

Workshop
Portuguese presence in Goa in digital perspective

Historical Terminology

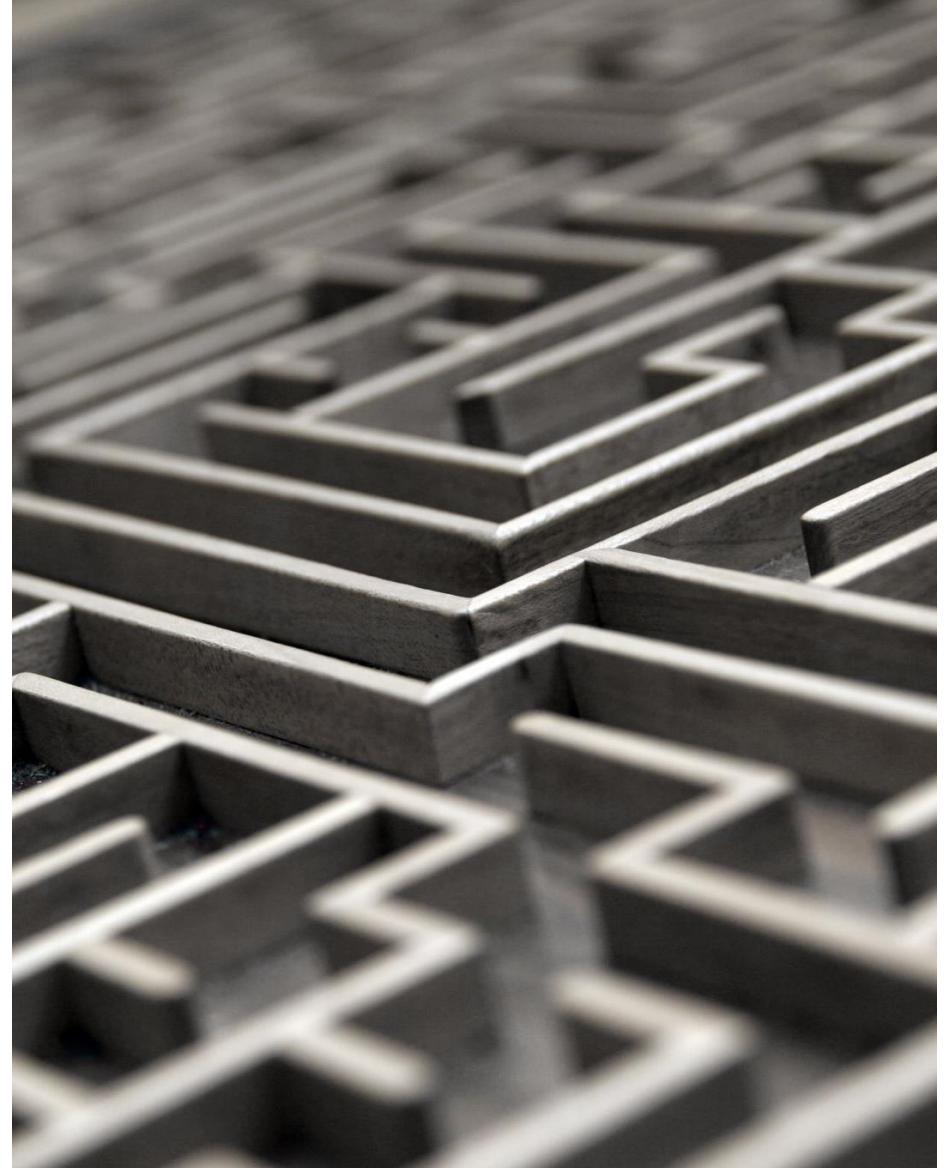
Experiences with historical medical documents

Leonardo Zilio (FAU Erlangen-Nürnberg, Germany)

Maria José B. Finatto (UFRGS, Brazil)

Outline

- Introduction to the Historical Terminology project
- Working with historical documents
 - Transcription
 - Named-entity recognition
 - Complexity of historical documents
 - Normalisation of documents
- Future experiments





The Historical Terminology project

The Historical Terminology project

Menu

PÁGINA INICIAL

SEMEDO (1707)

SEMEDO (1721) E +

COMO TRANSCREVER?

NORMA TRANSCRIÇÃO
(dez 2019)

AMOSTRA INICIAL

EXERCÍCIO TRANSCRIÇÃO

MATERIAIS PARA ESTUDO

ACESSE: CORPUS-AMOSTRA

ACESSE: LIVRO COMPLETO

EQUIPE

<< TEXTECC

Apoio


Universidade Federal do Rio Grande do Sul

Bem vindo à página - Terminologia Histórica



Terminologia Histórica

SEJA BEM-VINDO AO PROJETO TERMINOLOGIA HISTÓRICA



Estamos construindo e ampliando nosso corpus-amostra composto por **obras médicas impressas em português do século XVIII**.

