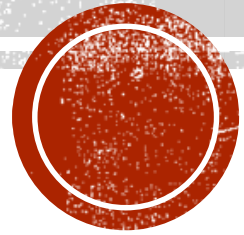
KEYWORD SPOTTING (IN THEORY)
QUERY BY EXAMPLE

Ioannis Pratikakis
Konstantinos Zagoris



DEMOCRITUS UNIVERSITY OF THRACE
Department of Electrical and Computer Engineering
Visual Computing Group

WORD SPOTTING – QUERY BY EXAMPLE

Word Spotting by Example can be conceived as the task of identifying locations on a document image which have high probability to correspond to an instance of a queried word-image, without explicitly recognizing it.

Segmentation-based **Segmentation-free**

Fundamental difference concerns the search space which could be either a set of **segmented word images** (segmentation-based approach) or the **complete document image** (segmentation-free approach)

The selection of the segmentation-based strategy is preferred when the **layout is simple enough to correctly segment the words** while the segmentation-free strategy performs better when there is **considerable degradation** on the document

SEGMENTATION-BASED VS. SEGMENTATION-FREE

Segmentation-based:

Suggested:

- the layout is simple enough to correctly segment the words

Advantages:

- Improved word-spotting effectiveness (when the layout is simple)
- Fast retrieval times

Disadvantages:

- Cannot handle degraded or complicated documents
- Detect only words

Segmentation-free:

Suggested:

- there is considerable degradation on the document

Advantages:

- Good handling of complex document layouts
- Ability to match partial words or phrases
- It can locate not only words but also symbols.

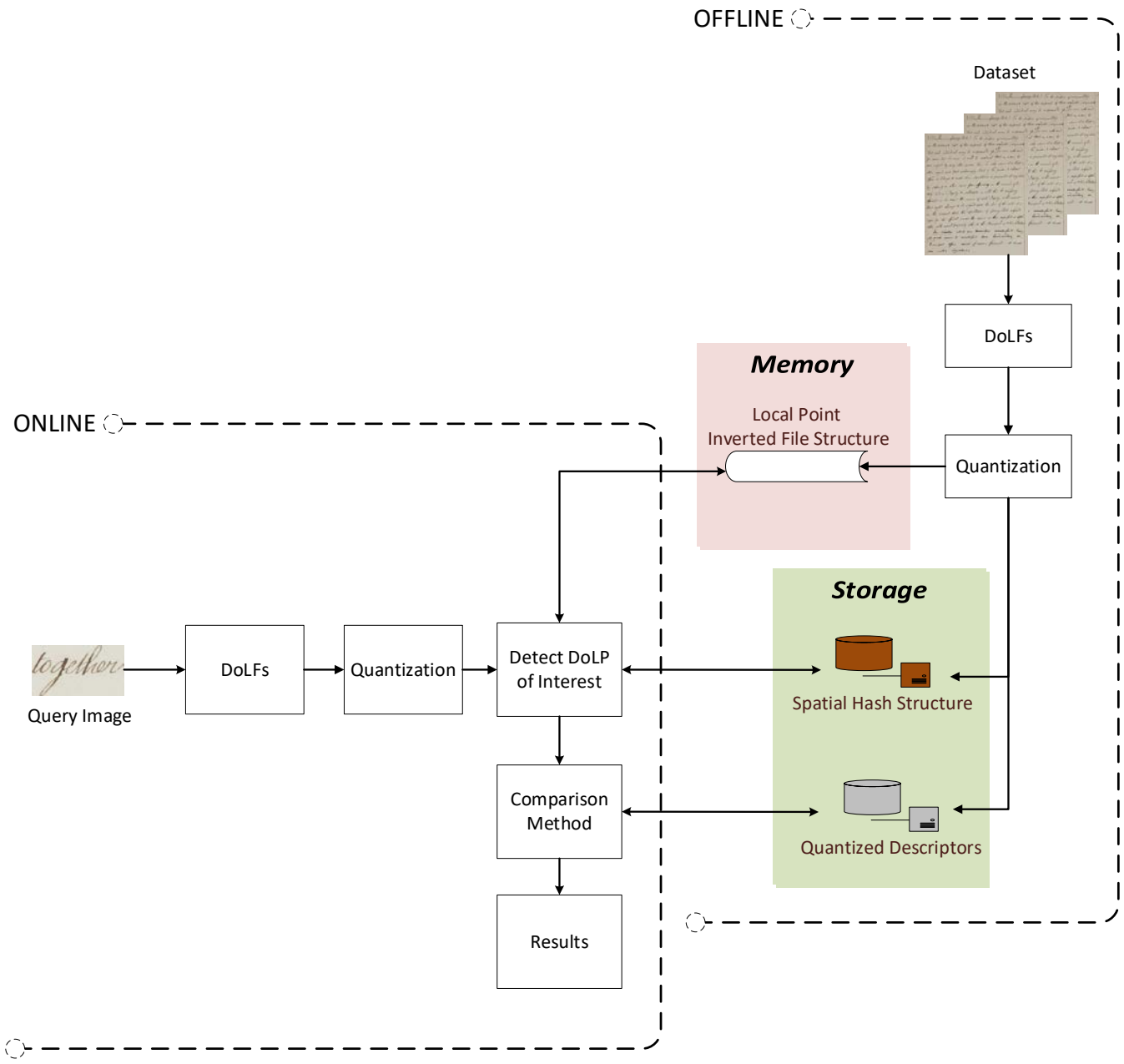
Disadvantages:

