DETECTION OF HISTORICAL TEXT REUSE

From a research question to the right model for detecting Historical Text Reuse

Marco Büchler (with contributions from Greta Franzini, Emily Franzini & Maria Moritz)

eTRAP
Electronic Text Reuse Acquisition Project

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

TABLE OF CONTENTS

1. Who am I?

2. What is text reuse?

3. Research on the characteristics

4. ACID paradigm

5. Comparison of Luke & Mark

6. Reuse process

7. Results

8. Automatic vs. manual text simplification

WHO AM I?

WHO AM I?

- 2001-2002: Head of Quality Assurance department in a software company;
- 2006: Diploma in Computer Science on big scale co-occurrence analysis;
- 2007: Consultant for several SMEs in IT sector;
- 2008: Technical project management of the eAQUA project;
- 2011: PI and project manager of the eTRACES project;
- 2013: PhD in Digital Humanities on Text Reuse;
- 2014: Head of Early Career Research Group eTRAP at the University of Göttingen.
- 2017: Head of Digital Historical Research at Leibnitz Institute of European History.

ABOUT ETRAP

Electronic Text Reuse Acquisition Project (eTRAP)

Interdisciplinary Early Career Research Group funded by the German Ministry of Education & Research (BMBF).

Budget: €1.6M.
Duration: March 2015 - February 2019.
Team: 4 core staff + ca. 4-5 research & student assistants (Bachelor, Masters and PhD theses).

WHAT IS TEXT REUSE?

TEXT REUSE

Text Reuse:

- spoken and written repetition of text across time and space.

For example:

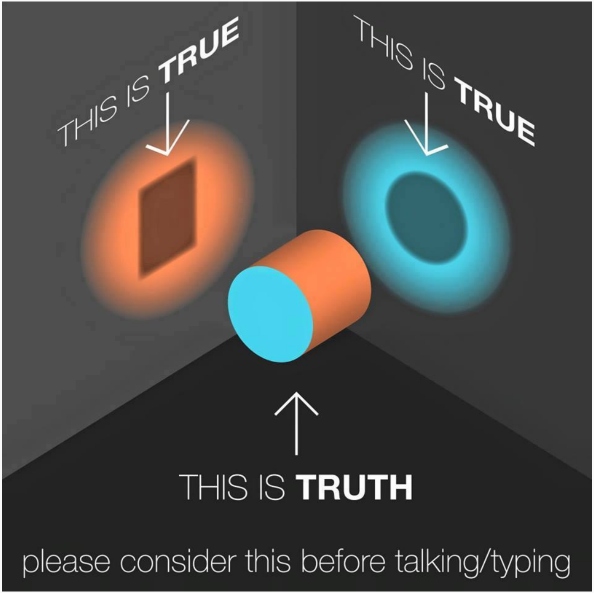
- citations, allusions, and translations.

Detection methods are needed to support scholarly work.

- E.g., they help to ensure clean libraries or identify fragmentary authors.

Text is often modified during the reuse process.

WHAT DO YOU ASSOCIATE WITH TEXT REUSE AND INTERTEXTUALITY?



EXPECTATIONS OF A COMPUTER SCIENTIST: OVERSIMPLIFICATION

It was the best of times, it
was the worst of times,
the age of wisdom, it w
age of foolishness, it
epoch of belief, it was
epoch of incredulity, it was
the season of Light, it was
the season of Darkness, it
It was the best of times,
it was the worst of times,
it wait was the age of
om, it was|

EXPECTATIONS OF A HUMANIST: OVERSIMPLIFICATION

TEXT REUSE FOR HUMANITIES AND COMPUTER SCIENCE

Question:
Why is text reuse detection relevant for Humanities and Computer Science?

- Humanities:
  - Lines of transmission and textual criticism.
  - Transmissions of ideas & thoughts under different circumstances and conditions.
- Computer Science:
  - Text decontamination for stylometry and authorship attribution, dating of texts.
  - Text Mining, Corpus Linguistics.

ETRAP'S OBJECTIVE

Title: eTRAP - electronic Text Reuse Acquisition Project

Premise: Language is a changing system. Compared to biometry the volatility is much higher.

- Research on the characteristics
  - What are good characteristics?
  - Which characteristics are stable and which are volatile and therefore not helpful in the detection process?
- Research on the reuse process
  - Begins with: Why do we quote what we quote?
  - Passes by: If changes in the reuse process happen, why do they happen and what is the model behind (if one exists)?
  - Ends with: Understanding paraphrases and allusions

ABOUT ETRAP

Electronic **T**ext **R**euse **A**cquisition **P**roject (eTRAP)

Interdisciplinary Early Career Research Group funded by the German Ministry of Education & Research (BMBF).

Budget: €1.6M.
Duration: March 2015 - February 2019.
Team: 4 core staff + ca. 4-5 research & student assistants (Bachelor, Masters and PhD theses).

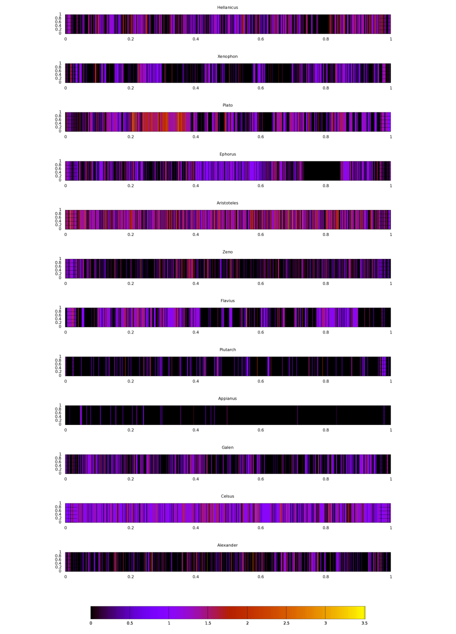
BIG (HUMANITIES) DATA

Ulrike Rieß (*Big Data bestimmt die IT-Welt*):

- Large amounts of data that can't be processed and analysed manually;
- Less structured data, e.g. in comparison to databases and data warehouse systems;
- Heterogeneous and distributed data across resources.

Information overload = large amounts of data (Big Data).
Information poverty = noisy, fragmentary (Humanities Data).

TEMPERATURE MAP



RESEARCH ON THE CHARACTERISTICS

TEXT REUSE CORE ELEMENTS

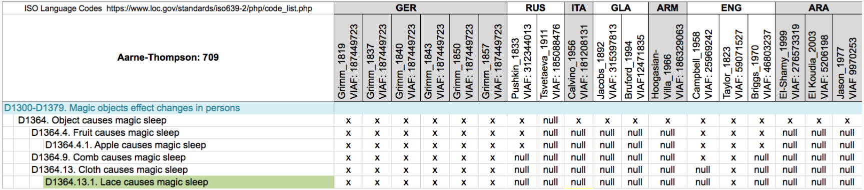
Motif: *"1. A minimal thematic unit"* (Prince, 2003, p. 55), set of core elements.

Core elements from an interdisciplinary standpoint:

- **Literature**: tracing MOTIFS
- **Cultural Studies**: tracing MEMES
- **Linguistics**: tracing PATTERNS
- **Computer Science**: tracing FEATURES
- **Forensics**: tracing MINUTIAE
- **Cognitive Psychology & Literature Studies**: tracing FIGURES OF MEMORY

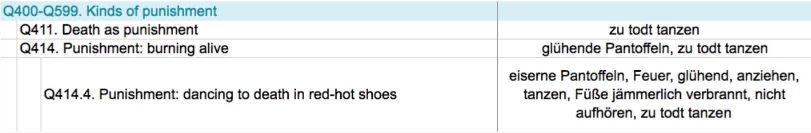
DATA COLLECTION AND CURATION

Tasks: Verify presence of motifs in different collections and record their "base form" as text reuse training data.



ISO Language Codes https://www.loc.gov/standards/iso639-2/php/code_list.php	GER						RUS		ITA	GLA		ARM	ENG			ARA		
Aarne-Thompson: 709	Grimm_1819 VIAF: 187449723	Grimm_1837 VIAF: 187449723	Grimm_1840 VIAF: 187449723	Grimm_1843 VIAF: 187449723	Grimm_1850 VIAF: 187449723	Grimm_1857 VIAF: 187449723	Pushkin_1833 VIAF: 312344013	Tsvetaeva_1911 VIAF: 185088476	Calvino_1956 VIAF: 181208131	Jacobs_1892 VIAF: 315397813	Bruford_1994 VIAF12471835	Hoogasian- Villa_1966 VIAF: 186329063	Campbell_1958 VIAF: 25969242	Taylor_1823 VIAF: 59071527	Briggs_1970 VIAF: 46803237	El-Shamy_1999 VIAF: 276573319	El Koudia_2003 VIAF: 5206198	Jason_1977 VIAF 9970253
D1300-D1379. Magic objects effect changes in persons																		
D1364. Object causes magic sleep	x	x	x	x	x	x	x	null	x	x	x	x	x	x	x	x	x	x
D1364.4. Fruit causes magic sleep	x	x	x	x	x	x	x	null	null	null	null	null	x	x	x	null	null	null
D1364.4.1. Apple causes magic sleep	x	x	x	x	x	x	x	null	null	null	null	null	x	x	x	null	null	null
D1364.9. Comb causes magic sleep	x	x	x	x	x	x	null	null	null	null	null	null	x	x	null	null	null	null
D1364.13. Cloth causes magic sleep	x	x	x	x	x	x	null	null	null	null	null	null	null	x	null	null	null	null
D1364.13.1. Lace causes magic sleep	x	x	x	x	x	x	null	null	null	null	null	null	null	x	null	null	null	null

Figure 1: Microsoft Excel matrix of motifs. Left column lists AT motifs in *Snow White* (AT 709); top row lists languages and collections covered.



