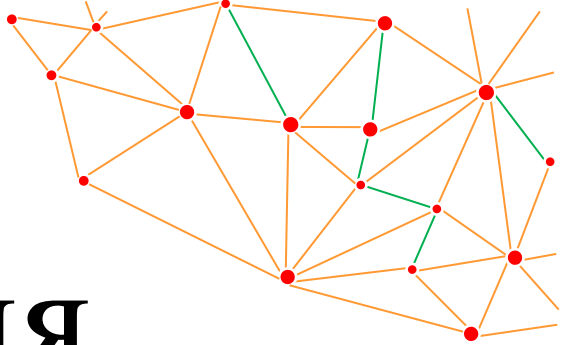


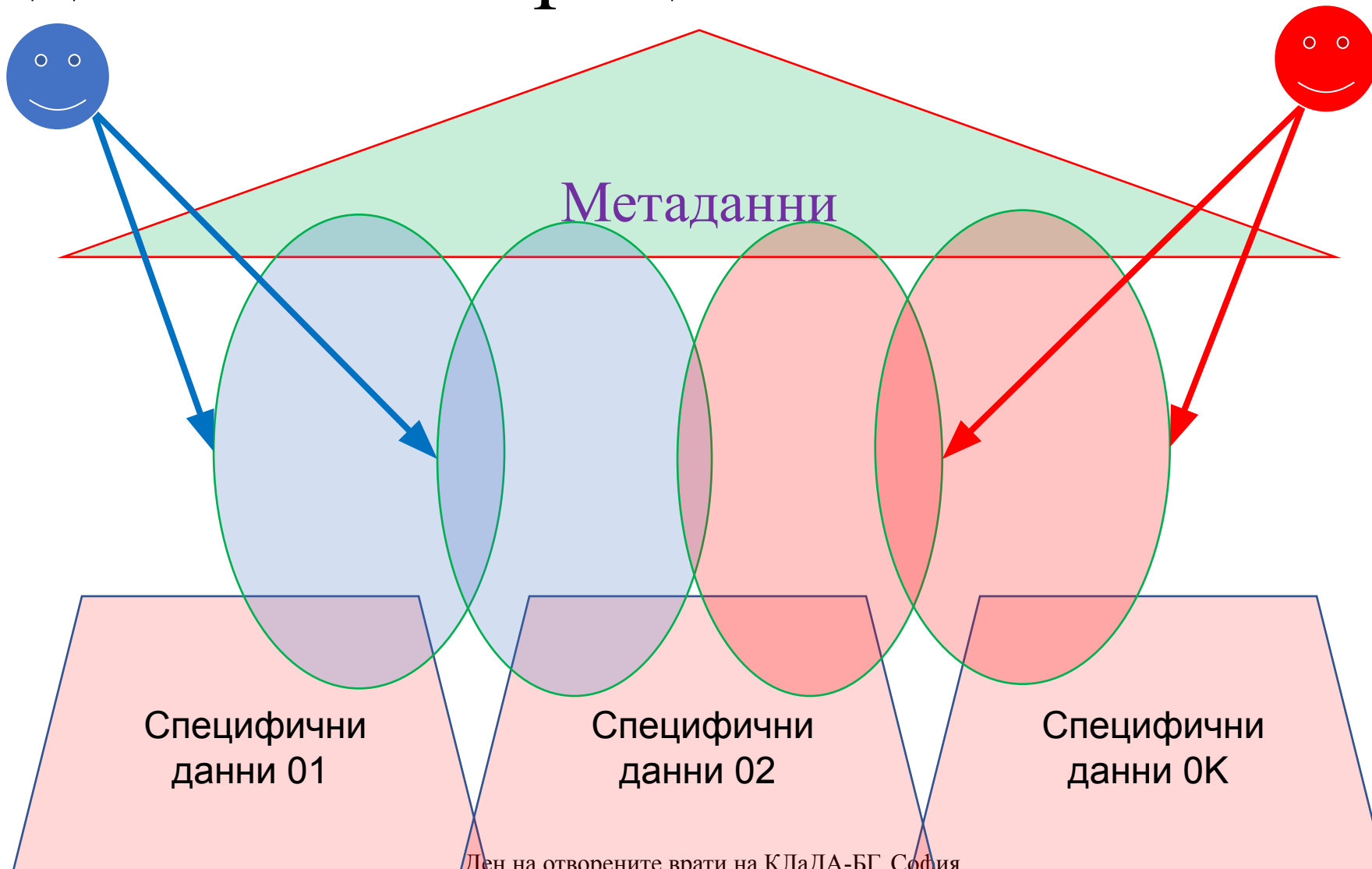


Мрежа от знания, услуги и големи езикови модели за извличане на знание

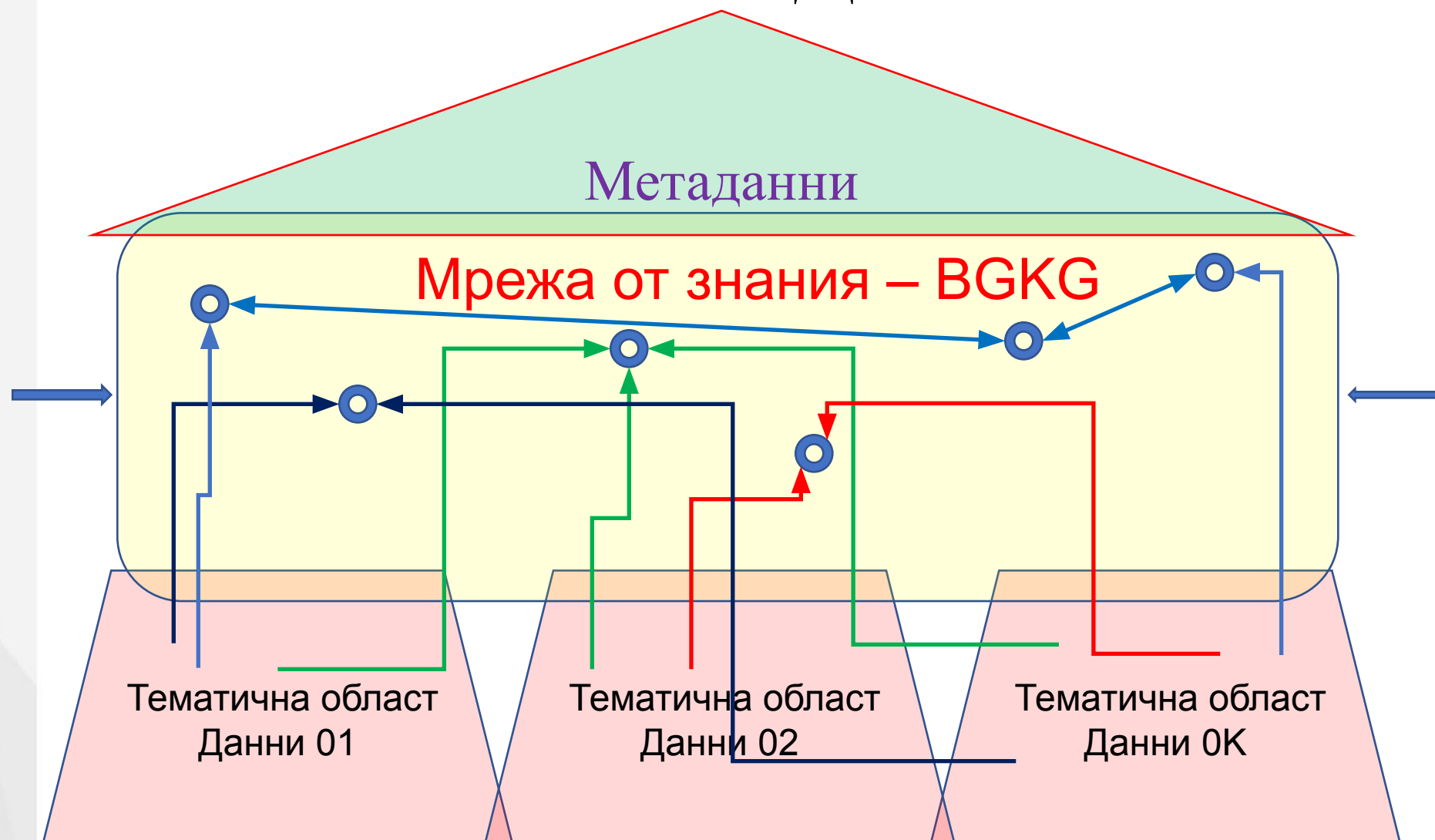
Кирил Симов, Петя Осенова, Живко Ангелов, Николай Паев, Стефан Маринов, Кристин Димитрова, Цветелина Александрова, Теодор Вълчев

