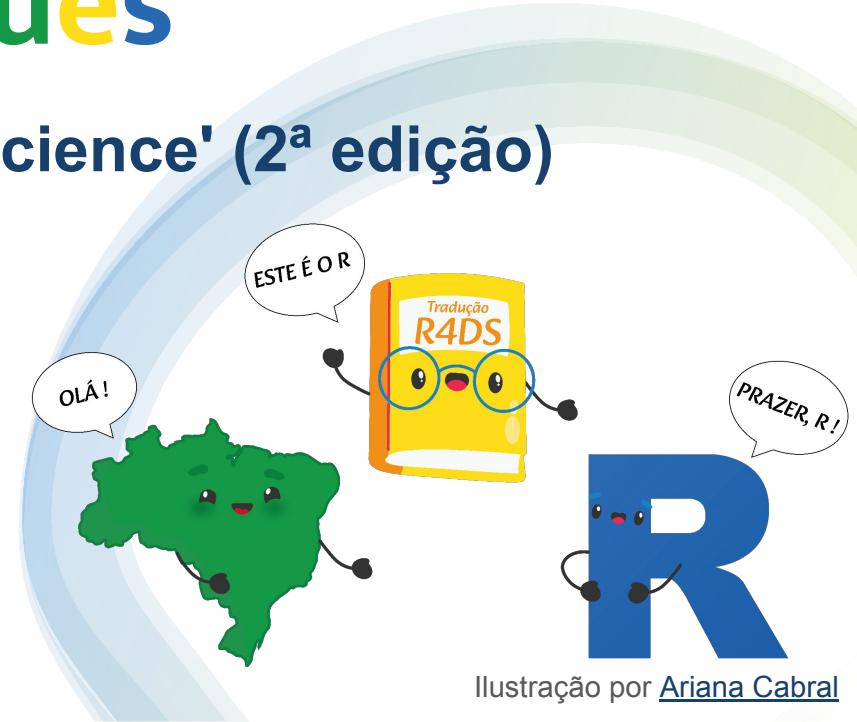


Democratizando a Ciência de Dados em Português

Uma Tradução de 'R for Data Science' (2^a edição)
Liderada pela Comunidade



Motivação

Motivação

A comunidade R oferece muitos recursos gratuitos sobre R e ciência de dados em inglês (blogs, tutoriais, livros, vídeos, etc).

Big Book of R

Welcome :)

AUTHOR
Oscar Baruffa

PUBLISHED
2024-10-20

1 Your last-ever bookmark

Thanks for stopping by. If you're like me, you can't help but bookmark every R-related programming book you find in the hopes that one day you, or someone you know, might find it useful.

Hopefully this is the only bookmark you'll need in future :).

When I initially released this collection in late August 2020, it contained about 100 books that I'd been collecting over the previous two years. Since then I've found a few more and there have been contributions from many people. The collection now stands at over 400 free, open-sourced books.

Most of these are free. Some are paid but usually quite affordable.

1.1 Searching

If there's something specific you're looking for, use the menu or search using the magnifying glass icon at the top of the screen.

1.2 Supporting

Table of contents

- 1 Your last-ever bookmark**
 - 1.1 Searching
 - 1.2 Supporting
 - 1.3 Contributing
 - 1.4 Big thanks to Fathom Data
 - 1.5 License
 - 1.6 Live stats
 - 1.7 About me



Exemplo: **Big Book of R**
+ de 400 livros gratuitos

Motivação: Contexto Brasileiro

- População: > 212 milhões de habitantes
- País de baixa proficiência em inglês -> maioria da população sem nível intermediário ou superior*



* Fonte: [EF English Proficiency Index EPI 2023 - Ranking of Countries and Regions](#)

