す



Elizebeth Friedman



William F. Friedman

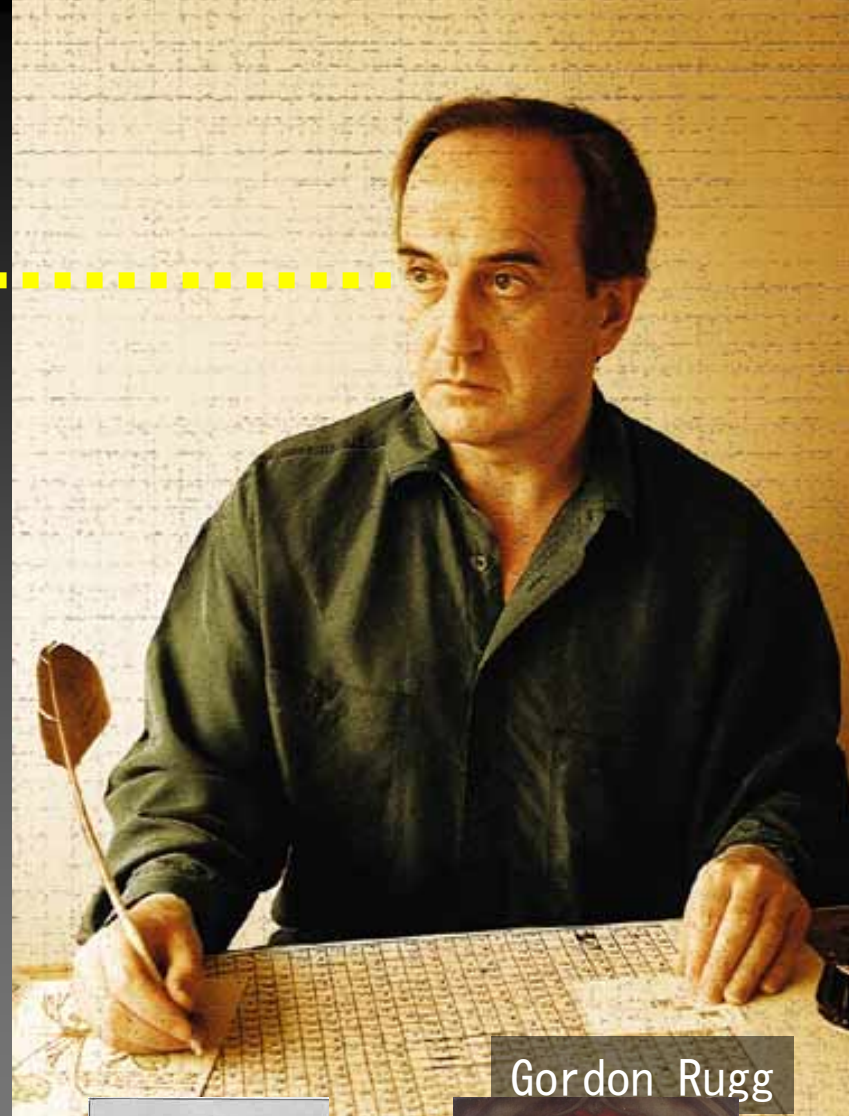
言語学的な分析

- ・ 1998年 G. LandiniとR. Zandbergen
Zipfの法則に従うことを検証
- ・ 2000年 Zandbergen
エントロピー分析ではラテン語と同程度の複雑さを持つ＝冗長暗号ではない

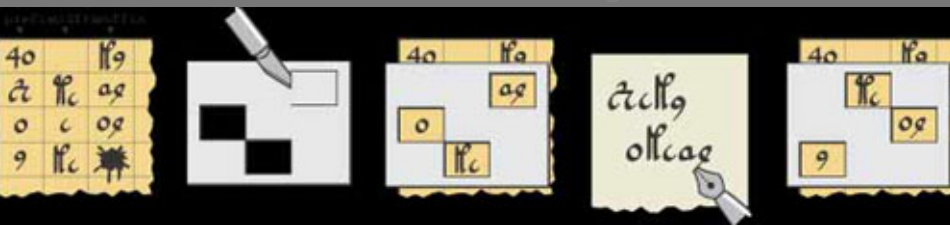
→ 真正性への根拠

写本を捏造とする研究

- ・ 2004年キール大学の Gordon Ruggによる復元実験
- ・ 似た特徴の文書をカルダーノ・グリルで作成が容易に可能
- ・ 16世紀の錬金術師で John Deeの弟子？
Edward Kellyの捏造



Gordon Rugg



Edward Kelly



John Dee



カルダーノ・グリル

- ・ 中世のイタリアの数学者ジェロラモ・カルダーノ (Girolamo Cardano) の考案
- ・ 英語読みでカルダングリル (Cardan grille) と呼ばれる

Sir John regards you well and spekes again that
all as rightly 'sails him is yours now and ever.
May he 'tore for past d'lays with many charms.



John DeeとEdward Kelly

- ・ John Dee
 - 16世紀エリザベス女王の下で相談役
 - ・ 有名なコード名
 - 錬金術師、数学者
- ・ Edward Kelly
 - 錬金術師、霊媒師
- ・ DeeとKellyの日記
 - エノク語＝人工言語？