ИИКТ-БАН

Исследовательский процесс в CH&X

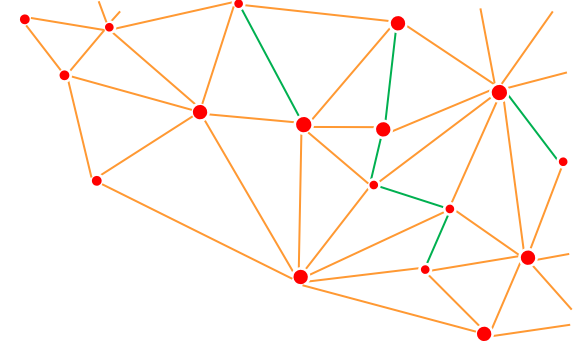


Йерархия от данни в КЛаДА-БГ – BGKG



Ден на отворените врати на КЛаДА-БГ, София

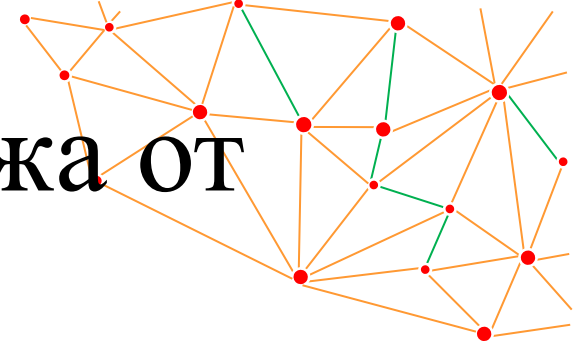
Българска мрежа от знания



КЛаДА-БГ ще постави началото на изграждане на **Българска мрежа за взаимосвързани знания** (Bulgaria-centric Knowledge Graph – **BGKG**) - база за контекстуализацията в КЛаДА-БГ

- **Хора** – техните биографии – важни събития в живота им, роли
- **Обекти** – географски, артефакти и други
- **Събития** – място, време и участници, връзки с други събития
- **Документи** – автори, съдържание, мнения и отношение към хора, събития, обекти – **колко надеждно е представеното знание**
- **Текстът** като *основен източник на информация*
- **Технология** – свързани бази от отворени данни (Linked Open Data) – мрежа от знания (Knowledge graph), ЕЕ (

Архитектура за изграждане на Мрежа от знания



- Създаване на бази от специфични данни (неструктурирани, полуструктурирани и структурирани данни)
- Анализ на данните – извличане на важни елементи от тях: терминология и описания на обекти, представителни за областта
- Концептуализация на резултата от анализа на данните
- Съвързване на концептуализацията с резултата от анализа на данните и от там със самите данни
- Формално представяне на цялостния резултат (RDF), построяване на Мрежата от знания и имплементация на приложения

Резултати от проекти

Къща №	Домакинство №	ФАМИЛИЯ	Име	Възраст	Родство	Професия	Грамотност	Бележка	Общ брой
ЦДА,	ф. 246К, оп.	9,а.е. 524, л. 1–185. Печатно.	Ръкопис.						
1	1	МАНДИЧЕВ	Георгий	33	баща	свещеник	Да		1.
			Илина	27	майка	-	Да		1.
			Иван	9	син	ученик	Да		1.
			Асен	7	син	ученик	Да		1.
			Любица	4	дъщеря	-	Не		1.
			Йосиф	1	син	-	Не		1.
2	2	ДИМИТРОВ	Божин	33	баща	учител	Да		1.
			Траянка	29	майка	-	Да		1.
			Лазар	8	син	ученик	Не		1.
			Димитър	6	син	ученик	не		1.
			Кирил	1	син	-	-		1.
	3	ПОП ТАШОВА	Магдалина	55	майка-вдовица	уч[илищна] слугиня	не		1.
			Тома	30	син	надничар	не		1.
3	4	БЪЧВАРОВ	Димитър	45	баща	готвач	да		1.
			Султана	33	майка	-	да		1.
			Георгий	10	син	ученик	не		1.
			Марика	8	дъщеря	ученичка	да		1.

Публикации

ДИМИТЪР ГЕНКОВ ФИЛОВ

(Роден в 1847 г. — Калофер; убит в 1887 г. — Русе)

Един от първите офицери, получили до Освобождението военно образование в Русия. Участник в Руско-турската война (1877/78) в Българското опълчение. След Освобождението на служба в Източнорумелийската милиция; участва в подготовката и провъзгласяването на Съединението; член на „Привременното правителство“. Като военен комендант на Русе взема участие в Русенския бунт (1887), когато бил убит.

Източници

За фондообразувателя: Анчев, А. Българското опълчение. С., 1968, с. 16; Бендерев, Ан. История на Българското опълчение и Освобождението на България 1877—1878. С., 1930, с. 79; Дренски, Ив. Олимпий Панов. С., 1964, с. 126; Начов, Н. Калофер в миналото. С., 1927 492—493; Радев, С. Строителите... Т. 2, 1911, 696—718; НБКМ — БИА, ф. 209, арх. ед. 40 (Вж. Обзор на архивните фондове, колекции и единични постъпления, съхранявани в Български исторически архив. Кп. 4. С., 1973, 91—99).

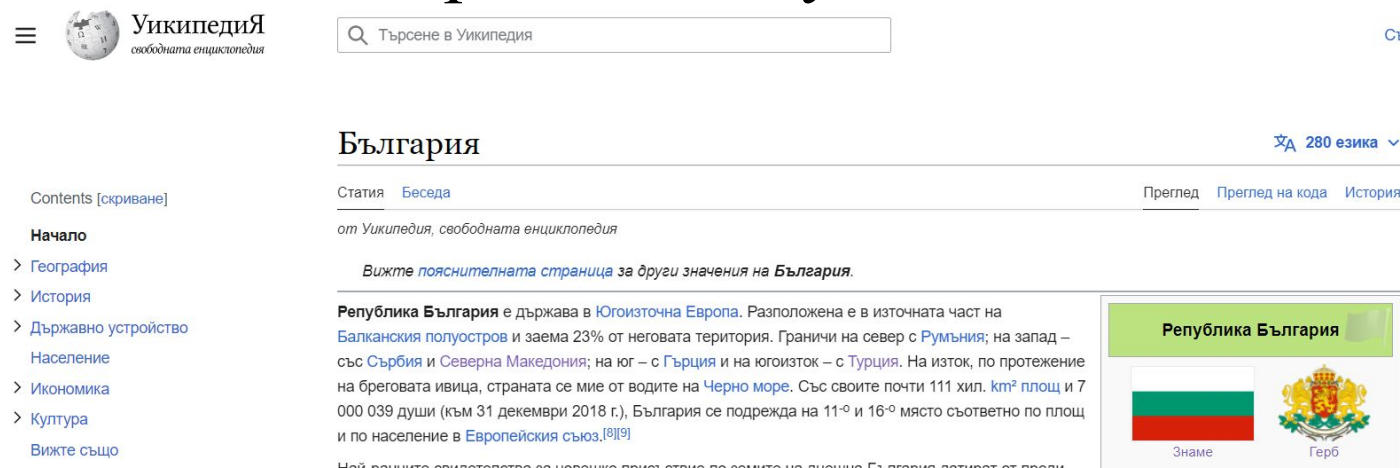
Исторически корпуси 1800-1950, създадени от Народната библиотека „Иван Вазов“ - Пловдив

