The Voynich Manuscript: A New Perspective on a 600-Year-Old Mystery

Chris Kunferman

Go directly to the text of the paper

Abstract

Known to a great many as "The Holy Grail" of cryptology, MS-408, aka The Voynich Manuscript, is and has remained the same since its rediscovery: Unreadable. With no shortage of great minds and thinkers throughout history, all available means of decoding or translating it have been nearly exhausted, with little concrete evidence to support any current or past theory pertaining to its contents.

In this paper, the author approaches the document from a new perspective. The reverse processing of the images has led to the identifying of stroke patterns, and distinct characters that can be used in a number of ways for comparative research, and also computational analysis carried out by machine. The final alphabet to be presented has been derived using a deductive grid method. This grid method was created by a separation of characters in a three-fold manner, controlled by a common "imaginary" median line. The new set of characters will additionally be shown as a singular and combined alphabet, bearing resemblances to known natural languages, which the author feels confident will ultimately lead to unraveling a full translation of the text.

Le Manuscrit de Voynich : Une Nouvelle Approche pour un Mystère Datant de Six-cents Ans

Chris Kunferman

Résumé

Connu par un grand nombre comme le « Saint Graal » de la cryptologie, MS-408, à savoir le Manuscrit de Voynich, est, et est resté depuis sa redécouverte, identique à lui-même : illisible. Bien qu'un grand nombre de penseurs et de grands esprits s'y soient attelés, toutes les formes de décodage ou de traduction connues ont été pratiquement épuisées, avec peu de preuves concrètes à l'appui de toutes les théories, présentes ou passées, relatives à son contenu.

Dans cet article, l'auteur aborde ce manuscrit sous un angle nouveau. La numérisation inversée des images de ce manuscrit a conduit à l'identification de motifs de touches et de caractères uniques, qui peuvent être utilisés de nombreuses manières pour une recherche comparative, tout comme pour une analyse informatique. L'alphabet final qui sera présenté a été dérivé par le biais d'une méthode de « grille déductive ». Cette grille a été réalisée par une séparation des caractères d'une manière triple, contrôlée par une ligne médiane « imaginaire » commune. Le nouveau jeu de caractères apparaît en outre comme un alphabet singulier et combiné, ayant des ressemblances à des langues naturelles connues, de telle manière que l'auteur pense qu'il aboutira finalement à une traduction intégrale du texte.

El Manuscrito de Voynich: Una Nueva Perspectiva Sobre Un Misterio de 600 Años

Chris Kunferman

Resumen

Conocido por muchos como "El Santo Grial" de la criptología, MS-408, también conocido como el Manuscrito Voynich, es y ha permanecido igual desde su redescubrimiento: ilegible. Sin escasez de grandes mentes y pensadores a través de la historia, todos los medios disponibles para decodificarlos o traducirlos han sido casi agotados, con poca evidencia concreta para respaldar cualquier teoría actual o pasada relacionada con su contenido.

En este artículo, el autor aborda el documento desde una nueva perspectiva. El procesamiento inverso de las imágenes ha llevado a la identificación de patrones de trazo y caracteres distintos que se pueden usar de varias maneras para la investigación comparativa, y también al análisis computacional realizado por máquina. El alfabeto final que se presentará se ha derivado usando un método de cuadrícula deductiva. Este método de cuadrícula fue creado por una separación de caracteres de manera triple, controlada por una línea media común "imaginaria". Además, el nuevo conjunto de caracteres se mostrará como un alfabeto singular y combinado, que se asemejará a los lenguajes naturales conocidos, que el autor se siente confiado llevará a resolver una traducción completa del texto.

O Manuscrito Voynich: uma nova perspectiva sobre um mistério de 600 anos

Chris Kunferman

Sumário

Conhecido por muitos como "O Santo Graal" da criptologia, o MS-408, também conhecido como "O Manuscrito Voynich", é e continua sendo o mesmo desde a sua redescoberta: ilegível. Sem escassez de grandes mentes e pensadores ao longo da história, todos os meios disponíveis de decodificação ou tradução foram quase esgotados, com poucas evidências concretas para apoiar qualquer teoria atual ou passada pertencente ao seu conteúdo.

Neste artigo, o autor aborda o documento numa nova perspectiva. O processamento reverso das imagens levou à identificação de padrões de traçado e à caracteres distintos que podem ser usados de várias maneiras para pesquisa comparativa, e também para análise computacional realizada por máquina. O alfabeto final a ser apresentado foi derivado usando um método de grade dedutiva. Esse método de grade foi criado por uma separação de caracteres de uma maneira tripla, controlada por uma linha mediana comum "imaginária". O novo conjunto de caracteres será adicionalmente mostrado como um alfabeto singular e combinado, semelhante as linguagens naturais conhecidas, em que o autor se sente confiante que levará a desvendar uma tradução completa do texto.