Para saber as **NOVIDADES de 2022** do nosso projeto - agora com contrastes de obras de médicos, cirurgiões e enfermeiros, acesse [nosso site](#).

Partindo de diferentes obras do médico português **João Curvo Semedo** (1635-1719), queremos:

- divulgar a importância dos *corpora* históricos - com destaque para os em português - para diferentes tipos de pesquisas em Letras e em áreas afins no cenário acadêmico brasileiro;
- subsidiar professores e alunos com *corpora* - em formato de amostra - e indicação de recursos computacionais para o seu estudo em termos de ESTUDOS DE VOCABULÁRIOS (incluindo **terminologias**);
- indicar *corpora* de grandes dimensões pré-existentes para aprofundamentos de estudos que iniciam aqui;
- sugerir atividades de ensino para uma introdução ao **estudo diacrônico do vocabulário e das terminologias técnico-científicas** empregadas em documentos antigos.
- fornecer dados históricos sobre complexidade ou acessibilidade textual e terminológica em Saúde, incluindo dados sobre a história da Medicina em português.

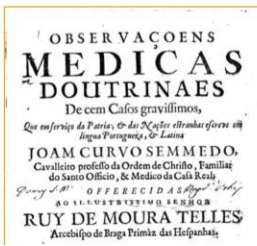


Imagem da obra de João Curvo SEMEDO (1707) página 01.
Observações medicas doutrinaes de cem casos gravissimos(...)
Digitalização Biblioteca Pública de Évora - Portugal

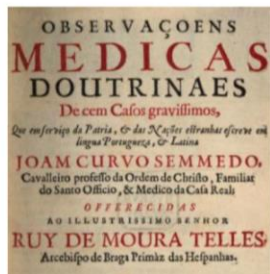
<https://www.ufrgs.br/textecc/terminologia/>

The Historical Terminology project

- Project lead by Prof Maria José B. Finatto
 - Federal University of Rio Grande do Sul
- Allows the download of transcribed and (some) normalised historical medical documents from the 18th century

Para comparar obras de ENFERMEIROS e de MÉDICOS – veja abaixo nossas primeiras transcrições produzidas:

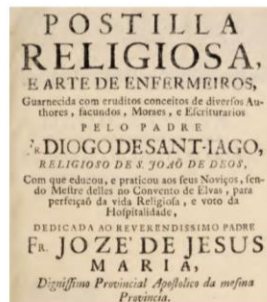
Textos apenas no formato .txt para uso em ferramentas de processamento computacional



Observações Medicas Doutrinaes (1707)

Arquivos de transcrições do livro de João Curvo Semedo estão disponíveis aqui

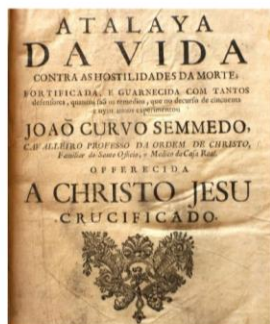
[Acessar](#)



Postilla Religiosa, e Arte de Enfermeiros (1741)

Arquivos de transcrições desse livro estão disponíveis aqui

[Acessar](#)



Atalaya da vida, contra as hostilidades da morte (1721)

Arquivos de transcrições do livro de João Curvo Semedo estão disponíveis aqui

[Acessar](#)



Aviso à gente do mar sobre sua saúde (1794)

Arquivos de transcrições desse livro estão disponíveis aqui

[Acessar](#)

Currently available corpora

<https://sites.google.com/view/projeto38597>

João Curvo Semedo (1639-1719)



- 1707
- Observações medicas doutrinaes de cem casos gravíssimos
(Medical and doctrinal observations of a hundred severe cases)
- 1721 (*post mortem*)
- Atalaya da vida, contra as hostilidades da morte
(Observatory of life, against the hostilities of death)

POSTILLA RELIGIOSA,

E ARTE DE ENFERMEIROS,

Guarnecida com eruditos conceitos de diversos Au-
thores, facundos, Moraes, e Escriurarios

PELO PADRE

FR. **DIOGO DE SANT-IAGO,**

RELIGIOSO DE S. JOÃO DE DEOS,

Com que educou, e praticou aos seus Noviços, sen-
do Mestre delles no Convento de Elvas, para
perfeição da vida Religiosa, e voto da
Hospitalidade,

DEDICADA AO REVERENDISSIMO PADRE

FR. **JOZE' DE JESUS
M A R I A,**

*Dignissimo Provincial Apostolico da mesma
Provincia.*



LISBOA OCCIDENTAL.