Q400-Q599. Kinds of punishment	
Q411. Death as punishment	zu todt tanzen
Q414. Punishment: burning alive	glühende Pantoffeln, zu todt tanzen
Q414.4. Punishment: dancing to death in red-hot shoes	eiserne Pantoffeln, Feuer, glühend, anziehen, tanzen, Füße jämmerlich verbrannt, nicht aufhören, zu todt tanzen

Figure 2: Grimm motifs reduced to keywords.

DETECTION OF CROSS-LINGUAL MOTIFS

Train an (adapted) Named Entity Recognition (NER) tagger, ideally as language-independent as possible, to automatically annotate further fairy tales and texts.

EXAMPLE CASE STUDY: SNOW WHITE

RQ: How to computationally detect a motif despite its variants?

For example:

- DE [Grimm][1]: ***Schneewittchen und die sieben Zwerge***
- EN [Briggs][2]: ***Snow White and the three robbers***
- IT [Calvino][3]: ***Bella Venezia e i dodici ladroni***
- SQ [von Hahn][4]: ***Schneewittchen und die vierzig Drachen***
- RU [Pushkin][5]: Сказка о мертвой царевне и о семи богатырях
- ...

THE NRC EMOTION LEXICON

The NRC (National Research Council Canada) **Emotion Lexicon:**

- The Roget Thesaurus
- 14,182 words types

Emotions: (Plutchik, 1980)

anger
anticipation
disgust
fear
joy
sadness
surprise
trust

Sentiments:

negative emotions
positive emotions

TAGGING EMPATHY

Classroom Questionnaires

- Empathy
- Identification
- Transportation

- Six- and ten-year-old children
- Y-Labor

- Data set

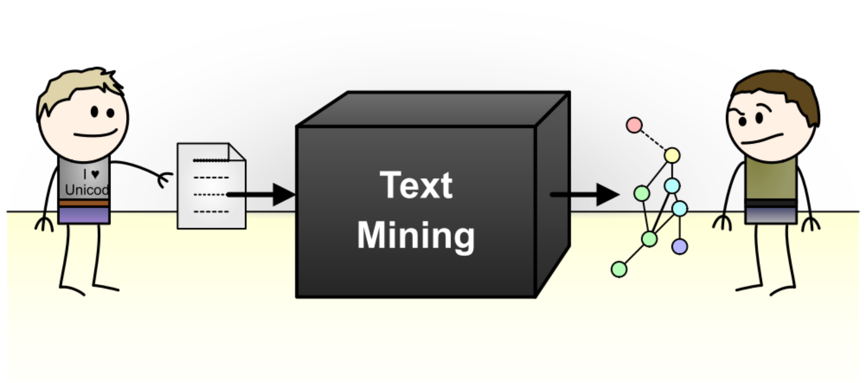
ACID PARADIGM

ACID PARADIGM

ACID for the Digital Humanities:

- **A**cceptance
- **C**omplexity
- **I**nteroperability
- **D**iversity

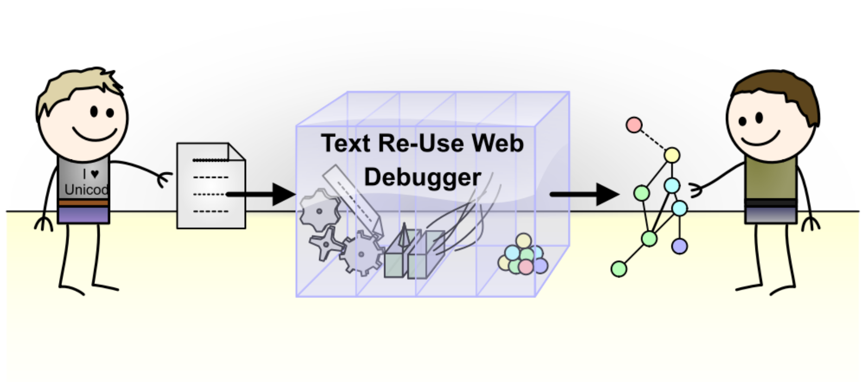
ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE I



ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE II

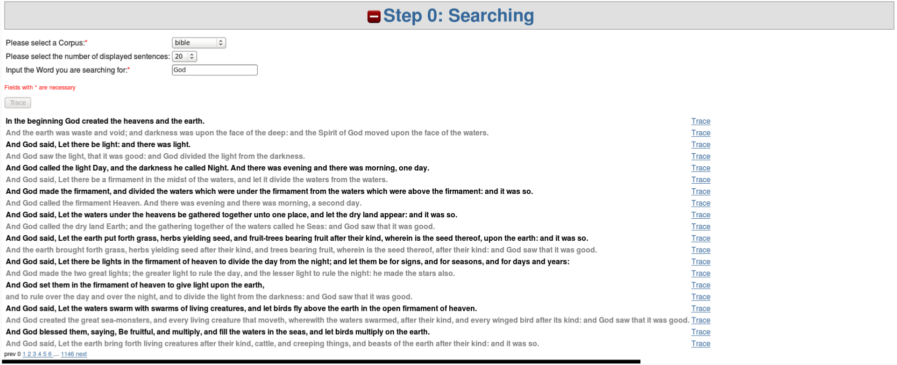
How to be accepted by humanists if text mining is a black box we can’t look into?

ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE III



Transparency: How to provide user-friendly insights into complex mining techniques and machine learning?

ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE IV



Step 0: Searching

Please select a Corpus:* bible

Please select the number of displayed sentences: 20

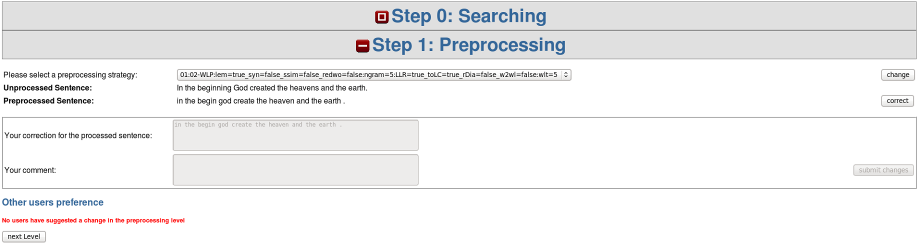
Input the Word you are searching for:* God

Fields with * are necessary

In the beginning God created the heavens and the earth.	Trace
And the earth was waste and void; and darkness was upon the face of the deep: and the Spirit of God moved upon the face of the waters.	Trace
And God said, Let there be light: and there was light.	Trace
And God saw the light, that it was good: and God divided the light from the darkness.	Trace
And God called the light Day, and the darkness he called Night. And there was evening and there was morning, one day.	Trace
And God said, Let there be a firmament in the midst of the waters, and let it divide the waters from the waters.	Trace
And God made the firmament, and divided the waters which were under the firmament from the waters which were above the firmament: and it was so.	Trace
And God called the firmament Heaven. And there was evening and there was morning, a second day.	Trace
And God said, Let the waters under the heavens be gathered together unto one place, and let the dry land appear: and it was so.	Trace
And God called the dry land Earth; and the gathering together of the waters called he Seas: and God saw that it was good.	Trace
And God said, Let the earth put forth grass, herbs yielding seed, and fruit-trees bearing fruit after their kind, wherein is the seed thereof, upon the earth: and it was so.	Trace
And the earth brought forth grass, herbs yielding seed after their kind, and trees bearing fruit, wherein is the seed thereof, after their kind: and God saw that it was good.	Trace
And God said, Let there be lights in the firmament of heaven to divide the day from the night; and let them be for signs, and for seasons, and for days and years:	Trace
And God made the two great lights; the greater light to rule the day, and the lesser light to rule the night: he made the stars also.	Trace
And God set them in the firmament of heaven to give light upon the earth,	Trace
and to rule over the day and over the night, and to divide the light from the darkness: and God saw that it was good.	Trace
And God said, Let the waters swarm with swarms of living creatures, and let birds fly above the earth in the open firmament of heaven.	Trace
And God created the great sea-monsters, and every living creature that moveth, wherewith the waters swarmed, after their kind, and every winged bird after its kind: and God saw that it was good.	Trace
And God blessed them, saying, Be fruitful, and multiply, and fill the waters in the seas, and let birds multiply on the earth.	Trace
And God said, Let the earth bring forth living creatures after their kind, cattle, and creeping things, and beasts of the earth after their kind: and it was so.	Trace

prev 0 1 2 3 4 5 6 ... 1146 next

ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE V



Step 0: Searching

Step 1: Preprocessing

Please select a preprocessing strategy: 01:02-WLP:lem=true_syn=false_ssim=false_redwo=false:ngram=5:LLR=true_toLC=true_rDia=false_w2wl=false:wlt=5

Unprocessed Sentence: In the beginning God created the heavens and the earth.

Preprocessed Sentence: in the begin god create the heaven and the earth .

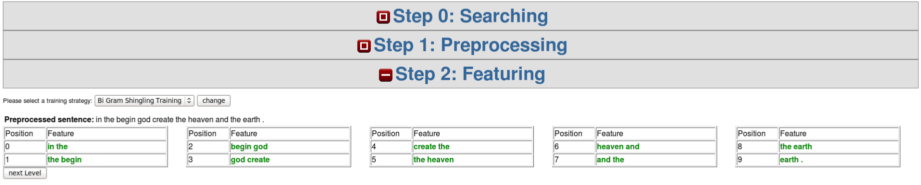
Your correction for the processed sentence: in the begin god create the heaven and the earth .

Your comment:

Other users preference

No users have suggested a change in the preprocessing level

ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE VI



Step 0: Searching

Step 1: Preprocessing

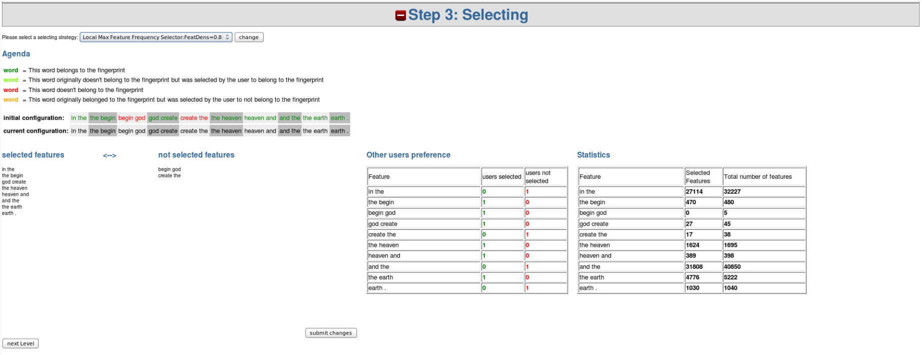
Step 2: Featuring

Please select a training strategy: Bi Gram Shingling Training

Preprocessed sentence: in the begin god create the heaven and the earth .