Корпус от издания за София, създаден от РИМ София

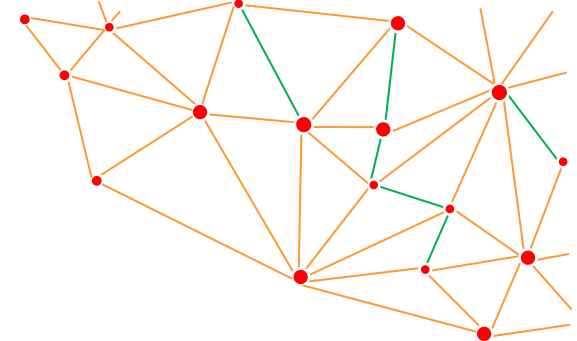


Свободно достъпни ресурси

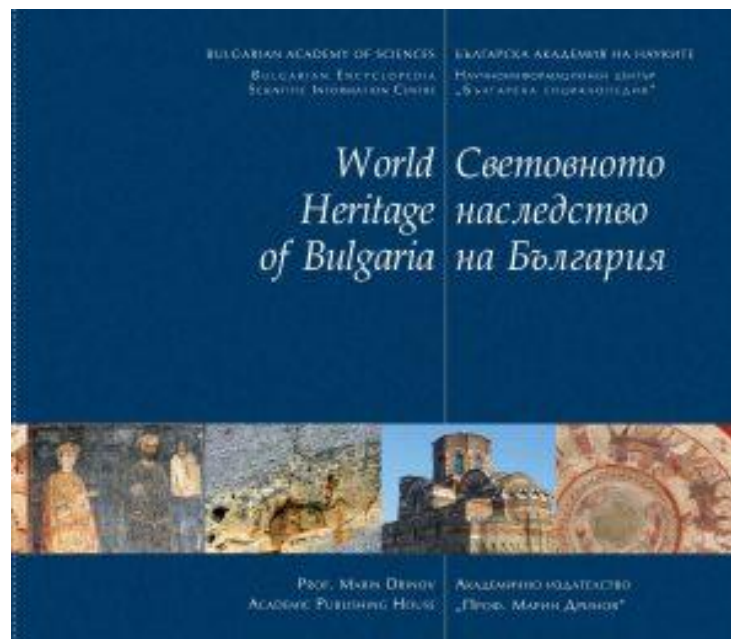
- Обикновено резултат от “crowdsourcing” типични примери:
 - Уикипедия
 - Wikidata
- Достъпност и покритие – включва знанието на много хора
- Липса на надеждност – отразява заблудите на „тълпата“



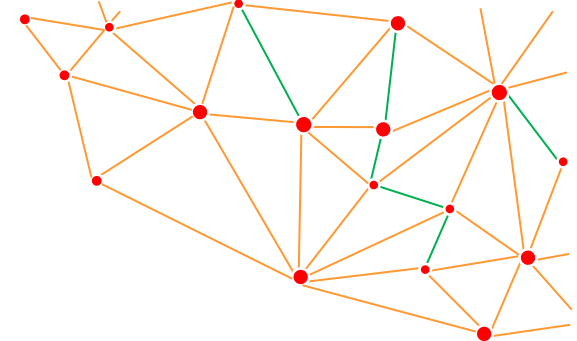
Професионално създадени ресурси



- Научни публикации
- Енциклопедии, терминологични речници и други справочници
- Гарантирано качество
- Ограничен достъп

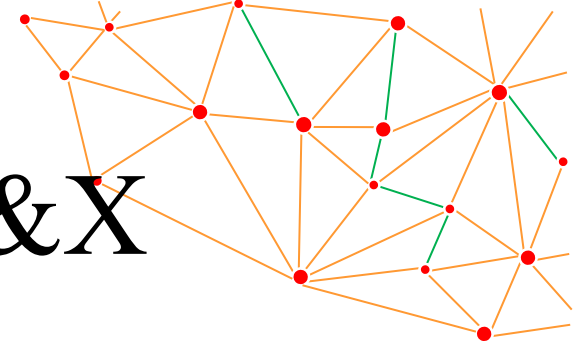


Концептуализация в КЛаДА-БГ



- Избор на основни онтологии, свързани с CH&X – CIDOC-CRM
- Мрежата от думи, съпоставена с избраните онтологии, терминологични речници
- Модели на обекти, създадени от партньорите в консорциума.
Например, модел на *Тип управление* (Режим на пълномощията (27 април 1881 – 6 септември 1883))
- Концептуализация на базата на резултатите от анализа на данни

Езикови ресурси за областта на СН&Х



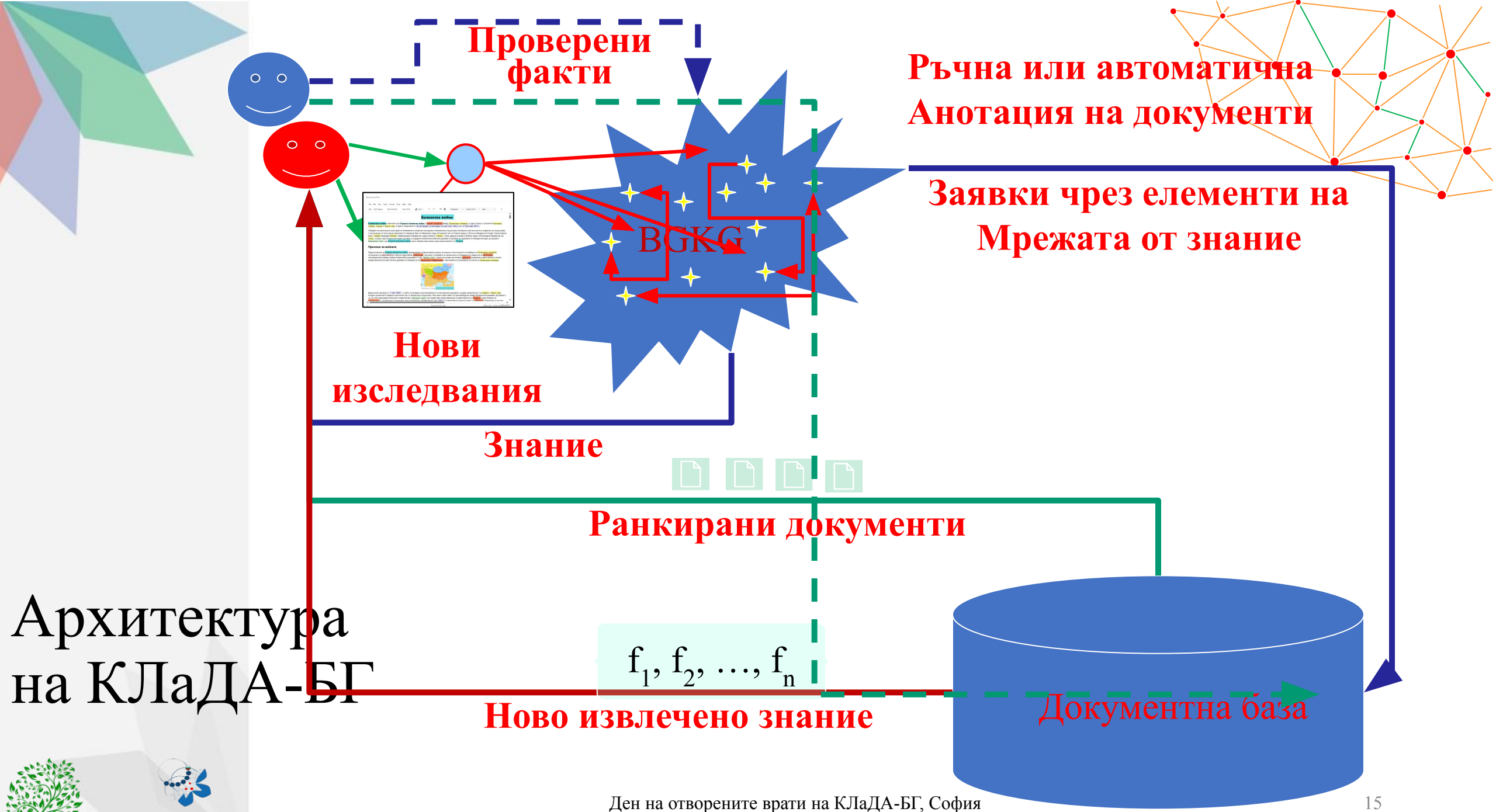
- Тези ресурси са необходими за автоматична обработка на документи в областта за индексирането им и за извличане на знания
- Анотация на морфосинтактично ниво на текстове от дефинициите, примерите в Мрежата от думи и от Уикипедията
- Терминологични речници – свързани с Мрежата от думи
- Специализирана анотационна схема и Корпус, аотиран със събития
- Интегриране на валентен речник с Мрежата от думи

