



&DICOLAB

MILESTONE 4 - G6 - PI 24/25

Team: André Dora, António Alberto, Gonçalo Sousa,
Hugo Sousa, José Oliveira

Advisors: Osvaldo Rocha Pacheco, João Paulo Silvestre

Table of Content

01 Context

03 Functional Requirements

05 Edicolab before and after

02 State of Art

04 System's Architecture

06 Development challenges

Table of Content

- 07 Acknowledgements 08 Future Work
- 09 Conclusions

Context

- **iFORAL**

Project that aims to historical and linguistic analyse Portuguese forais.

- **Edicolab**

Platform for digital writing of critical editions of medieval manuscripts.

- **TEI-XML**

Standard based on XML for the digital representation of historical texts.

The problem

- **Lack of validation**
Edicolab lacked validation for its off-standard tagging
- **Furthermore**
Current Edicolab users asked us to improve editing experience

State of Art

	Oxygen XML Editor	Menota	Patrimonivm Editor	Edicolab V1	Edicolab V2
TEI Editor	✓	✗	✓	✓	✓
HD document viewer	✗	✓	✓	✓	✓
Glossary	✗	✗	✓	✓	✓
Paleographic and Critic view	✗	✗	✗	✓	✓
Custom Tags	✗	✗	✗	✓	✓
Semantic analysis	✗	✗	✓	✗	✓
Automatic Tag Generation	✗	✗	✗	✗	✓
Customizable Validation Schema	✗	✗	✗	✗	✓

Functional Requirements



Validation Schemas

Provide custom validation schemas when creating a document.



Automatically create tags

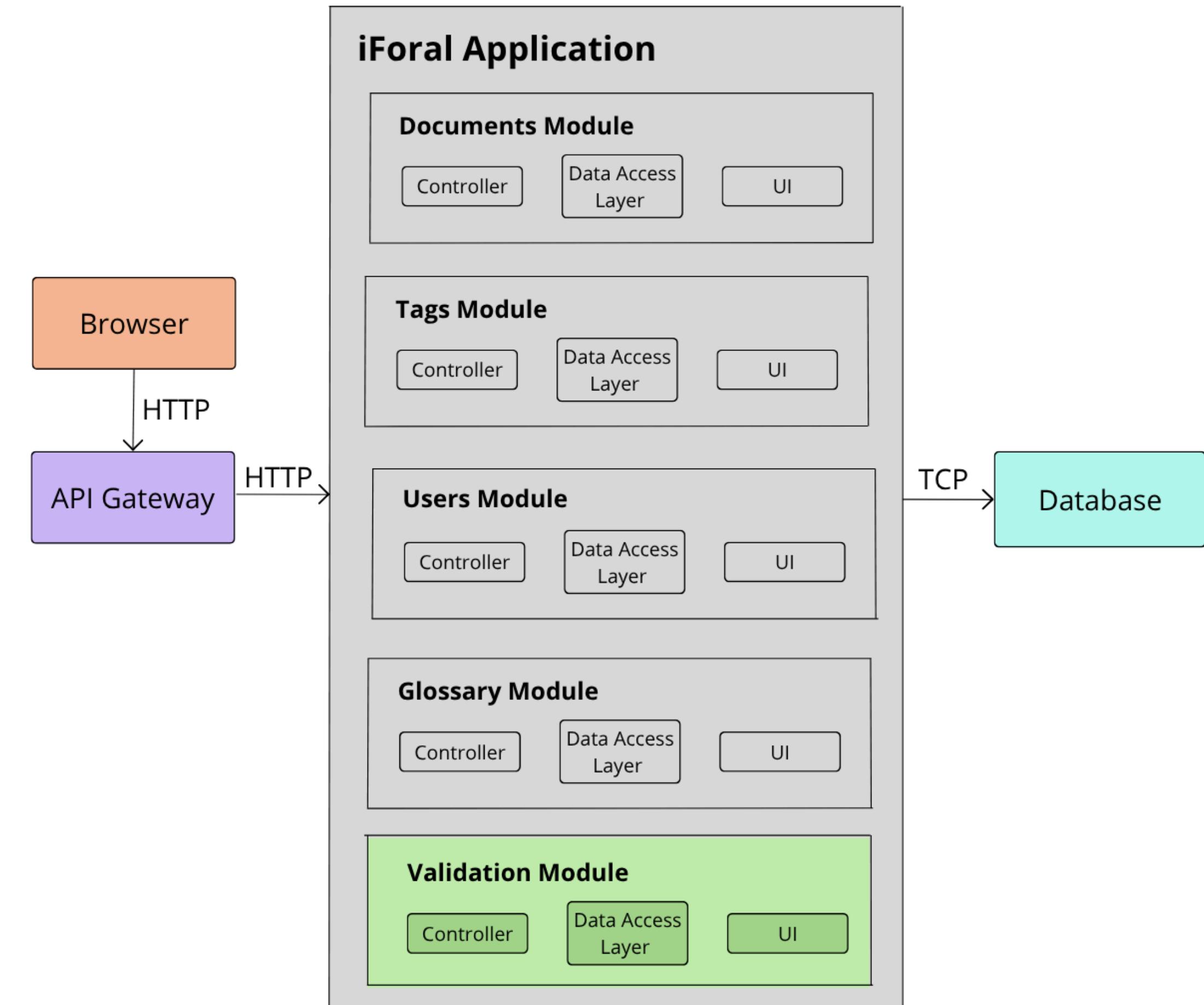
Automatically create tags that structurally depend on others.