Position	Feature
0	in the
1	the begin
2	begin god
3	god create
4	create the
5	the heaven
6	heaven and
7	and the
8	the earth
9	earth .

ACID FOR THE DIGITAL HUMANITIES: ACCEPTANCE VII



Step 3: Selecting

Please select a selecting strategy: Local Max Feature Frequency Selector:FeatDens=0.8

Agenda

word = This word belongs to the fingerprint
word = This word originally doesn't belong to the fingerprint but was selected by the user to belong to the fingerprint
word = This word doesn't belong to the fingerprint
word = This word originally belonged to the fingerprint but was selected by the user to not belong to the fingerprint

initial configuration:	in the	the begin	begin god	god create	create the	the heaven	heaven and	and the	the earth	earth .
current configuration:	in the	the begin	begin god	god create	create the	the heaven	heaven and	and the	the earth	earth .

selected features <--> not selected features

selected features	not selected features
in the the begin god create the heaven heaven and and the the earth earth .	begin god create the

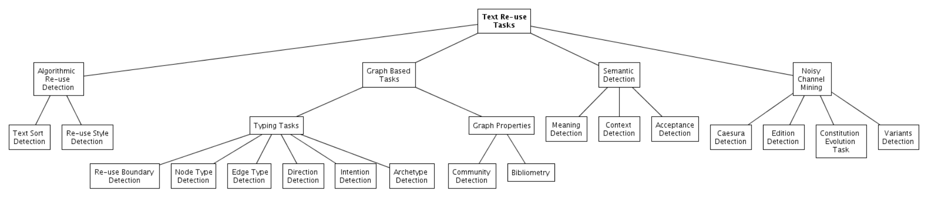
Other users preference

Feature	users selected	users not selected
in the	0	1
the begin	1	0
begin god	1	0
god create	1	0
create the	0	1
the heaven	1	0
heaven and	1	0
and the	0	1
the earth	1	0
earth .	0	1

Statistics

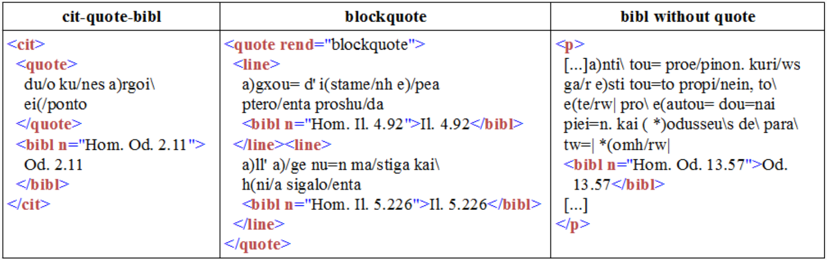
Feature	Selected Features	Total number of features
in the	27114	32227
the begin	470	480
begin god	0	5
god create	27	45
create the	17	38
the heaven	1624	1695
heaven and	389	398
and the	31808	40850
the earth	4776	5222
earth .	1030	1040

ACID FOR THE DIGITAL HUMANITIES: COMPLEXITY



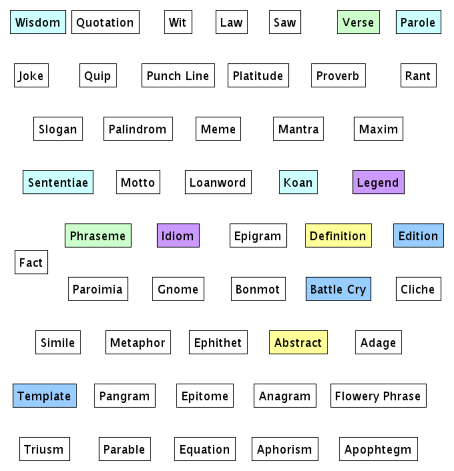
- **Text Re-use Tasks**
  - Algorithmic Re-use Detection
    - Text Sort Detection
    - Re-use Style Detection
  - Graph Based Tasks
    - Typing Tasks
      - Re-use Boundary Detection
      - Node Type Detection
      - Edge Type Detection
      - Direction Detection
      - Intention Detection
      - Archetype Detection
    - Graph Properties
      - Community Detection
      - Bibliometry
  - Semantic Detection
    - Meaning Detection
    - Context Detection
    - Acceptance Detection
  - Noisy Channel Mining
    - Caesura Detection
    - Edition Detection
    - Constitution Evolution Task
    - Variants Detection

ACID FOR THE DIGITAL HUMANITIES: INTEROPERABILITY



cit-quote-bibl	blockquote	bibl without quote
<cit> <quote> du/o ku/nes a)rgoi\ ei(/ponto </quote> <bibl n="Hom. Od. 2.11"> Od. 2.11 </bibl> </cit>	<quote rend="blockquote"> <line> a)gxou= d' i(stame/nh e)/pea ptero/enta proshu/da <bibl n="Hom. Il. 4.92">Il. 4.92</bibl> </line><line> a)ll' a)/ge nu=n ma/stiga kai\ h(ni/a sigalo/enta <bibl n="Hom. Il. 5.226">Il. 5.226</bibl> </line> </quote>	<p> [...]a)nti\ tou= proe/pinon. kuri/ws ga/r e)sti tou=to propi/nein, to\ e(te/rw| pro\ e(autou= dou=nai piei=n. kai (*)odusseu\s de\ para\ tw=| *(omh/rw| <bibl n="Hom. Od. 13.57">Od. 13.57</bibl> [...] </p>

DIVERSITY (REUSE TYPES)



Wisdom Quotation Wit Law Saw Verse Parole

Joke Quip Punch Line Platitude Proverb Rant

Slogan Palindrom Meme Mantra Maxim

Sententiae Motto Loanword Koan Legend

Phraseme Idiom Epigram Definition Edition

Fact

Paroimia Gnome Bonmot Battle Cry Cliche

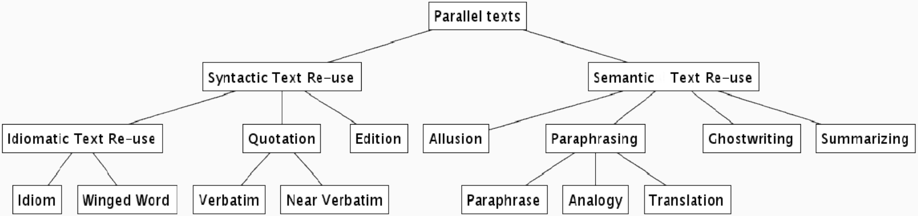
Simile Metaphor Ephithet Abstract Adage

Template Pangram Epitome Anagram Flowery Phrase

Triusm Parable Equation Aphorism Apophtegm

- **Stability** (yellow)
- **Purpose** (green)
- **Size of text reuse** (blue)
- **Classification** (light blue)
- **Degree of distribution** (purple)
- Written and oral transmission

DIVERSITY (REUSE STYLES)



- Parallel texts
  - Syntactic Text Re-use
    - Idiomatic Text Re-use
      - Idiom
      - Winged Word
    - Quotation
      - Verbatim
      - Near Verbatim
    - Edition
  - Semantic Text Re-use
    - Allusion
    - Paraphrasing
      - Paraphrase
      - Analogy
      - Translation
    - Ghostwriting
    - Summarizing

KEY PROBLEM

Question:

The distribution of **Reuse Types** and **Reuse Styles** is often unknown - which model(s) should be chosen?

TRACER: DISSEMINATION

Webpage: `http://www.etrap.eu/research/tracer`
Repository: `http://vcs.etrap.eu/tracer-framework/tracer.git`
Upcoming tutorials:

- **DATeCH 2017** (May 2017): pre-conference workshop, Göttingen, Germany.
- No further TRACER tuturials in 2017!

COMPARISON OF LUKE & MARK

TRACER: OVERVIEW

TRACER: suite of 700 algorithms developed by Marco Büchler.
Command line environment with no GUI.

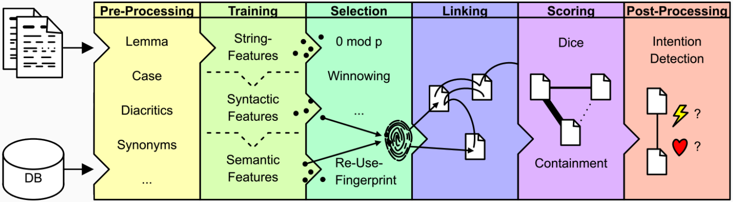


Figure 3: Detection task in six steps. More than 1M permutations of implementations of different levels are possible.

TRACER is **language-independent**. Tested on: Ancient Greek, Arabic, Coptic, English, German, Hebrew, Latin, Tibetan.

REUSE PROCESS

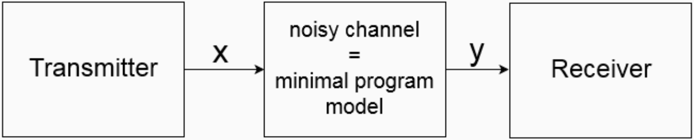
STATE OF THE ART

Paraphrasing and non-literal reuse challenges many approaches:

- Alzahrani et al. (2012)
  - study n-gram-, syntax-, and semantic-based detection approaches;
  - they find: as soon as reuse is slightly modified (words changed) most approaches fail.
- Barrón-Cedeño et al. (2013)
  - experiment with paraphrasing to improve plagiarism detection;
  - they found that complex paraphrasing with a high density challenges plagiarism detection, and
  - that lexical substitution is the most frequent plagiarism technique.