Das Voynich-Manuskript: ein 600 Jahre altes geheimverschlüsseltes Dokument

Chris Kunferman

Zusammenfassung

Vielen ist das Voynich -Manuskript alias Schrift Nummer MS-408 als der Heilige Gral der Kryptologie bekannt. Dieser Schrift ist auch seit seiner Wiederentdeckung unverändert bzw. unlesbar geblieben.

Viele große Geister und Denker versuchten das Manuskript zu entschlüsseln; alle zur Verfügung stehenden Dekodierungen oder Übersetzungen wurden nahezu ausgeschöpft. Jeder brauchbare Hinweis, der eine aktuelle oder vergangene Theorie unterstützen könnte, wurde geprüft.

In dieser Abhandlung behandelt der Verfasser das Dokument aus einem neuen Blickwinkel. Eine andere Methode, indem man die Abbildungen spiegelbildlich liest, ergab, dass man gewisse Muster und bestimmte Zeichen identifizieren konnte. Diese sollten für verschiedene vergleichende Untersuchungen und Computer gesteuerte Analysen geeignet sein. Das endgültige Alphabet, das vorgestellt wird, ergab sich aus einer Methode mit deduktivem Raster. Dieses Raster wurde durch Trennung der Zeichen anhand einer dreifachen Methode, gesteuert durch einen gemeinsamen "imaginären" Mittelstreifen, erhalten. Der neue Buchstabensatz ist zusätzlich präsentiert als ein Alphabet mit eigenständigen Zeichen und Ligaturen mehrerer Zeichen, die eine Ähnlichkeit mit bekannten natürlichen Sprachen zeigen. Der Autor ist überzeugt, hiermit eine Lösung gefunden zu haben, die letztendlich zu einer kompletten Übersetzung des Textes führen könnte.

Introduction

Throughout its history, the art and study of deciphering forgotten languages has been of great interest to Rosicrucians, philosophers, historians, and many other disciplines. To once again read and understand what was lost and forgotten opens our perception to times and places that could only be speculated on previously.

Without key works such as the Rosetta Stone, the land of ancient Egypt would still be shrouded in deep mystery. Without it, we would still be struggling helplessly to understand what the hieroglyphic images left behind had to say.

MS-408 remains to this day without such a key to deciphering. Much like the hieroglyphic images once did, the script and images call out to the observer, and the desire to understand exactly what the writer had to say sparks the imagination and ignites a passion to embark on a quest to find the elusive stone or key to the mystery within.

It has held such a great fascination for generations that the limitations of literary analysis and cryptology have been nearly exhausted, while additional areas of science and technology have provided but a few concrete answers.

With its own language referred to as "Voynichese,"¹, it has been alphabetized numerous times, but a large debate still remains as to the validity of each attempt, as a singular character alphabet has yet to be presented.

The longest working hypothesis that has been a flagship in the quest to decipher MS-408 is that it contains a "letter-based cipher." This belief was wildly popular among "code breakers" and "cryptologists," both professional and amateur. During World War I and also World War II, both British and American cryptographers utilized leisure time to focus on it in an effort to discover the algorithm and key.² World renowned cryptographers, NSA cryptographers, and countless scores of scholars and research groups have tackled the manuscript, yet still it remains a mystery. While they excitedly began to decode what appeared simple and systematic,^{3,4} armed with examples and fully decoded ciphers from the past, it eventually became apparent that if this was a cipher, it played by rules that only the original writer could fully comprehend.⁵

With the lack of any known existing characters for comparison, the language itself has also been looked at as an invention or creative endeavor made for profit. Critics of this argue that there appears to be far too much structure and detail for it to be a language of invention,⁶ and the author is inclined to agree with this assessment in light of the new findings.

1. Methods

A. Gaining the new perspective

Initial investigation towards a new perspective in which to interpret MS-408 began with the experimentation of period-appropriate technology that could be used to manifest the document. When viewing the manuscript in digital format,⁷ a clearly visible feature is the slight transparency of the vellum. In the bulk of the manuscript, what is written on the back of the page is often visible. Using this as evidence that the document was plausibly traced or copied, the use of mirrors and early other forms of projection were explored. With a proper method of tracing, and a feasible device determined, the author could then begin reverse engineering the same process digitally, and in theory reproduce the original document before it was encoded, obscured, and rendered unreadable.

The most plausible method derived was the use of the "camera obscura" device, which could be used to project an original document onto the back of the vellum, providing the writer with an easily traceable image that was both backwards and upside down.



Figure 1: Artist rendition of alternate views depicting MS-408 process of encoding using a camera obscura device.