Motivação: Tradução para PT-BR

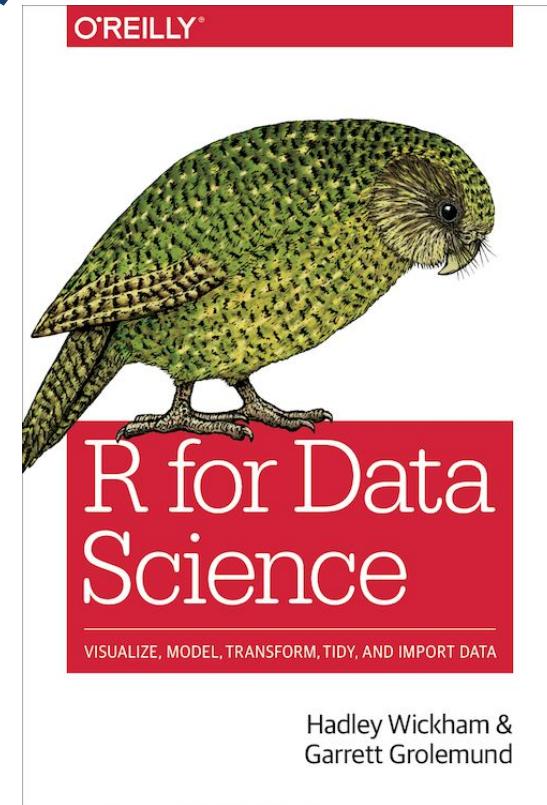
- Ampliar o acesso a conteúdos técnicos,
- Democratizar a disseminação de conhecimento,
- Fortalecer a colaboração entre grupos diversos,
- Criar oportunidades para aproveitar avanços em ciência de dados.



Ilustração por [Ariana Cabral](#)

R for Data Science (1ed)

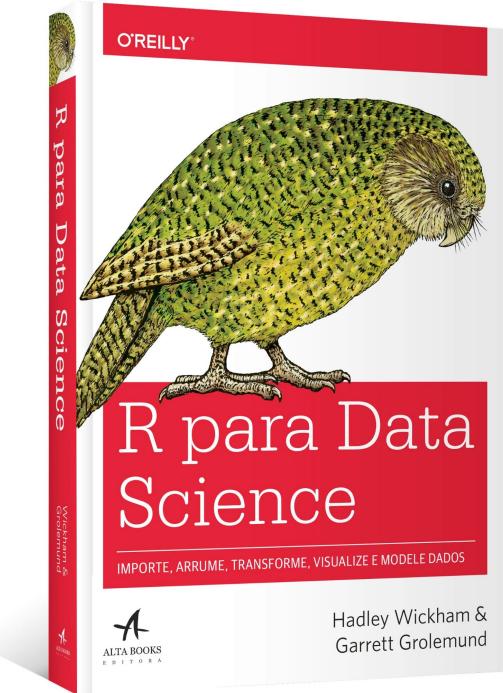
- Excelente recurso sobre ciência de dados com R e o tidyverse.
- Disponível gratuitamente online e também à venda em formato físico.



R para Data Science (1 edição)

Primeira edição traduzida para o português por outra editora:

- Não disponibilizada gratuitamente online



R para Ciencia de Datos!

Primeira edição traduzida para o espanhol por voluntários da comunidade latino-americana: gratuita e online (<https://es.r4ds.hadley.nz/>).



Colaboração



Pacote {dados} - 2020

- > Disponível no CRAN
- > Coordenação:
 - Beatriz Milz
 - Riva Quiroga

dados 0.1.9000 Reference Changelog



dados 

Este pacote disponibiliza a tradução de conjuntos de dados em inglês originalmente disponíveis em outros pacotes de R. Os dados traduzidos são utilizados nos exemplos do livro *R for Data Science* de Hadley Wickham & Garrett Grolemund, cuja [versão em português](#) foi publicada por AltaBooks & O'Reilly.

Instalação

Este pacote pode ser instalado a partir do CRAN:

```
install.packages("dados")
```

A versão de desenvolvimento pode ser instalada a partir do GitHub:

```
# install.packages("remotes")
remotes::install_github("cienciadedados/dados")
```

Traduções

As traduções disponíveis dentro de `dados` são as seguintes:

Nome	Título	Pacote	Dataset
casas	Dados de habitação em Ames	AmesHousing	ames_raw
bebés	Nomes de bebês	babynames	babynames
velho_fiel	Dados do gêiser Velho Fiel (Old Faithful)	datasets	faithful
dados_iris	Dados de espécies de flor de íris por Edgar Anderson	datasets	iris
mtcars	Testes de estrada para automóveis	datasets	mtcars

Links

- [View on CRAN](#)
- [Browse source code](#)
- [Report a bug](#)

License

- [Full license](#)
- [GPL-3](#)

Citation

- [Citing dados](#)