Корпус, аотиран със събития - имена

Label	Description
DOC	Various texts, including documents, excluding juridical documents – see JUR
EVT	Named events like Second World Wars
JUR	Juridical documents: laws, regulations, etc.
LOC	Locations/places — natural or man-made like mountains, lakes, etc., geopolitical units are excluded – see LOC-GPE
LOC-GPE	Geopolitical units (countries, regions, cities, cantons, etc.)
MSC	Miscellaneous names that not included in the other categories
MSR	Measurements with expressed quantity
ORG	Organizations of any kind
PER	People (existing in reality or fictional ones)
PER-GPE	Nationalities (Bulgarian), the birth place, or the place where people live
PER-GRP	Groups of people that cannot be described as PER-GPE or PER-LOC (Slavs, etc.)
PER-LOC	People that are related to geographical region, but not PER-GPE
PRO	Products — tangible and intangible (DOC and JUR excluded)
REF	Bibliographical references, citations of them, links.
SUM	Amounts of money — a subclass of MSR
TIME	Time points or periods

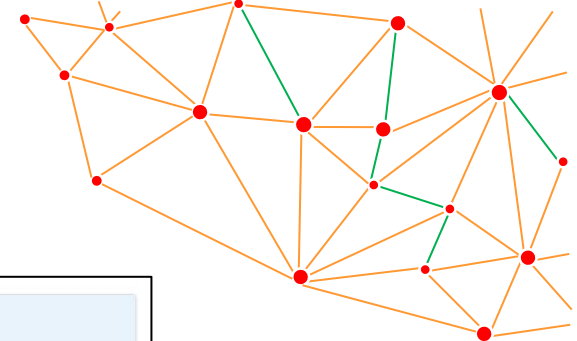
Корпус, аотиран със събития - събития

Event	Roles
Donation	donor (person or organization) recipient (person or organization) theme (object) mediator (person or organization, it could be fund) period-of-iterations (time: the length of time from when the event denoted by the target began to be repeated to when it stopped) goal (situation: the goal for which the donor gives the theme to the recipient) time place
Giving-Birth	brought-into-life (the new born person) parents (the mother and father expressed together, for example “his parents” or “Penka and Toncho Ivanovi”) mother father place (the birth place — usually the name of a city, country or hospital) time (the time of birth — usually it’s a date, but can include hours, or it’s just month and year)
Moving-in-Place	agent (a person) or theme (another type of object) coagent (another person or group of people the agent is moving with) move-from (the place from which the agent or the time moves) move-to (the place where the agent or the theme moves to) time/beginning/end/duration purpose (a situation or another event which causes the moving) goal (a situation/event to be achieved with the moving)



Архитектура на КЛаДА-БГ

Мрежа от знания



1 of 823 row(s) selected [Clear selection](#)

Create New Record

Щастливеца

Select	id	url	text	Actions
☒	2535	https://bg.wikipedia.org/wiki/Алеко_Константинов	Алеко Иваницов Константинов (познат с псевдонима Щастливеца) е български писател, адвокат, общественик и с	
☐	9531	https://bg.wikipedia.org/wiki/Костадин_Щерков	Костадин Николов Щерков е български учител и революционер, деец на Вътрешната македоно-одринска револю	
☐	9563	https://bg.wikipedia.org/wiki/Кочо_Колищърков	Кочо Колищърков е български учител и революционер, деец на Вътрешната македоно-одринска революционна с	
☐	18848	https://bg.wikipedia.org/wiki/Стойчевци	Стойчевци е село в Северна България. То се намира в община Елена, област Велико Търново. Стойчевци е разпол	
☐	18501	https://bg.wikipedia.org/wiki/Рибново	Р?бново е село в Югозападна България. То се намира в община Гърмен, област Благоевград. Известно е в цяла Бъ	
☐	20045	https://bg.wikipedia.org/wiki/Гепидско_кралство	Гепидското кралство е европейска феодална държава от Ранното средновековие, просъществувала със столица С	

Rows per page 50 1-50 of 823

Визуализиране на подтаблица PROPERTIES

NAMES PROPERTIES LINKS DESCRIPTION

Create New Record

id	url	text	category	s.category	comment
15887	2535	Роден	1 януари 1863 г. Свищов, Османска империя	0	
15888	2535	Починал	11 май 1897 г. Радилово, Княжество България	0	
15889	2535	Възраст	(34 г.)	0	
15890	2535	Погребан	Централни софийски гробища, София, Република България	0	
15891	2535	Етнос	българи	0	
15892	2535	Националност	България	0	

Edit Row

id 2535

url https://bg.wikipedia.org/wiki/Алеко_Константинов

text Алеко Иваницов Константинов (познат с псевдонима Щастливеца) е български писател, адвокат, общественик и основоположник на организираното туристическо движение в България.