Figure 1 displays the following: (A) Front of device. (B) Back of device with stretched vellum displaying obscured manuscript for tracing from view of the author. (C)Seating position of author to perform tracing. (D) Pinhole entrance where light travels through and is projected onto tracing material. (E) Position of original document, reflecting a light source that will be obscured by the pinhole.

To reproduce the original document for analysis, this process was then reversed using digital photography editing software. [Figure 2]

This reverse imaging process was then applied to each page of the digitized document scans provided by the Beinecke Rare Book Library so that the entire document could be viewed and analyzed from the new proposed perspective.



Figure 2: Reverse camera obscura digital editing process: Digitally scanned images of the manuscript were first reflected vertically along the X-axis and then horizontally reflected along the Y-axis.

B. Structural analysis of writing formation

Paragraphs, words, and sentences were analyzed to deduce a more readily recognizable formation of text flow commonly found in other manuscripts. Spacing, indentation, titling, labeling, and embellishments were all taken under consideration.

C. Writing difficulty

The directional plotting of pen strokes to produce the original characters was carried out first by manually transcribing paragraphs and sentences by hand. This was done in order to see if the writing style still maintained an extraneous difficulty, or if the hand movements involved in writing decreased in difficulty.

D. Digital reproduction of characters

After determining that there was a large decrease in handwriting difficulty, many isolated or easily discernible characters were digitally recreated using vector paths, based mainly on the analysis of manual stroke direction.

The type of visual assistance seen in Figure 3 is used widely as a means to educate persons in writing in languages they are unfamiliar with and can be produced at very early ages in learning institutions to teach basic handwriting.⁸



Figure 3: Proposed stroke path of original manuscript.

With the visual assistance and directional plotting complete, paths could then be produced in vector drawing software, placing points along the directional paths. [Figure 4]



Figure 4: Stroke path vector reconstruction.

E. Sample Comparison

In order to narrow down a large portion of characters, folios of the manuscript were selected that contained a large sample of characters for comparison to the stroke paths.

To confirm that the stroke paths observed produced an accurate representation of the text, a brush style was applied to the stroke paths in order to simulate the pen or quill characteristics and traits that commonly produce calligraphic style writing. [Figure 5]



Figure 5: Stroke path vector reconstruction with calligraphy brush style

F. Reducing Character Glyph

Each character was assessed using the stroke pattern to deduce if recurring strokes were present within some of the more complex characters.

In order to establish a clear distinction between glyphs, a grid was used to reduce the proposed glyph set into three categories, based on the stroke path's relation to the imaginary median line (which the author will herein refer to as the "IML").

Singular pen strokes used to construct glyphs were found in the following three categories of separation:

- Upper: Glyphs that rest above the IML.
- Lower: Glyphs that rest mostly below the IML.
- Full: Glyphs that extend above and below the IML.

With the single stroke characters discovered, separate grids were used in order to produce combinations of pen strokes, providing a lengthy set of glyphs for deductive comparison. The author will herein refer to the overall method as the "MLG" or "Median Line Grid."

Lower

No multiple stroke combinations were found for glyphs below the IML; thus, they formed the entirety of the lower glyph set.

MLG Table 1: Upper

All single stroke glyphs that rest above the IML were duplicated and then placed across the top (X) as well as along the left column (Y) of the grid. By combining pen strokes into the corresponding cells, an extended upper glyph set is created.

MLG Table 2: Full

Results from MLG 1 were then placed into MLG 2 as X, along with the initial "Full" category glyphs that were deduced prior to the MLG as single stroke glyphs without further combination. "Lower" category glyph results placed as Y.

Results

1. General appearance after reversing the camera obscura process:

After applying the reverse imaging process to the entire document, many features of common writing structure and formatting become visible. Certain traits in terms of writing have become clearer, and pen strokes have become easier to decipher.

The following traits can now be observed:

A. The language now appears much less extraneous and easier on the eyes of the reader. **B.** Paragraph forms, sentence structures, embellishments, and labels are now placed in different order and alignment.

C. Indentations and paragraph lead-ins have taken on new positioning.

D. Pen strokes have become more pronounced.

E. Naturally occurring differences and imperfections in pen strokes are now clearly visible, making the manuscript easier to decipher in terms of a recurring alphabet.

Guideline Limitations

A limitation evident from experimentation in duplicating the original was that to avoid running out of space and severely overlapping characters without a guideline becomes increasingly difficult and time consuming. Without any evidence of this guideline, it would in theory require the author to make very specific measurements in order to bring the text to follow an embellished character flush with the extremities as exhibited in the original scans. This limitation is lifted by a downward arching observable in the sentence formation, as it now presents these characters being written in a way that avoids characters already present before and above each line, paragraph, and label. All the characters that originally rose above the imaginary median line now fall below the line, denoting that the character set is largely comparable to a "small" or "lower case" alphabet.