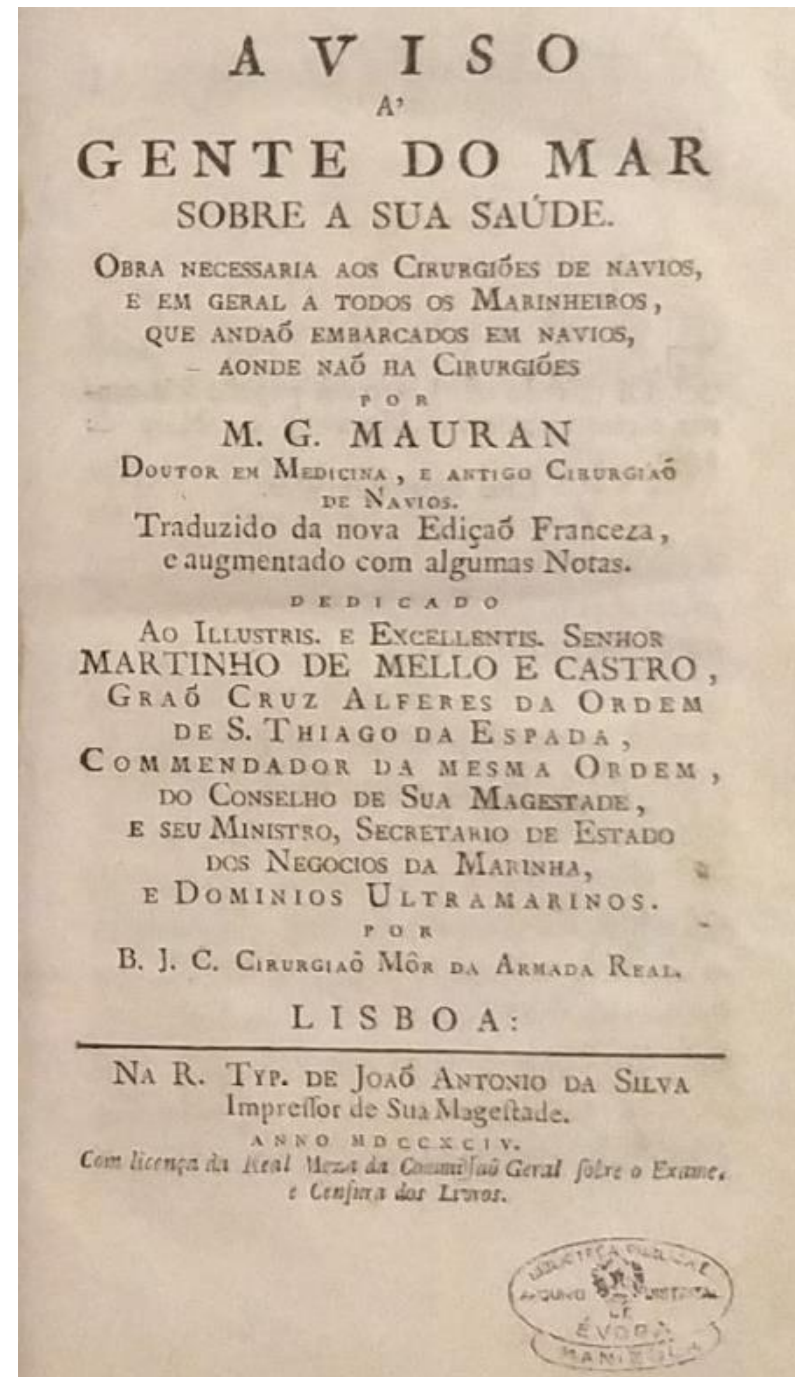
L. OF. J. MICHEL MANESCAL DA COSTA

Diogo de Sant-lago

- 1741
- Postilla religiosa, e arte de enfermeiros
(Religious Postil, and Art of Nurses)
- First of its kind ever written in Portuguese