- memory and computational power requirements

DUTH KEYWORD SPOTTING FRAMEWORK

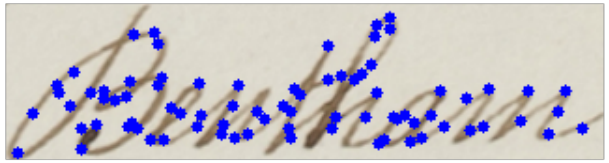
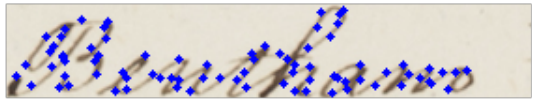
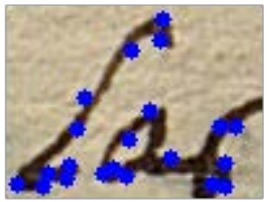
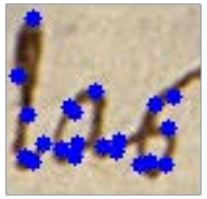
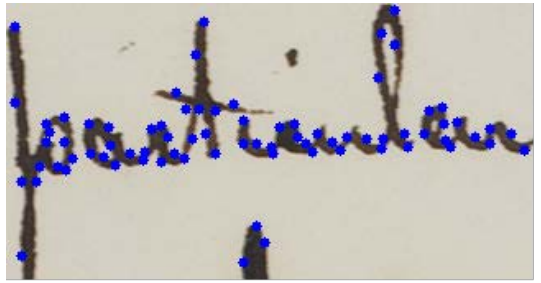
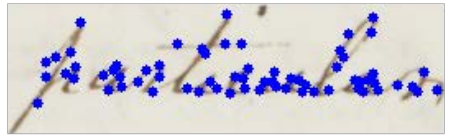
- The focus of the work is about minimizing memory and computational power requirements in a segmentation-free context which it would enable us to search in large document collections
- Does not need any training data
- It provides consistency between different handwritten writing variations.
- Use of the same operational pipeline in both segmentation-based and segmentation-free scenarios (**segmented word images** (segmentation-based approach) or the **complete document image** (segmentation-free approach))

ARCHITECTURE



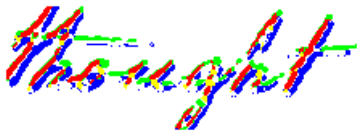
DOCUMENT-ORIENTED LOCAL FEATURES*

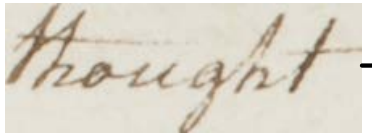
- Use of local features that takes into consideration the handwritten documents particularities. Therefore, it is able to detect meaningful points of the characters that reside in the documents independently of its scaling.
- It provides some consistency between different handwritten writing variations.

* K. Zagoris, I. Pratikakis and B. Gatos, "Unsupervised Word Spotting in Historical Handwritten Document Images Using Document-Oriented Local Features," in IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 4032-4041, Aug. 2017