Show markup errors

Error highlighting and explanation as each document is transcribed

System's Architecture



EDICOLAB



iForal Documentos Glossário

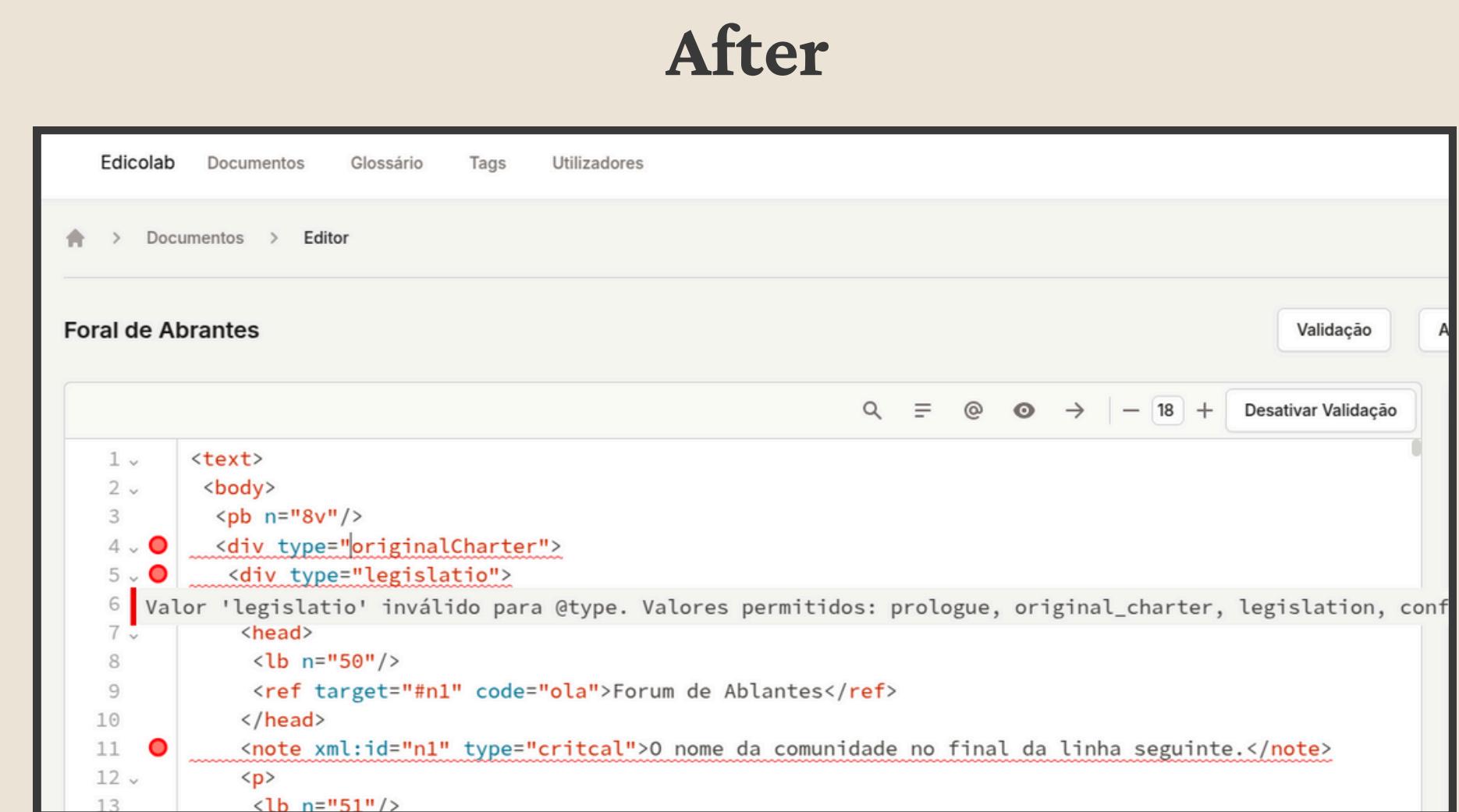
Documentos > Editor

Foral de Abrantes

Estrutura ▾ Acidentes ▾ Substituição ▾ Editor ▾ Abreviatura ▾ Nota ▾

```
1 <text>
2 <body>
3 <pb n="8v"/>
4 <div type="originalCharter">
5   <div type="legislatio">
6     <cb n="2"/>
7   <head>
8     <lb n="50"/>
9     <ref target="#n1">Forum de Ablantes</ref>
10    </head>
11    <note xml:id="n1" type="critcal">O nome da comunidade no final da linha seguinte.</note>
12    <p>
13      <lb n="51"/>
```