APPROACH

- Inspired by
  - Noisy channel model: given a "scrambled" word or sentence, guess the intended version of that sentence (Brill, 2000),
  - Kolmogorov Complexity: describes the length of the shortest program that produces an output string (Li and Vitáni, 2008),
- we study Ancient text reuse to understand how text is transferred.
  - Identify operations to characterize morphological & semantic changes
  - Design an algorithm which applies these OPs to our datasets
  - Transform one text excerpt into another by a minimum OP set



DATA-SETS - ANCIENT GREEK AND LATIN DATA-SET

“Salvation for the Rich”
Clement of Alexandria
Christian theologian, 2nd cent.

- Known for his retelling of biblical excerpts
- Reuse annotated by Biblindex team (Mellerin, 2014; Mellerin, 2016)
- We obtain 199 verse-reuse-pairs
- Pointing to 15 Bible books

Extracts from 12 works & 2 collections
Bernard of Clairvaux
French abbot, 12th cent.

- Known for his influence on the Cistercian order and his work in biblical studies
- Reuse extracted by Biblindex team (Mellerin, 2014; Mellerin, 2016)
- We obtain 162 verse-reuse-pairs
- Pointing to 31 Bible books

The data was tokenized and punctuation was kept but ignored in the analyses.

BIBLICAL REUSE EXAMPLES

more literal	Bible verse	Bernard reuse
Proverbs 18 3	**impius cum in profundum venerit** peccatorum **contemnit** sed sequitur eum ignominia et obprobrium (*When the wicked man is come into the depth of sins, also contempt comes but ignominy and reproach follow him)*	**Impius** , **cum venerit in profundum** malorum , **contemnit** (*When the wicked man is come into the depth of evil*)

less literal	Bible verse	Clement reuse
1Cor 13 13	νυνὶ **δὲ μένει πίστις** , **ἐλπίς** , **ἀγάπη** , **τὰ τρία ταῦτα μείζων δὲ τούτων ἡ ἀγάπη** (*And now remain faith, hope, love, these three; but the greatest of those is love.*)	**πίστει** καὶ **ἐλπίδι** καὶ **ἀγάπῃ** (*faith, and hope, and love* - in dative case)
		ἀγάπην , **πίστιν** , **ἐλπίδα** (*love, faith, hope* - in accusative case)
		μένει δὲ τὰ τρία ταῦτα , **πίστις** , **ἐλπίς** , **ἀγάπη** · **μείζων δὲ** ἐν **τούτοις ἡ ἀγάπη** *(and remain these three, faith, hope, love; but the greatest among them is love)*

non-literal	Bible verse	Clement reuse
Mt 12 35	ὁ ἀγαθὸς ἄνθρωπος ἐκ τοῦ ἀγαθοῦ θησαυροῦ ἐκβάλλει ἀγαθά , καὶ ὁ **πονηρὸς** ἄνθρωπος ἐκ τοῦ πονηροῦ θησαυροῦ ἐκβάλλει **πονηρά** . (*A good man out of good storage brings out good things , and an evil man out of the evil storage brings evil things .*)	Ψυχῆς , τὰ δὲ ἐκτός , κἂν μὲν ἡ ψυχὴ χρῆται καλῶς , καλὰ καὶ ταῦτα δοκεῖ , ἐὰν δὲ **πονηρῶς** , **πονηρά** , ὁ κελεύων ἀπαλλοτριοῦν τὰ ὑπάρχοντα (*[are whitin the] soul, and some are out, and if the soul uses them good, those things are also thought of as good, but if [they are used as] bad, [they are thought of as] bad; he who commands the renouncement of possessions*)

LINGUISTIC SUPPORT – LEMMA RESOURCES

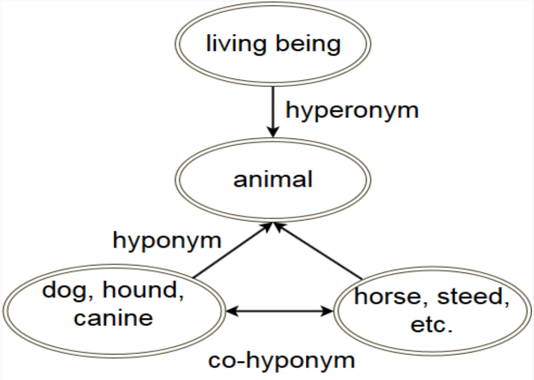
We aggregate:

- Biblindex' Lemma Lists
  - 65,537 Biblical Greek entries
  - 315,021 Latin entries
- Classical Language Tool Kit (CLTK) (Johnson et al., 2014)
  - 953,907 Ancient Greek words
  - 270,228 Latin words
- Greek New Testament of the Society of Biblical Literature[1] & Septuaginta (Rahlfs, 1935a; UPenn) 59,510 word-lemma-pairs

[1] Logos Bible Software `http://sblgnt.com/about/`

LINGUISTIC SUPPORT – ANCIENT GREEK WORDNET (AGWN)

99K synsets
of which 33K contain Ancient Greek and 27K Latin words
(Bizzoni et al., 2014; Minozzi, 2009)



TRANFORMATION OPERATIONS

Table 1: Operation list for the automated approach

operation	description	example
NOP(reuse_word, orig_word)	Original and reuse word are equal.	*NOP(maledictus,maledictus)*
upper(reuse_word, orig_word)	Word is lowercase in reuse and uppercase in original.	*upper(kai,Kai)* - in Greek
lower(reuse_word, orig_word)	Word is uppercase in reuse and lowercase in original.	*lower(Gloriam,gloriam)*
lem(reuse_word, orig_word)	Lemmatization leads to equality of reuse and original.	*lem(penetrat,penetrabit)*
repl_syn(reuse_word, orig_word)	Reuse word replaced with a synonym to match original word.	*repl_syn(magnificavit,glorificavit)*
repl_hyper(reuse_word, orig_word)	Word in Bible verse is a hyperonym of the reused word.	*hyper(cupit,habens)*
repl_hypo(reuse_word, orig_word)	Word in Bible verse is a hyponym of the reused word.	*hypo(dederit,tollet)*
repl_co-hypo(reuse_word, orig_word)	Reused word and original have the same hyperonym.	*repl_co-hypo(magnificavit,fecit)*
NOPmorph(reuse_tags, orig_tags)	Case or PoS did not change between reused and original word.	*NOPmorph(na,na)*
repl_pos(reuse_tag, orig_tag)	Reuse and original contain the same cognate, but PoS changed.	*repl_pos(n,a)*
repl_case(reuse_tag, orig_tag)	Reuse and original have the same cognate, but the case changed.	*repl_case(g,d)* - cases genitive, dative
lemma_missing(reuse_word, orig_word)	Lemma unknown for reuse or original word.	*lemma_missing(tentari, inlectus)*
no_rel_found(reuse_wword, orig_word)	Relation for reuse or original word not found in AGWN.	*no_rel_found(gloria,arguitur)*

QUALITATIVE COMPLEMENT

We manually analyze:

- 60 Ancient Greek & 100 Latin instances
- 192 &. 224 replacements
- Using `ins(word)`, `del(word)` and replacements:
  - `NOP`, `lem`, `repl_syn`, `repl_hyper`, `repl_hypo`, `repl_co-hypo`
- We assign morphological categories from Perseus' tag-set (Bamman and Crane 2011)
  - E.g., `repl_case_a_g` `repl_num_s_p`

Table 2: Excerpt from Perseus' tag-set

Category	Value	Tag
person	first person	1
	second person	2
	third person	3
number	singular	s
	plural	p
	dual	d
tense	present	p
	imperfect	i
	perfect	r
	pluperfect	l
	future perfect	t
	future	f
	aorist	a

RESULTS

LITERAL SHARE OF THE REUSE (RQ1)

What is the extent of non-literal reuse in our datasets?

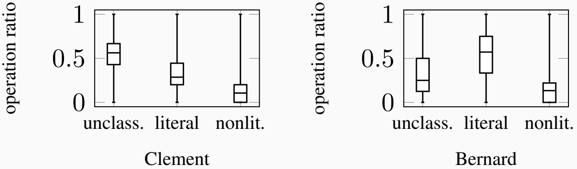


Figure 4: Ratios of operations in reuse instances. literal: NOP, lem, lower, etc.; nonlit: syn, hyper, etc.

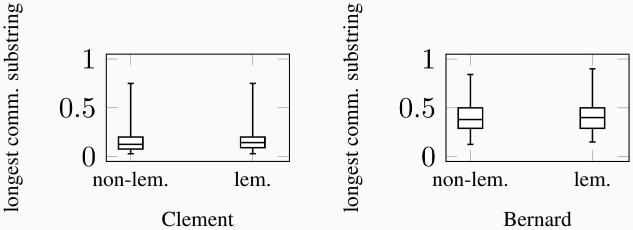


Figure 5: Ratios of literal overlap between reuse instances and originals.

AUTOMATED APPROACH (RQ2.1)

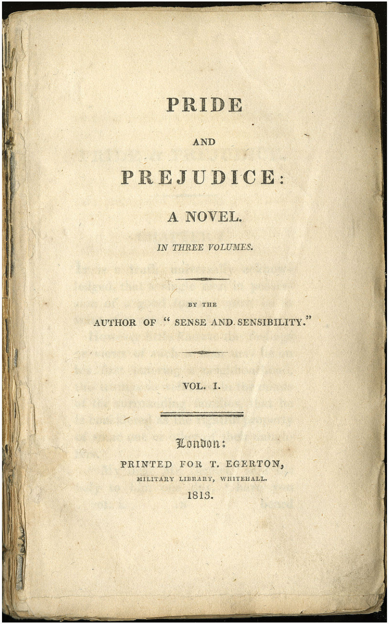
How is the non-literally reused text modified in our datasets? (RQ2)
How can linguistic resources support the discovery of non-literal reuse? (RQ2.1)

Table 3: Absolute numbers of operations identified automatically.

	literal				non-literal				unclassified		
	NOP	**upper**	**lower**	**lem**	**syn**	**hyper**	**hypo**	**co-hypo**	**no_rel_found**	**lem_missing**	**total**
Greek	337	6	0	356	153	20	14	101	563	639	2189
Latin	587	0	44	102	60	14	28	68	347	85	1335

AUTOMATIC VS. MANUAL TEXT SIMPLIFICATION

JANE AUSTEN'S PRIDE & PREJUDICE



PRIDE

AND

PREJUDICE:

A NOVEL.