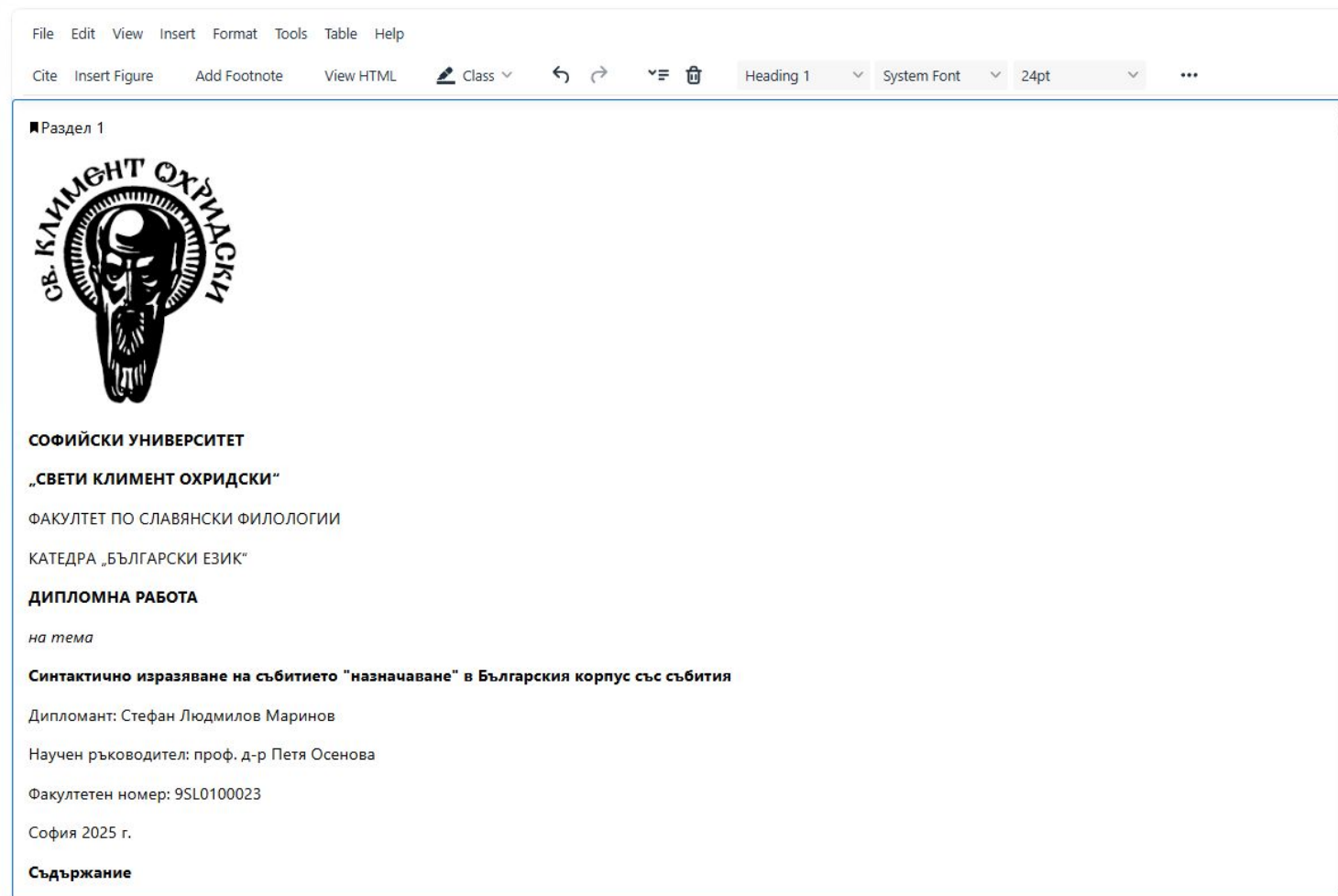
category Личност

s.category писател, адвокат и общественик

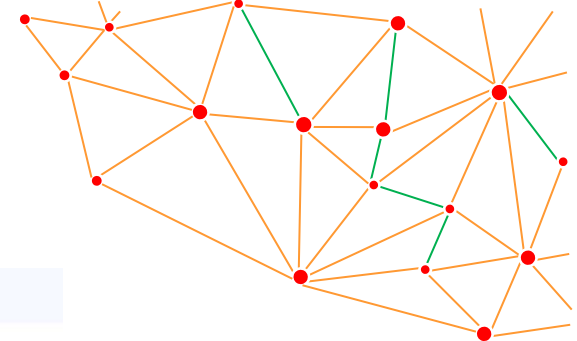
comment

Cancel Save

CLaDA-BG Editor



CLaDA-BG Editor – Анотации



CLaDA-BG TinyMCE Editor

Ньойски договор - част 1

File Edit View Insert Format Tools Table Help

Cite Insert Figure Add Footnote View HTML Class Paragraph System Font 24pt B I U S

Удроу Уилсън настоява да се зачита прогласеният от него принцип за етническо самоопределение на териториите при следвоенното устройство, дори българска граница по линията Мидия – Родосто твърдо е подкрепена от САЩ, но Франция и Великобритания отхвърлят разумния подход на САЩ. В отговор на това президентът напуска конференцията и оставя само някои свои служители да участват в нея.

На 19 септември сутринта, в 10 часа и 25 минути, в министерството на външните работи на Франция Клемансо връчва официално на българската делегация условията за мир. В 25-дневен срок българската страна трябва да представи своите писмени възражения на Върховния съвет, който, след като ги разгледа, ще определи нов срок за окончателния общ договор на България.

Подписване



Стамболийски подписва

Договорот е подписан на 27 ноември 1919 година от министър-председателя Александър Стамболийски от българска страна. След подписването, според някои медии, той счупва писалката заради тежките клаузи. Според други домакините му предложили да я вземе за спомен, но той отказал.

От страна на Антантата страни по договора са Съединените американски щати, Британската империя, Франция, Италия и Япония, посочени като основни сили, както и Белгия, Китай, Куба, Гърция, Хиджаз, Полша, Португалия, Румъния, Кралството на сърби, хървати и словенци, Сиам и Чехословакия.

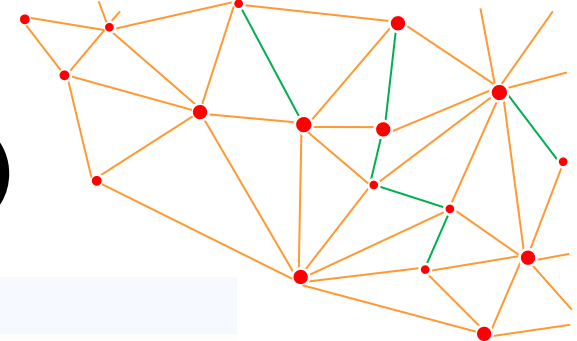
Ньойският договор влиза в сила на 9 август 1920 година.

Условия

Отнети земи

Според договора България трябва да предаде на Кралството на сърби, хървати и словенци Западните покрайнини – областите около Босилеград, Цариброд и Струмица, както и няколко села в Кулско с предимно влашко население. Антантата поема под управление Беломорска Тракия, която скоро след това е дадена на Гърция.

CLaDA-BG Editor – Анотации (филтър)



CLaDA-BG TinyMCE Editor

Ньойски договор - част 1

File Edit View Insert Format Tools Table Help

Cite Insert Figure Add Footnote View HTML Class ↶ ↷ ≡ ☒ Paragraph System Font 24pt

както и става през 1939 г. Президентът **Удроу Уилсън** настоява да се зачита прогласеният от него принцип за етническо самоопределение на териториите при следвоенното устройство, дори българска граница по линията **Мидия – Родосто** твърдо е подкрепена от **САЩ**, но **Франция** и **Великобритания** отхвърлят разумния подход на **САЩ**. В отговор на това президентът напуска конференцията и оставя само някои свои служители да участват в нея.

На 19 септември сутринта, в 10 часа и 25 минути, в министерството на външните работи на **Франция** **Клемансо** връчва официално на българската делегация условията за мир. В 25-дневен срок българската страна трябва да представи своите писмени възражения на **Върховния съвет**, който, след като ги разгледа, ще определи нов срок за окончателния общ договор на **България**.

Подписване



Стамболийски подписва договора

Договорът е подписан на 27 ноември 1919 година от министър-председателя **Александър Стамболийски** от българска страна. След подписването, според някои медии, той счулва писалката заради тежките клаузи. Според други домакините му предложили да я вземе за спомен, но той отказал.

От страна на Антантата страни по договора са Съединените американски щати, Британската империя, Франция, Италия и Япония, посочени като основни сили, както и Белгия, Китай, Куба, Гърция, Хиджаз, Полша, Португалия, Румъния, Кралството на сърби, хървати и словенци, Сиам и Чехословакия.

Ньойският договор влиза в сила на 9 август 1920 година.

Условия

Отнети земи

Според договора **България** трябва да предаде на **Кралството на сърби, хървати и словенци** Западните покрайнини – областите около **Босилеград, Цариброд** и

Annotations

Settings

Zoom Tracking

Zoom out

☐ On/Off All annotations

☒ PER

☐ PER-GRP

☐ PER-GPE

☐ PER-LOC

☐ PER-LOC-GPE

☐ NAME

☒ ORG

☒ TIME

☒ LOC-GPE

☐ COREF

☒ EVT

☐ LANG

☐ PUBLISHER

☐ EDITOR

☐ SOURCE

☐ PAGES

☐ DATE

☐ ISBN

☐ MSC

☐ MSR

☐ SUM

☐ PRO

☐ REF

☐ AUTHOR

☐ TITLE

☐ TYPE

☐ LINK

☐ ...