$\begin{smallmatrix} 2 & b \\ b & 1 \end{smallmatrix}$	$\begin{smallmatrix} G \\ b & b \end{smallmatrix}$	$\begin{smallmatrix} q \\ b \end{smallmatrix}$	$\begin{smallmatrix} B \\ 2 & 2 \end{smallmatrix}$	$\begin{smallmatrix} 2.4.6 & b & b \\ b & b & b \\ 1 & 4 & 6 \end{smallmatrix}$	$\begin{smallmatrix} L \\ b \end{smallmatrix}$	$\begin{smallmatrix} B \\ rog \end{smallmatrix}$	$\textcircled{B}$
$\begin{smallmatrix} 8 & b \\ b & 2 \end{smallmatrix}$	$\begin{smallmatrix} ff \\ b & 8 \end{smallmatrix}$	$\begin{smallmatrix} G \\ b \end{smallmatrix}$	$\begin{smallmatrix} G & G \\ b \end{smallmatrix}$	$\begin{smallmatrix} 152 \\ b \end{smallmatrix}$	$\begin{smallmatrix} 152 \\ b \end{smallmatrix}$	$\begin{smallmatrix} 52 \\ BBB \end{smallmatrix}$	$\begin{smallmatrix} B' \\ -I- \\ B \end{smallmatrix}$
$\begin{smallmatrix} D \\ q & B \\ q & 1 \end{smallmatrix}$	$\begin{smallmatrix} b & o \\ o & o \end{smallmatrix}$	$\begin{smallmatrix} B \\ 7 & 9 \end{smallmatrix}$	$\begin{smallmatrix} b & b & b \\ b & b & b \\ b & b & b \end{smallmatrix}$	$\begin{smallmatrix} 11 \\ B \\ 5 \end{smallmatrix}$	$\begin{smallmatrix} b & b \\ b & b \end{smallmatrix}$	$\begin{smallmatrix} b & b \\ b \end{smallmatrix}$	$\begin{smallmatrix} b \\ 8 & 3 \\ b \end{smallmatrix}$
$\begin{smallmatrix} b & b \\ b & b \\ b & b \end{smallmatrix}$	$\begin{smallmatrix} b & b \\ b & 15 \\ b & b \end{smallmatrix}$	$\begin{smallmatrix} b & M \\ 1 & 6 & 6 \end{smallmatrix}$	$\begin{smallmatrix} 7 \\ \triangle & b \end{smallmatrix}$	$\begin{smallmatrix} \textcircled{b} \\ 5 \end{smallmatrix}$	$\begin{smallmatrix} G \\ M \end{smallmatrix}$	$\begin{smallmatrix} B \end{smallmatrix}$	$\begin{smallmatrix} b & A \\ 1556 \end{smallmatrix}$
$\begin{smallmatrix} b \\ 1 \end{smallmatrix}$	$\begin{smallmatrix} B \\ 1 & 2 & 3 \end{smallmatrix}$	$\begin{smallmatrix} b \end{smallmatrix}$	$\begin{smallmatrix} T \\ b \end{smallmatrix}$	$\begin{smallmatrix} 4 \\ BB \\ 9 \end{smallmatrix}$	$\begin{smallmatrix} BBB \\ 6 \end{smallmatrix}$	$\begin{smallmatrix} b & b \\ 72 \\ F \end{smallmatrix}$	$\begin{smallmatrix} b \end{smallmatrix}$

Chetham Library



Ruggの研究の反響

- ・ 近年のヴォイニッチ研究の中では最も注目を集めている
 - Nature Science Update他、一般誌を含む数多くの雑誌に取り上げられ、日本語にも訳される

ヴォイニッチ手稿の謎

中世に書かれ、「世界で最も不可解な書物」と呼ばれてきた謎の手稿がある
奇妙な文字でつづられており、解読の試みはことごとく失敗……
しかし最近の研究によって、巧妙な“でっち上げ”の線が強まった

G. ラグ (英キール大学)

1912年、稀こう本を扱う米国の古書商ヴォイニッチ (Wilfrid Voynich) はローマ近郊にあるイエズス会系大学の図書館で一世一代の掘り出し物を見つけた。約230ページからなる1冊の手稿だ。見慣れない文字でつづられており、植物や天体、入浴する女性など

稿は広く知られるところとなった。
この手稿はいったい何か、どこでどのように生まれたのか？ 解読の試みがことごとく失敗したため、そもそも略号文ではないかもしれないという疑いが浮上した。意味のあるメッセージは含まれておらず、人を一杯食わせる

とする古くからの説を強く支持するものだ (皇帝は手稿を600ダカット、現在の貨幣価値にして5万ドル相当で買い上げたと伝えられている)。

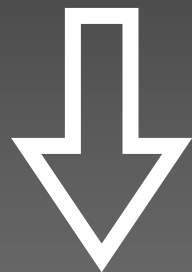
私がヴォイニッチ手稿の謎の分析に用いた一連の手法は、別の分野の難問にも適用できると思われる。この古び

Ruggの研究の検証

- ・ Andreas Schinner. The Voynich Manuscript: Evidence of the Hoax.
 - Cryptologia, vol. 31, no. 2, p. 95–107, 2007
 - ・ Voynich写本は何らかの言語を暗号化したものではなく確率過程によって生成された可能性が高い
- =ねつ造説およびRuggの仮説を支持

既往研究のまとめと本研究の目的

- ・ 数多くの解読の試み→解読の成功例はない
- ・ 2004年Ruggの研究以降真贋論争が再燃



- ・ ヴォイニッチ写本の部分文書をクラスタリングしたときに、挿図との対応が見られるか
＝文書構造を持つか、デタラメな文書かを判定

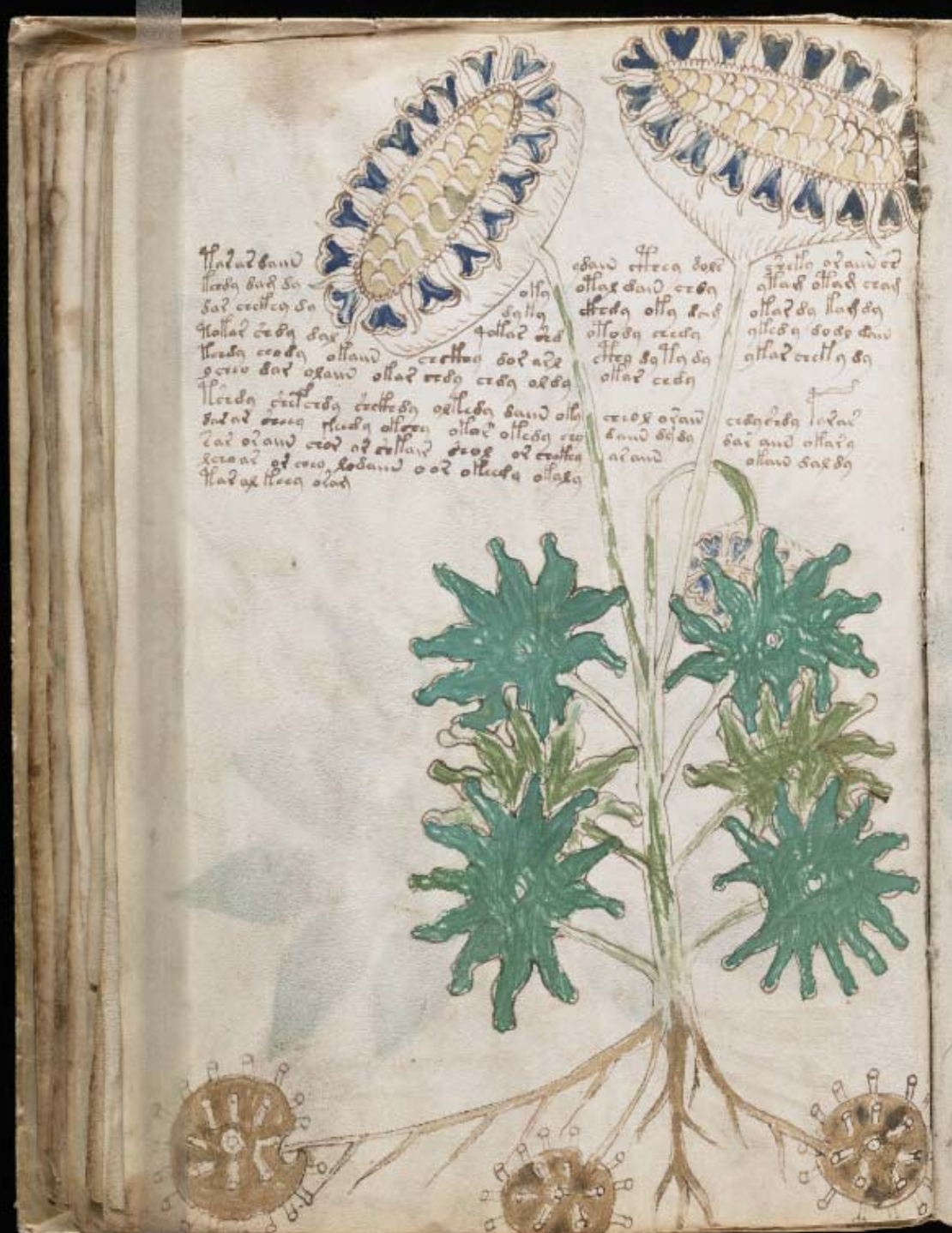
- ・ 部分文書の単位：ページ単位
- ・ 図との対応：
 - 図から推測されたセクション