IN THREE VOLUMES.

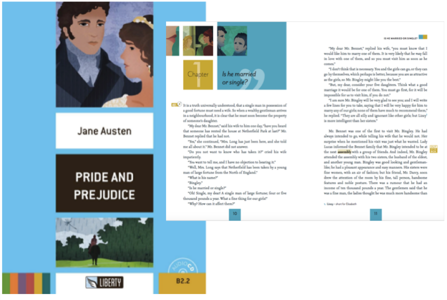
BY THE

AUTHOR OF "SENSE AND SENSIBILITY."

VOL. I.

London:
PRINTED FOR T. EGERTON,
MILITARY LIBRARY, WHITEHALL.
1813.

GRADED READER



Chapter 1

Is he married or single?

It is a truth universally understood, that a single man in possession of a good fortune must need a wife. So when a wealthy gentleman arrives in a neighbourhood, it is clear that he must soon become the property of someone's daughter.

"My dear Mr. Bennet," said his wife to him one day, "have you heard that someone has rented the house at Netherfield Park at last?" Mr. Bennet replied that he had not.

"Yes," she continued, "Mrs. Long has just been here, and she told me all about it." Mr. Bennet did not answer.

"Do you not want to know who has taken it?" cried his wife impatiently.

"You want to tell me, and I have no objection to hearing it."

"Well, Mrs. Long says that Netherfield has been taken by a young man of large fortune from the North of England."

"What is his name?"

"Bingley."

"Is he married or single?"

"Oh! Single, my dear! A single man of large fortune; four or five thousand pounds a year. What a fine thing for our girls!"

"Why? How can it affect them?"

10

IS HE MARRIED OR SINGLE?

"My dear Mr. Bennet," replied his wife, "you must know that I would like him to marry one of them. It is very likely that he may fall in love with one of them, and so you must visit him as soon as he comes."

"I don't think that is necessary. You and the girls can go, or they can go by themselves, which perhaps is better, because you are as attractive as the girls, so Mr. Bingley might like you the best."

"But, my dear, consider your five daughters. Think what a good marriage it would be for one of them. You must go first, for it will be impossible for us to visit him, if you do not."

"I am sure Mr. Bingley will be very glad to see you; and I will write a few lines for you to take, saying that I will be very happy for him to marry any of our girls; none of them have much to recommend them," he replied. "They are all silly and ignorant like other girls; but Lizzy[1] is more intelligent than her sisters."

Mr. Bennet was one of the first to visit Mr. Bingley. He had always intended to go, while telling his wife that he would not. Her surprise when he mentioned his visit was just what he wanted. Lady Lucas informed the Bennet family that Mr. Bingley intended to be at the next assembly with a group of friends. And indeed, Mr. Bingley attended the assembly with his two sisters, the husband of the eldest, and another young man. Bingley was good looking and gentleman-like; he had a pleasant appearance and easy manners. His sisters were fine women, with an air of fashion; but his friend, Mr. Darcy, soon drew the attention of the room by his fine, tall person, handsome features and noble posture. There was a rumour that he had an income of ten thousand pounds a year. The gentlemen said that he was a fine man, the ladies thought he was much more handsome than

[1] Lizzy – short for Elizabeth

11

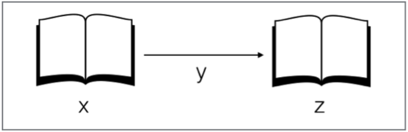
Definition:

Graded readers are "simplified books written at varying levels of difficulty for second language learners", which "cover a huge range of genres ranging from adaptation of classic works of literature to original stories, to factual materials such as biographies, reports and so on" [Waring 2012].

RESEARCH

To computationally analyse the process Y and classifying the changes:

- Do the changes follow strict rules?
- Do they form patterns?
- Can they be computationally reproduced?



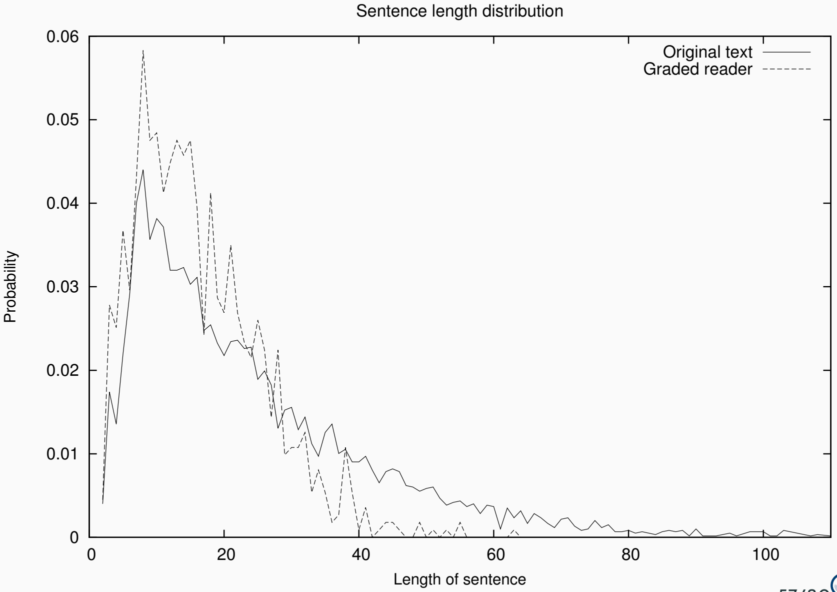
TYPES OF CHANGES

1. Structural changes:
    - I do not wish to be too hasty.
    - We must not conceal it.

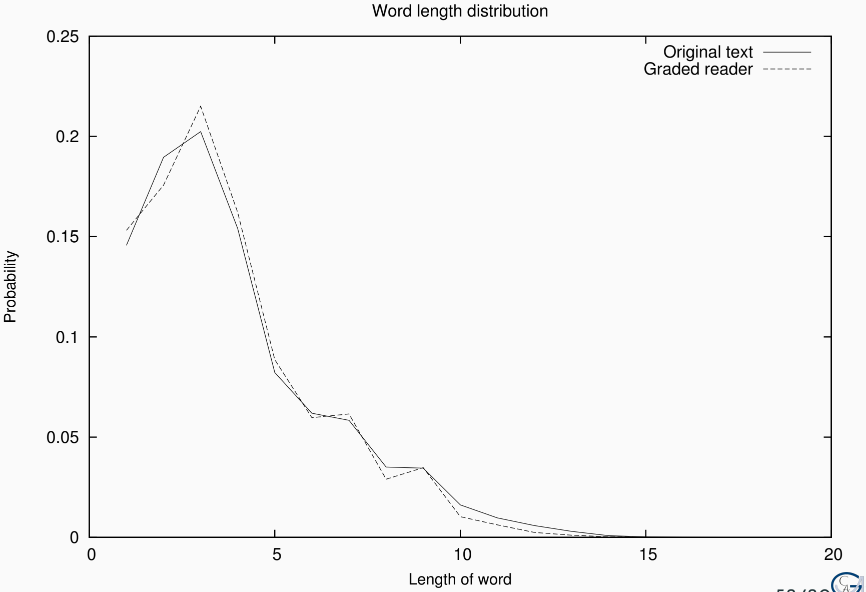
2. Cognitive changes:
    - … Soon after this event, Elizabeth received a visit…

3. Structural & cognitive changes:
    - Elizabeth is exceedingly handsome.

COMPARISON OF SENTENCE LENGTH



COMPARISON OF WORD LENGTH



THE POWER OF COGNITIVE CHANGES

Stylistic analyses of the original novel compared to an automatic text simplification (ATS) and to a human-made graded reader.

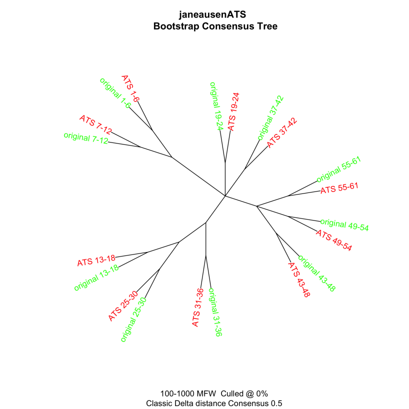


Figure 6: Dendrogram of the ON compared to ATS.

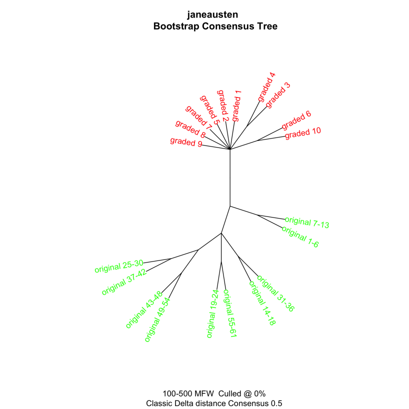
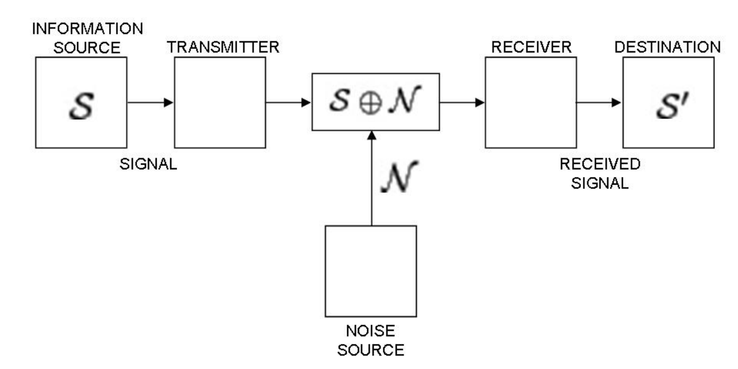


Figure 7: Dendrogram of the ON compared to the GR.