G. Mauran

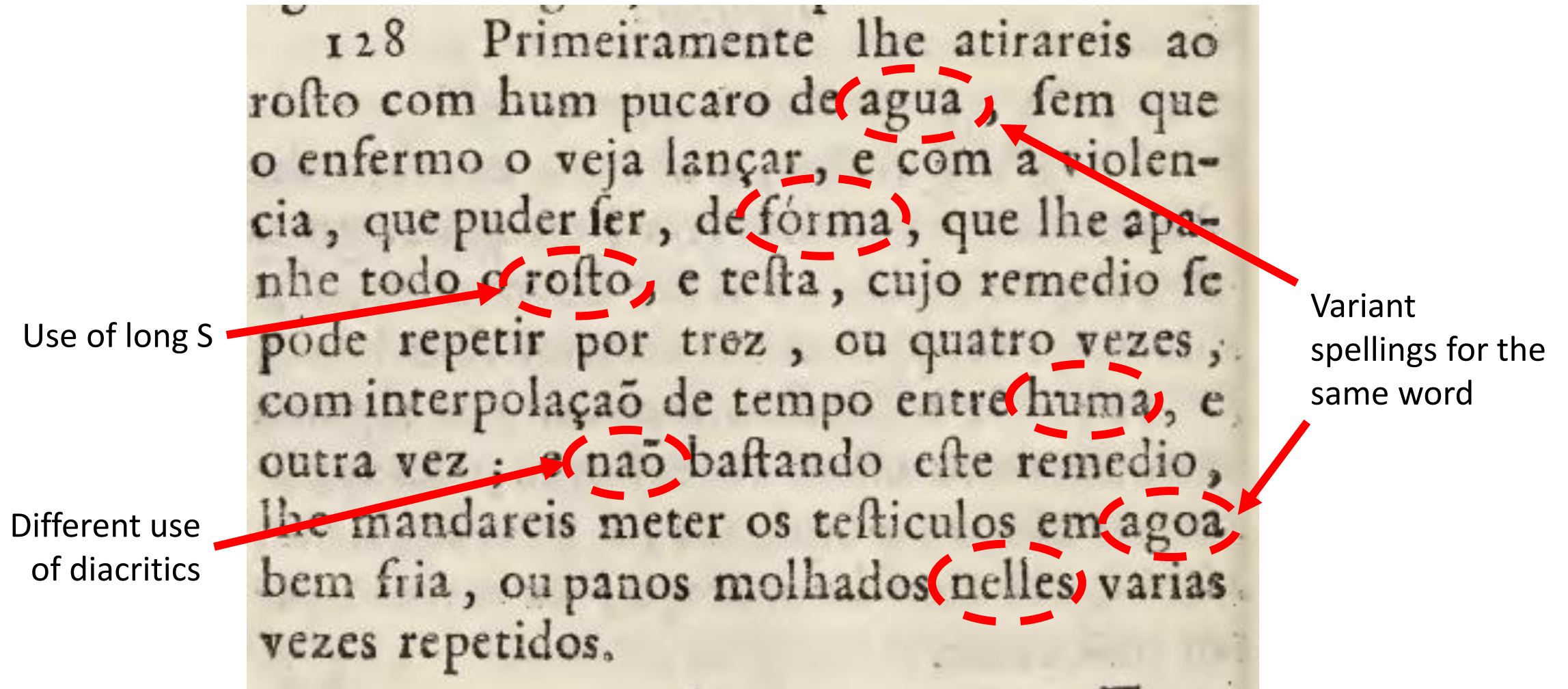
- 1786 – Translated into Portuguese in 1794
- Aviso a' gente do mar sobre sua saúde
(Advice to Sea People about their Health)
- Book translated from French by the High Surgeon of the Royal Armada, Bernardo José de Carvalho





Working with historical documents

Challenges for transcribing documents



Challenges for transcribing documents

- These are problems that appear when you can actually read the text...

*medici exaltabit caput iñim. &
in conspectu magnarum collam-
dabitur.*

...continuamente estão perdendo na accão do tra-
balho, he possível pois que possão resistir a
tantas fadigas? os homens os mais robustos,
bem depressa se vêm exauridos, suas forças
se diminuem, seu corpo se enfraquece, accres-

238 **S**E o Medico não determinar a
hora, que o enfermo ha de to-
mar a lexandria, lhe dareis a comer o af-
fucar rosado na cresecença do dia, das oito
para

» ção nestes vasos. Tambem se pode servir
» de pipas forradas interiormente de laminas
» de chumbo, ou de estanho, que não deixão

to o humor maligno estivesse dentro no corpo) pade-
ceria o doente as queixas referidas.

*ta causa fosse desprezada de todos; nesta desconfiança,
& tenor me alentou muito a lembrança de que desde o*

