

iAligner : a tool for syntax-based intra-language alignment

Tariq Yousef & Chiara Palladino
University of Leipzig
Classical Philology Goes Digital
Potsdam
February 17, 2017

What is text alignment?

- Text alignment is the comparison of two or more parallel texts
- It tries to define correspondences/similarities and divergences/variants
- One of the most important tasks in Natural Language Processing: it can be performed automatically through algorithmic and dynamic programming methods

W1	This morning	the cat	observed	little	birds in the	trees	.		
W2		The cat	was observing		birds in the	little	trees	this morning, it observed birds for two hours	.

Intra-Language alignment: alignment of texts in the same language

iAligner [Home](#) [Instructions](#) [Bibliography](#) [Contact](#)

Alignment options: **Ignore Nonalphabetical** | **Case sensitive** | **Ignore Diacritics** | **Levenshtein Distance**

- 1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.
- 1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

Length: **21** | **Aligned-complete 21**

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

The longest common substring

- 2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judah and his brethren;
- 2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judas and his brethren;

2 Abraham begat Isaac ; and Isaac begat Jacob ; and Jacob begat Judah and his brethren ;

2 Abraham begat Isaac ; and Isaac begat Jacob ; and Jacob begat Judas and his brethren ;

Length: **19** | **Aligned-complete 18** | **notAligned 1**

2 Abraham begat Isaac ; and Isaac begat Jacob ; and Jacob begat Judah and his brethren ;

The longest common substring

2 Abraham begat Isaac ; and Isaac begat Jacob ; and Jacob begat

2 Abraham begat Isaac ; and Isaac begat Jacob ; and Jacob begat

- 3 and Judah begat Perez and Zerah of Tamar ; and Perez begat Hezron ; and Hezron begat Ram ;
- 3 And Judas begat Phares and Zara of Thamar; and Phares begat Esrom; and Esrom begat Aram;

3 and Judah begat Perez and Zerah of Tamar ; and Perez begat Hezron ; and Hezron begat Ram ;

3 And Judas begat Phares and Zara of Thamar ; and Phares begat Esrom ; and Esrom begat Aram ;

Length: **21** | **Aligned-complete 12** | **Aligned-case 1** | **notAligned 8**

3 and Judah begat Perez and Zerah of Tamar ; and Perez begat Hezron ; and Hezron begat Ram ;

The longest common substring

3 and

3 And

- 4 and Ram begat Amminadab; and Amminadab begat Nahshon; and Nahshon begat Salmon;
- 4 And Aram begat Aminadab; and Aminadab begat Naasson; and Naasson begat Salmon;

4 and Ram begat Amminadab ; and Amminadab begat Nahshon ; and Nahshon begat Salmon ;

4 And Aram begat Aminadab ; and Aminadab begat Naasson ; and Naasson begat Salmon ;

Length: **17** | **Aligned-complete 11** | **Aligned-case 1** | **notAligned 6**

4 and Ram begat Amminadab ; and Amminadab begat Nahshon ; and Nahshon begat Salmon ;

The longest common substring

4 and

4 And

Cross-language alignment: alignment of texts in different languages

- Cross-language alignment is difficult to perform automatically
- It still needs training data from manual alignment

HYE

Յաղագս թէ որպէս կայսրն Վաղէս հալածեալ զամենայն ուղղափառ հաւատացեալս ,և կամ որպէս կամեցաւ յանդիմանութիւն առնել կայսրն ի մէջ ճշմարտութեանն հաւատացելոց ,և ի մէջ ծախող աղանդոյն արիանոսաց .

(29) 100% HYE - DEU

DEU

Darüber ,wie der Kaiser Valēs alle rechtläubigen Gläubigen verfolgte ,und noch wie der Kaiser beabsichtigte ,ein Streitgespräch unter den Gläubigen der Wahrheit und unter der bösen Sekte der Arianer zu machen .

(35) 100% HYE - DEU

Armenian-German aligned text, Alignment is done manually by Uta Koschmieder

CollateX (<http://collatex.net/>)

- A software for expressing textual variance with a graph-based data model
- It offers a preprocessing stage (tokenization, normalization, regularization) that is configurable by the user according to issues in language, orthography, punctuation
- Uses graphs and tables as graphic outputs

CollateX – Software for Collating Textual Sources

"In a language, in the system of language, there are only differences." – Jacques Derrida

» Home » Download » Documentation » Demo » About

CollateX is a software to

1. read **multiple** (≥ 2) **versions** of a text, splitting each version into parts (**tokens**) to be compared,
2. **identify similarities and differences between the versions** (including moved/transposed segments) by aligning tokens, and
3. output the alignment results in a **variety of formats for further processing**, for instance
4. to support the **production of a critical apparatus** or the stemmatic analysis of a text's genesis.

It resembles software used to compute differences between files (e.g. **diff**) or tools for **sequence alignment** which are commonly used in Bioinformatics. While CollateX shares some of the techniques and algorithms with those tools, it mainly aims for a flexible and configurable approach to the problem of finding similarities and differences in texts, sometimes trading computational soundness or complexity for the user's ability to influence results.

As such it is primarily designed for use cases in disciplines like **Philology** or – more specifically – the field of **Textual Criticism** where the assessment of findings is based on interpretation and therefore can be supported by computational means but is not necessarily computable.

Latest Version

» **collatex-tools-1.7.1.jar**

Please refer to the **documentation** for detailed information about CollateX like its underlying concepts or usage instructions.

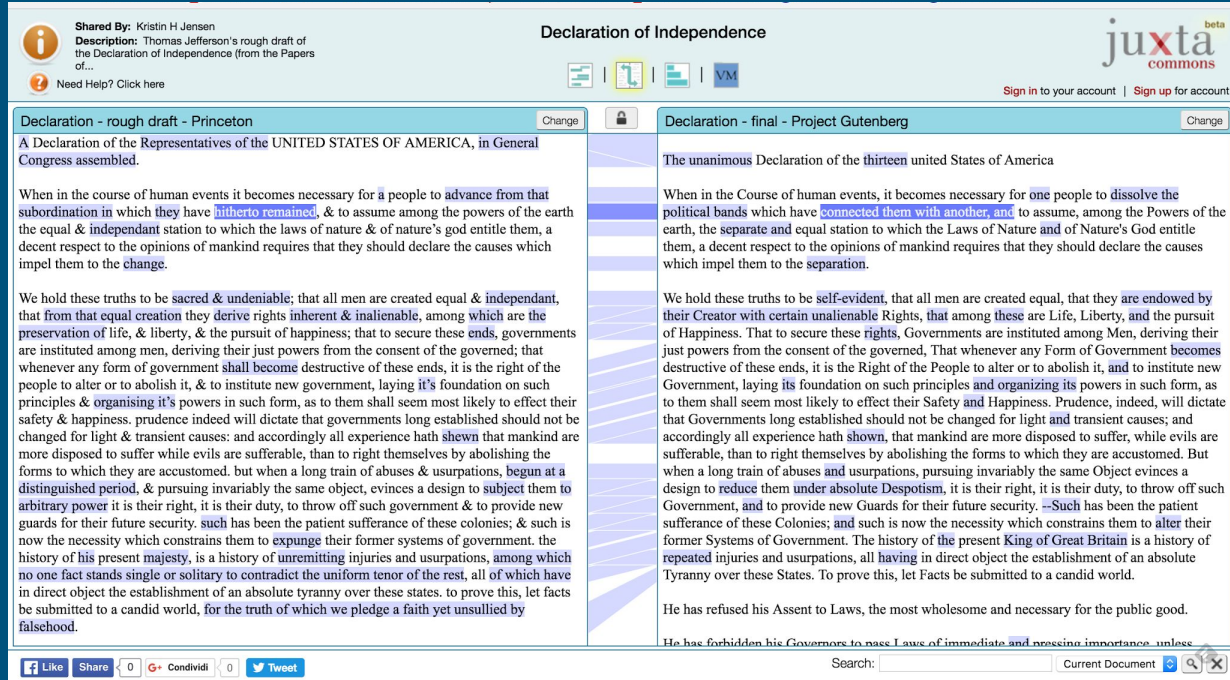
For alternative packages and license terms, please read the **download section**.