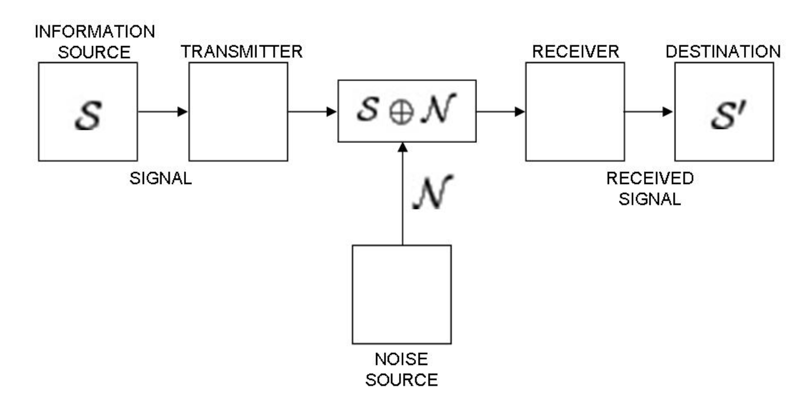
AUTOMATIC EVALUATION

Basic idea: Embed historical text reuse in Shannon's **Noisy Channel** theorem.

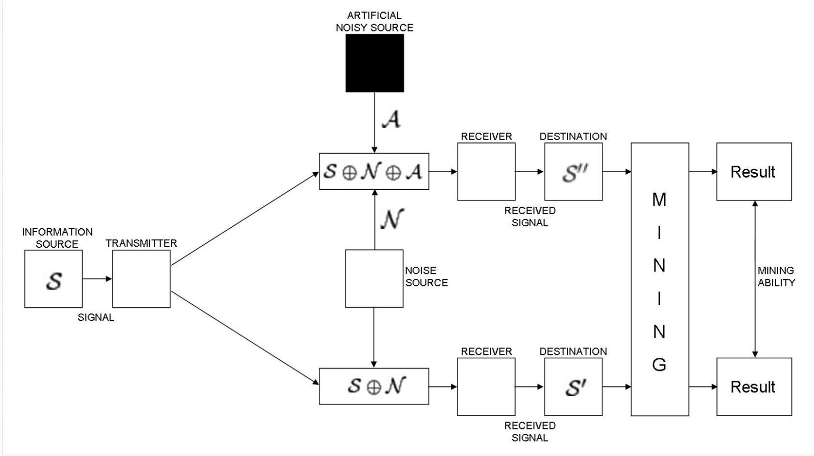


METHODOLOGY

Basic idea: Embed historical text reuse in Shannon's **Noisy Channel** theorem.



METHODOLOGY: NOISY CHANNEL EVALUATION I



Hint: The results are ALWAYS compared between the natural texts and the randomised texts as a whole.

METHODOLOGY: NOISY CHANNEL EVALUATION II

Signal-Noise-Ratio *adapted* from signal- and satellite techniques:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

Signal-Noise-Ratio *scaled*, unit is dB:

$$SNR_{db} = 10.log_{10}\left(\frac{P_{signal}}{P_{noise}}\right)$$

Mining Ability (in dB): The Mining Ability describes the power of a method to make distinctions between natural-language structures/patterns and random noise given a model with the same parameters.

$$L_{Quant}(\Theta) = 10.log_{10}\frac{|E_{D_s,\phi_\Theta}|}{max(1, |E_{D_s^m}, \phi_\Theta|)} dB$$

METHODOLOGY: NOISY CHANNEL EVALUATION III

Motivation for randomisation by **Word Shuffling**:

1. Syntax and distributional semantics are randomised and ”destroyed”.
2. Distributions of words and sentence lengths remain unchanged; changes JUST and ONLY depend on destruction of 1) and are not induced by changes of distributions.
3. Easy measurement of ”randomness” of the randomising method with the entropy test:

$$\Delta H^n = H_{max} - H^n$$

Die Wahl von $n \in [180, 183]$ sichert eine Genauigkeit von $\Delta H^n \leq 10^{-3}$ Bit für den Entropietest.

METHODOLOGY: TEXT REUSE COMPRESSION

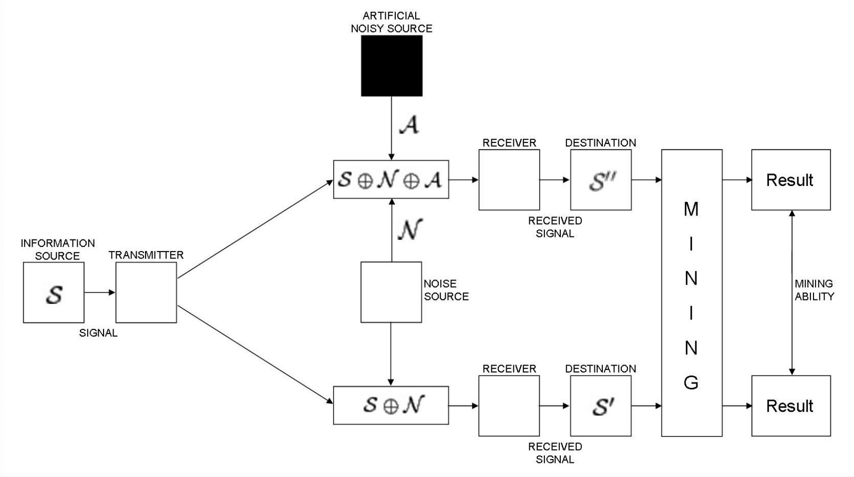
1. eTRAP works on text reuse.
2. eTRAP works on text reuse.
3. eTRAP works on text reuse.
4. eTRAP works on text reuse.
5. eTRAP works on text reuse.
6. ...

$$
\begin{array}{c|ccccc}
 & s_1 & s_2 & s_3 & s_4 & s_5 \\
\hline
s_1 & 0.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
s_2 & 1.00 & 0.00 & 1.00 & 1.00 & 1.00 \\
s_3 & 1.00 & 1.00 & 0.00 & 1.00 & 1.00 \\
s_4 & 1.00 & 1.00 & 1.00 & 0.00 & 1.00 \\
s_5 & 1.00 & 1.00 & 1.00 & 1.00 & 0.00
\end{array}
$$

$$\mathcal{C}_\Theta = \frac{n \cdot (n-1)}{n^2} = 1 - \frac{1}{n}$$

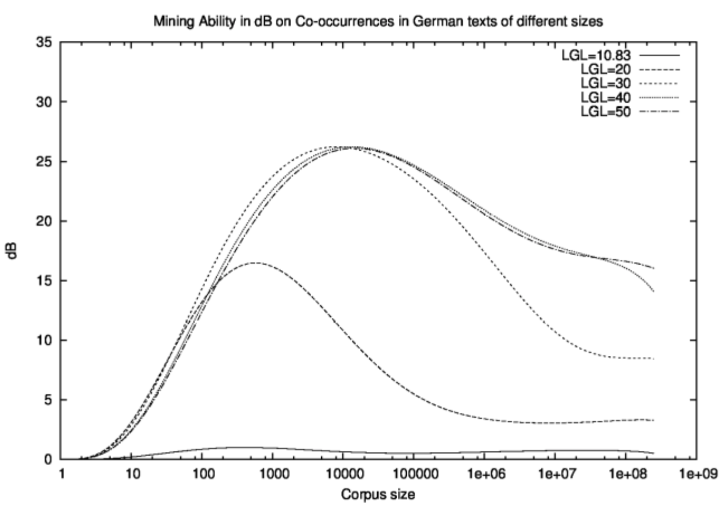
$$C_\Theta = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} \theta_\Theta(s_i, s_j)}{n \cdot m}$$

RANDOMNESS & STRUCTURE



Question: Why is the result of a randomised Digital Library typically not empty?

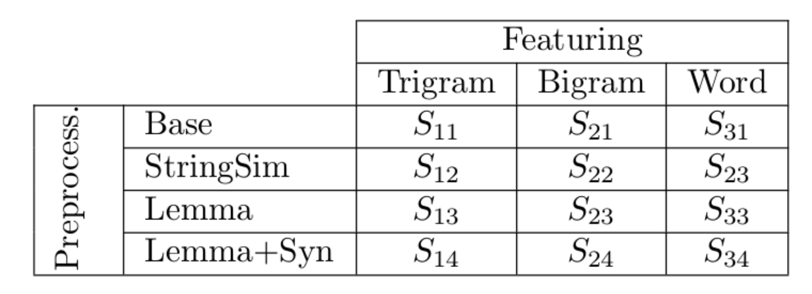
RANDOMNESS & STRUCTURE: IMPACT



Corpus size in sentences (average sentence length is ca. 18 words). LGL is the threshold for the Log-Likelihood-Ratio.

TEXT REUSE IN ENGLISH BIBLE VERSIONS: SETUP

Segmentation: disjoint and verse-wise segmentation.



		Featuring		
		Trigram	Bigram	Word
Preprocess.	Base	S_{11}	S_{21}	S_{31}
	StringSim	S_{12}	S_{22}	S_{23}
	Lemma	S_{13}	S_{23}	S_{33}
	Lemma+Syn	S_{14}	S_{24}	S_{34}

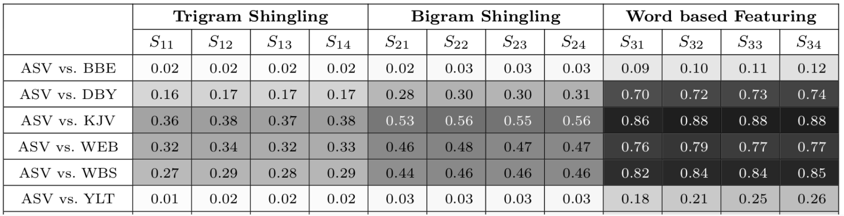
Selection: max pruning with a Feature Density of 0.8;

Linking: Inter- Digital Library Linking (different Bible editions);

Scoring: *Broder's Resemblance* with a threshold of 0.6;

Post-processing: not used.