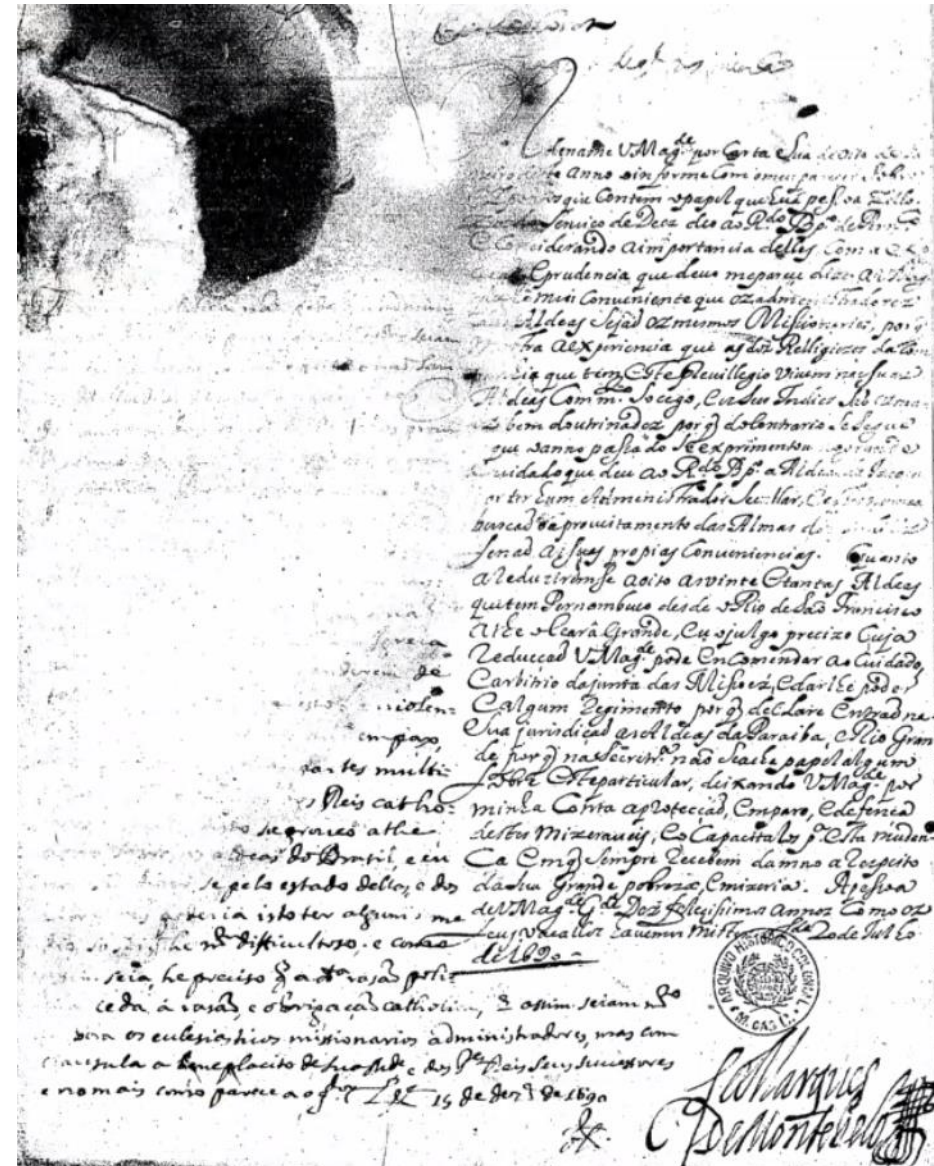
*palavras, & no rigor, & feveridade, com que trata-
dos doentes fastientos. Confesso que não posso sofrer*

*sua cura em vinte e quatro horas lavar bem a
ferida com agoa, chegar bem as bordas hu-*

*aço à Madre Dona Francisca de Sá Coutinho, Reli-
giosa no Convento de Santa Monica: & à mulher de
hum entalhador, morador na rua da Oliveira. Duzen-
tos banhos dei a hũa Religiosa do Salvador, filha do:*

- And we are still only looking at printed texts.

Challenges for transcribing documents



Named-entity recognition

- 2nd DhandNLP – 2022
- Zilio, L., Finatto, M.J.B. and Vieira, R., 2022. Named Entity Recognition Applied to Portuguese Texts from the XVIII Century. *Proceedings of the Second Workshop on Digital Humanities and Natural Language Processing* (2nd DHandNLP 2022), p. 1-10.
- Main goal: test off-the-shelf NER models on historical documents

Named-entity recognition

- We created a small corpus: 3 texts with original and normalised / modernised spellings
 - The modernised version was manually annotated with NEs
- We tested three NER models on these texts and compared the results
- BERT-CRF had the best performance in both original and modernised texts
- We fully annotated Semedo's book with BERT-CRF
 - https://github.com/uebelsetzer/NER_for_Portuguese_XVIII-Century_Texts/



Complexity of historical documents

- Domínios de Lingu@gem – 2023
- Zilio, L., Finatto, M.J.B., Vieira, R. and Quaresma, P. 2023. A natural language processing approach to complexity assessment of 18th-century health literature. *Domínios de Lingu@agem*, v. 17.
- Main goal: exploratory analysis of the relative textual complexity of an historical medical document

Complexity of historical documents

- We again used a small corpus: 3 texts with original and normalised / modernised spellings
- NILC-Metrix was used to analyse the complexity of each text:
 - 149 complexity metrics
- Semedo's text was found to have similar complexity to a sermon by Fr. Antonio Vieira, who is an exponent of the baroque writing style
- Information and data:
 - https://github.com/uebelsetzer/complexity_18th-century_texts_PT/

Normalisation of historical documents



Why normalise?

Applying NLP tools to raw transcribed historical documents usually yield poor results



What does it look like?

Normalisation of historical documents

Source: *Observações medicas doutrinaes de cem casos gravíssimos*, by João Curvo Semedo, 1707

OBSERVAÇÃO XLV.