☐ On/Off Events

☐ обучение

☐ събитие

☐ край

☐ назначаване

☐ к-събитие

☐ смърт

☐ раждане

☐ занимание

☐ основаване

☐ преместване

☐ промяна

☐ характеристика

☐ унищожение

☐ каузация

☐ напускане

☐ Track Sentence

CLaDA-BG Editor – Мрежата от знание

CLaDA-BG TinyMCE Editor

Ньойски договор - част 1

File Edit View Insert Format Tools Table Help

Cite Insert Figure Add Footnote View HTML

Class

Paragraph

System Font

24pt

B I U S ...

1939 г. Президентът Удроу Уилсън настоява да се зачита прогласеният от него принцип за етническо самоопределение на териториите при следвоенното устройство, дори българска граница по линията Мидия – Родосто твърдо е подкрепена от САЩ, но Франция и Великобритания отхвърлят разумния подход на САЩ. В отговор на това президентът напуска конференцията и оставя само някои свои служители да участват в нея.

На 19 септември сутринта, в 10 часа и 25 минути, в министерството на външните работи на Франция Клемансо връчва официално на българската делегация условията за мир. В 25-дневен срок българската страна трябва да представи своите писмени възражения на Върховния съвет, който, след като ги разгледа, ще определи нов срок за окончателния общ договор на България.

Подписване



Стамболийски подписва договора

Договорът е подписан на 27 ноември 1919 година от министър-председателя Александър Стамболийски от българска страна. След подписването, според някои медии, той счулва писалката заради тежките клаузи. Според други домакините му предложили да я вземе за спомен, но той отказал.

От страна на Антантата страни по договора са Съединените американски щати, Британската империя, Франция, Италия и Япония, посочени като основни сили, както и Белгия, Китай, Куба, Гърция, Хиджаз, Полша, Португалия, Румъния, Кралството на сърби, хървати и словенци, Сиам и Чехословакия.

Ньойският договор влиза в сила на 9 август 1920 година.

Условия

Отнети земи

Според договора България трябва да предаде на Кралството на сърби, хървати и словенци Западните покрайнини – областите около Босилеград, Цариград и Струмица, както

Info

Австрия

Александър_Стамболийски

Class:: български политик

Class:: Стамболийски в Париж през 1921 г. Източник: ДА „Архиви“

Роден: 1 март 1879 г.Славица, Източна Румелия

Починал: 14 юни 1923 г.Славица, България

Възраст: (44 г.)

Етнос: българи

Образование: Мюнхенски университет

Class:: Политика

Партия: Български земеделски народен съюз

Class:: 20-и министър-председател на България

Class:: 6 октомври 1919 – 9 юни 1923

Човек представител в УИМ ОНС В

Search

CLaDA-BG Editor – Свързани документи

CLaDA-BG TinyMCE Editor

Александър Стамболийски - част 1

File Edit View Insert Format Tools Table Help

Cite Insert Figure Add Footnote View HTML Class Paragraph System Font 24pt B I U S

Александър Стамболийски

Александър Стоименов Стамболийски (1 март 1879 – 14 юни 1923) е български политик, водач на Българския земеделски народен съюз (БЗНС). Той е министър-председател на България в правителството на БЗНС (1919 – 1923).

Биография

Ранни години

Роден е на 1 март 1879 г. в Славовица, Пазарджишко. Александър Стамболийски учи в земеделското училище в Садово (1893 – 1895) и завършва Лозаро-винарското училище в Плевен (1895 – 1897). В Плевен е ученик на основателя на Българския земеделски съюз (БЗС) Янко Забунов. През 1899 участва в учредителния конгрес на БЗНС. През следващите години учи философия в Хале и агрономия в Мюнхен, но прекъсва образованието си, поради заболяване от туберкулоза.



Стамболийски като студент в Хале

Навлизване в политиката

След завръщането си в България се занимава с политическа дейност. През 1908 година Александър Стамболийски е избран за народен представител от БЗНС. Той се превръща във фактическия водач на БЗНС и под негово влияние съюзът е преобразуван от селскостопанска организация в политическа партия. Народен представител е XIV (1908 – 1911) и XVI-XV (1911 – 1914) ОНС.