Further experimentation in hand transcription has shown that to avoid the horizontal collision of characters due to space, the language is most likely written from right to left. Evidence of this can be seen by the vertical alignment of sentences along the right side of pages in contrast to the staggered appearance of words on the left side of pages.

Embellishments

Embellishments that were previously seen as at the beginning of pages are now seen at the end, providing characteristic punctuation at the end of pages and/or paragraphs, the absence of which would normally denote that the information and context are continued on the next page.

Occasionally these embellishments now appear to emphasize words or small phrases. On other occasions they take on the appearance of a signature.

Glyph analysis:

Singular characters without combination of strokes were found in all three categories of the IMLG. The following glyphs can be used in combination to manifest all of the remaining characters found in the document.

Full Category Glyphs

Stroke Recreations *The IML has been placed as a dashed line for visual reference.*

c s c x z (x

Vector Paths with calligraphy brush shape.



Upper Category Glyphs

Stroke Recreations

The IML has been placed as a dashed line for visual reference.

 $6 \rightarrow 0 \rightarrow 0 \rightarrow 0 \rightarrow 0$

Vector Paths with calligraphy style brush shape

_, 0

Lower Category Glyphs

Stroke Recreations

The IML has been placed as a dashed line for visual reference.



Vector Paths with calligraphy style brush shape



MLG Results

MLG Table 1: Upper Category Combination Glyphs

	- 1 	0	Ţ.	1	6
- 1 +5- +	- 1 - 1 				
)	- 1 - 5	Q>	<u>0</u> >	1>	<u>6</u> >
0	2	00	2.0	<i>L.</i> .0	6.0
Ŋ.	:	09	<u>99</u>	<u>1</u> 9	<u>(a. 1</u> 2
1	21	0.1	<u>2/</u>	11	6/
6		a6	2.6	1.6	66

MLG Table 2: Full Combination Glyphs

Some distinct areas of combinations resulting from the use of the grids do not appear in the manuscript and can be eliminated from the final alphabet set. Spaces in the following grid have been intentionally left blank.



This method has produced a singular alphabet that can be quite easily assigned to standard language fonts and keyboards for word processing and transcription.

Entire Glyph Set

	-	* •	• , 6	2 8 2	- 8 4
6 x 3	◦	6 2		26 6. 60	م 4 م
~ ~ * *	۴ ۲ ۴ ۴	# #	ᠼ	ม (ม	ૠ ૠ ભ
6_{-0} 2_{-0} 2_{-} 2_{-}	°₹ 6 ₀ ° <u>~</u> °				

Using the stroke patterns presented as guides, a transcription by hand becomes easy after a small amount of practice, adding further evidence to the idea that a natural language is present. Scalable vector graphics of the results will be made available at the author's website upon request.⁹

Discussion

Other Content

There still remains a large portion of mysteries in addition to the alphabet and language. It should be noted that the illustrations have also changed in perspective, and now bear more resemblance to biological depictions rather than botanical depictions, and should also be revisited for interpretation.

Conclusion

This singular and combined alphabet has shown promising leads, but will require additional papers and an expert in languages of the time period for an accurate translation.

Armed with the new perspective and characters that have resulted, along with a solid history of prior research that can be reapplied, a full unraveling of the mystery within MS-408 seems within sight. The author looks forward to seeing the progression of this new perspective and a new quest into this amazing manuscript.

Endnotes

¹ Jim Reeds, "William F. Friedman's Transcription of the Voynich Manuscript" (PDF), Accessed January 4, 2019. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.7619.

² Melissa Hogenboom, "Mysterious Voynich Manuscript has genuine message," Accessed January 4, 2019. https://www.bbc.com/news/science-environment-22975809.

³ René Zandbergen, "The Voynich Manuscript – Text Analysis," Accessed January 4, 2019. http://www.voynich.nu/analysis.html.

⁴ "Languedoc Topics - Mysteries of the Languedoc - The Voynich Manuscript," Accessed January 4, 2019. http://www.languedocmysteries.info/voynich.htm.

⁵ Nick Pelling (August 27, 2009). "Cipher Mysteries - Voynich Cipher Structure," Accessed January 4, 2019. https://ciphermysteries.com/2009/08/27/voynich-cipher-structure.

⁶ Gerry Kennedy and Rob Churchill, *The Voynich Manuscript: The Mysterious Code That Has Defied Interpretation for Centuries*, (Vermont: Inner Traditions, 2004).

⁷ "Cipher Manuscript," Accessed January 4, 2019.

https://brbl-dl.library.yale.edu/vufind/Record/3519597.

⁸ Pam Versfeld, "Teaching Handwriting – A Stroke Based Approach," Accessed January 4, 2019. https://skillsforaction.com/handwriting/stroke-based-approach.

⁹ "The Voynich Manuscript – Reopened," Accessed January 4, 2019.

http://voynich.egdflabs.com.