Snippet of a Variant Graph produced by CollateX

Copyright © 2010-2013 **The Interedition Development Group**. All rights reserved.

Juxta Commons (<http://juxtacommons.org/>)

- A group of Softwares for performing various alignment tasks
- Enables users to see changes in texts up to 20 versions
- Provides three visualization outputs: overlapping versions, parallel panels, bar histograms



The screenshot displays the Juxta Commons web interface for comparing two versions of the Declaration of Independence. The interface is split into two main panels: "Declaration - rough draft - Princeton" on the left and "Declaration - final - Project Gutenberg" on the right. Each panel shows the text of the document with highlighted changes. The left panel shows the rough draft with changes like "hitherto remained" and "showed". The right panel shows the final version with changes like "connected them with another, and" and "shown". The interface includes a top navigation bar with the Juxta Commons logo, a search bar, and a "Need Help? Click here" link. The bottom of the interface features social media sharing buttons (Facebook, Twitter, etc.) and a search bar.

Shared By: Kristin H Jensen
Description: Thomas Jefferson's rough draft of the Declaration of Independence (from the Papers of...)

Declaration of Independence

Declaration - rough draft - Princeton

A Declaration of the Representatives of the UNITED STATES OF AMERICA, in General Congress assembled.

When in the course of human events it becomes necessary for a people to advance from that subordination in which they have hitherto remained, & to assume among the powers of the earth the equal & independant station to which the laws of nature & of nature's god entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the change.

We hold these truths to be sacred & undeniable; that all men are created equal & independant, that from that equal creation they derive rights inherent & inalienable, among which are the preservation of life, & liberty, & the pursuit of happiness; that to secure these ends, governments are instituted among men, deriving their just powers from the consent of the governed; that whenever any form of government shall become destructive of these ends, it is the right of the people to alter or to abolish it, & to institute new government, laying it's foundation on such principles & organising it's powers in such form, as to them shall seem most likely to effect their safety & happiness. prudence indeed will dictate that governments long established should not be changed for light & transient causes; and accordingly all experience hath shewn that mankind are more disposed to suffer while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed. but when a long train of abuses & usurpations, begun at a distinguished period, & pursuing invariably the same object, evinces a design to subject them to arbitrary power it is their right, it is their duty, to throw off such government & to provide new guards for their future security. such has been the patient sufferance of these colonies; & such is now the necessity which constrains them to expunge their former systems of government. the history of his present majesty, is a history of unremitting injuries and usurpations, among which no one fact stands single or solitary to contradict the uniform tenor of the rest, all of which have in direct object the establishment of an absolute tyranny over these states. to prove this, let facts be submitted to a candid world, for the truth of which we pledge a faith yet unswilled by falsehood.

Declaration - final - Project Gutenberg

The unanimous Declaration of the thirteen united States of America

When in the Course of human events, it becomes necessary for one people to dissolve the political bands which have connected them with another, and to assume, among the Powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the pursuit of Happiness. That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed, That whenever any Form of Government becomes destructive of these ends, it is the Right of the People to alter or to abolish it, and to institute new Government, laying its foundation on such principles and organizing its powers in such form, as to them shall seem most likely to effect their Safety and Happiness. Prudence, indeed, will dictate that Governments long established should not be changed for light and transient causes; and accordingly all experience hath shown, that mankind are more disposed to suffer, while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed. But when a long train of abuses and usurpations, pursuing invariably the same Object evinces a design to reduce them under absolute Despotism, it is their right, it is their duty, to throw off such Government, and to provide new Guards for their future security. --Such has been the patient sufferance of these Colonies; and such is now the necessity which constrains them to alter their former Systems of Government. The history of the present King of Great Britain is a history of repeated injuries and usurpations, all having in direct object the establishment of an absolute Tyranny over these States. To prove this, let Facts be submitted to a candid world.

He has refused his Assent to Laws, the most wholesome and necessary for the public good.

He has forbidden his Governors to pass Laws of immediate and pressing importance, unless

Search: Current Document

The Versioning Machine

(<http://v-machine.org/>)

- The comparison is visualized on text panels
- Allows comparison of diplomatic transcriptions or editions of the same text
- Provides a user space for annotations and bibliographical references, and an image viewer for manuscripts

The screenshot displays the 'Orchard Farming' interface within the Versioning Machine. The top navigation bar includes a logo, the title 'Orchard Farming', and tabs for '4 Total Versions', 'Line numbers', 'Bibliographic panel', and 'Critical introduction'. The main content area is divided into four panels:

- Bibliographic Information:** Contains the title 'Orchard Farming' by Baroness Elsa von Freytag-Loringhoven, the 'Original Source' (Papers of Elsa von Freytag-Loringhoven), a 'Witness List' (First, Second, Third, Fourth), and 'Textual Notes' regarding the poem's drafts.
- Version va1: First:** Displays a list of 24 lines of text, with some words highlighted in yellow (e.g., 'TO', 'MOURNERS', 'SNOTTOY').
- Version va2: Second:** Displays a list of 13 lines of text, with some words highlighted in yellow (e.g., 'TO:', 'MOURNERS', 'SNOTTOY').
- Images/litnss-003275-0001.jpg:** A thumbnail image of a manuscript page with handwritten text.

At the bottom, there is a status bar indicating 'In attesa di risposta da v-machine.org...'.

eComparatio (www.eComparatio.net)

- An ongoing project still at the experimental stage
- Will allow comparison of an arbitrary number of texts in Classical languages
- The visualization output will be in different panels with variations highlighted
- Aims at the creation of an apparatus of variants for born-digital critical editions

Menü: Textvervollständigung eComparatio Alles Keilschrift Schreiben > Hilfe / ? 11

Breadcrumbs: Sie sind im Menü "Fragmenttool / eComparatio / gleiche Editionen (Arbeitsfiter)"

Serverauslastung: Prozessorlast (1, 5, 15 Min): 0,00, 0,01, 0,05 %

[Anaximanderb1](#)
[IosvonHalikarnassosHistoriaeRomanae](#)
[Gilgamesh](#)
[LincolnGettysburg](#)
[AntiquitatesRomanaebook1](#)
[Livius](#)
[Hipparchus](#)
[demotu](#)
[++](#)
[URN](#)

[Asulanus Franciscus \(Hrsg.\)](#)
[Diels Hermannus](#)
[Ritter H. Preller L.](#)
[Mansfeld Jaap](#)
[Fortenbaugh William W.](#)
[Huby Pamela M.](#)
[Sharpley Robert W.](#)
[Gutas Dimitri](#)