Info

Авторитаризъм

Next

Агрономия

Next

Александър_Оббов

Next

Александър_Стамболийски

Next

Антанта

Асен_Стамболийски

Атентат

Балканска_федеративна
ублика

Борис_III

Българи

България

Info

Class:: 6 октомври 1919 – 9 юни 1923

Народен представител в:: XIV ОНС V
ВНС XVI ОНС XVII ОНС XVIII ОНС XIX ОНС
XX ОНС

Class:: Семейство

Съпруга: Милена Даскалова

Деца: 2 Асен Стамболийски, Надежда
Стамболийска

Class:: Подпис

Class:: Александър Стамболийски в
Общомедия

аса: ccc

Mentioned in other documents

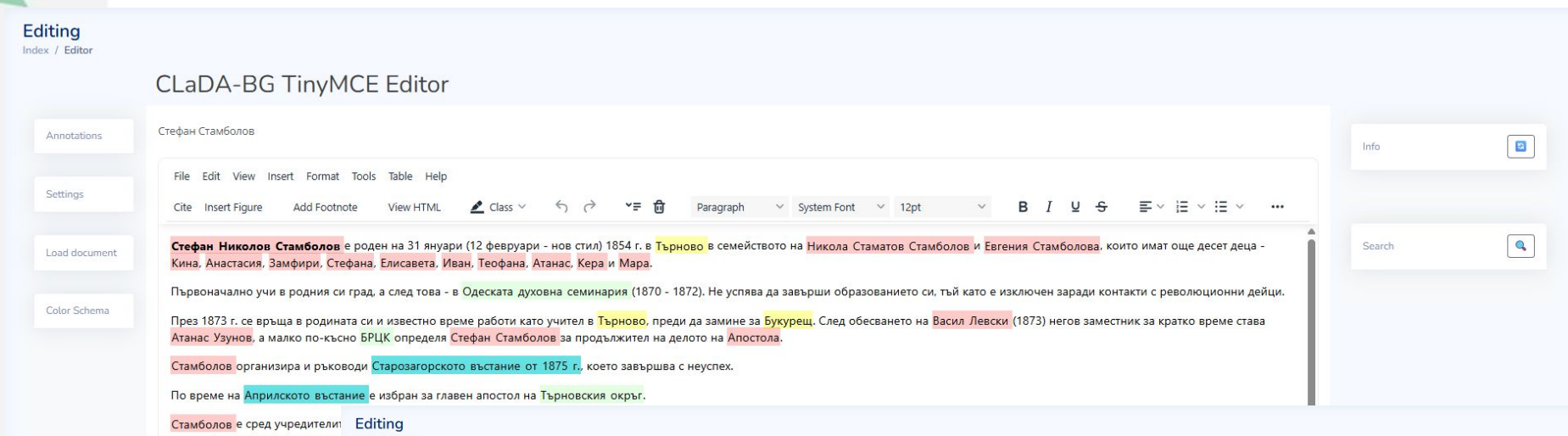
Александър Стамболийски - част 1

Български земеделски народен
съюз (БЗНС) - част 1

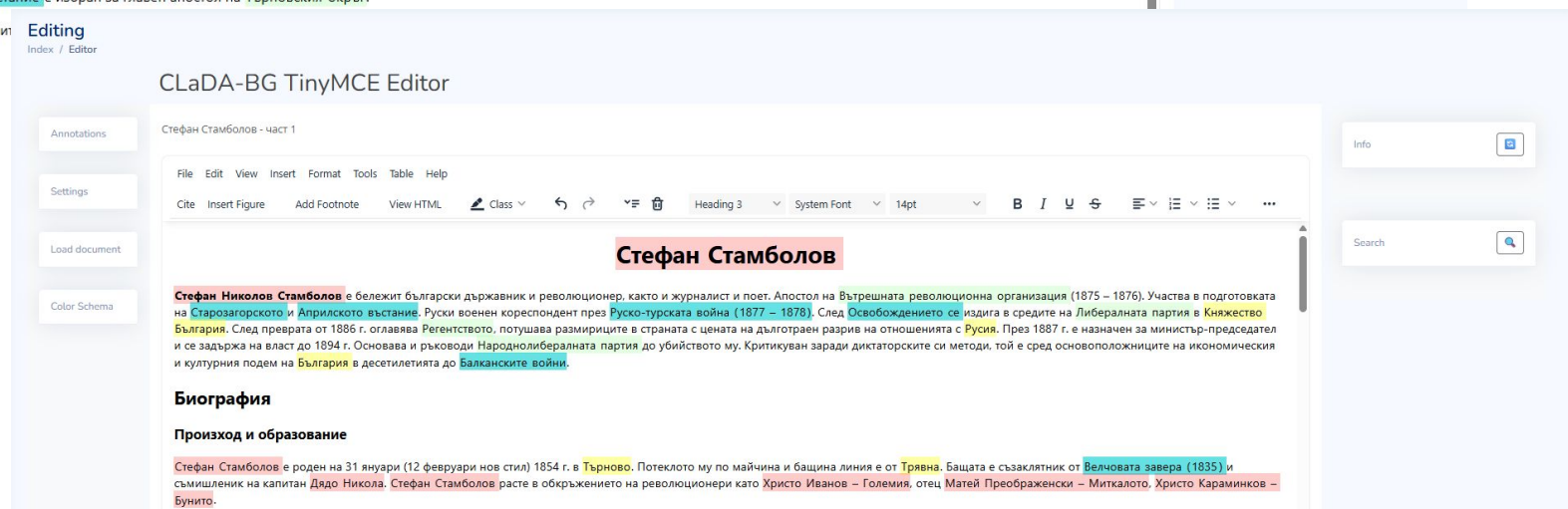
Добруджански въпрос

Георги Димитров - част 1

CLaDA-BG Editor – Свързани документи



The image shows the CLaDA-BG TinyMCE Editor interface. On the left, there is a sidebar with buttons for 'Annotations', 'Settings', 'Load document', and 'Color Schema'. The main editing area displays a document titled 'Стефан Стамболов'. The document text is annotated with various tags and colors. For example, 'Стефан Николов Стамболов' is highlighted in red, and 'Търново' is highlighted in yellow. The document content includes a paragraph about his birth and family, a paragraph about his education and early career, and a paragraph about his role in the April Uprising. On the right side of the editor, there are buttons for 'Info' and 'Search'.



The image shows the CLaDA-BG TinyMCE Editor interface. On the left, there is a sidebar with buttons for 'Annotations', 'Settings', 'Load document', and 'Color Schema'. The main editing area displays a document titled 'Стефан Стамболов - част 1'. The document text is annotated with various tags and colors. For example, 'Стефан Николов Стамболов' is highlighted in red, and 'Търново' is highlighted in yellow. The document content includes a paragraph about his birth and family, a paragraph about his education and early career, and a paragraph about his role in the April Uprising. On the right side of the editor, there are buttons for 'Info' and 'Search'.

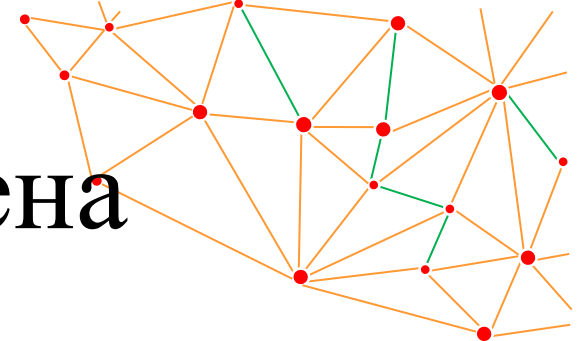
Разпознаване на собствени имена

- Разпознаване на имена и категориите им в текст
- Обучен е BERT (енкодер) модел
- Индексиране на документи спрямо имената в тях
- F1 е 96 %

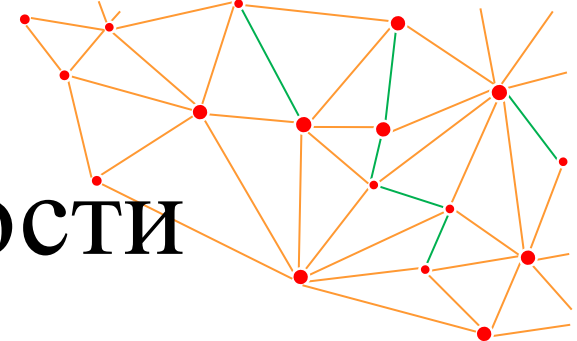
word	Портретът	на	Левски	виси	на	стената	.
UPOS	NOUN	ADP	PROPN	VERB	ADP	NOUN	PUNCT
NER	O	O	B-PER	O	O	O	O

word	Левски	отново	се	класира	за	Лига	Европа	.
UPOS	PROPN	ADV	PRON	VERB	ADP	NOUN	PROPN	PUNCT
NER	B-ORG	O	O	O	O	B-OTH	I-OTH	O

word	Влакът	пристига	в	град	Левски	привечер	.
UPOS	NOUN	VERB	ADP	NOUN	PROPN	NOUN	PUNCT
NER	O	O	O	O	B-LOC	O	O

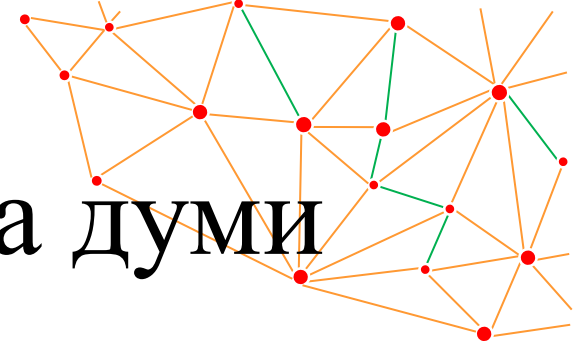


Свързване на наименованите същности



- Анотиране на наименованите същности с идентификатори от Мрежата от знания
- Използване на ранкиращ модел, който създава контекстно представяне на имената и го съпоставя с описанията на същностите в Мрежата от знания
- Точността при определяне на правилния идентификатор на първа позиция е 83 %
- При поставянето му сред първите три възможности точността е 93 %

Анотиране със значения от Мрежата думи



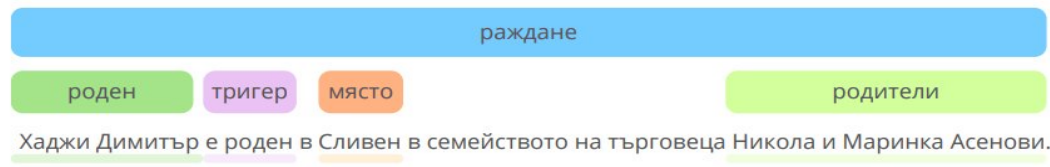
- Определяне на значението на думите
- Свързване с Мрежата от думи (ВТВ-WordNet)
- На 2 стъпки:
 - Разпознаване на възможни основни форма на думата чрез речник и/или модел
 - Откриване на правилното значение чрез модел
- Точност от над 86% върху многозначни думи

Анотиране на събития

- Извличане и класификация на събитията в текста, техните участници и категориите им
- Анотирано множество за обучение - Български корпус със събития
- Използване на T5 модел, който генерира структурирано описание на събитията в дадено изречение
- Точност при определяне на типа от 86 % и F1 – 77 %