KEYPOINT DETECTION



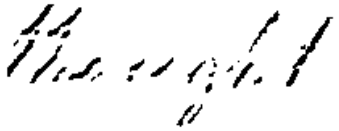


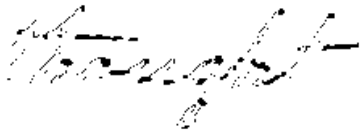
Gradient Orientation

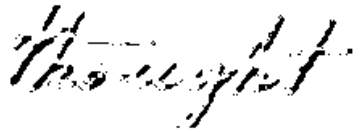
Quantization

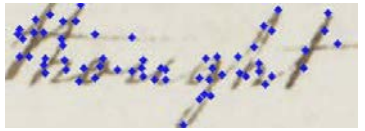
Connected Components

Final Keypoints

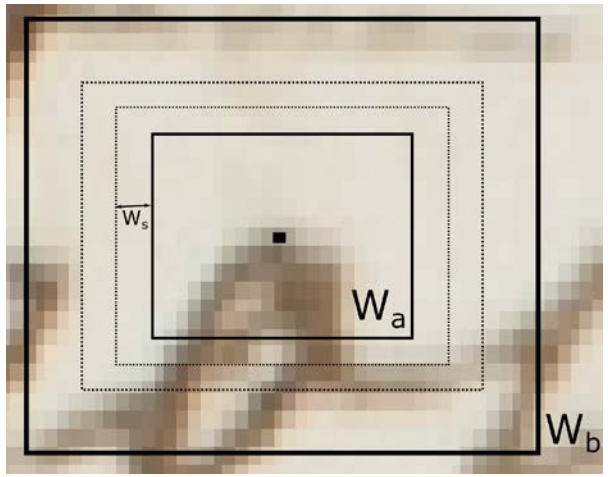


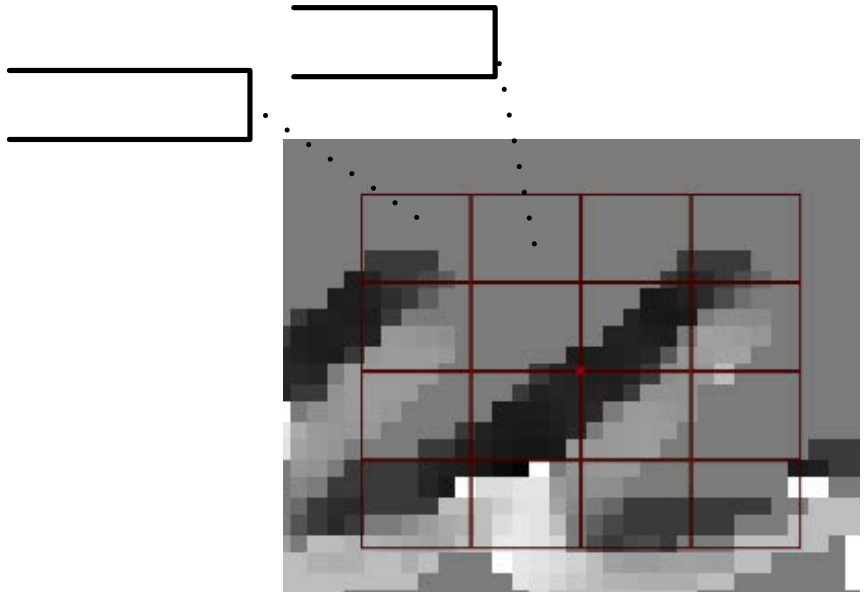






DESCRIPTOR CREATION





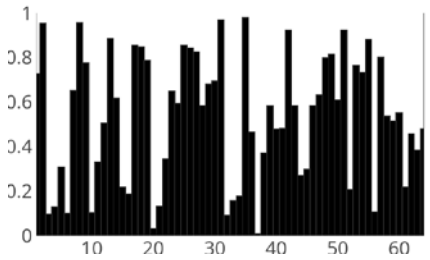
4 Elements ... 4 Elements

[0.1 , 0.34, 0 , 0, 0.1 ..... 0.3, 0.45, 0.9]

4x16 = 64 Elements

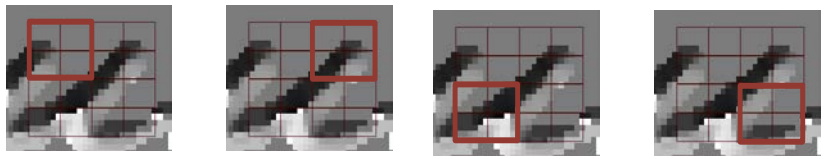
QUANTIZATION

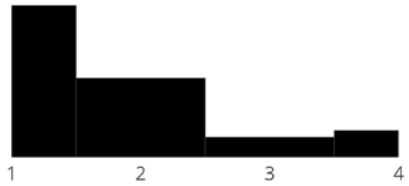
Descriptor



[0.02, 0.1.....0.2, 0.1]

BoW1 BoW2 BoW3 BoW4





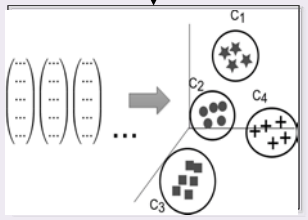
4 Different Bag of Visual Words with **K** Words