[Kirk Geoffrey S.](#)
[Raven John E.](#)
[Schofield Malcolm](#)
[Kirk](#)
[Graham Daniel W.](#)
[Woehle Georg \(Hrsg.\)](#)


[Paralell-Darstellung](#)
[Detail-Darstellung](#)
[Buch-Darstellung](#)
[Bild-Darstellung](#)
[Exportieren](#)

Asulanus. Franciscus. (Hrsg.); Venetiis 1526	Diels. Hermannus; Berlin 1903	Ritter. H. . Preller. L.; Gotha 1934	M
0 τῶν δὲ ἐν καὶ κινούμενον καὶ ἄπειρον	0 τῶν δὲ ἐν καὶ κινούμενον καὶ ἄπειρον	0	0
1 λεγόντων ἀναξίμανδρος μὲν πραξιάδου	1 λεγόντων ἀναξίμανδρος μὲν πραξιάδου	1	1
2 μιλῆσιος θαλοῦ γενόμενος διάδοχος καὶ	2 μιλῆσιος θαλοῦ γενόμενος διάδοχος καὶ	2	2
3 μαθητῆς ἀρχὴν τε καὶ στοιχείων εἴρηκε	3 μαθητῆς ἀρχὴν τε καὶ στοιχείων εἴρηκε	3 ἀρχὴν τε καὶ στοιχείων εἴρηκε	3
4 τῶν ὄντων τὸ ἄπειρον, πρῶτος τοῦτο	4 τῶν ὄντων τὸ ἄπειρον, πρῶτος τοῦτο	4 τῶν ὄντων τὸ ἄπειρον, πρῶτος τοῦτο	4
5 τοῦνομα κομίσας τῆς ἀρχῆς. λέγει δ' αὐτὴν	5 τοῦνομα κομίσας τῆς ἀρχῆς. λέγει δ' αὐτὴν	5 τοῦνομα κομίσας τῆς ἀρχῆς. λέγει δ' αὐτὴν	5
6 μήτε ὕδωρ μήτε ἄλλο τῶν καλουμένων	6 μήτε ὕδωρ μήτε ἄλλο τῶν καλουμένων	6 μήτε ὕδωρ μήτε ἄλλο τῶν καλουμένων	6
7 εἶναι στοιχείων, ἀλλ' ἑτέραν τινὰ φύσιν	7 εἶναι στοιχείων, ἀλλ' ἑτέραν τινὰ φύσιν	7 εἶναι στοιχείων, ἀλλ' ἑτέραν τινὰ φύσιν	7
8 ἄπειρον, ἐξ ἧς ἅπαντας γίνεσθαι τοὺς	8 ἄπειρον, ἐξ ἧς ἅπαντας γίνεσθαι τοὺς	8 ἄπειρον, ἐξ ἧς ἅπαντας γίνεσθαι τοὺς	8
9 οὐρανούς καὶ τοὺς ἐν αὐτοῖς κόσμους ἐξ	9 οὐρανούς καὶ τοὺς ἐν αὐτοῖς κόσμους ἐξ	9 οὐρανούς καὶ τοὺς ἐν αὐτοῖς κόσμους ἐξ	9
10 ὧν δὲ ἡ γένεσις ἐστὶ τοῖς οὐσι, καὶ τὴν	10 ὧν δὲ ἡ γένεσις ἐστὶ τοῖς οὐσι, καὶ τὴν	10 ὧν δὲ ἡ γένεσις ἐστὶ τοῖς οὐσι, καὶ τὴν	10
11 φθορὰν εἰς ταῦτα γίνεσθαι κατὰ τὸ χρεών.	11 φθορὰν εἰς ταῦτα γίνεσθαι κατὰ τὸ χρεών.	11 φθορὰν εἰς ταῦτα γίνεσθαι κατὰ τὸ χρεών.	11

iAligner

www.ialigner.com

https://github.com/OpenGreekAndLatin/ILA_python

 **iAligner**

[Home](#) [Get Started](#) [Bibliography](#) [Contact](#)

The tool performs automatic syntax-based intra-language alignment. It performs automatic alignment of different versions of a text. Its concept is based on a modified version of the Needleman-Wunsch algorithm (for more information, see the Bibliography).
The tool allows automatic alignment between parallel texts in the same language. Its purpose is to display various degrees of textual variants based on syntactic alignment.

Upload your file

Select File

Nessun file selezionato

Supported formats: **txt and csv**

☐ Ignore Nonalphabetical ☐ Ignore Diacritics ☐ Case sensitive
☐ Levenshtein Distance

Enter Your Text

☐ Ignore Nonalphabetical ☐ Ignore Diacritics ☐ Case sensitive
☐ Levenshtein Distance

[View on Github \(Python\)](#) [View on Github \(Php\)](#)

Implemented by Tariq Yousef

A tool for automatic intra-language alignment

- **Automatic:** it is performed with algorithmic methods to reduce human intervention in the mechanical process of comparison.
- **Intra-language:** works with texts in the same language.
- **Pairwise or multiple:** works with two texts or with an unlimited number of multiple texts.

Algorithmic methods to produce alignment

The Needleman-Wunsch algorithm

- used in bioinformatics to align protein or nucleotide sequences.
- it uses Dynamic Programming to find the optimal alignment.
- divides a large problem into a series of smaller problems and uses the solutions to the smaller problems to reconstruct a solution to the larger problem.
- uses a score function and similarity matrix to represent all possible combinations of tokens and their resulting score.

The Needleman-Wunsch algorithm

- Aligning Bible Text John 1:1

NLT: In the beginning the Word already existed.

KJB: In the beginning was the Word

$$\begin{aligned}
 M[0][j] &= j * \text{indel} \quad \text{where } j \in [0, n], \\
 M[i][0] &= i * \text{indel} \quad \text{where } i \in [0, m]
 \end{aligned}$$

$$M[i][j] = \max \begin{cases} M[i-1][j-1] + m \searrow & \text{Matching} \\ M[i-1][j-1] + \text{mis} \searrow & \text{Mismatching} \\ M[i][j-1] + \text{indel} \downarrow & \text{Gap in S1} \\ M[i-1][j] + \text{indel} \rightarrow & \text{Gap in S2} \end{cases}$$

Where $0 < i \leq m, 0 < j \leq n$

The used score function (Matching = 5, Mismatching = -5, In/Del = -2)

		In	the	beginning	the	Word	already	existed	.
	0	-2→	-4→	-6→	-8→	-10→	-12→	-14→	-16→
In	-2↓								
the	-4↓								
beginning	-6↓								
was	-8↓								
the	-10↓								
Word	-12↓								
,	-14↓								

$$M[0][j] = j * \text{indel} \quad \text{where } j \in [0, n],$$

$$M[i][0] = i * \text{indel} \quad \text{where } i \in [0, m]$$

$$M[i][j] = \max \begin{cases} M[i-1][j-1] + m \searrow & \text{Matching} \\ M[i-1][j-1] + \text{mis} \searrow & \text{Mismatching} \\ M[i][j-1] + \text{indel} \downarrow & \text{Gap in S1} \\ M[i-1][j] + \text{indel} \rightarrow & \text{Gap in S2} \end{cases}$$

$$\text{Where } 0 < i \leq m, 0 < j \leq n$$

The used score function (Matching = 5, Mismatching = -5, In/Del = -2)