フォリオ番号	セクション	ページ数
f. 1r~65v	植物 (Plant)	116
f. 66r~73 v	天文 (Astoro.)	26
f. 75r~84 v	生物 (Bio.)	20
f. 85r1~85v6	十二宮図 (Zodiac)	8
f. 87r~102v2	薬草 (Herb.)	32
f. 103r~116r	レシピ (Recipe)	23

植物

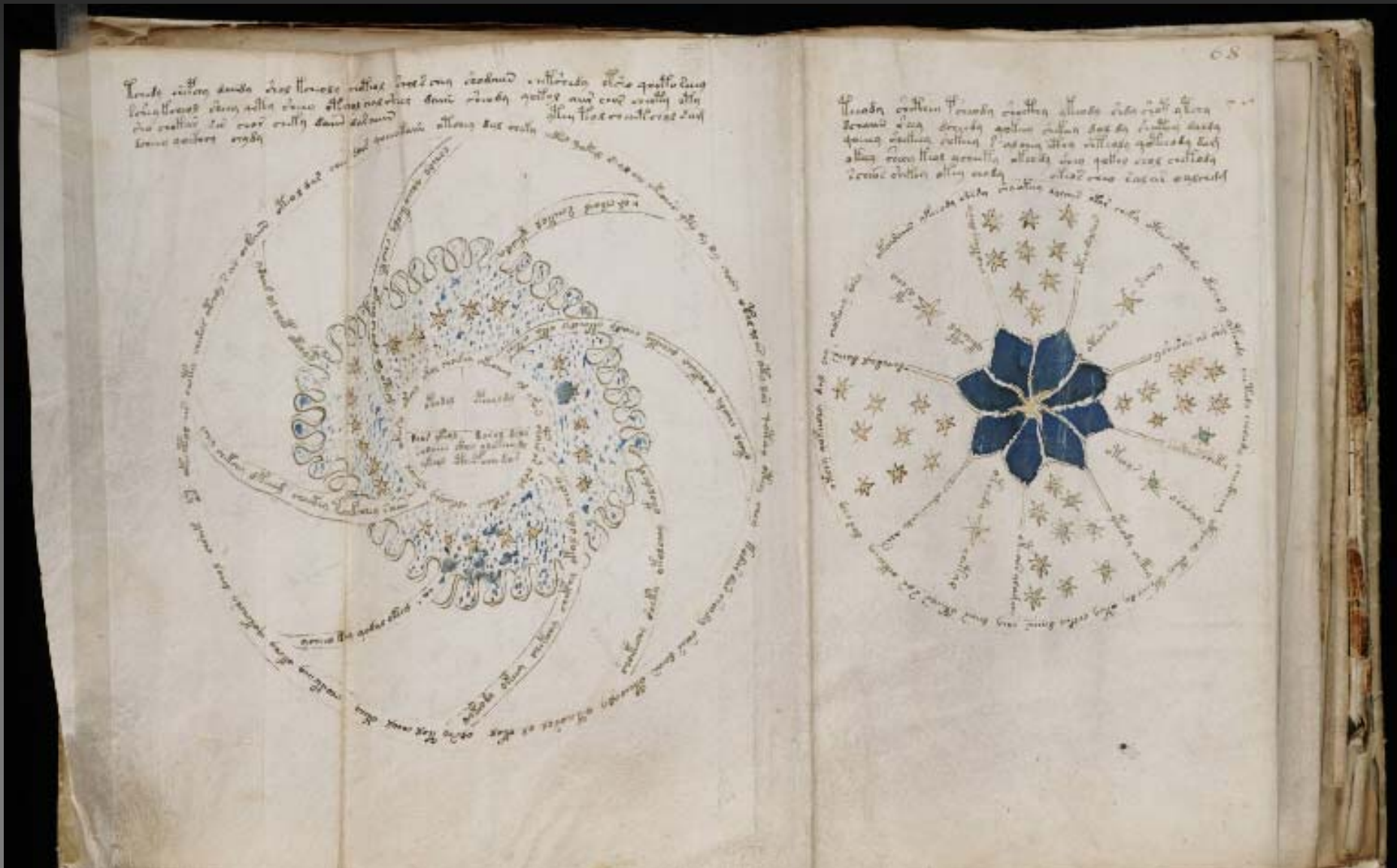
セクション

- ・ 未知の植物の図
- ・ 説明文らしきもの
- ・ 分量が最も多い



天文 セクション

- ・ テキストは比較的少ない
- ・ 星雲が描かれているという説も



生物

セクション

- ・ テキストは比較的多い
- ・ 小さい裸の女性が描かれている