De hum mercador, a quem repentinamente assaltou hum
ma dor de colica tão intoleravel, que estando na mesa
sacramental para commungar, o não pode fazer; &
sendo eu chamado, conheci dos grandíssimos ardores,
& continuos desejos de urinar, & vomitar, das picadas
da bexiga, & do adormecimento da perna direita,
que a tal dor era nephritica; para cujo remedio appliquei
quei hum vomitorio de tres onças de agua benedicta
vigorada, & tres ajudas feitas de cozimento de rim de
vacca, misturando em cada hum, hum onça de terebinthina
de beta lavada em agua de malvas, até se fazer
muito branca, com hum gema de ovo crua; metendo
depois disto ao doente em hum banho de agua, & primeiro
fosse cozida com hum arratel de amendoas doces bem
pisadas; & foram estes remedios tão maravilhosamente
sucedidos, que dentro de quatro horas deitou muitas
pedras redondas do tamanho de grãos de pimenta, & no mesmo dia se tiraram as dores, & ficou
são.

Original

1 OBSERVAÇÃO XLV.

2

3 De um mercador, a quem repentinamente assaltou uma
dor de cólica tão intolerável, que estando na fé
sacramental para comungar, o não pode fazer; e sendo
eu chamado, conheci dos grandíssimos ardores; e
contínuos desejos de urinar, e vomitar, das picadas
da bexiga, e do adormecimento da perna direita, que a
tal dor era nefrítica; para cujo remédio apliquei um
vomitório de três onças de água benedita vigorada, e
três ajudas feitas de cozimento de rim de vaca,
misturando em cada uma, uma onça de terebintina de
beta lavada em água de malvas, até se fazer muito
branca, com uma gema de ovo crua; metendo depois
disto ao doente em um banho de água, que primeiro
fosse cozida com um arrátel de amêndoas doces bem
pisadas; e foram estes remédios tão maravilhosamente
sucedidos, que dentro de quatro horas deitou muitas
pedras redondas do tamanho de grãos de pimenta, e no
mesmo dia se tiraram as dores, e ficou são.

Normalised

Normalisation of historical documents

OBSERVAÇAM XLV.

De hum mercador , a quem repentinamente assaltou huma dor de colica taõ intoleravel , que estando na fé sacramental para commungar , o não pode fazer ; e sendo eu chamado , conheci dos grandissimos ardores ; e continuos desejos de ourinar , e vomitar , das picadas da bexiga , e do adormecimento da perna direita , que a tal dor era nephritica ; para cujo remedio appliquei hum vomitorio de tres onças de agua benedicta vigorada , e tres ajudas feitas de cozimento de rim de vacca , [...]

Source: *Observações medicas doutrinaes de cem casos gravíssimos*, by João Curvo Semedo, 1707