		In	the	beginning	the	Word	already	existed	.
	0	-2→	-4→	-6→	-8→	-10→	-12→	-14→	-16→
In	-2↓	5↘	3→	1→	-1→	-3→	-5→	-7→	-9→
the	-4↓	3↓							
beginning	-6↓	1↓							
was	-8↓	-1↓							
the	-10↓	-3↓							
Word	-12↓	-5↓							
,	-14↓	-7↓							

$$\begin{aligned}
 M[0][j] &= j * \text{indel} \quad \text{where } j \in [0, n], \\
 M[i][0] &= i * \text{indel} \quad \text{where } i \in [0, m] \\
 M[i][j] &= \max \begin{cases} M[i-1][j-1] + m \searrow & \text{Matching} \\ M[i-1][j-1] + \text{mis} \searrow & \text{Mismatching} \\ M[i][j-1] + \text{indel} \downarrow & \text{Gap in S1} \\ M[i-1][j] + \text{indel} \rightarrow & \text{Gap in S2} \end{cases} \\
 \text{Where } & 0 < i \leq m, \quad 0 < j \leq n
 \end{aligned}$$

The used score function (Matching = 5, Mismatching = -5, In/Del = -2)

		In	the	beginning	the	Word	already	existed	.
	0	-2→	-4→	-6→	-8→	-10→	-12→	-14→	-16→
In	-2↓	5 ↘	3→	1 →	-1 →	-3 →	-5 →	-7 →	-9 →
the	-4↓	3↓	10 ↘	8 →	6 ↘	4 →	2 →	0 →	-2 →
beginning	-6↓	1↓	8 ↓						
was	-8↓	-1↓	6 ↓						
the	-10↓	-3↓	4 ↓						
Word	-12↓	-5↓	2 ↓						
,	-14↓	-7↓	0 ↓						

$$\begin{aligned}
 M[0][j] &= j * \text{indel} \quad \text{where } j \in [0, n], \\
 M[i][0] &= i * \text{indel} \quad \text{where } i \in [0, m] \\
 M[i][j] &= \max \begin{cases} M[i-1][j-1] + m \searrow & \text{Matching} \\ M[i-1][j-1] + \text{mis} \searrow & \text{Mismatching} \\ M[i][j-1] + \text{indel} \downarrow & \text{Gap in S1} \\ M[i-1][j] + \text{indel} \rightarrow & \text{Gap in S2} \end{cases} \\
 \text{Where } & 0 < i \leq m, \quad 0 < j \leq n
 \end{aligned}$$

The used score function (Matching = 5, Mismatching = -5, In/Del = -2)

		In	the	beginning	the	Word	already	existed	.
	0	-2→	-4→	-6→	-8→	-10→	-12→	-14→	-16→
In	-2↓	5↘	3→	1→	-1→	-3→	-5→	-7→	-9→
the	-4↓	3↓	10↘	8→	6↘	4→	2→	0→	-2→
beginning	-6↓	1↓	8↓	15↘	13→	11→	9→	7→	5→
was	-8↓	-1↓	6↓	13↓					
the	-10↓	-3↓	4↓	11↓					
Word	-12↓	-5↓	2↓	9↓					
,	-14↓	-7↓	-0↓	7↓					

$$M[0][j] = j * \text{indel} \quad \text{where } j \in [0, n],$$

$$M[i][0] = i * \text{indel} \quad \text{where } i \in [0, m]$$

$$M[i][j] = \max \begin{cases} M[i-1][j-1] + m \quad \text{Matching} \\ M[i-1][j-1] + \text{mis} \quad \text{Mismatching} \\ M[i][j-1] + \text{indel} \downarrow \quad \text{Gap in S1} \\ M[i-1][j] + \text{indel} \rightarrow \quad \text{Gap in S2} \end{cases}$$

Where $0 < i \leq m, 0 < j \leq n$

The used score function (Matching = 5, Mismatching = -5, In/Del = -2)

New Living Translation	In	the	beginning		the	Word	already	existed	.	
King James Bible	In	the	beginning	was	the	Word				,

		In	the	beginning	the	Word	already	existed	.
	0	-2→	-4→	-6→	-8→	-10→	-12→	-14→	-16→
In	-2↓	5↘	3→	1→	-1→	-3→	-5→	-7→	-9→
the	-4↓	0↓	10↘	8→	6↘	4→	2→	0→	-2→
beginning	-6↓	-2↓	8↓	15↘	13→	11→	9→	7→	5→
was	-8↓	-4↓	6↓	13↓	11↓	9→	7→	5→	3→
the	-10↓	-8↓	4↓	11↓	18↘	16→	14→	12→	10→
Word	-12↓	-10↓	2↓	9↓	16↓	23↘	21→	19→	17→
,	-14↓	-12↓	-0↓	7↓	14↓	21↓	19↓	17↓	15↓

The modification to the algorithm

$$k = 14, n = 157, m = 134$$

$$\text{Search space} = m * n = 21038$$

after modification

$$(2k+1) * m = 3886$$

1In the beginning God created the heavens and the earth. 2The earth was formless and empty, and darkness covered the deep waters. And the Spirit of God was hovering over the surface of the waters. 3 Then God said, "Let there be light," and **there** was light. 4And God saw that the light was good. Then he separated the light from the darkness. 5God called the light "day" and the darkness "night." And evening passed and morning came, marking the first day.

6 Then God said, "Let there be a space between the waters, to separate the waters of the heavens from the waters of the earth." 7And that is what happened. God made this space to separate the waters of the earth from the waters of the heavens. 8God called the space "sky."

1In the beginning God created the heavens and the earth. 2The earth was without form, and void; and darkness was on the face of the deep. And the Spirit of God was hovering over the face of the waters. 3 **Then God said, "Let there be light"; and there was light.** 4**And God saw the light, that it was good; and God divided the light from the darkness.** 5 God called the light Day, and the darkness He called Night. So the evening and the morning were the first day.

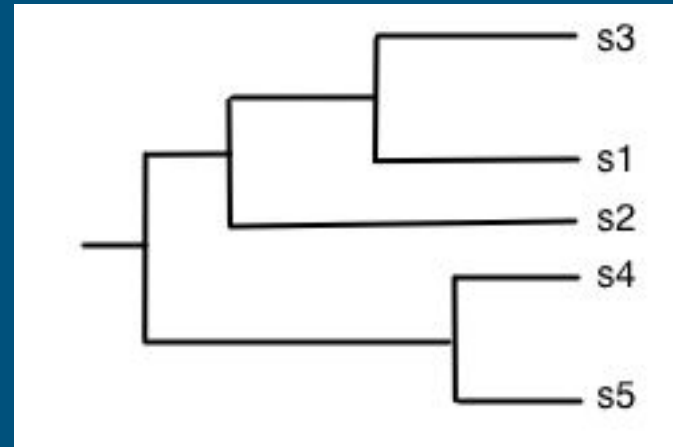
6 Then God said, "Let there be a firmament in the midst of the waters, and let it divide the waters from the waters." 7 Thus God made the firmament, and divided the waters which were under the firmament from the waters which were above the firmament; and it was so. 8And God called the firmament Heaven. So the evening and the morning were the second day.

