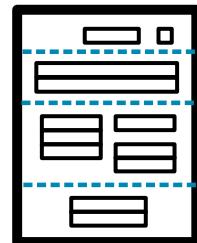


Document information

- * Text
- * Position and size of each line and column

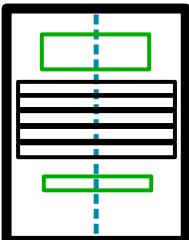
Split page in blocks



New block
New block
New block

Find centered blocks

Center



Use Regex to assign title level to centered blocks

Text...

Title I

Chapter I

Level 1

Level 2

Level 3

Walk the file in reading order and assemble the tree

Root

Root

Root

Root

ORGANIC

...

ORGANIC

...

Title I

ORGANIC

...

Title I

Chapter I

Article 1

The present
section ...