十二宮図 セクション

- ・ テキストはほとんどない
- ・ 複数枚で一つ
- ・ 数は最も少ない



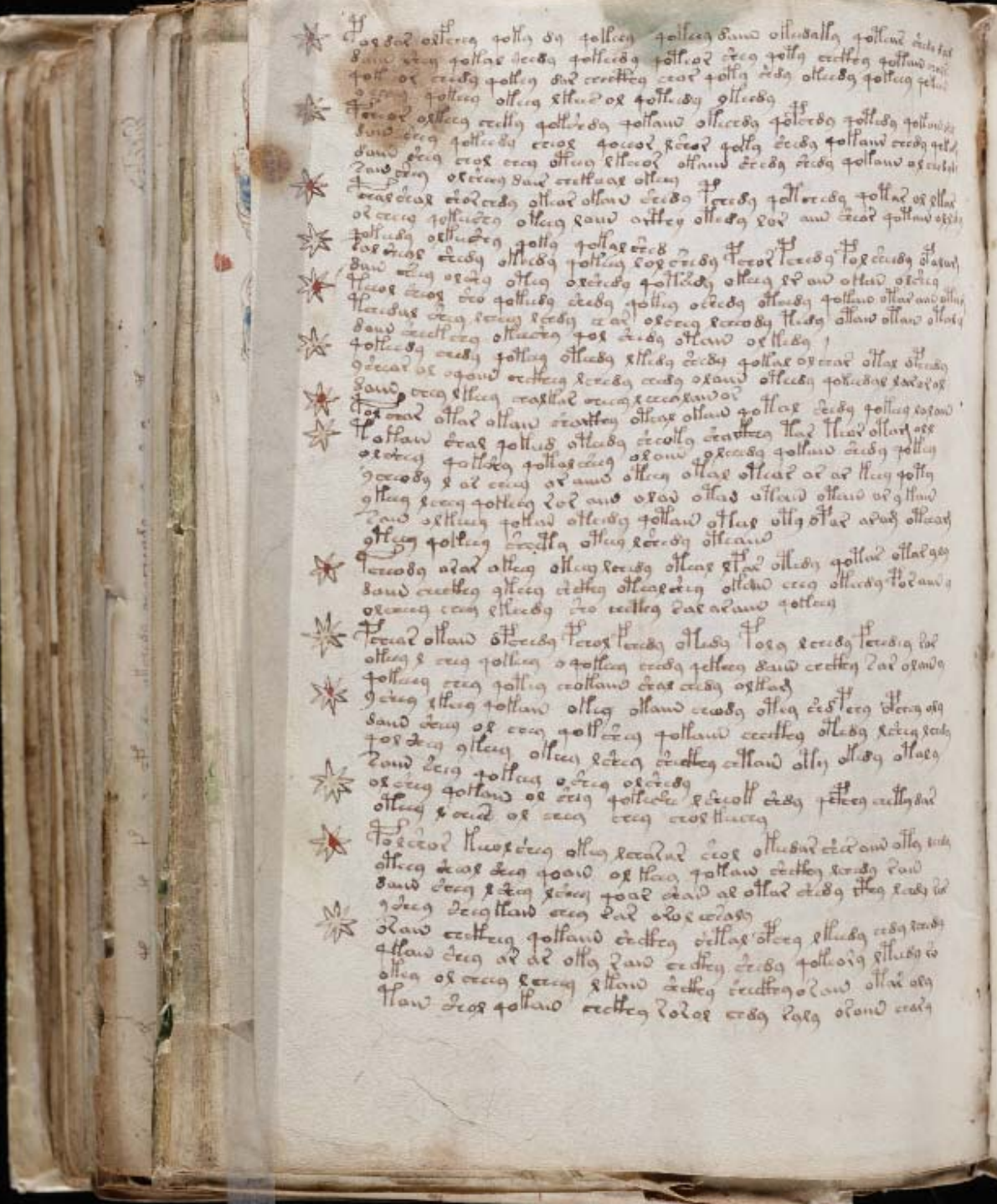
薬草セクション

- ・ 絵柄は植物セクションに似ている
- ・ 薬草壺らしき物体



レシピ セクション

- ・ 行頭文字のよ
うな挿図のみ
- ・ ほとんどテキ
スト

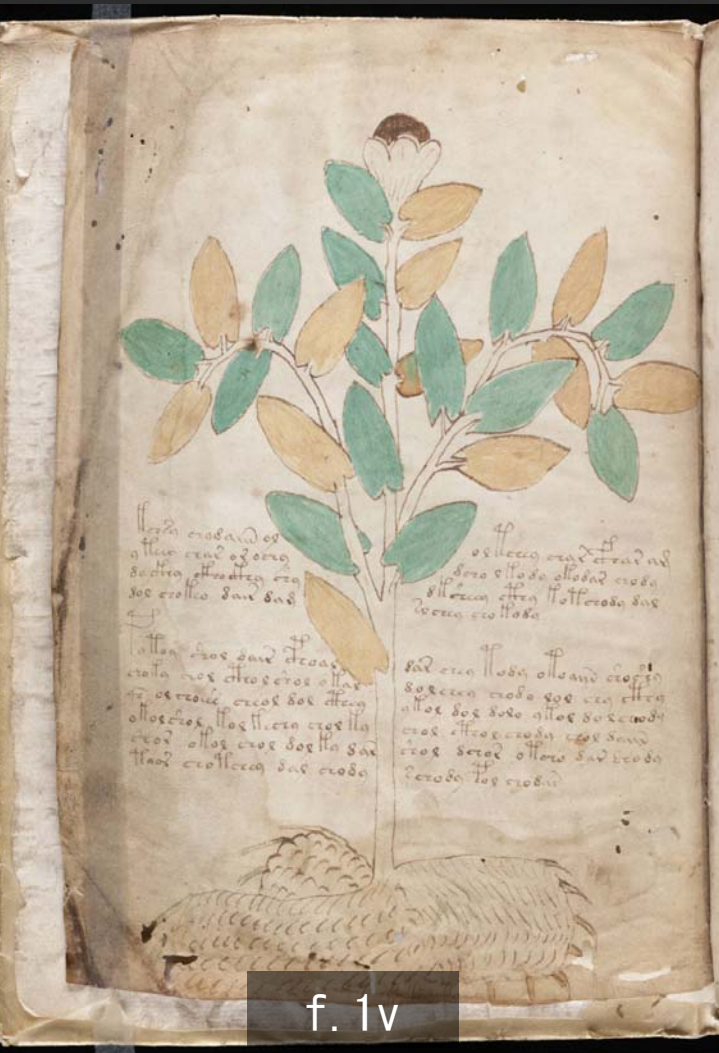


トランスクリプションの必要性

- ・ 未解読文書であるため、テキスト処理の
にはトランスクリプションが必要
- ・ EVA : European Voynich Alphabet
 - Zandbergenらが考案
 - 26文字の基本セット