Multiple Sequence Alignment (In progress)

- Progressive alignment

builds up a final MSA by combining pairwise alignments beginning with the most similar pair and progressing to the most distantly related, it requires two stages:

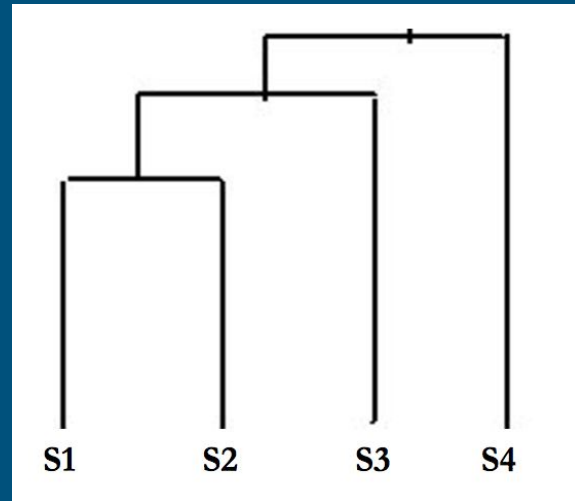
- creating the guide tree (clustering)
- adding the sequences sequentially to the growing MSA according to the guide tree



Multiple Sequence Alignment (In progress)

- Iterative alignment

The aim is to reduce the problem of a multiple alignment to an iteration of pairwise alignments.



How to align your texts with iAligner: copy your text on the editor

1 The book of the generation of Jesus Christ, the son of Davi
1 The book of the generation of Jesus Christ, the son of Davi

2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat
2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat

3 and Judah begat Perez and Zerah of Tamar; and Perez begat H
3 And Judas begat Phares and Zara of Thamar; and Phares begat

4 and Ram begat Amminadab; and Amminadab begat Nahshon; and N
4 And Aram begat Aminadab; and Aminadab begat Naasson; and Na

5 and Salmon begat Boaz of Rahab; and Boaz begat Obed of Ruth
5 And Salmon begat Booz of Rachab; and Booz begat Obed of Rut

6 and Jesse begat David the king. And David begat Solomon of
6 And Jesse begat David the King; and David the king begat So

7 and Solomon begat Rehoboam; and Rehoboam begat Abijah; and
7 And Solomon begat Roboam; and Roboam begat Abia; and Abia b

8 and Asa begat Jehoshaphat; and Jehoshaphat begat Joram; and
8 And Asa begat Josaphat; and Josaphat begat Joram; and Joram

9 and Uzziah begat Jotham; and Jotham begat Ahaz; and Ahaz be
9 And Ozias begat Joatham; and Joatham begat Achaz; and Achaz

Enter Your Text

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judah and his brethren;

2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judas and his brethren;

3 and Judah begat Perez and Zerah of Tamar; and Perez begat Hezron; and Hezron begat Ram;

3 And Judas begat Phares and Zara of Thamar; and Phares begat Esrom; and Esrom begat Aram;

4 and Ram begat Amminadab; and Amminadab begat Nahshon; and Nahshon begat Salmon;

☐ Ignore Nonalphabetical ☐ Ignore Diacritics ☐ Case sensitive ☐ Levenshtein Distance

Reset

Align

The text has to be parsed in sentences first

...Or upload it

Currently supports .txt and .csv files

Upload your file

Select File

bible test.txt

Supported formats: **txt** and **csv**

- ☒ Ignore Nonalphabetical
- ☐ Ignore Diacritics
- ☐ Case sensitive
- ☐ Levenshtein Distance

Refinement criteria

Enter Your Text

1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.
1 The book of the generation of Jesus Christ, the son of David, the son of Abraham.

2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judah and his brethren;
2 Abraham begat Isaac; and Isaac begat Jacob; and Jacob begat Judas and his brethren;

3 and Judah begat Perez and Zerah of Tamar; and Perez begat Hezron; and Hezron begat Ram;
3 And Judas begat Phares and Zara of Thamar; and Phares begat Esrom; and Esrom begat Aram;

4 and Ram begat Amminadab; and Amminadab begat Nahshon; and Nahshon begat Salmon;

☐ Ignore Nonalphabetical ☐ Ignore Diacritics ☐ Case sensitive ☐ Levenshtein Distance

Reset

Align

Upload your file

Select File

Scegli file bible test.txt

Supported formats: **txt** and **csv**

Upload

☒ Ignore Nonalphabetical ☐ Ignore Diacritics ☐ Case sensitive
☐ Levenshtein Distance

- **Ignore non-alphabetical:** ignores symbols, such as punctuation and numbers, anything that is not an alphabetical character
- **Case sensitive:** if activated, detects variation across words according to the case
- **Ignore diacritics:** ignores any type of diacritical character (including punctuation marks)
- **Levenshtein distance:** applies a revised version of the Levenshtein algorithm and increases the tolerance threshold on the alignment of similar words.

The Levenshtein distance

The Levenshtein distance between two words is the minimum number of single-character edits (i.e. insertions, deletions or substitutions) required to change one word into the other. e.g

<code>lev(George, Georg) = 1</code>	Normalised Distance = $1 - 1/6 = 0.833$
<code>lev(Prague, Prag) = 2</code>	Normalised Distance = $1 - 2/6 = 0.666$
<code>lev(Rome, Rom) = 1</code>	Normalised Distance = $1 - 1/4 = 0.75$

Mathematically, the Levenshtein distance between two strings $\{a, b\}$ (of length $|a|$ and $|b|$ respectively) is given by $\text{lev}_{a,b}(|a|, |b|)$

$$\text{lev}_{a,b}(i, j) = \begin{cases} \max(i, j) & \text{if } \min(i, j) = 0, \\ \min \begin{cases} \text{lev}_{a,b}(i-1, j) + 1 \\ \text{lev}_{a,b}(i, j-1) + 1 \\ \text{lev}_{a,b}(i-1, j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{otherwise.} \end{cases}$$

Modified Levenshtein Distance

Levenshtein distance is not very helpful in our case, because it is binary and there is no tolerance with errors produced by OCR or Transcription.

the distance between letters is not binary, but it is on scale. The cost of insertion or deletion depends on:

- Letter position
- Letter type (vowel or consonant)

Modified Levenshtein Distance

OCR		
Group elements		Distance within group elements
t	l	0.25
v	u	0.25
6	8	0.25
c	e	0.25
b	h	0.25
in	m	0.25
o	e	0.25
i	l	0.25
i	j	0.25

Transliteration	
Group elements	Distance within group elements
Vowel letters	0.25
C K Q	0.3
S C	0.2

`lev(Georgee, Georg) = 0.25`

Normalised Distance = $1 - 0.25/6 = 0.958$

`lev(Praguee, Prag) = 0.5`

Normalised Distance = $1 - 0.5/6 = 0.916$

A Greek text with no refinement criteria

τῆς δ' ἀσίας ἀπὸ κανώβου ἕως ταναΐδος ποταμοῦ μετὰ τῶν κόλπων ὁ παράπλους, σταδίων μυριάδων δ καὶ ρια·
τῆς δὲ ἀσίας, ἀπὸ κανώβου, ἕως ταναΐδος ποταμοῦ· μετὰ τῶν κόλπων ὁ παράπλους· στάδια Δ ρια·