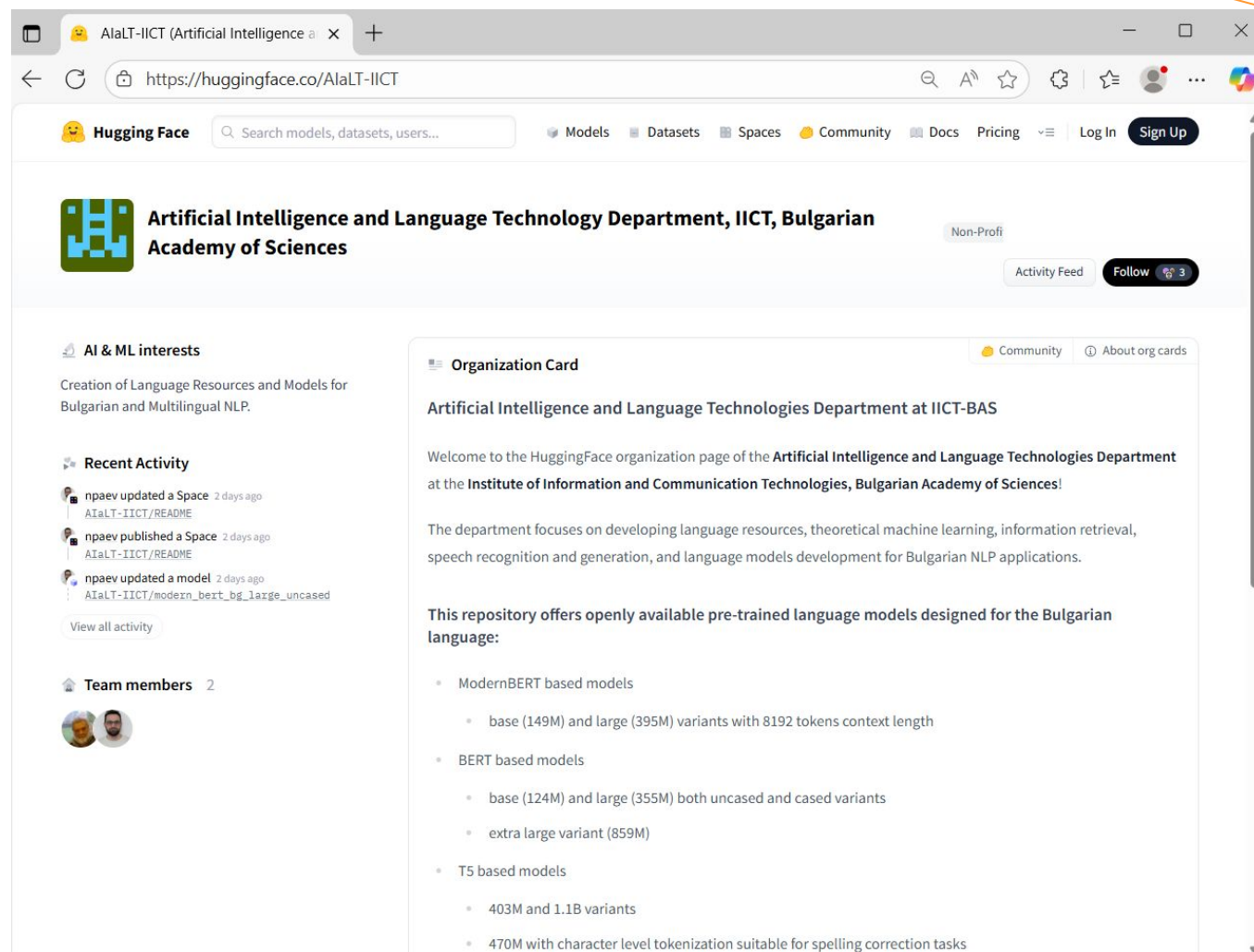
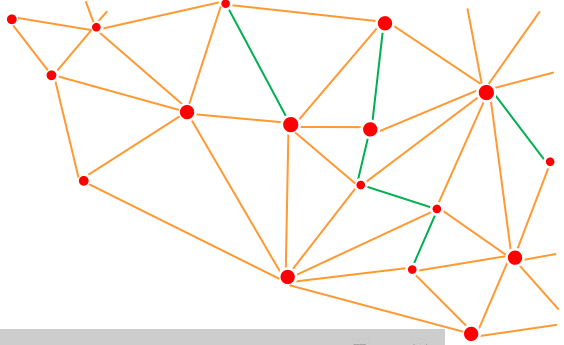
Token_id	Token	Ne_1	Ne_2	Ev_1	Ro_11	Ro_12
29	Той	C<		з<	а<	
30	е			а	т<	
31	министър-председател			н	р	п<
32	на			и		
33	България	L<		м	в<	
34	в			а		
35	правителството	O<		н	в<	
36	на	R		и	ъ	
37	БЗНС	G	O<	е	з	
38	(				л	
39	1919		T<		о	
40	-				ж	
41	1923		T<		и	
42	)				т	
43	.					

```
[{
  "type": "occupation",
  "text": "He is a prime-minister of
Bulgaria in the government of BZNS
(1919 - 1923).",
  "roles": [
    ["agent", "He"],
    ["position", "prime-minister"],
    ["assignor", "the government
of BZNS"],
    ["beginning", "1919"],
    ["end", "1923"],
    ["assignor", "Bulgaria"],
    ["trigger", "is a prime-minister"]
  ]
}]
```



Свободно достъпни предварително обучени модели

<https://huggingface.co/AIaLT-IICT>



The screenshot shows the Hugging Face profile page for the organization AIaLT-IICT. The page includes a header with the organization's name and a 'Non-Profit' badge. Below the header, there is a section for 'AI & ML interests' with the description 'Creation of Language Resources and Models for Bulgarian and Multilingual NLP.' and a 'Recent Activity' section showing updates to Spaces and models. The 'Team members' section lists two members. The 'Organization Card' section provides a welcome message and details about the department's focus on developing language resources and models for Bulgarian NLP applications. It also lists the types of models available: ModernBERT based models, BERT based models, and T5 based models.

AI & ML interests
Creation of Language Resources and Models for Bulgarian and Multilingual NLP.

Recent Activity

- npaev updated a Space 2 days ago
AIaLT-IICT/README
- npaev published a Space 2 days ago
AIaLT-IICT/README
- npaev updated a model 2 days ago
AIaLT-IICT/modern_bert_bg_large_uncased

[View all activity](#)

Team members 2

Organization Card

Artificial Intelligence and Language Technologies Department at IICT-BAS

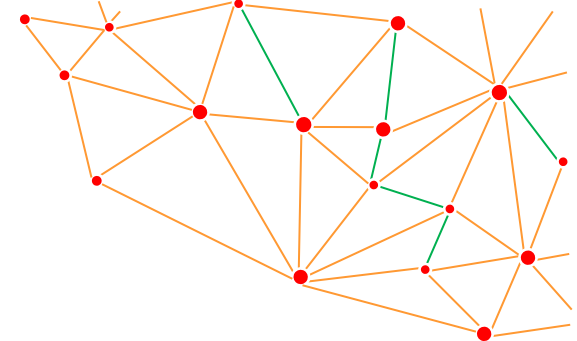
Welcome to the HuggingFace organization page of the **Artificial Intelligence and Language Technologies Department** at the **Institute of Information and Communication Technologies, Bulgarian Academy of Sciences**!

The department focuses on developing language resources, theoretical machine learning, information retrieval, speech recognition and generation, and language models development for Bulgarian NLP applications.

This repository offers openly available pre-trained language models designed for the Bulgarian language:

- ModernBERT based models
 - base (149M) and large (395M) variants with 8192 tokens context length
- BERT based models
 - base (124M) and large (355M) both uncased and cased variants
 - extra large variant (859M)
- T5 based models
 - 403M and 1.1B variants
 - 470M with character level tokenization suitable for spelling correction tasks

Демо



- Система за аотиране и свързване на наименовани същности и събития
- Търсене в аотирани документи и Мрежата от знание
- Системата за автоматичен анализ на текстове