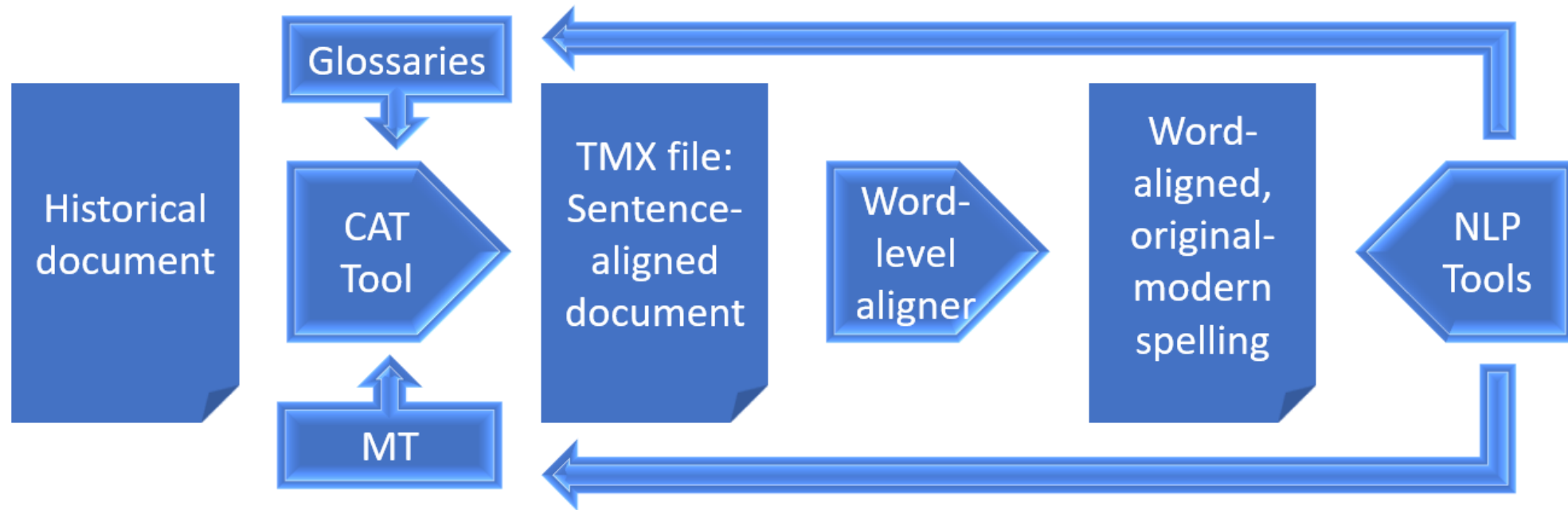
Normalisation of historical documents

OBSERVAÇÃO XLV.

De **um** mercador, a quem repentinamente assaltou **uma** dor de **cólica tão intolerável**, que estando na fé sacramental para **comungar**, o **não** pode fazer; e sendo eu chamado, conheci dos **grandíssimos** ardores; e **contínuos** desejos de **urinar**, e vomitar, das picadas da bexiga, e do adormecimento da perna direita, que a tal dor era **nefrítica**; para cujo **remédio apliquei um vomitório** de **três** onças de água **benedita** vigorada, e **três** ajudas feitas de cozimento de rim de **vaca**, [...]

Source: *Observações medicas doutrinaes de cem casos gravíssimos*, by João Curvo Semedo, 1707