Developers

- Riva Quiroga
Author, maintainer 
- Sara Mortara
Author 
- Beatriz Milz
Author 
- Andrea Sánchez-Tapia
Author 
- Alejandra Andrea Tapia Silva
Author 
- Beatriz Maurer Costa
Author
- Jean Prado
Author 
- Renata Hirota
Author 
- William Amorim
Author 
- Emmanuelle Rodrigues Nunes



```
# Para instalar o pacote
install.packages("dados")
# Para usar o pacote
library(dados)

# Exemplo dos dados da base Pinguins
> pinguins
# A tibble: 344 × 8
  especie      ilha comprimento_bico profundidade_bico comprimento_nadadeira
  <fct>       <fct>     <dbl>            <dbl>                  <int>
1 Pinguim-de-adé... Torg...      39.1            18.7                  181
2 Pinguim-de-adé... Torg...      39.5            17.4                  186
3 Pinguim-de-adé... Torg...      40.3             18                   195
4 Pinguim-de-adé... Torg...       NA              NA                   NA
5 Pinguim-de-adé... Torg...      36.7            19.3                  193
6 Pinguim-de-adé... Torg...      39.3            20.6                  190
7 Pinguim-de-adé... Torg...      38.9            17.8                  181
8 Pinguim-de-adé... Torg...      39.2            19.6                  195
9 Pinguim-de-adé... Torg...      34.1            18.1                  193
10 Pinguim-de-adé... Torg...      42              20.2                 190
# i 334 more rows
# i 3 more variables: massa_corporal <int>, sexo <fct>, ano <int>
# i Use `print(n = ...)` to see more rows
```

Pacote {dados}

> 11 pessoas participaram



Clube do livro: R Ladies SP (2021)

- Reuniões para apresentar capítulos do livro, tirar dúvidas, conversar sobre o conteúdo
- A Bea 🕵️ viu no GitHub que a segunda edição do livro estava sendo desenvolvida 💬



Perguntamos ao Hadley... (2021)

Translation of R4DS 2ed in Portuguese #955



beatrizmilz opened this issue on May 21, 2021 · 34 comments



beatrizmilz commented on May 21, 2021

...

Hi Hadley! I meant to send this on Twitter but I can't send you DM's. Let me know if there is a better place to send this message other than an GitHub issue.

Are there any plans for the translations for the 2nd edition yet? <https://r4ds.hadley.nz/>

I would like to know if there is a chance that this edition could be translated to Brazilian Portuguese by the community (when the book is ready).

The first edition was translated by Alta Books, but there are some downsides on this version:

- is not available free and online.
- all the graphic elements are black and white, so reading the chapter for ggplot2 is not really clear
- I did not buy this version, I have the English version by O'Reilly, but people of the community said that some parts of the translations are not really clear/accurate.

I'm part of the group that translated the datasets used in the book to Brazilian Portuguese: <https://cienciadedatos.github.io/dados/>. This group (and more people) are really interested in helping to translate to Portuguese in a way that the book could be available for free online, in a similar way that the community translated to Spanish (@rivaquiroga <3).

Let me know what are your thoughts, and thank you so much for your attention.



[github.com/
hadley/r4ds/
issues/955](https://github.com/hadley/r4ds/issues/955)

2021 - 2022



hadley commented on Dec 1, 2021

Owner ...

Unfortunately O'Reilly owns the translation rights, and there's not much I can do about the first edition. However, I can see if they might be open to a community translation for the second edition.



2



4



hadley commented on Feb 22, 2022

Owner ...