・ トランスクリプション

- 高橋健によるトランスクリプションを利用



kchsy. chadai in. ol-
[plant] oltchey. char. cfhar. am-
yteey. char. or. ochy-
[plant] dcho. lkody. okodar. chody-
d!o. ckhy. ckho. ckhy. shy-
[plant] dksheey. cthy. kotchody. dal-
dol. chokeo. dair. dam-
[plant] sochey. chokody=
potoy. shol. dair. cphoal-
[plant] dar. chey. tody. otoai in. shoshy-
choky. chol. cthol. shol. okal-
[plant] dolchey. chodo. lol. chy. cthy-
qo. ol. choeee. cheol. dol. cthey-
[plant] ykol. dol. dolo. ykol. do. !!ch!!ody-
okol. shol. kol. kechy. chol. ky-
[plant] chol. cthol. chody. chol. dai in-
shor. okol. chol. dol. ky. dar-
[plant] shol. dchor. otcho. dar. shody-
taor. chotchey. dal. chody-
[plant] schody. pol. chodar=

- ・ 単語の切り分け：空白と改行で行う
- ・ 注釈の除去：
 - {} で囲まれた注釈は除去

異なり単語数	7907 語
平均単語数／ページ	166.0 語
平均単語長	5.0 文字
総ページ数	225 ページ

- ・ 単語の重み付け : TF・IDF法

$$w_{ij} = tf_{ij} \cdot idf_j = \frac{f_{ij}}{\max_{j=1, \dots, M} f_{ij}} \cdot \log \left(\frac{N}{n_j} \right)$$

- ・ 部分文書同士類似度 : キャンベラ距離
 - 距離が近いほど、類似度が高い

$$d_{ij} = \sum_{k=1}^n \frac{|x_{ik} - x_{jk}|}{|x_{ik}| + |x_{jk}|}$$

	Plant_f1r	Plant_f1v	Plant_f2r	Plant_f2v	Plant_f3r	Plant_f3v	Plant_f4r
Plant_f1r	0	7907	7907	7907	7907	7907	7907
Plant_f1v	7907	0	7615.696	7619.004	7570.843	7635.833	7713.638
Plant_f2r	7907	7615.696	0	7417.186	7448.504	7456.909	7636.899
Plant_f2v	7907	7619.004	7417.186	0	7401.165	7642.471	7556.286
Plant_f3r	7907	7570.843	7448.504	7401.165	0	7345.488	7502.947
Plant_f3v	7907	7635.833	7456.909	7642.471	7345.488	0	7566.793
Plant_f4r	7907	7713.638	7636.899	7556.286	7502.947	7566.793	0
Plant_f4v	7907	7710.788	7386.276	7313.975	7642.974	7346.136	7645.943
Plant_f5r	7907	7563.289	7370.537	7384.411	7250.004	7424.785	7582.286
Plant_f5v	7907	7686.23	7563.814	7577.542	7037.572	7554.82	7771.522
Plant_f6r	7907	7719.043	7664.163	7472.549	7175.165	7584.594	7515.726
Plant_f6v	7907	7660.002	7588.82	7370.369	7617.167	7591.089	7628.562
Plant_f7r	7907	7579.024	7165.719	7436.899	7564.423	7449.303	7625.256
Plant_f7v	7907	7647.892	7660.676	7595.512	7403.558	7372.305	7713.869