τῆς	δ'	ἀσίας	ἀπὸ	κανώβου	ἕως	ταναΐδος	ποταμοῦ	μετὰ	τῶν	κόλπων	ὁ	παράπλους	,	σταδίων	μυριάδων	δ	καὶ	ρια	·	
τῆς	δὲ	ἀσίας	·	ἀπὸ	κανώβου	·	ἕως	ταναΐδος	ποταμοῦ	·	μετὰ	τῶν	κόλπων	ὁ	παράπλους	·	στάδια	Δ	ρια	·

Length: 24 Aligned-complete 13 notAligned 5 Gap 5 Aligned-case 1

The longest common substring

μετὰ	τῶν	κόλπων	ὁ
μετὰ	τῶν	κόλπων	ὁ

The same text with additional refinement criteria applied

τῆς δ' ἀσίας ἀπὸ κανώβου ἕως ταναΐδος ποταμοῦ μετὰ τῶν κόλπων ὁ παράπλους, σταδίων μυριάδων δ καὶ ρια·
τῆς δὲ ἀσίας, ἀπὸ κανώβου, ἕως ταναΐδος ποταμοῦ· μετὰ τῶν κόλπων ὁ παράπλους· στάδια Δ ρια·

τῆς	δ'	ἀσίας	ἀπὸ	κανώβου	ἕως	ταναΐδος	ποταμοῦ	μετὰ	τῶν	κόλπων	ὁ	παράπλους	,	σταδίων	μυριάδων	δ	καὶ	ρια	·	
τῆς	δὲ	ἀσίας	·	ἀπὸ	κανώβου	·	ἕως	ταναΐδος	ποταμοῦ	·	μετὰ	τῶν	κόλπων	ὁ	παράπλους	·	στάδια	Δ	ρια	·

Length: 24 Aligned-complete 13 notAligned 1 Gap 5 Aligned-levenshtein 2 Aligned-removed diacritics 1 Aligned-removed NonAlphanumeric 1 Aligned-case 1

The longest common substring

μετὰ	τῶν	κόλπων	ὁ	παράπλους	,	σταδίων
μετὰ	τῶν	κόλπων	ὁ	παράπλους	·	στάδια

Alignment output: a table-graph

iAligner displays all the nuances of variants according to a color-key:

- Completely aligned tokens (deep green)
- Tokens aligned by excluding case sensitivity or punctuation detection (light green)
- Gaps (yellow)
- Divergences (red)
- Tokens aligned by applying Levenshtein distance (blue-green)

τῆς δ' Ἀσίας ἀπὸ κανώβου ἕως ταναΐδος ποταμοῦ μετὰ τῶν κόλπων ὁ παράπλους, σταδίων μυριάδων δ καὶ ρία·
τῆς δὲ Ἀσίας. ἀπὸ κανώβου. ἕως ταναΐδος ποταμοῦ· μετὰ τῶν κόλπων ὁ παράπλους· στάδια Δ ρία·

τῆς	δ'	Ἀσίας		ἀπὸ	κανώβου		ἕως	τανάιδος	ποταμοῦ		μετὰ	τῶν	κόλπων	ὁ	παράπλους	,	σταδίων	μυριάδων	δ	καὶ	ρία	·
τῆς	δὲ	Ἀσίας	.	ἀπὸ	κανώβου	.	ἕως	ταναΐδος	ποταμοῦ	·	μετὰ	τῶν	κόλπων	ὁ	παράπλους	·	στάδια		Δ		ρία	·

Length: 24 Aligned-complete 13 notAligned 1 Gap 5 Aligned-Levenshtein 2 Aligned-removedDiacritics 1 Aligned-removedNonAlphanumeric 1 Aligned-case 1

The longest common substring

μετὰ	τῶν	κόλπων	ὁ	παράπλους	,	σταδίων
μετὰ	τῶν	κόλπων	ὁ	παράπλους	·	στάδια

What can you do with iAligner? Some case studies

Manuscript alignment

Three manuscripts of Plato's *Crito*
aligned (<http://i-alignment.com/crito/>)

Manuscript Alignment

Plato's Crito

Clark: A digital encoding of Ms Clark 39, 20v-26r, Oxford, University Bodleian Library

Paris1808: A digital encoding of Ms Grec 1808, 17r-21v, Paris, Bibliothèque Nationale

Tuebingen: A digital encoding of Ms Gr Mb 14, 21-38, Tübingen, Universität, Bibliothek

43 ▾ 44 ▾ 45 ▾ 46 ▾ 47 ▾ 48 ▾ 49 ▾ 50 ▾ 51 ▾ 52 ▾ 53 ▾ 54 ▾

46C

CLARK

#Σωκράτης καὶ τιμῶ, οὐσπερ καὶ
πρότερον· ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν
ἐν τῷ παρόντι, εὖ ἴσθι. ὅτι οὐ μὴ σοὶ
συγχαρήσω· οὐδ' ἂν πλείω τῶν νῦν
παρόντων, ἢ τῶν πολλῶν δύναιμι, ὥσπερ
παῖδας ἡμᾶς μορμολύττηται· δεσμούς· καὶ
θανάτους ἐπιτέμπουσα· καὶ χρημάτων
ἀφαιρέσεις· πῶς οὖν ἂν μετριώτατα
σκοποίμεθα αὐτά; εἰ πρῶτον μὲν τοῦτον
τὸν λόγον ἀναλάβοιμεν, ὃν σὺ λέγεις περὶ
τῶν δοξῶν· πότερον καλῶς ἐλέγετο
ἐκάστοτε, ἢ οὐ· ὅτι ταῖς μὲν δεῖ τῶν δοξῶν
προσέχειν τὸν νοῦν.

Paris1808

#Σωκράτης καὶ τιμῶ οὐσπερ καὶ
πρότερον· ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν
ἐν τῷ παρόντι· εὖ ἴσθι ὅτι οὐ μὴ σοὶ
συγχαρήσω. οὐδ' ἂν πλείω τῶν νῦν
παρόντων ἢ τῶν πολλῶν δύναιμι· ὥσπερ
παῖδας ἡμᾶς μορμολύττηται δεσμούς καὶ
θανάτους ἐπιτέμπουσα· καὶ χρημάτων
ἀφαιρέσεις· πῶς οὖν μετριώτατα
σκοποίμεθα αὐτά· εἰ πρῶτον μὲν τοῦτον
τὸν λόγον ἀναλάβοιμεν· ὃν σὺ λέγεις περὶ
τῶν δοξῶν· πότερον· καλῶς ἐλέγετο
ἐκάστοτε ἢ οὐ· ὅτι ταῖς . δεῖ τῶν δοξῶν
προσέχειν τὸν νοῦν.

Tuebingen

#Σωκράτης καὶ τιμῶ· οὐσπερ καὶ
πρότερον, ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν
ἐν τῷ παρόντι. εὖ ἴσθι. ὅτι οὐ μὴ σοὶ
συγχαρήσω· οὐδ' ἂν πλείω τῶν νῦν
παρόντων· ἢ τῶν πολλῶν δύναιμι, ὥσπερ
παῖδας ἡμᾶς μορμολύττηται· δεσμούς καὶ
θανάτους ἐπιτέμπουσα· καὶ χρημάτων
ἀφαιρέσεις· πῶς οὖν ἂν μετριώτατα,
σκοποίμεθα αὐτά, εἰ πρῶτον μὲν τοῦτον
τὸν λόγον ἀναλάβοιμεν· ὃν σὺ λέγεις περὶ
τῶν δοξῶν· πότερον καλῶς ἐλέγετο
ἐκάστοτε, ἢ οὐ· ὅτι ταῖς μὲν, δεῖ τῶν δοξῶν
προσέχειν τὸν νοῦν.