Bag of Visual Words – Codebook Creation

Dataset

Feature Extraction

Clustering



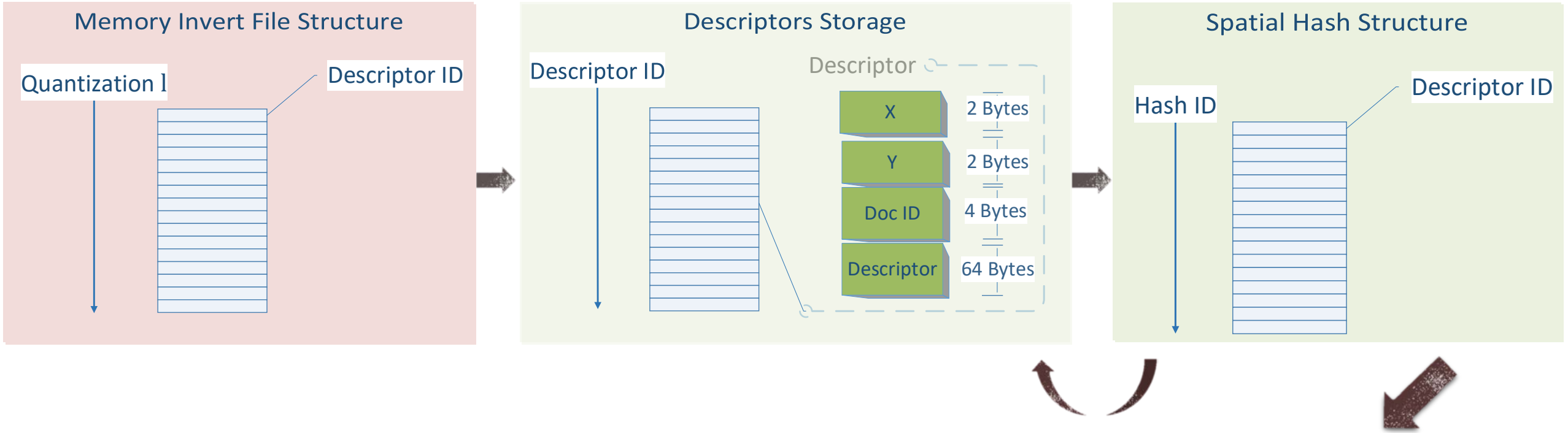
Use the Codebook for Visual Words Assignment

Quantization Hash Function

$$\mathbf{L} = BOW_1 * K^3 + BOW_2 * K^2 + BOW_3 * K + BOW_4$$

MEMORY AND STORAGE STRUCTURES AND THEIR RELATIONSHIPS

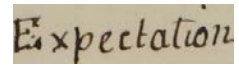
OFFLINE STEP FOR POPULATING STRUCTURES - ONLINE STEP FOR USING THE STRUCTURES



Spatial Hash Function

$$h(x, y, Doc\ ID) = d * A^2 + \mathrm{y} * A + \mathrm{x}\ where\ A > x_{max}\ \&\ A > y_{max}$$

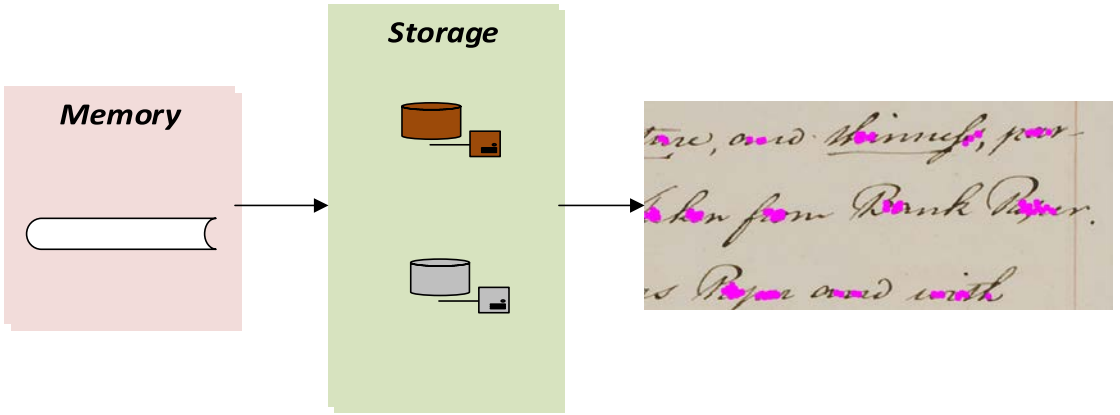
RETRIEVAL PROCEDURE



Query Image

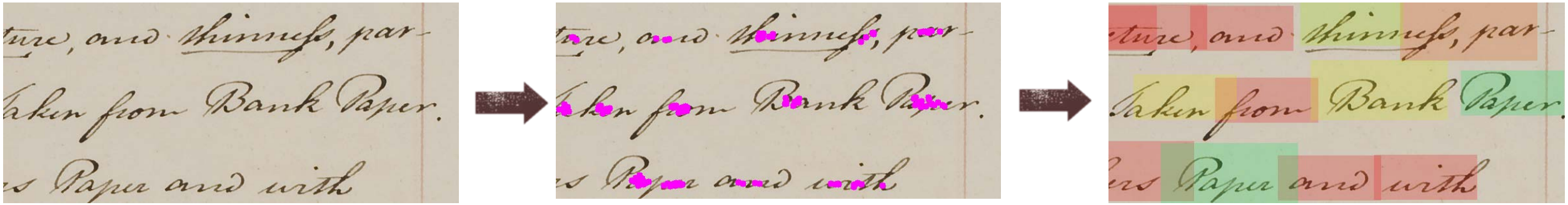
Quantization

Retrieve Local Points of Interest



Spatial Nearest Neighbor Search

SPATIAL NEAREST NEIGHBOR SEARCH



DUTH KEYWORD SPOTTING DEMONSTRATOR

The KeyWord Spotting Demonstrator supports the following main tasks:

- Creation (Indexing) of new Datasets
- User interactive word image query selection
- Presentation of the spotted words

Moreover, the communication between the front-end and the back-end is defined by a **REST API** which is freely available at:

https://github.com/transkribus/DUTH/WSBackend-API

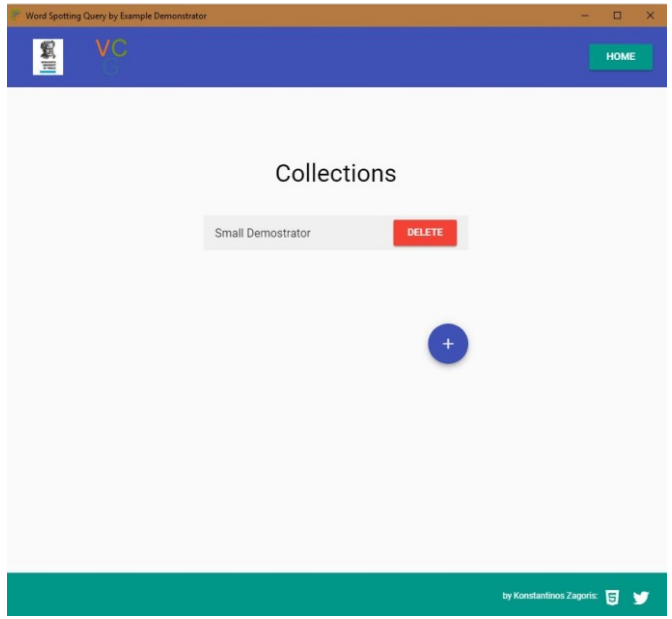
DUTH KEYWORD SPOTTING DEMONSTRATOR

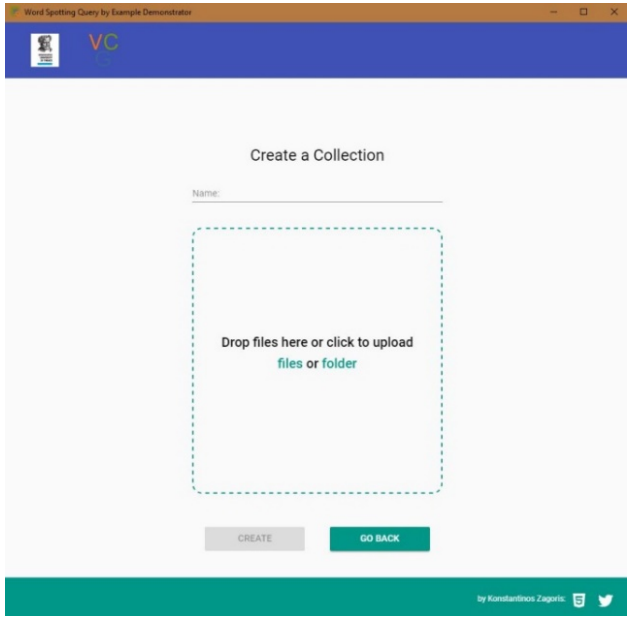
In order to showcase the above segmentation-free word spotting method, a web word-spotting application was created.

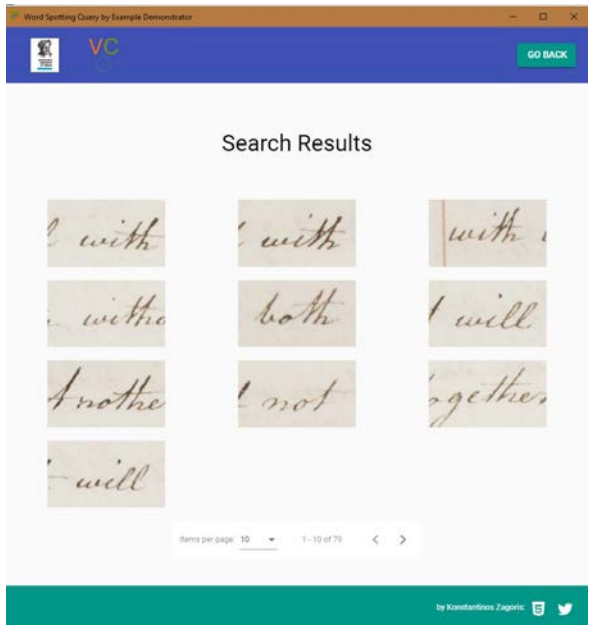
It is based on Angular 5 and Material Design for the front-end (GUI) and the back-end is created by the C#/.NET Core framework.