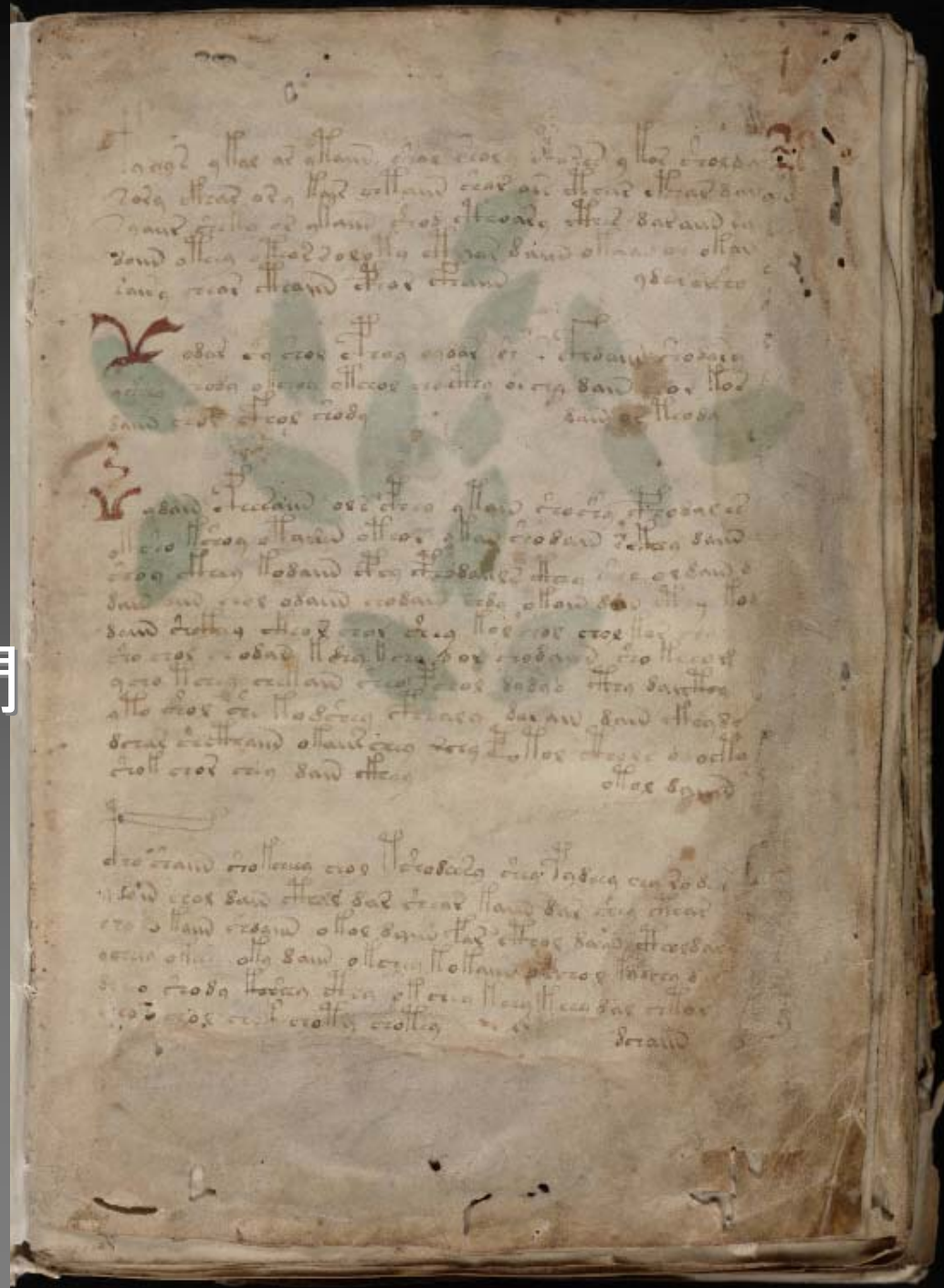
* 距離行列の一部

最初のページは他のページと共通する単語が全く出現していない

- ・全てのページ同士の組に対して類似度を算出し、類似度行列を作成
 - f. 1rには他のページに出現する語が全く出現しない
 - 他のページと性格が異なる
 - 後世の誰かが捏造し追加？
 - 暗号のキー？

folio.1r

- ・サインが消された跡
- ・ Jacobj `a Tepenece
=Jakub Horcicky of
Tepenec (薬草の専門
家)
- ・蔵書？作成者？
- ・サイン自体がねっ
造？



- ・ 類似度行列からセクションごとに平均を算出

セクション内外のページ同士の類似度の平均

	植物	天文	生物	十二宮図	薬草	レシピ
植物	7612.1	7715.7	7653.8	7666.0	7659.6	7686.3
天文	7715.7	7619.9	7638.9	7645.6	7670.2	7646.7
生物	7653.8	7638.9	6942.7	7302.9	7575.8	7242.6
十二宮図	7666.0	7645.6	7302.9	7350.7	7596.7	7346.2
薬草	7659.6	7670.2	7575.8	7596.7	7568.2	7593.4
レシピ	7686.3	7646.7	7242.6	7346.2	7593.4	7148.8