#Σωκράτης καὶ τιμῶ , οὐσπερ καὶ πρότερον · ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν ἐν τῷ παρόντι , εὖ ἴσθι . ὅτι οὐ μὴ σοὶ συγχαρήσω

#Σωκράτης καὶ τιμῶ οὐσπερ καὶ πρότερον · ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν ἐν τῷ παρόντι · εὖ ἴσθι ὅτι οὐ μὴ σοὶ συγχαρήσω

#Σωκράτης καὶ τιμῶ · οὐσπερ καὶ πρότερον , ὦν ἐὰν μὴ βελτίῳ ἔχωμεν λέγειν ἐν τῷ παρόντι . εὖ ἴσθι . ὅτι οὐ μὴ σοὶ συγχαρήσω

συγχαρήσω · οὐδ' ἂν πλείω τῶν νῦν παρόντων , ἢ τῶν πολλῶν δύναιμι , ὥσπερ παῖδας ἡμᾶς μορμολύττηται · δεσμούς · καὶ θανάτους

συγχαρήσω . οὐδ' ἂν πλείω τῶν νῦν παρόντων ἢ τῶν πολλῶν δύναιμι : ὥσπερ παῖδας ἡμᾶς μορμολύττηται δεσμούς καὶ θανάτους

συγχαρήσω · οὐδ' ἂν πλείω τῶν νῦν παρόντων · ἢ τῶν πολλῶν δύναιμι , ὥσπερ παῖδας ἡμᾶς μορμολύττηται · δεσμούς καὶ θανάτους

ἐπιτέμπουσα · καὶ χρημάτων ἀφαιρέσεις · πῶς οὖν ἂν μετριώτατα σκοποίμεθα αὐτά ; εἰ πρῶτον μὲν τοῦτον τὸν λόγον ἀναλάβοιμεν ,

ἐπιτέμπουσα · καὶ χρημάτων ἀφαιρέσεις · πῶς οὖν μετριώτατα σκοποίμεθα αὐτά · εἰ πρῶτον μὲν τοῦτον τὸν λόγον ἀναλάβοιμεν ·

ἐπιτέμπουσα · καὶ χρημάτων ἀφαιρέσεις · πῶς οὖν ἂν μετριώτατα , σκοποίμεθα αὐτά , εἰ πρῶτον μὲν τοῦτον τὸν λόγον ἀναλάβοιμεν ·

OCR output alignment

Example of Aligning OCR-outputs:

coo.31924054869700_ocr/024.txt	hvd.32044019207893_ocr/024.txt	Alignment
δυνάμει πεπυργωμένην, καὶ οἷον ἀμάχῳ φλογὶ ταῖς	δυνάμει πεπυργωμένην, καὶ οἷον ἀμάχῳ φλογὶ ταῖς	δυνάμει πεπυργωμένην, καὶ οἷον ἀμάχῳ φλογὶ ταῖς
ἄνωθεν εὐμενεῖαις διεζωσμένην. Τρέχει τοίνυν ὁ ἅγ-	ἄνωθεν εὐμενεῖαις διεζωσμένην. Τρέχει τοίνυν ὁ ἅγ-	ἄνωθεν εὐμενεῖαις διεζωσμένην. Τρέχει τοίνυν ὁ ἅγ-
γελος ἐκμετρῆσαι τὴν Ἱερουσαλὴμ, καὶ οἷον ὀρίσαι	γελος ἐκμετρῆσαι τὴν Ἱερουσαλὴμ, καὶ οἷον ὀρίσαι	γελος ἐκμετρῆσαι τὴν Ἱερουσαλὴμ, καὶ οἷον ὀρίσαι
πλατεῖαν αὐτὴν καὶ εὐμηκεσάτην. Καὶ ταυτὶ μὲν	πλατεῖαν αὐτὴν καὶ εὐμηκεσάτην. Καὶ ταυτὶ μὲν	πλατεῖαν αὐτὴν καὶ εὐμηκεσάτην. Καὶ ταυτὶ μὲν
ἱπορικῶς.	ἱστορικῶς. η	ἱπορικῶς. η
Φαίην δ' ἂν ὅτι καὶ ἐπ' αὐτῆς τῆς Ἐκκλησίας	. Φαίην δ' ἂν ὅτι καὶ ἐπ' αὐτῆς τῆς Ἐκκλησίας	Φαίην δ' ἂν ὅτι καὶ ἐπ' αὐτῆς τῆς Ἐκκλησίας
Χριστοῦ νοοῖτο ἂν εἰκότως ἡ ὄρασις. Τετυραννεύκει	Χριστοῦ νοοῖτο ἂν εἰκότως ἡ ὄρασις. Τετυραννεύκει	Χριστοῦ νοοῖτο ἂν εἰκότως ἡ ὄρασις. Τετυραννεύκει
μὲν γὰρ κατὰ πάντων ὁ Σατανᾶς τῶν ὄντων ἐπὶ τῆς	μὲν γὰρ κατὰ πάντων ὁ Σατανᾶς τῶν ὄντων ἐπὶ τῆς	μὲν γὰρ κατὰ πάντων ὁ Σατανᾶς τῶν ὄντων ἐπὶ τῆς
γῆς, καὶ γεγόναμεν αἰχμάλωτοι, τοῖς ἐκείνου ζυγοῖς	γῆς, καὶ γεγόναμεν αἰχμάλωτοι, τοῖς ἐκείνου ζυγοῖς	γῆς, καὶ γεγόναμεν αἰχμάλωτοι, τοῖς ἐκείνου ζυγοῖς

Alignment of two OCR outputs from the *Patrologia Graeca*. The third column shows the overlapping sections and offers the user the choice between two variants where the two texts diverge.

OCR output alignment

auctoritatem , sub significatione nativitatis , proprietas naturalis ostensa
 sit : Patrem suum dicebat Deum , aequalem se faciens Deo . Anne
 naturalis nativitas non est , ubi per nomen patris proprii , naturae
 aequalitas demonstratur ? Non enim ambigitur , quin aequalitas nihil
 differat . Quis porro dubitabit , quin indifferentem naturam nativitas
 consequatur ? Hinc enim est sola illa quae vere esse possit aequalitas :
 quia naturae aequalitatem sola possit praestare nativitas . Aequalitas
 vero numquam ibi esse credetur , ubi unio est : nec tamen illic reperietur ,
 ubi differt . Ita similitudinis aequalitas nec solitudinem habet , nec
 diversitatem : quia omnis aequalitas nec diversa , nec sola sit . 16 . Filii
 nativitas et cum Patre aequalitas ex ipsius

auctoritatem , sub significatione nativitatis , proprietas naturalis ostensa
 sit : Patrem suum dicebat Deum , qualem se faciens Deo . Anne naturalis
 nativitas non est , ubi per nomen patris proprii , naturae qualitas
 demonstratur ? Non enim ambigitur , quin qualitas nihil differat . Quis
 porro dubitabit , quin indifferentem naturam nativitas consequatur ? Hinc
 enim est sola illa quae vere esse possit qualitas : quia naturae qualitatem sola
 possit praestare nativitas . qualitas vero numquam ibi esse credetur , ubi
 unio est : nec (c) tamen illic reperietur , ubi differt . Ita similitudinis
 qualitas nec solitudinem habet , nec diversitatem : quia omnis qualitas
 nec diversa , nec sola sit . 16 . Filii nativitas et cum Patre qualitas ex
 ipsius

Length: 127 || Aligned-complete (109) || Gap (3) || Aligned-levenshtein (14) || notAligned (1) ||

Patrologia Latina: OCR output vs. correct version: <http://www.i-alignment.com/pl/index.html>

OCR output alignment

SANCTI AMBROSII MEDIOLANENSIS EPISCOPI

LIBER PRIMUS.