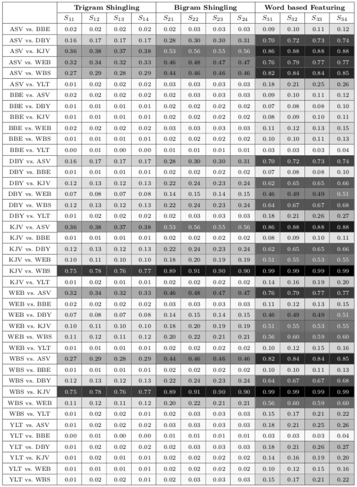
TEXT REUSE IN ENGLISH BIBLE VERSIONS: RESULTS – RECALL



	Trigram Shingling				Bigram Shingling				Word based Featuring			
	S_{11}	S_{12}	S_{13}	S_{14}	S_{21}	S_{22}	S_{23}	S_{24}	S_{31}	S_{32}	S_{33}	S_{34}
ASV vs. BBE	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.09	0.10	0.11	0.12
ASV vs. DBY	0.16	0.17	0.17	0.17	0.28	0.30	0.30	0.31	0.70	0.72	0.73	0.74
ASV vs. KJV	0.36	0.38	0.37	0.38	0.53	0.56	0.55	0.56	0.86	0.88	0.88	0.88
ASV vs. WEB	0.32	0.34	0.32	0.33	0.46	0.48	0.47	0.47	0.76	0.79	0.77	0.77
ASV vs. WBS	0.27	0.29	0.28	0.29	0.44	0.46	0.46	0.46	0.82	0.84	0.84	0.85
ASV vs. YLT	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.18	0.21	0.25	0.26

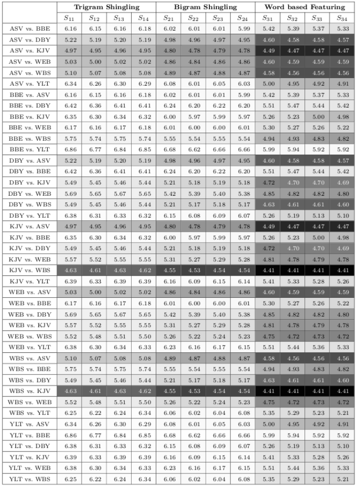
TEXT REUSE IN ENGLISH BIBLE VERSIONS: RECALL VS. TEXT REUSE COMPRESSION

With



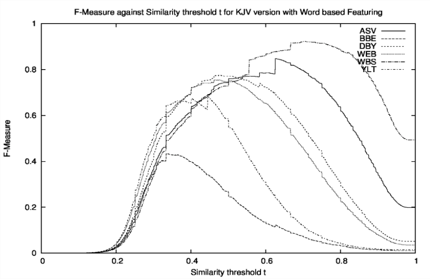
	Trigram Shingling				Bigram Shingling				Word based Featuring			
	S_{11}	S_{12}	S_{13}	S_{14}	S_{21}	S_{22}	S_{23}	S_{24}	S_{31}	S_{32}	S_{33}	S_{34}
ASV vs. BBE	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.09	0.10	0.11	0.12
ASV vs. DBY	0.16	0.17	0.17	0.17	0.28	0.30	0.30	0.31	0.70	0.72	0.73	0.74
ASV vs. KJV	0.36	0.38	0.37	0.38	0.53	0.56	0.55	0.56	0.86	0.88	0.88	0.88
ASV vs. WEB	0.32	0.34	0.32	0.33	0.46	0.48	0.47	0.47	0.76	0.79	0.77	0.77
ASV vs. WBS	0.27	0.29	0.28	0.29	0.44	0.46	0.46	0.46	0.82	0.84	0.84	0.85
ASV vs. YLT	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.18	0.21	0.25	0.26
BBE vs. ASV	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.09	0.10	0.11	0.12
BBE vs. DBY	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.07	0.08	0.08	0.10
BBE vs. KJV	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.08	0.09	0.10	0.11
BBE vs. WEB	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.11	0.12	0.13	0.15
BBE vs. WBS	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.10	0.10	0.11	0.13
BBE vs. YLT	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.04
DBY vs. ASV	0.16	0.17	0.17	0.17	0.28	0.30	0.30	0.31	0.70	0.72	0.73	0.74
DBY vs. BBE	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.07	0.08	0.08	0.10
DBY vs. KJV	0.12	0.13	0.12	0.13	0.22	0.24	0.23	0.24	0.62	0.65	0.65	0.66
DBY vs. WEB	0.07	0.08	0.07	0.08	0.14	0.15	0.14	0.15	0.46	0.49	0.49	0.51
DBY vs. WBS	0.12	0.13	0.12	0.13	0.22	0.24	0.23	0.24	0.64	0.67	0.67	0.68
DBY vs. YLT	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.18	0.21	0.26	0.27
KJV vs. ASV	0.36	0.38	0.37	0.38	0.53	0.56	0.55	0.56	0.86	0.88	0.88	0.88
KJV vs. BBE	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.08	0.09	0.10	0.11
KJV vs. DBY	0.12	0.13	0.12	0.13	0.22	0.24	0.23	0.24	0.62	0.65	0.65	0.66
KJV vs. WEB	0.10	0.11	0.10	0.10	0.18	0.20	0.19	0.19	0.51	0.55	0.53	0.55
KJV vs. WBS	0.75	0.78	0.76	0.77	0.89	0.91	0.90	0.90	0.99	0.99	0.99	0.99
KJV vs. YLT	0.01	0.02	0.01	0.01	0.02	0.02	0.02	0.02	0.14	0.16	0.19	0.20
WEB vs. ASV	0.32	0.34	0.32	0.33	0.46	0.48	0.47	0.47	0.76	0.79	0.77	0.77
WEB vs. BBE	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.11	0.12	0.13	0.15
WEB vs. DBY	0.07	0.08	0.07	0.08	0.14	0.15	0.14	0.15	0.46	0.49	0.49	0.51
WEB vs. KJV	0.10	0.11	0.10	0.10	0.18	0.20	0.19	0.19	0.51	0.55	0.53	0.55
WEB vs. WBS	0.11	0.12	0.11	0.12	0.20	0.22	0.21	0.21	0.56	0.60	0.59	0.60
WEB vs. YLT	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.10	0.12	0.15	0.16
WBS vs. ASV	0.27	0.29	0.28	0.29	0.44	0.46	0.46	0.46	0.82	0.84	0.84	0.85
WBS vs. BBE	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.10	0.10	0.11	0.13
WBS vs. DBY	0.12	0.13	0.12	0.13	0.22	0.24	0.23	0.24	0.64	0.67	0.67	0.68
WBS vs. KJV	0.75	0.78	0.76	0.77	0.89	0.91	0.90	0.90	0.99	0.99	0.99	0.99
WBS vs. WEB	0.11	0.12	0.11	0.12	0.20	0.22	0.21	0.21	0.56	0.60	0.59	0.60
WBS vs. YLT	0.01	0.02	0.02	0.01	0.02	0.03	0.03	0.03	0.15	0.17	0.21	0.22
YLT vs. ASV	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.18	0.21	0.25	0.26
YLT vs. BBE	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.04
YLT vs. DBY	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.18	0.21	0.26	0.27
YLT vs. KJV	0.01	0.02	0.01	0.01	0.02	0.02	0.02	0.02	0.14	0.16	0.19	0.20
YLT vs. WEB	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.10	0.12	0.15	0.16
YLT vs. WBS	0.01	0.02	0.02	0.01	0.02	0.03	0.03	0.03	0.15	0.17	0.21	0.22