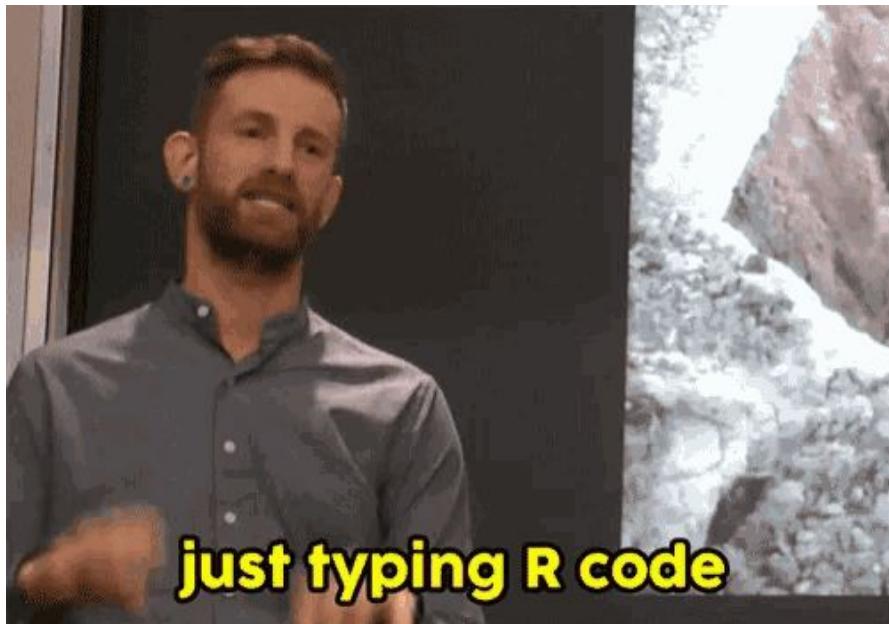
I have confirmed with O'Reilly that we can do a community translation of the 2nd edition 😊



8

github.com/hadley/r4ds/issues/955

Escrevendo o livro.... 2021-2023



<https://beamilz.com/posts/2022-08-12-rstudio-conf-pt/>

2023



hadley commented on Jun 9, 2023

Owner ...

R4DS 2e is now finished and off to the printers! Would you like me to get permission from O'Reilly for an open-access Portuguese translation?



hadley commented on Sep 4, 2023

Owner ...

Just got permission! 😊

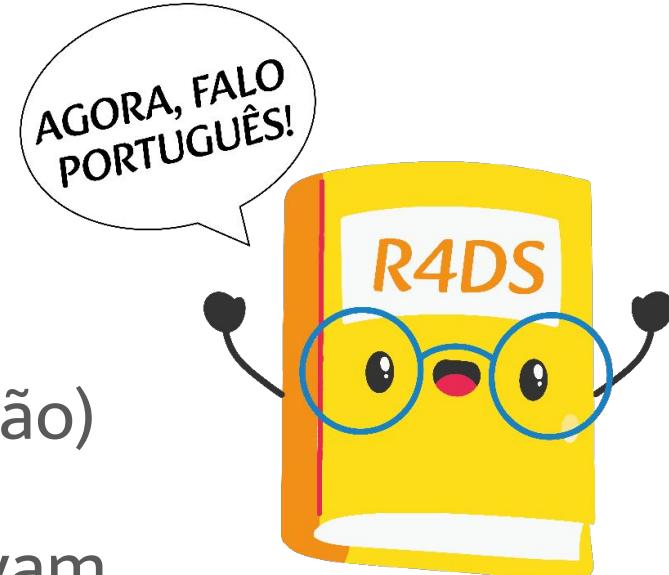


github.com/hadley/r4ds/issues/955

Projeto de tradução

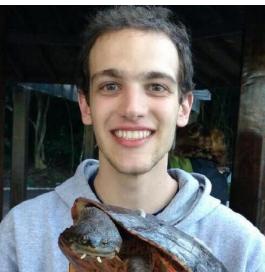
Tradução - A partir de Set/2023

- Coordenação: Beatriz Milz
- Orientação: Riva Quiroga
- > 20 participantes (tradução, revisão)
- 1 pessoa traduzia, 2 pessoas revisavam



Equipe: pessoas de diversas áreas!

> 22 pessoas participaram

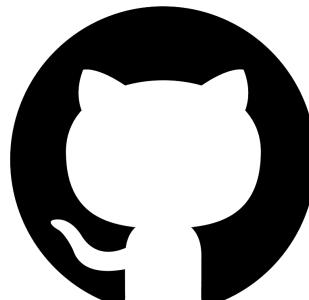


Equipe: pessoas de diversas áreas!