1113 [i ↕]. Deduximus, fratres dilectissimi, hostiam
meam, hostiam incontaminatam, hostiam Deo placentem,

[b] domum et fratrem meum Satyrum.
Memineram esse mortalem, nec fefellit opinio,

a Scripti anno 579.

b Omnes edit. ac mss. nonnum. *dominum et fra-*
trem; ali mss. plures et potiores, *domnum et fratrem*.
Observat autem Monardus noster in cap. 70 Regulae
S. Bened. vocem *Domnum* regiam et imperatoriam
esse, nec ea salutari voluisse Augustum Caesarem:
postea autem a scriptoribus, ut humiliori vocabulo
uterentur, ex too *dominus* factam esse vocem *dom-*
nus per synopen. Quod ubi exemplis e Sev. Sulpi-
tio, et aliunde petitis probavit, addit apud Graecos

PATROL. XVI.

41

A

[a] sed superabundavit gratia. Itaque nihil habeo quod
querar, et habeo ✓ in quo Deo gratias agam; quia semper
optavi, ut si quae in quo Deo gratias agam Ecclesiam vel
me manerent, in me potius ac meam deciderent
domum. Deo igitur gratias, quia in hoc omnium
metu, cum omnia motibus sint suspecta Barbaricis,
communem moerorem privato dolore transegi, et in
[nic ↕] conversum est [b] quidquid timebam omnibus.
Atque utinam hic consummatum sit, ut dolor [meas ↕]

A

[a] sed superabundavit gratia. Itaque nihil habeo quod
querar, et habeo [inquo Deo gratias ↕] agam; quia semper
optavi, ut si quae perturbationes [vet ↕] Ecclesiam vel
me manerent, in me potius ac meam deciderent
domum. Deo igitur gratias, quia in hoc omnium
metu, cum omnia motibus sint suspecta Barbaricis,
communem moerorem privato dolore transegi, et in
[nic ↕] conversum est [b] quidquid timebam omnibus.
Atque utinam hic consummatum sit, ut dolor [meas ↕]
publici doloris redemptio [sili 1 ↕]
2. Nihil quidem habui, fratres charissimi, in rebus
humanis tanto fratre [pretiositas ↕], [hihil ↕] amabilius,
nihil charius; sed praestant privatis publica. Ipsius
quoque sententiam si quis sciscitaretur, [c] mallet
[occidi ↕] pro [aliis ↕], quam sibi vivere; propterea enim

B

pro omnibus secundum carnem Christus est mortuus,
ut [itos ↕] non solis nobis vivere disceremus [il ↕] Cor.
v, 15).

edit illud quod [ingl atus ↕] divinitati esse non
n : laetandum est [euim ↕] magis quod talem [fratreni ↕]
im, quam dolendum quod fratrem amiserim;
im munus, hoc debitum est. Itaque
[Unctus ↕] sum, quamdiu [licuit ↕], 1114 commisso
[anore ↕]: qui deposuit pignus, recepit. Nihil
t utrum abjures depositum, an doleas restitutum.
que fidei ambiguum, [a vitse ↕] periculum
si pecuniam neget, culpa est: si hostiam
[r] pietas est? cum pecuniae fenerator illudi
naturae auctor, et necessitudinis creditor
i non queat. Itaque quanto uberior fenoris

Alignment of editions

- Ζεὺς μὲν ἀφίκτωρ ἐπίδοι προφρόνως
- Ζεὺς μὲν ἀφίκτωρ ἐπίδοι προφρόνως

Ζεὺς	μὲν	ἀφίκτωρ	ἐπίδοι	προφρόνως
Ζεὺς	μὲν	ἀφίκτωρ	ἐπίδοι	προφρόνως

Length: 5 Aligned-complete 5

- στόλον ἡμέτερον νάιον ἀρθέντ'
- στόλον ἡμέτερον νάιον ἀρθέντ'

στόλον	ἡμέτερον	νάιον	ἀρθέντ'
στόλον	ἡμέτερον	νάιον	ἀρθέντ'

Length: 4 Aligned-complete 4

- ἀπὸ προσομίῳν λεπτοψαμάθων
- ἀπὸ προσομίῳν λεπτοψαμάθων

ἀπὸ	προσομίῳν	λεπτοψαμάθων
ἀπὸ	προσομίῳν	λεπτοψαμάθων

Length: 3 Aligned-complete 3

Vv. 1-3

- καὶ στασίαρχος τάδε πεσσονομῶν
- καὶ στασίαρχος τάδε πεσσονομῶν,

καὶ	στασίαρχος	τάδε	πεσσονομῶν
καὶ	στασίαρχος	τάδε	πεσσονομῶν ,

Length: 6 Aligned-complete 4 Gap 2

- κύδιστ' ἀχέων ἐπέκρανε,
- κύδιστ' ἀχέων ἐπέκρανεν,

κύδιστ'	ἀχέων	ἐπέκρανε	,
κύδιστ'	ἀχέων	ἐπέκρανεν	,

Length: 5 Aligned-complete 4 notAligned 1

Vv. 12-13

- πνεύματι χώρας· ἀρσενοπληθῆ δ'
- πνεύματι χώρας· ἀρσενοπληθῆ δ'

πνεύματι	χώρας	·	ἀρσενοπληθῆ	δ'
πνεύματι	χώρας	:	ἀρσενοπληθῆ	δ'

Length: 5 Aligned-complete 4 notAligned 1

- ἔσμον ὕβριστήν Αἰγυπτογενῆ,
- ἔσμον ὕβριστήν Αἰγυπτογενῆ,

ἔσμον	ὕβριστήν	Αἰγυπτογενῆ	,
ἔσμον	ὕβριστήν	Αἰγυπτογενῆ	,

Length: 5 Aligned-complete 5

- πρὶν πόδα χέρσῳ τῇ δ' ἐν ἀσώδει
- πρὶν πόδα χέρσῳ τῇδ' ἐν ἀσώδει

πρὶν	πόδα	χέρσῳ	τῇ	δ'	ἐν	ἀσώδει
πρὶν	πόδα	χέρσῳ	τῇδ'	ἐν	ἀσώδει	

Length: 7 Aligned-complete 5 Gap 1 notAligned 1

Vv. 29-31

Three excerpted editions of Aeschylus' *Supplices* aligned.
www.i-alignment.com/Aeschylus

Future work

- Import and export options
- Language dependent options for Latin, Greek, Arabic
- Handling crossings and transpositions

Thanks for the attention!

Contact:

Tariq Yousef (tariq.yousef@dh.uni-leipzig.de)
Chiara Palladino (chiarapalladino1@gmail.com)