The DUTH Keyowrd Spotting Demonstrator is available at:

http://orpheus.ee.duth.gr/word-spotting-demonstrator/

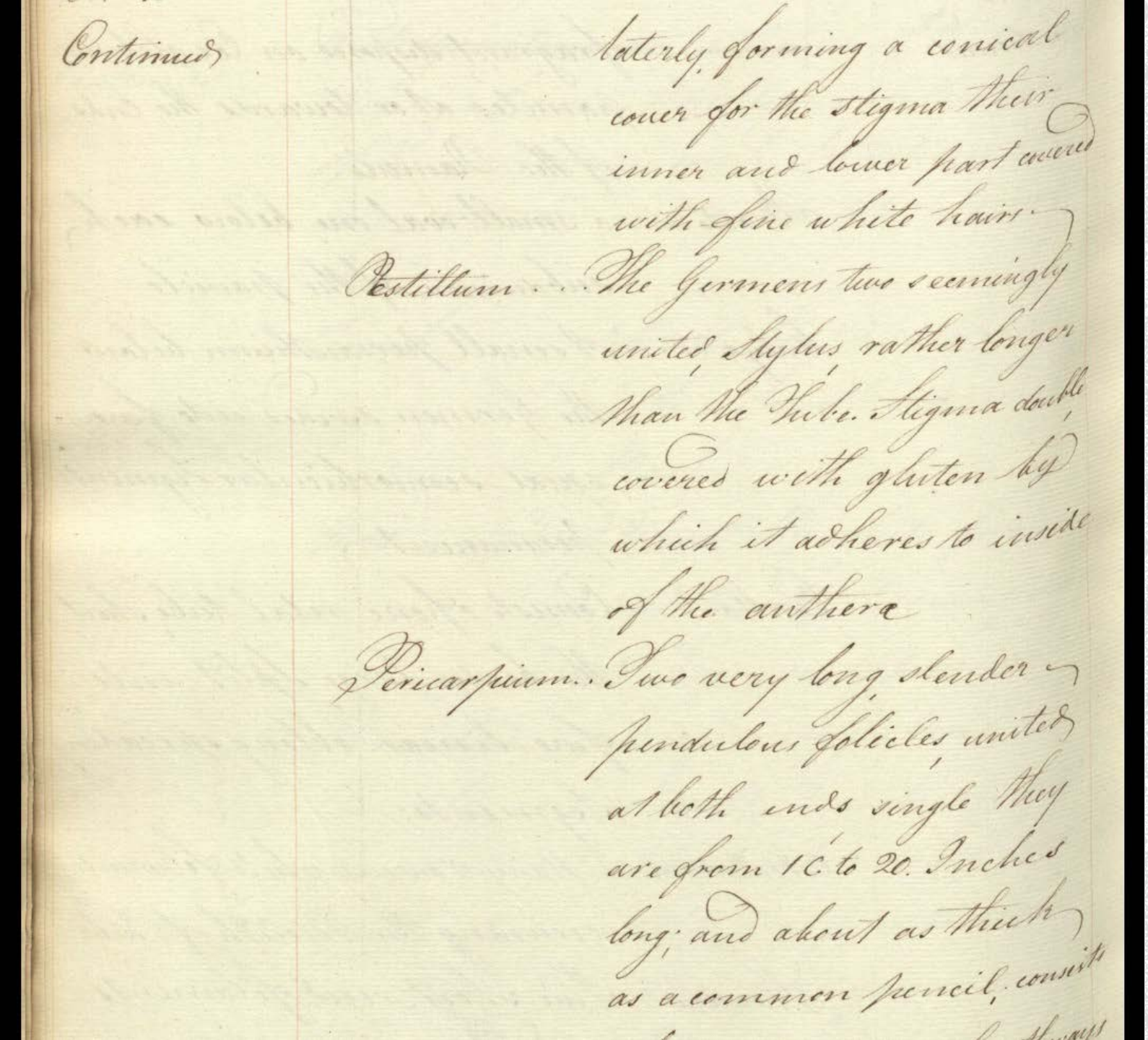






EXPERIMENTAL RESULTS

- English, German, Finnish Dataset
- The punctuation marks and capitals are considered in the ground truth corpora.
- Queries is every word with length greater than 3 and frequency greater than 2.
- English Dataset Queries (4790 words)
- German Dataset Queries (7119 words)
- Finnish Dataset Queries (5731 words)