> 22 pessoas participaram



Ferramentas



Repositório (fork do repositório do livro)

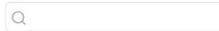
Issues (cada capítulo tem uma issue)

Pull requests (envio da tradução)

Projects (acompanhamento)

Guia de tradução

Guia de Tradução de conteúdos técnicos sobre R para Português Brasileiro



- 1 Introdução
- 2 Equipes de tradução
- 3 Participe de projetos de tradução/revisão
- 4 Bases de dados
- 5 Orientações para tradução
- Traduzindo via GitHub
 - 6 Conceitos importantes sobre Git e GitHub
 - 7 Revisando via GitHub
 - 8 Preparando o GitHub para realizar a tradução
 - 9 Configurando o ambiente local com o RStudio
 - 10 Traduzindo um capítulo
 - 11 Incorporando as sugestões da revisão no capítulo traduzido
 - 12 Sugerindo correções
 - Traduzindo via Weblate
 - 13 Weblate

Traduzindo via GitHub > 7 Revisando via GitHub

7 Revisando via GitHub

Nesta seção, explicaremos como as revisões podem ser feitas via GitHub.

A etapa de revisão é muito mais simples do que a tradução! Portanto, caso você esteja contribuindo com o projeto pela primeira vez, recomendamos que você comece revisando um capítulo.

É importante que você leia as orientações em [Capítulo 5](#), para que as traduções sigam os mesmos padrões.

Primeiramente, você precisará de uma conta no GitHub. Caso você ainda não tenha, [crie uma conta](#).

7.1 Escolhendo um capítulo para revisar

Para iniciar a revisão, você precisará escolher um capítulo para revisar. Para isso, você pode verificar a página de Pull Requests do repositório.

Author	Label	Milestones	Reviews	Assignee	Sort
lepr			5		
lidianvalada			18		
hallan-vilela			129		
neimarta			48		

<https://cienciadedados.github.io/guia-traducao-ptbr/revisando-github.html>

Guia de tradução facilitou a participação de pessoas que não possuíam experiência prévia com GitHub!

Livro R for Data Science 2ed

R para Ciência de Dados (2^a edição)



Boas-vindas

Prefácio da segunda edição

Introdução

Visão geral



1 Visualização de dados

2 Fluxo de Trabalho: básico

3 Transformação de dados

4 Fluxo de trabalho: estilo de código

5 Organização de dados (*data tidying*)

6 Fluxo de trabalho: scripts e projetos

7 Importação de dados

8 Fluxo de trabalho: obtendo ajuda

Visualizar



9 Camadas

10 Análise exploratória de dados

11 Comunicação

Transformar



12 Vetores lógicos

13 Números

R para Ciência de Dados (2^a edição)

Boas-vindas

Este é o website para a tradução em Português da [2^a edição do livro “R for Data Science”](#). Este livro vai te ensinar como fazer ciência de dados com R: você irá aprender como importar os seus dados para o R, arrumá-los em uma estrutura mais útil, transformá-los e visualizá-los.

Neste livro, você vai encontrar um conjunto de habilidades para a ciência de dados. Assim como estudantes de química aprendem a limpar tubos de ensaio e montar um laboratório, você vai aprender a limpar dados e fazer gráficos – além de muitas outras coisas. Essas são as habilidades que fazem a ciência de dados acontecer, e aqui você vai encontrar as melhores práticas para fazer cada uma dessas coisas com o R. Você vai aprender a usar a gramática dos gráficos, programação letrada (*literate programming*) e pesquisa reprodutível para economizar tempo. Você também vai aprender a gerenciar recursos cognitivos para facilitar descobertas ao lidar com, visualizar e explorar dados.

Este site é e sempre será gratuito, licenciado sob a Licença [CC BY-NC-ND 3.0](#). Se você gostaria de ter uma cópia física do livro, você pode comprar a versão em Inglês na [Amazon](#). Se você aprecia ler o livro gratuitamente e gostaria de retribuir, por favor faça uma doação para a [Kākāpō Recovery](#): o kākāpō (que aparece na capa do livro) é uma espécie de papagaio nativa da Nova Zelândia que está criticamente ameaçada de extinção; restam apenas 248 espécimes.

Você pode encontrar respostas sugeridas para exercícios no livro em <https://mine-cetinkaya-rundel.github.io/r4ds-solutions>.

Por favor, note que o R4DS utiliza um [Código de Conduta de Contribuidores](#). Ao contribuir para este livro, você concorda em seguir seus termos.

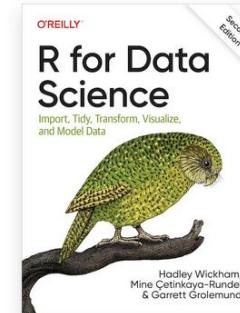
Índice

Boas-vindas

Agradecimentos

editar essa página

Criar uma issue



Como contribuir?