Normalisation pipeline

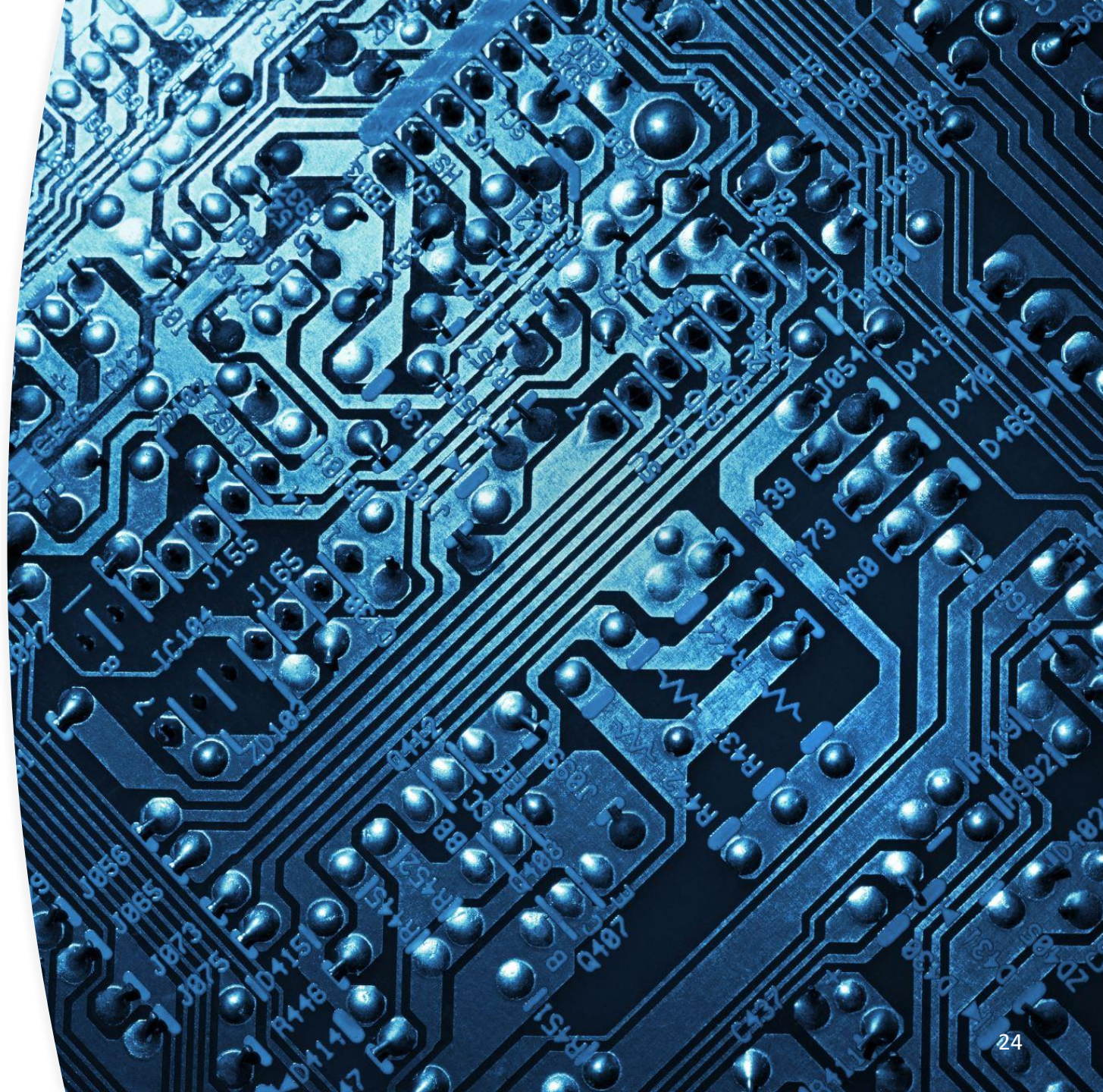


Current and future experiments



Neural Machine Translation for Normalisation

- Using a fairly small sample (20k tokens - ~300 sentences), fine-tune an NMT model for automatically normalising Portuguese texts
- Move from manual normalisation assisted by computers to post-editing NMT
- Promising preliminary results for medical documents



Hyperdictionary of 18th-century medicine

- Contextualising historical medical terms: lambedor, ajuda, rosálgar
- Linking 18th-century medicine to modern-day knowledge



Curiosities from the Brazilian arid region

News from 2018

Maria das Bonecas, a farmer from Mossoró*, teaches how to make a *lambedor* that helps with several health issues.

* City in Rio Grande do Norte, Brazil

g1 | ge | gshow | videos ASSINE JÁ MINHA CONTA E-MAIL

u g1 RIO GRANDE DO NORTE INTER TV

Conheça a receita do lambedor, remédio caseiro tradicional do sertão potiguar

Agricultora em Mossoró, dona Maria das Bonecas ensina a fazer um lambedor que ajuda a resolver vários problemas de saúde.

Por G1 RN
29/07/2018 09h40 · Atualizado há 3 anos

f t w i s



<https://g1.globo.com/rn/rio-grande-do-norte/noticia/2018/07/29/conheca-a-receita-do-lambedor-remedio-caseiro-tradicional-do-sertao-potiguar.ghtml>

Lambedor: a primitive kind of therapeutic food in sachets

ATPU

Alimento
Terapêutico
Pronto
para
Uso

RUTF

Ready-
to-
Use
Therapeutic
Food



Evidências confiáveis.
Decisões bem informadas.
Melhor saúde.

Busca...



Nossas evidências

Quem somos

Junte-se à Cochrane

News and jobs

Biblioteca Cochrane



Recursos sobre Coronavírus (COVID-19)

Alimento terapêutico pronto para uso no tratamento domiciliar de crianças de seis meses a cinco anos de idade com desnutrição aguda grave.

Publicada:
15 May 2019

Autores:
Schoonees A, Lombard MJ,
Musekiwa A, Nel E, Volmink J

Grupo de Revisão Principal:
Developmental, Psychosocial and
Learning Problems Group

Veja a revisão completa na
Biblioteca Cochrane



Imprimir



Introdução

As crianças desnutridas têm uma aparência muito magra e um alto risco de morrer ou de adoecer. Nas zonas rurais, o tratamento hospitalar de crianças com desnutrição aguda grave nem sempre é desejável ou prático. Nesses casos, o tratamento domiciliar pode ser melhor. O tratamento domiciliar pode ser feito de duas maneiras. A primeira consiste em dar a essas crianças alimentos preparados pelos responsáveis como mingau de farinha ou alimentos altamente energéticos ou ricos em nutrientes. A segunda maneira de tratar essas crianças é dando a elas um alimento terapêutico pronto para uso (ATPU) fornecido por uma unidade de saúde. O ATPU tem uma composição padronizada altamente calórica que é estabelecida pela Organização Mundial da Saúde (OMS). Os ingredientes habituais do ATPU incluem leite em pó, açúcar, manteiga de amendoim, óleo vegetal, vitaminas e minerais. Porém, os ingredientes dependem de fatores como aceitação, custos e disponibilidade local dos mesmos. O ATPU pode ser conservado sem refrigeração por um longo período e é servido sem nenhum preparo. Esta é uma atualização de uma revisão anterior que, baseada nos quatro estudos disponíveis na época, não chegou a conclusões definitivas sobre os efeitos dos ATPU.

Am score 18

Quem está falando sobre esse artigo?

Vídeo: O que são
Revisões Sistemáticas

Nossa evidência em
saúde - como ela pode
ajudar-lo?

Recursos em diferentes
idiomas

Hyperdictionary of 18th-century medicine

- Lambedor:
 - a mixture/potion to cure something;
 - a kind of therapeutic food



3rd DHandNLP

Third Workshop on Digital Humanities and Natural Language Processing

Co-located with PROPOR 2024

Santiago de Compostela, Spain

14-15 March 2024

<https://sites.google.com/view/dhandnlp-propor>

CfP deadline: 20 January 2024

Thanks for your
attention! 😊

leonardo.zilio@fau.de
mariafinatto@gmail.com