Continued

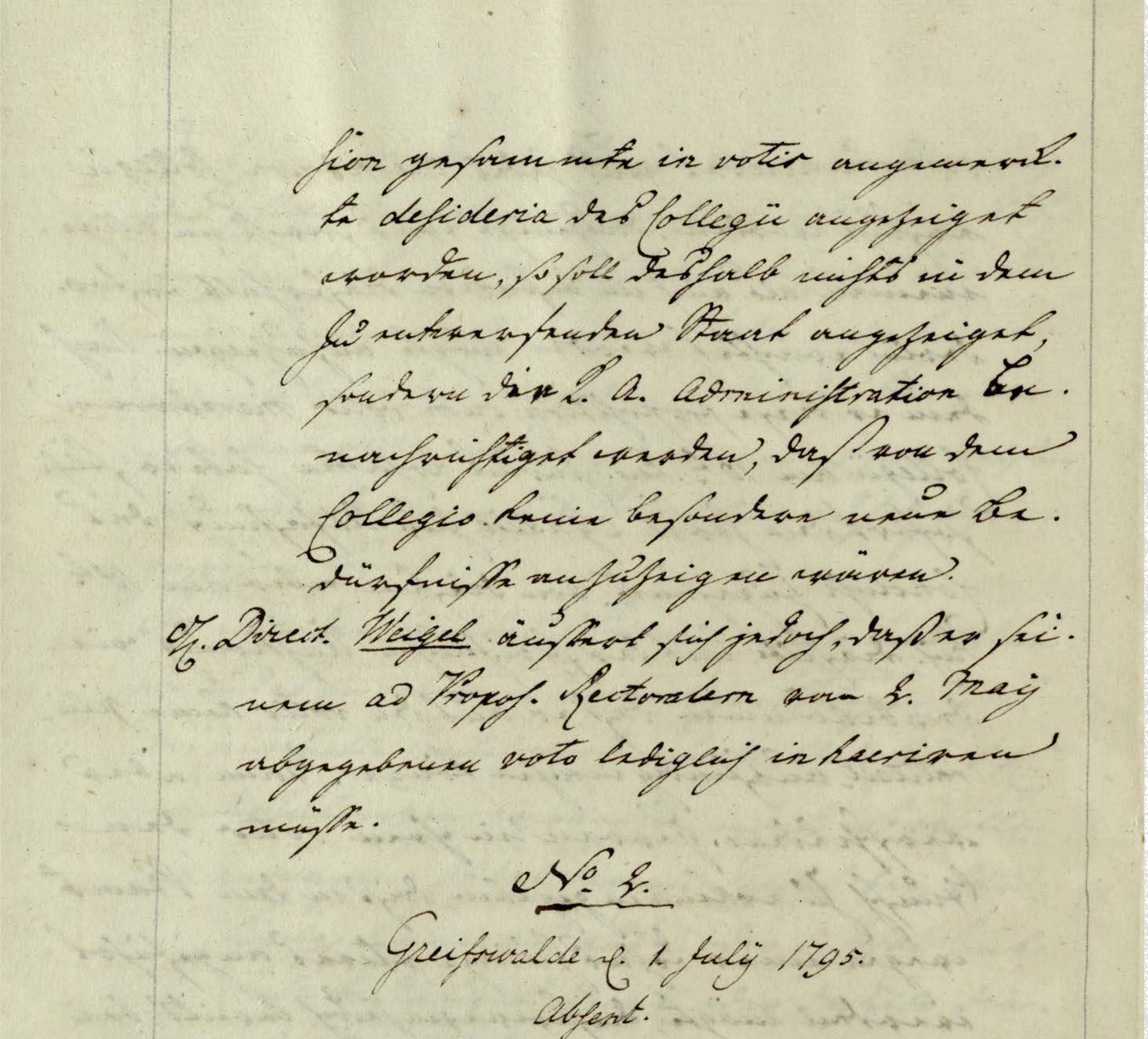
laterly, forming a conical
cover for the stigma their
inner and lower part covered
with fine white hairs.

Pestillum . . . The Germens two seemingly
united, Styles rather longer
than the Tube. Stigma double,
covered with gluten by
which it adheres to inside
of the anthera

Pericarpium. Two very long slender
pendulous follicles, united
at both ends single they
are from 16 to 20 Inches
long, and about as thick
as a common pencil, consist

ENGLISH DATASET

- 109 Pages
- 15 923 words



hier gesammelte in votis angemerk-
te desideria des Collegii angezeigt
worden, so soll deshalb nichts zu dem
Gesammt entworfenen Staat angezeigt,
sondern der K. A. Administration be-
nachrichtigt werden, daß von dem
Collegio keine besondere neue Be-
dürfnisse anzuführen wären.
H. Direct. Weigel äußerte sich jedoch, daß er sei-
nem ad Propos. Rectoralem vom 2. Maij
abgegebenen voto lediglich inhaeriren
müßte.
No. 2.
Greifswalde d. 1. July 1795.
Absent.