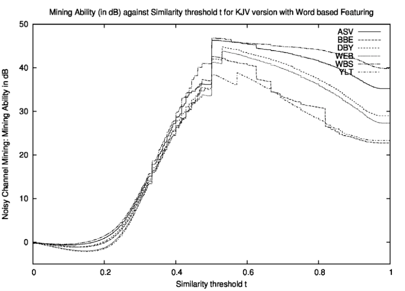
Without



	Trigram Shingling				Bigram Shingling				Word based Featuring			
	S_{11}	S_{12}	S_{13}	S_{14}	S_{21}	S_{22}	S_{23}	S_{24}	S_{31}	S_{32}	S_{33}	S_{34}
ASV vs. BBE	6.16	6.15	6.16	6.18	6.02	6.01	6.01	5.99	5.42	5.39	5.37	5.33
ASV vs. DBY	5.22	5.19	5.20	5.19	4.98	4.96	4.97	4.95	4.60	4.58	4.58	4.57
ASV vs. KJV	4.97	4.95	4.96	4.95	4.80	4.78	4.79	4.78	4.49	4.47	4.47	4.47
ASV vs. WEB	5.03	5.00	5.02	5.02	4.86	4.84	4.86	4.86	4.60	4.59	4.59	4.59
ASV vs. WBS	5.10	5.07	5.08	5.08	4.89	4.87	4.88	4.87	4.58	4.56	4.56	4.56
ASV vs. YLT	6.34	6.26	6.30	6.29	6.08	6.01	6.05	6.03	5.00	4.95	4.92	4.91
BBE vs. ASV	6.16	6.15	6.16	6.18	6.02	6.01	6.01	5.99	5.42	5.39	5.37	5.33
BBE vs. DBY	6.42	6.36	6.41	6.41	6.24	6.20	6.22	6.20	5.51	5.47	5.44	5.42
BBE vs. KJV	6.35	6.30	6.34	6.32	6.00	5.97	5.99	5.97	5.26	5.23	5.00	4.98
BBE vs. WEB	6.17	6.16	6.17	6.18	6.01	6.00	6.00	6.01	5.30	5.27	5.26	5.22
BBE vs. WBS	5.75	5.74	5.75	5.74	5.55	5.54	5.55	5.54	4.94	4.93	4.83	4.82
BBE vs. YLT	6.86	6.77	6.84	6.85	6.68	6.62	6.66	6.66	5.99	5.94	5.92	5.92
DBY vs. ASV	5.22	5.19	5.20	5.19	4.98	4.96	4.97	4.95	4.60	4.58	4.58	4.57
DBY vs. BBE	6.42	6.36	6.41	6.41	6.24	6.20	6.22	6.20	5.51	5.47	5.44	5.42
DBY vs. KJV	5.49	5.45	5.46	5.44	5.21	5.18	5.19	5.18	4.72	4.70	4.70	4.69
DBY vs. WEB	5.69	5.65	5.67	5.65	5.42	5.39	5.40	5.38	4.85	4.82	4.82	4.80
DBY vs. WBS	5.49	5.45	5.46	5.44	5.21	5.17	5.18	5.17	4.63	4.61	4.61	4.60
DBY vs. YLT	6.38	6.31	6.33	6.32	6.15	6.08	6.09	6.07	5.26	5.19	5.13	5.10
KJV vs. ASV	4.97	4.95	4.96	4.95	4.80	4.78	4.79	4.78	4.49	4.47	4.47	4.47
KJV vs. BBE	6.35	6.30	6.34	6.32	6.00	5.97	5.99	5.97	5.26	5.23	5.00	4.98
KJV vs. DBY	5.49	5.45	5.46	5.44	5.21	5.18	5.19	5.18	4.72	4.70	4.70	4.69
KJV vs. WEB	5.57	5.52	5.55	5.55	5.31	5.27	5.29	5.28	4.81	4.78	4.79	4.78
KJV vs. WBS	4.63	4.61	4.63	4.62	4.55	4.53	4.54	4.54	4.41	4.41	4.41	4.41
KJV vs. YLT	6.39	6.33	6.39	6.39	6.16	6.09	6.15	6.14	5.41	5.33	5.28	5.26
WEB vs. ASV	5.03	5.00	5.02	5.02	4.86	4.84	4.86	4.86	4.60	4.59	4.59	4.59
WEB vs. BBE	6.17	6.16	6.17	6.18	6.01	6.00	6.00	6.01	5.30	5.27	5.26	5.22
WEB vs. DBY	5.69	5.65	5.67	5.65	5.42	5.39	5.40	5.38	4.85	4.82	4.82	4.80
WEB vs. KJV	5.57	5.52	5.55	5.55	5.31	5.27	5.29	5.28	4.81	4.78	4.79	4.78
WEB vs. WBS	5.52	5.48	5.51	5.50	5.26	5.22	5.24	5.23	4.75	4.72	4.73	4.72
WEB vs. YLT	6.38	6.30	6.34	6.33	6.23	6.16	6.17	6.15	5.51	5.44	5.36	5.33
WBS vs. ASV	5.10	5.07	5.08	5.08	4.89	4.87	4.88	4.87	4.58	4.56	4.56	4.56
WBS vs. BBE	5.75	5.74	5.75	5.74	5.55	5.54	5.55	5.54	4.94	4.93	4.83	4.82
WBS vs. DBY	5.49	5.45	5.46	5.44	5.21	5.17	5.18	5.17	4.63	4.61	4.61	4.60
WBS vs. KJV	4.63	4.61	4.63	4.62	4.55	4.53	4.54	4.54	4.41	4.41	4.41	4.41
WBS vs. WEB	5.52	5.48	5.51	5.50	5.26	5.22	5.24	5.23	4.75	4.72	4.73	4.72
WBS vs. YLT	6.25	6.22	6.24	6.34	6.06	6.02	6.04	6.08	5.35	5.29	5.23	5.21
YLT vs. ASV	6.34	6.26	6.30	6.29	6.08	6.01	6.05	6.03	5.00	4.95	4.92	4.91
YLT vs. BBE	6.86	6.77	6.84	6.85	6.68	6.62	6.66	6.66	5.99	5.94	5.92	5.92
YLT vs. DBY	6.38	6.31	6.33	6.32	6.15	6.08	6.09	6.07	5.26	5.19	5.13	5.10
YLT vs. KJV	6.39	6.33	6.39	6.39	6.16	6.09	6.15	6.14	5.41	5.33	5.28	5.26
YLT vs. WEB	6.38	6.30	6.34	6.33	6.23	6.16	6.17	6.15	5.51	5.44	5.36	5.33
YLT vs. WBS	6.25	6.22	6.24	6.34	6.06	6.02	6.04	6.08	5.35	5.29	5.23	5.21

TEXT REUSE IN ENGLISH BIBLE VERSIONS: F-MEASURE VS. NOISY CHANNEL EVAL. I





F-Measure: WBS, ASV, DBY, WEB, YLT, BBE

NCE: WBS, ASV, DBY, WEB, BBE, YLT

CONTACT

Speaker

Marco Büchler.

Visit us

🌍 `http://www.etrap.eu`

✉ `contact@etrap.eu`

Stealing from one is plagiarism, stealing from many is research

(Wilson Mitzner, 1876-1933)

eTRAP

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

SPONSORED BY THE

Federal Ministry of Education and Research

LICENCE

INTERDISCIPLINARY CONCEPT OF ETRAP

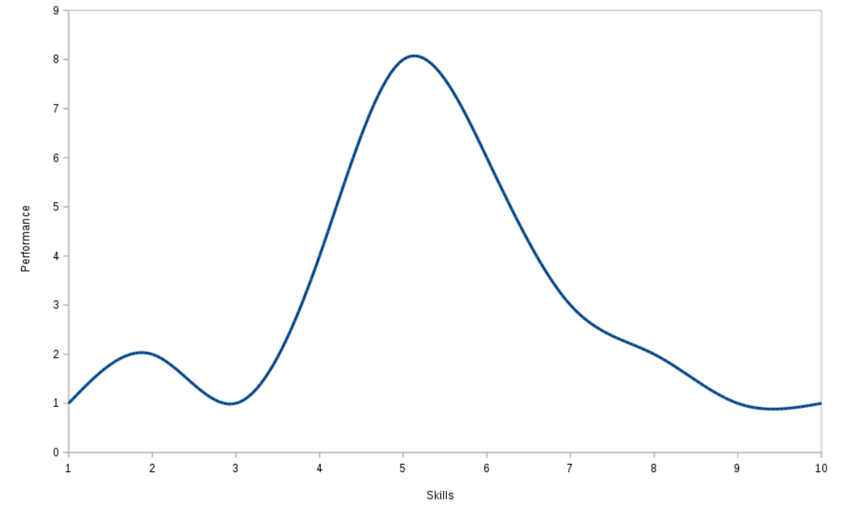
PROFESSIONAL TEAM COACHING OF ETRAP

Professional team coaching for effective group dynamic:

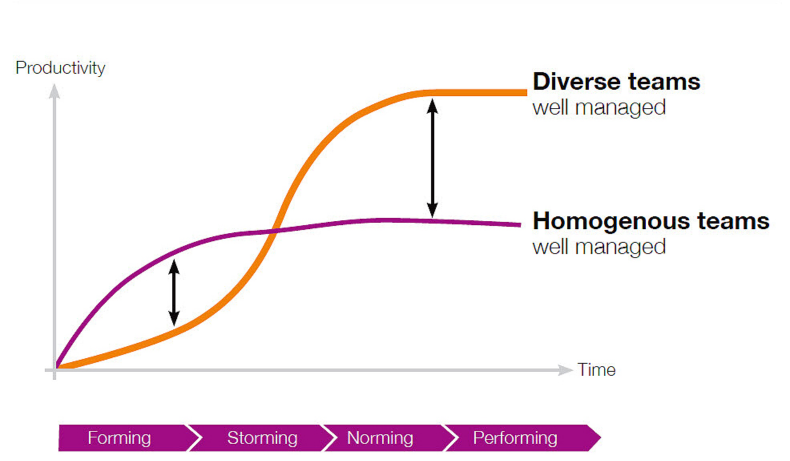
- Effective communication;
- Making the most of strengths;
- Effective delegation.

coach evolve lead
inspeyer

STRENGTHEN YOUR STRENGTHS OR YOUR WEAKNESSES?

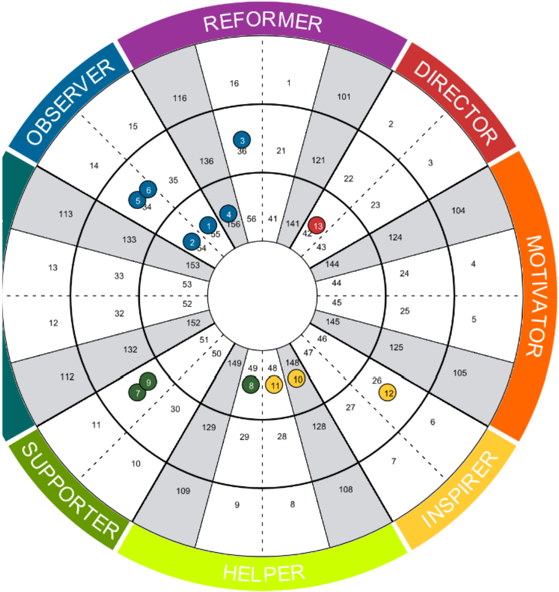


BUILDING A HIGH PERFORMANCE TEAM



TEAM TRAINING WITH PERSONALITY PROFILES

BUILDING A HIGH PERFORMANCE TEAM BY DIVERSITY OF SKILLS


REFORMER
DIRECTOR
MOTIVATOR
INSPIRER
HELPER
SUPPORTER
OBSERVER
16
1
101
2
3
104
4
5
105
6
7
108
8
9
109
10
11
112
12
13
113
14
15
116
36
21
121
22
23
124
24
25
125
26
27
128
28
29
129
30
132
32
33
133
34
35
136
56
41
141
42
43
144
44
45
145
46
47
148
48
49
149
50
51
152
52
53
153
54
55
156
1
2
3
4
5
6
7
8
9
10
11
12
13