Before



Edicolab Documentos Glossário Tags Utilizadores

Documentos > Editor

Foral de Abrantes

Validação Desativar Validação

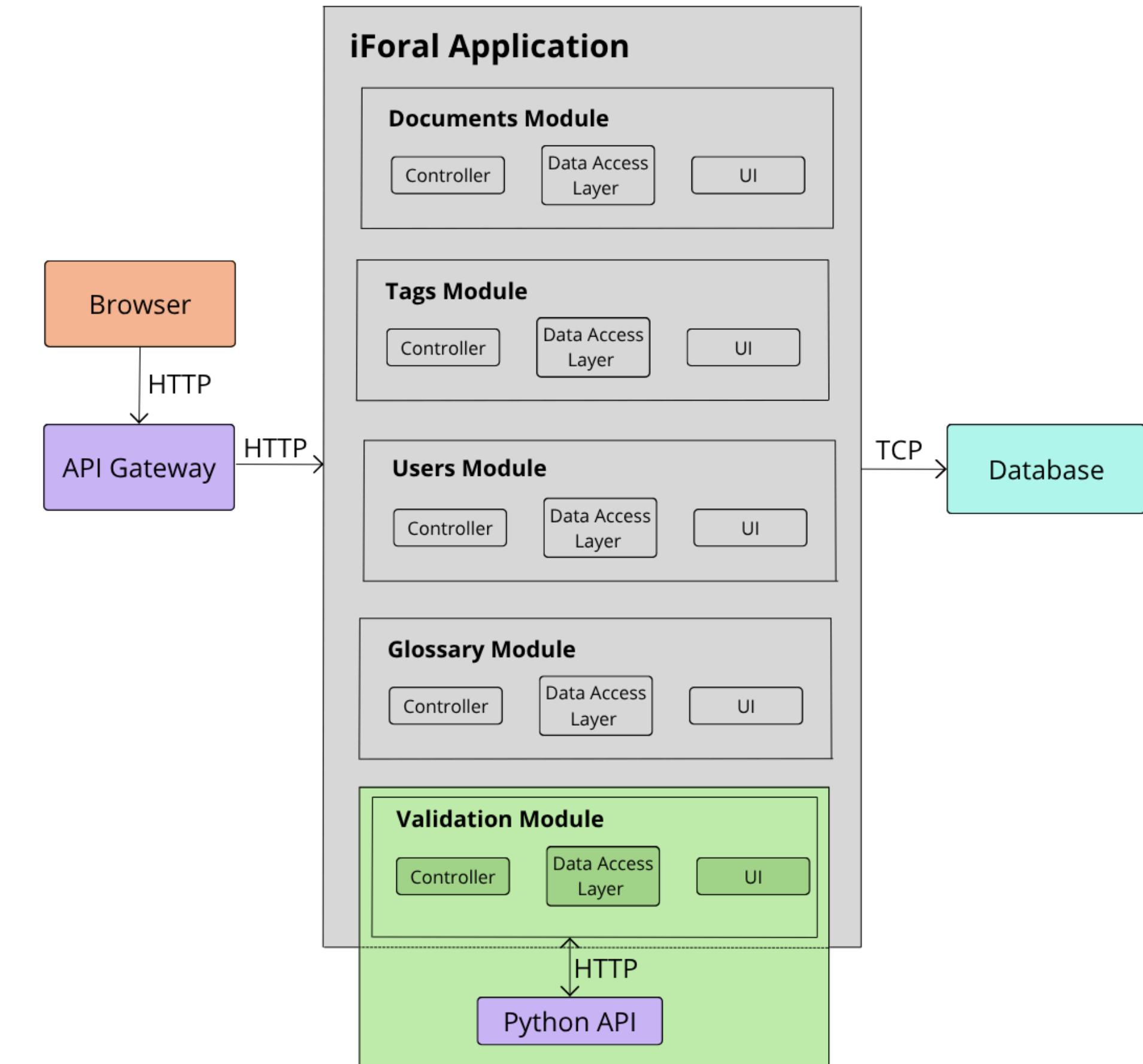
```
1 <text>
2 <body>
3 <pb n="8v"/>
4 <div type="originalCharter"> Valor 'legislatio' inválido para @type. Valores permitidos: prologue, original_charter, legislation, conf
5   <div type="legislatio">
6     <head>
7       <lb n="50"/>
8       <ref target="#n1" code="ola">Forum de Ablantes</ref>
9     </head>
10    <note xml:id="n1" type="critcal">O nome da comunidade no final da linha seguinte.</note>
11    <p>
12      <lb n="51"/>
```

After

Development challenges: first solution

External python API:

- Received the document;
- Applied the validation schema
- Returned the errors
- @codemirror/lint showed the errors



Development challenges: final solution

01 Database changes:

- Added “validationSchema: JSON” to Documents table
- Modified Tags entity to accomodate hierarchy

02 @codemirror/lint:

- Real-time validation

03 API updates:

- Endpoint added for validation schema customization
- POST request to save the validation schema as JSON for the specified document
- Updated Tags endpoints to accomodate the changes

Users Feedback

- Refinements in base validation schema
- Option to toggle real-time validation ON/OFF