KONZILSPROTOKOLLE – GERMAN DATASET

- 100 Pages
- 15 579 words

1812.

Rumjantseville lähetettävien kirjeiden kuljetusta Turusta Kuopioon ja sieltä Pietariin.

Kart. 10 N:o 60. Koskee kauppiaanpoika Nils Petter Cajanderin anomusta, että hänen Turussa omistamansa talo vapautettaisiin majoitusrasituksesta.

—"— "61 —"— eräorinewoksen leskeu Wilhelmina Forselleksen anomusta, että hänelle korvattaisiin 1808-09 v:n sodan vuoksi kärsimänsä (aineelliset) tappiot.

—"— "62 Puuttuu

—"— "63 —"— Suomessa ollutta katoa ja viljan jakelua varattomalle väestölle.

—"— "64 —"— eräästä Venäjän valtiolle kuuluvasta kuljetuslaivasta varastettua rautaerää.

—"— "65 —"— ilmoitusta, ettei kapteeni Bygoo komennussuntineen ole lähtenyt Kuopioon.

—"— "66 —"— eri aatelishenkilöitten anomuksia, että heidän sallittaisiin omistaa säteritiloja.

—"— "67 —"— Savonlinnan kaupungin raatimiehen (Radkewitsch) valitusta sen johdosta, että Kuopion lääninhallitus mikä laittomasti on pakkolunastanut hänelle kuuluvaa maata.

—"— "68 —"— postirahojen perimistä majuri Larshitskiltä, g:m

—"— "69 —"— turkulaisen kauppiaan Strömbergin anomusta erään hänelle kuuluvan Viipurissa takavarisuuden kauppamäärän vapauttamisesta takavarikosta

1812.

Kart. 11 N:o 70. Koskee majuri Abraham Joakim Mola musta hänelle 1808-09 v:n sodassa nava tulevan 913 riikintalarin erän suorittamisesta hänelle (Anon

—"— "71 —"— tiedon hankkimista Ruotsiin siirtyn arweijan ent. sotamieheltä fagotin Nils Eklundilta, haluaako hän s misiin erään Cecilia Mononen - nim sa asuvan naisen kanssa vai

—"— "72. —"— kauppias Johan Sjöblomin tekemä ruisjauhojen hankinnasta venäl. v. 1809, mitä sopimusta hän ei v tänyt, vaikka hän ennakolta on tämän jauhoerän hinnan 6,000

—"— "73 —"— Diskonttokonttorin johtokunnan kir se huomauttaa, ettei Ruotsin pankk nut hopeana eräistä panttia.

—"— "74 —"— hampurilaisen kauppiaan Gottfried Kivennavan pitäjästä löydettyä p

—"— "75 —"— turkulaisen kauppiaan Christian Be valitusta sen johdosta, että eräi ovat pahoinpidelleet hänen työm

—"— "76 —"— ruokouspäiviä ja ruokouspäivätekste

—"— "77 —"— eri henkilöiden anomuksia saa viljaa ja karjaa Venäjältä Suomee viljan tuontia ja vientiä koskevia

—"— "78 —"— Suomen ovranomaisten tiedusteluja palvelukseen otetuista suomalais

FINNISH DATASET

- 56 Double Pages
- Many transcription errors (non-existing words).
- Corrections are needed.

EXPERIMENTAL RESULTS FOR SEGMENTATION-FREE EVALUATION

Method	English		Konzilsprotokolle		Finnish	
	P@5	MAP	P@5	MAP	P@5	MAP
Original [ZAG2017]	0.35	0.22	0.59	0.42	0.58	0.43
Current	0.35	0.22	0.57	0.38	0.56	0.39

TIME, MEMORY AND STORAGE REQUIREMENTS FOR SEGMENTATION-FREE SCENARIO

Method	Retrieval Time per Query (sec)	Memory requirement per Document (KB)	Storage requirement per Document (KB)
Original [ZAG2017]	15.84	19800	19800
Current	0.66	49	1676

[ZAG2017] K. Zagoris, I. Pratikakis and B. Gatos, "Unsupervised Word Spotting in Historical Handwritten Document Images Using Document-Oriented Local Features," in IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 4032-4041, Aug. 2017

COMPARATIVE EVALUATION RESULTS FOR BIG DATASETS FOR SEGMENTATION-FREE SCENARIO

Dataset (Documents)	Retrieval Time per Query (sec)	Overall Memory requirement (MB)	Overall Storage requirement (MB)
CURRENT			
50	0.61	2.1	69
5000	0.89	213	2693
50000	1.1	448	5843

The results reveal that the retrieval time per query is increased in a non-linear manner so that make search feasible in terms of time consumption for large scale datasets.

Ευχαριστώ!
Thank You!