- > **Utilize, compartilhe, divulgue** o livro!
- > Caso encontre algum problema (erro de tradução, renderização, etc) **nos avise** para que possamos corrigir.

Desafios: atualização

- A versão original do livro tem recebido alterações (no geral, pequenas modificações)
- O objetivo é periodicamente utilizar as ferramentas do GitHub para comparar as versões (com a ferramenta `compare across forks`), enviar a contribuição (em inglês) para o repositório da tradução, e atualizar a versão em Português.

This branch is [54 commits ahead of](#), [74 commits behind](#) [hadley/r4ds:main](#).

 [Contribute](#) ▾

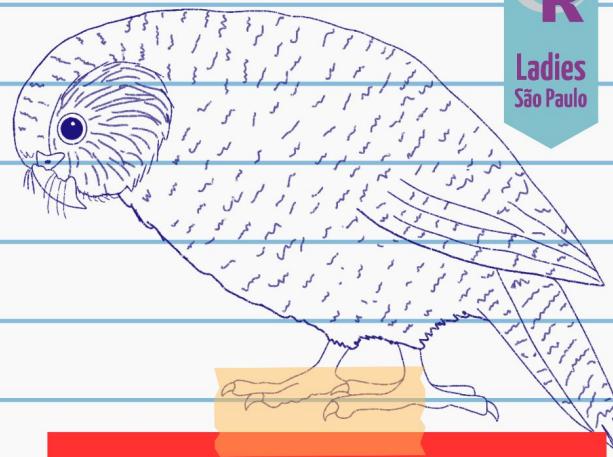
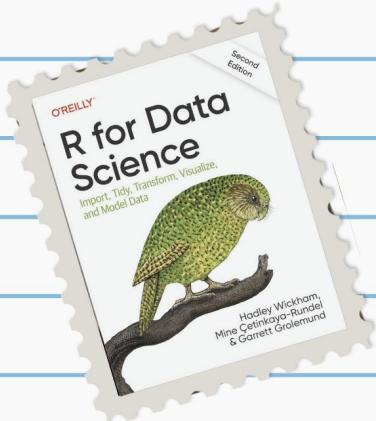
 [Sync fork](#) ▾

Continuidade e Divulgação

Clube do livro R-Ladies SP

**CLUBE DO LIVRO:
R FOR DATA SCIENCE 2ed**

Versão traduzida para Português



Gratuito e Online
Encontros mensais no Youtube

Clube do livro R-Ladies SP

Encontro 10: Visualização - Camadas (ggplot2) com Fernanda Peres - Clube do...
181 visualizações • Transmitido há 1 mês

Encontro 9: Importação de dados (pacotes tidyverse/readr) - Clube do Livro R for Data...
150 visualizações • Transmitido há 1 mês

Encontro 8: Organização de dados (pacote tidyrr) - Clube do Livro R for Data Science...
318 visualizações • Transmitido há 3 meses

Encontro 7: Transformação de dados (Pt. 3) - Pacote dplyr - Clube do Livro R for Data...
148 visualizações • Transmitido há 3 meses

Encontro 6: Transformação de dados (pacote dplyr) (Pt. 2) - Clube do Livro R for...
112 visualizações • Transmitido há 3 meses

Encontro 5: Transformação de dados (pacote dplyr) - Clube do Livro R for Data...
268 visualizações • Transmitido há 4 meses

Encontro 4: Fluxo de trabalho (Parte 2) - Clube do Livro R for Data Science (2ed)
179 visualizações • Transmitido há 5 meses

Encontro 3: Fluxo de trabalho - Clube do Livro R for Data Science (2ed)
298 visualizações • Transmitido há 6 meses

Encontro 2: Visualização - Clube do Livro R for Data Science (2ed) - Pacote ggplot2
479 visualizações • Transmitido há 7 meses

Encontro 1: Boas-vindas! Clube do Livro R for Data Science (2ed)
713 visualizações • Transmitido há 8 meses

<https://www.youtube.com/c/RLadiesSPaulostreams>

Perspectivas futuras

- Gerenciar e manter o projeto de forma colaborativa a longo prazo
- Disseminar a versão traduzida
- Ampliar o Guia de tradução para que seja usado em outros projetos de tradução
- Colaboração com outras traduções



Agradecemos pela
atenção!