Future Work

01

Control of document's version history & co-editing

Similar to Google Docs/Sheets/Slides

02

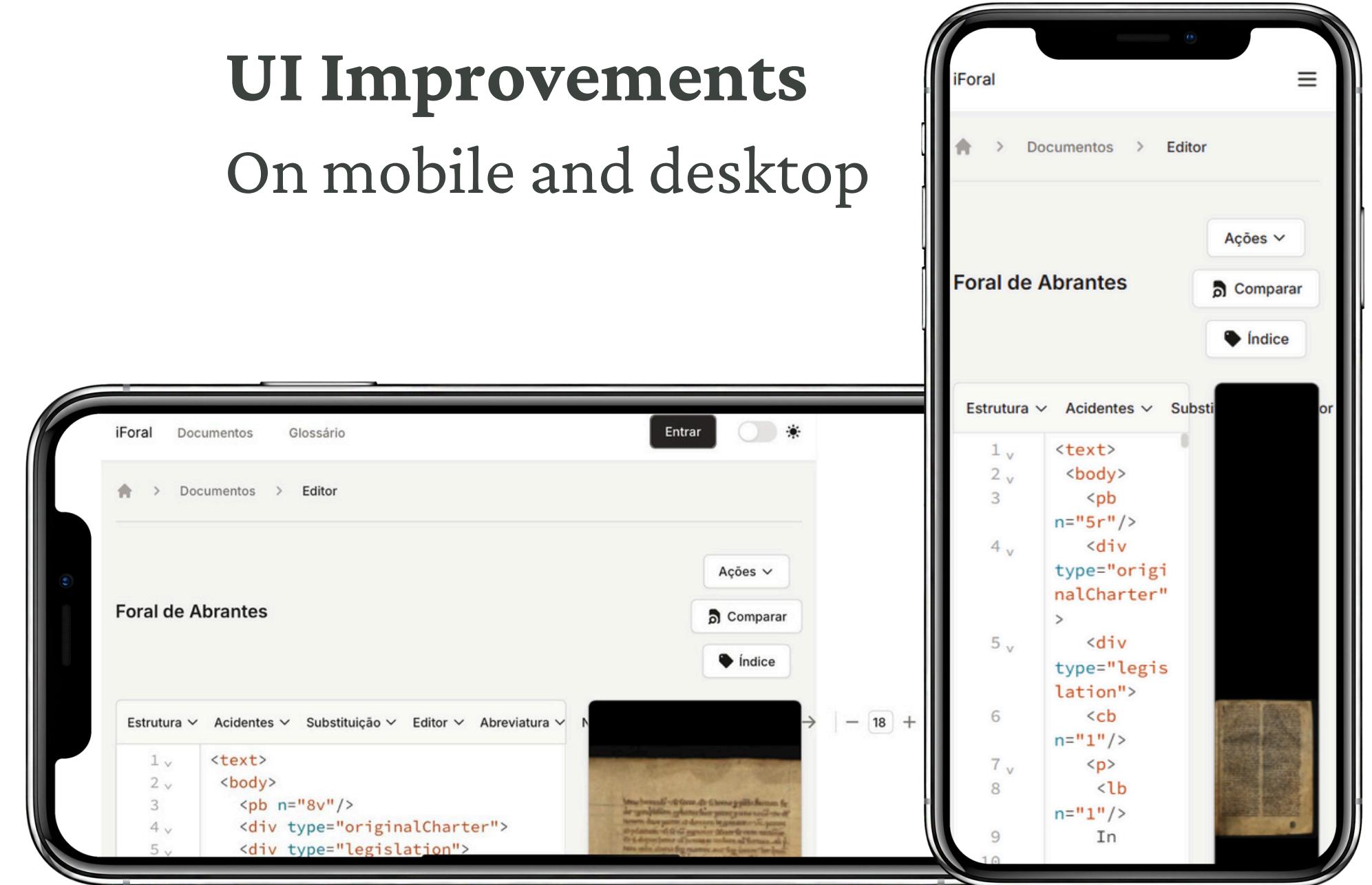
AI integration

To suggest tags based on context

Future Work

03

UI Improvements On mobile and desktop



Conclusions

■ Achieved Objectives

The objectives we set at the beginning of the project were successfully achieved

■ Lessons learned

Teamworking, learned new frameworks

Acknowledgements

- **Dr. Catarina Coelho**

Who explained us **Humanity linguists requirements** and the POV of our platform's usage

- **MSc Vasco Sousa**

Who explained us **how Edicolab was implemented** and helped us with **doubts regarding Software Engineering**

Thanks for
your attention



Documentation
Website