十二宮図以外は同じセクションのページ同士が異なるセクションよりも類似度が高い（距離が近い）⁴⁰

・ 類似度行列からクラスタ分析

- 代表的な手法

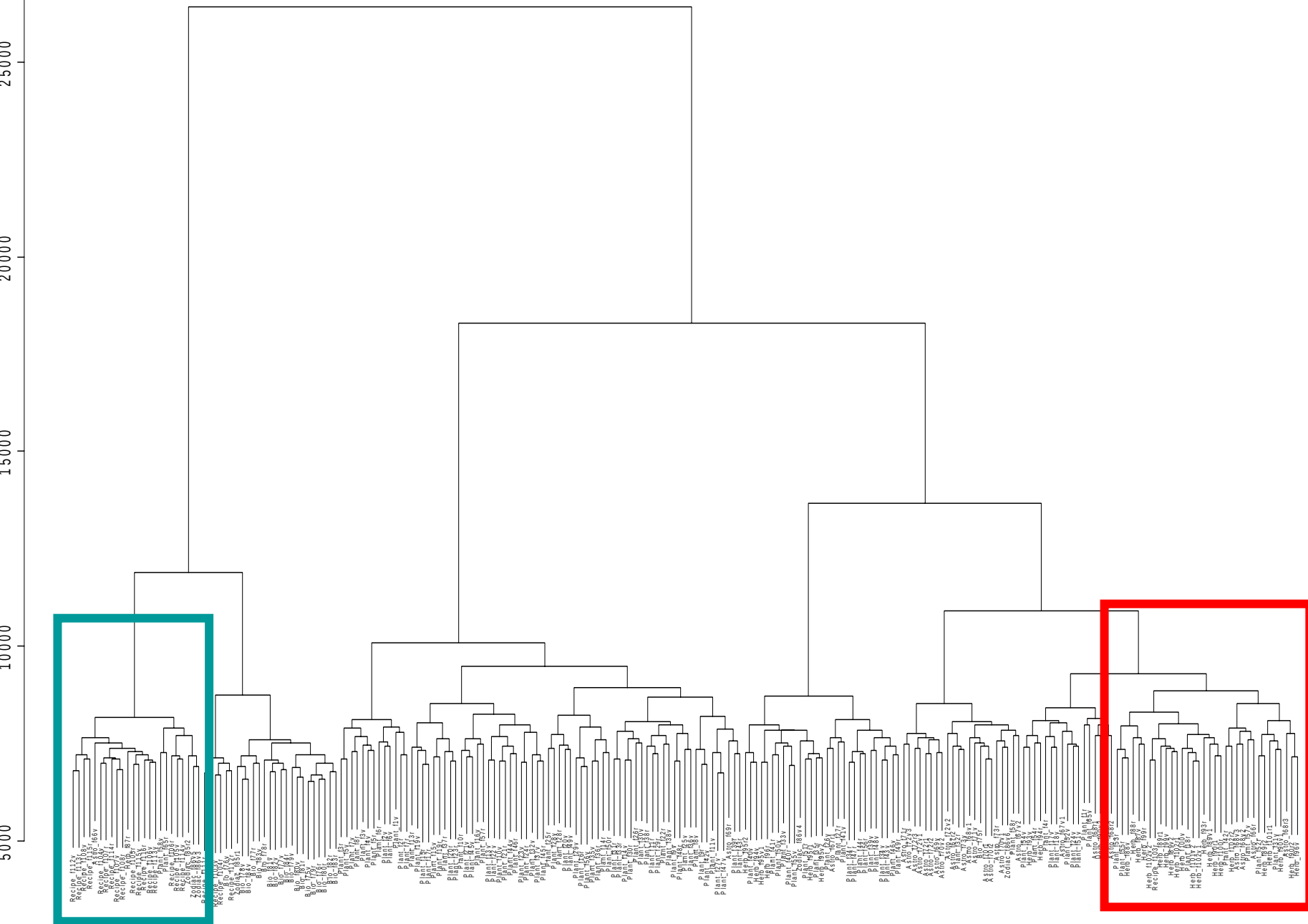
「最短距離法」

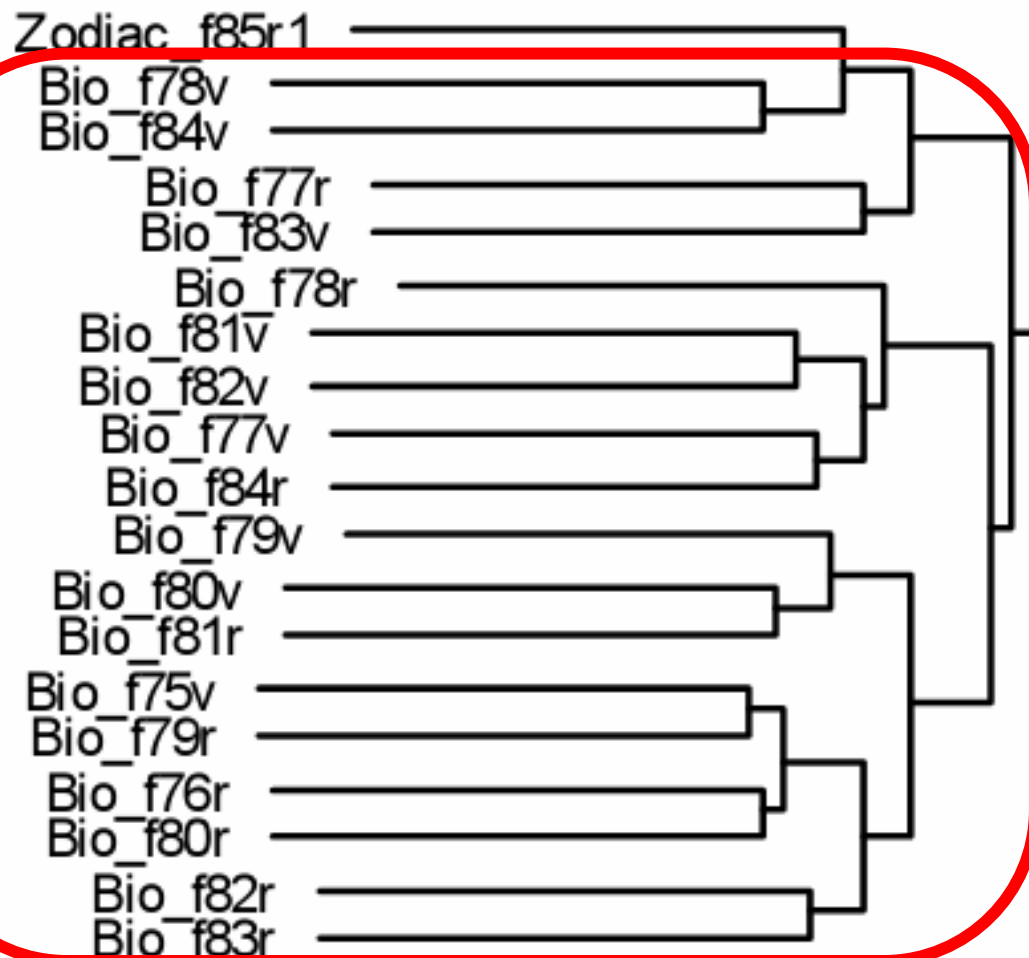
「最長距離法」

「群平均法」

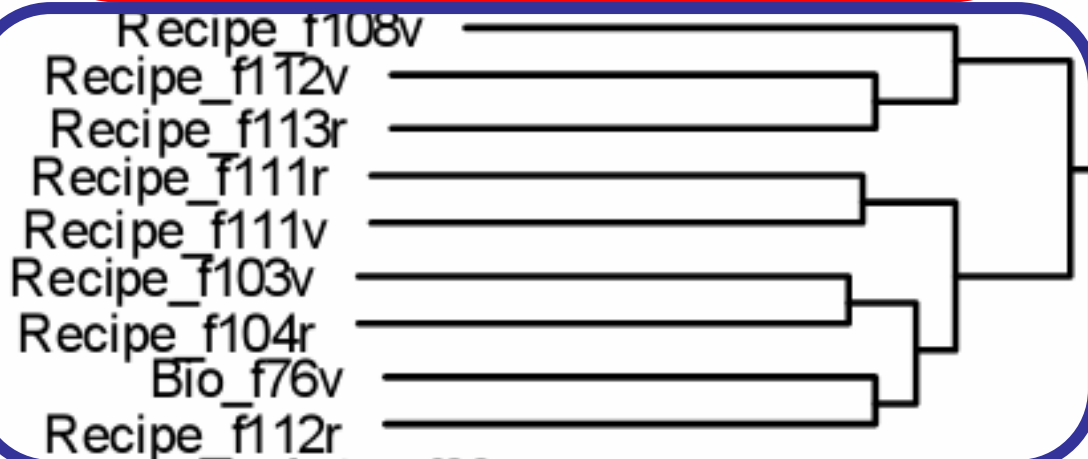
「ウォード法」

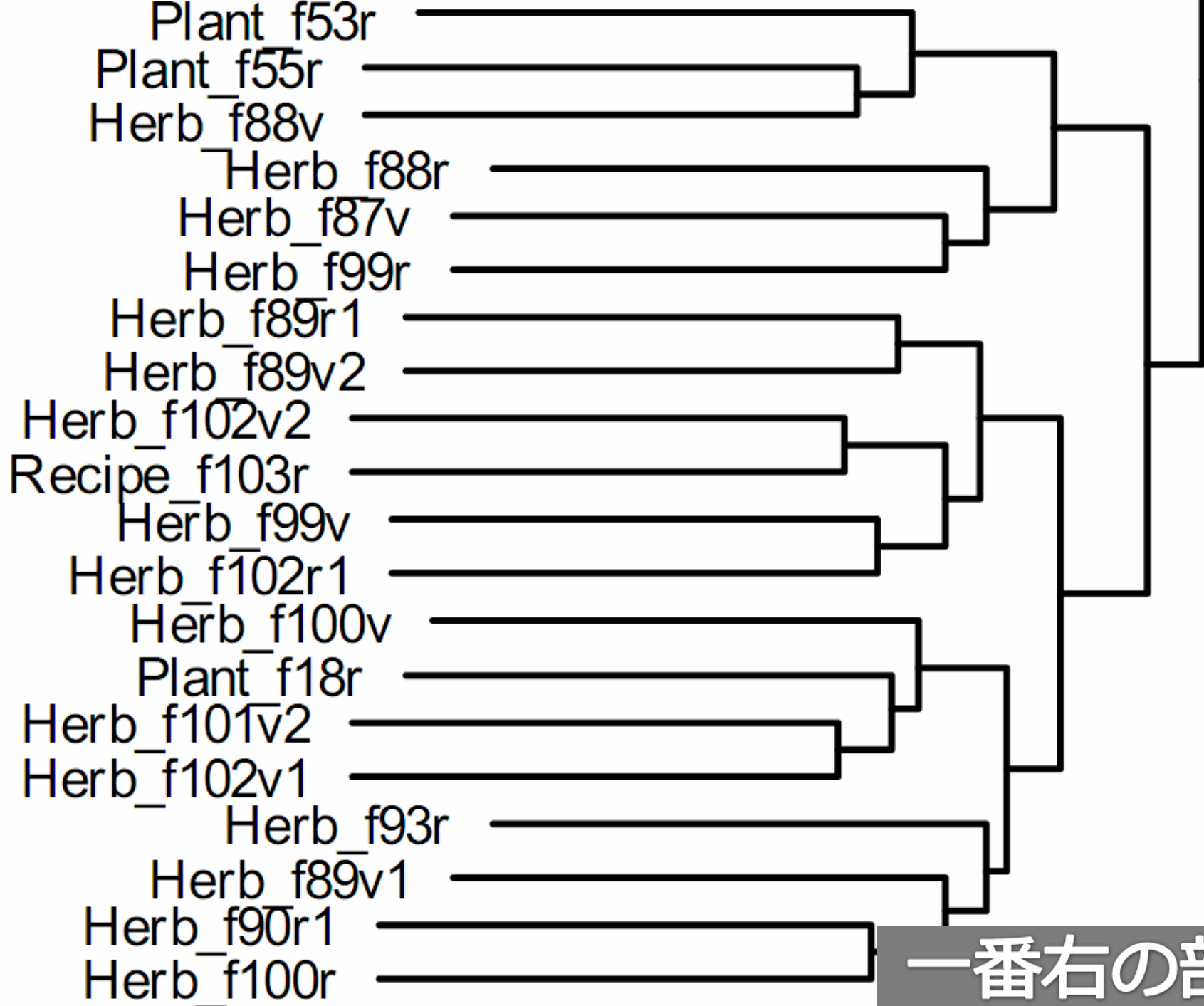
「ワード法」によるデンドログラム





図による
セクション
に対応し
ページ同士
がまとまる
(十二宮図以外)





一番右の部分

Voynich写本の 部分文書クラスタリング

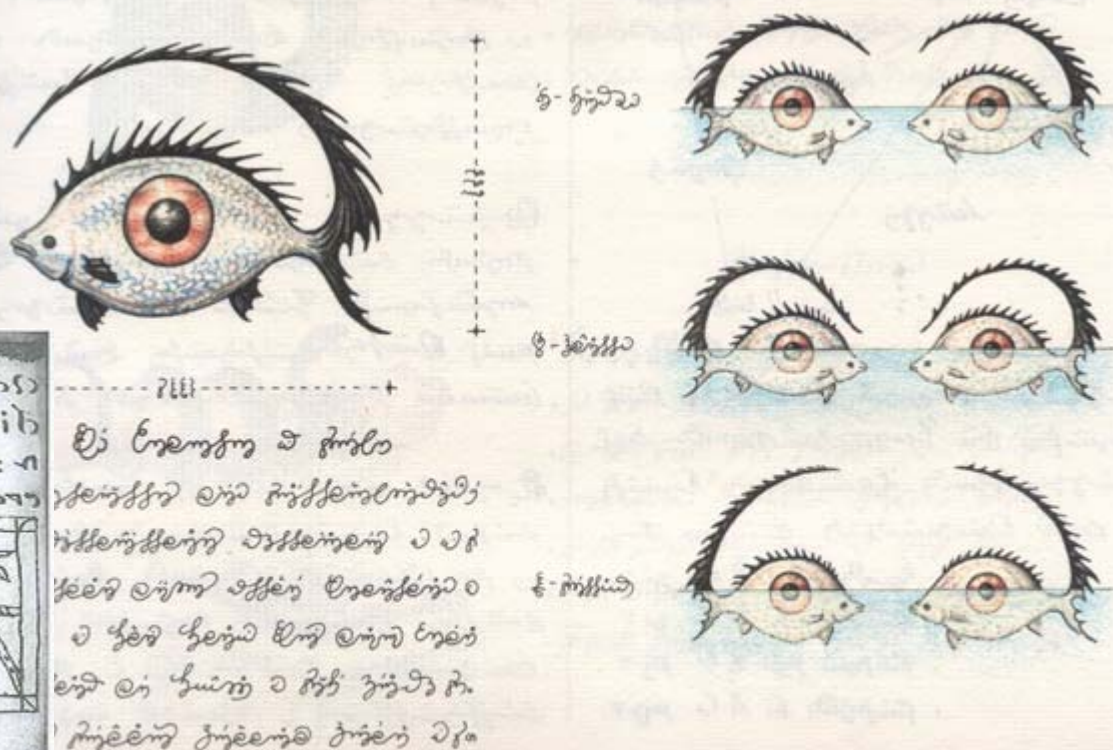
- ・ 図によるセクションに対応しページ同士がまとまる
 - カルダーノ・グリルでデタラメに作成された文書であれば図との対応関係はないはず
→ 捏造ではない
- ・ 暗号だとしても複雑な暗号（複式換字式）ではない → 暗号ではなく言語

まとめ

- ・ 仮説「未知の言語の文書に対しても既存のテキスト処理技術が応用可能」からの未解読文書の解読可能性の判定手法
- ・ Voynich写本に適用し、テキストから導き出した文書構造と挿図からのセクションの対応関係を示した
 - 図と対応した構造を持つ文書＝真正性が高い

今後の展開

- ・ クラスタリング結果の詳細な評価
- ・ Voynich写本と同様に解読不能文書
 - Luigi SerafiniによるCodex Seraphinianus
 - Rohonc Codex



未解読文書 の解読可能性の判定

－ 世界で最も謎に満ちた写本を例に －

安形輝（亜細亜大学）

[安形麻理（慶應義塾大学）]

2008／02／26 研究談話会

* 発表に使用したVoynich写本の画像は全てバイネッケ図書館がインターネット上で公開しているもの URL : <http://www.library.yale.edu/beinecke/>