СВЕЩЕНИЯТ СМЕТАЧ

ТОДОР АРНАУДОВ - ТОШ

АНЕЛИЯ

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ ИЗКУСТВЕН РАЗУМ И РАЗВИТИЕ НА ЧОВЕКА

ИСТОРИЯ ТЕОРИЯ И ПИОНЕРИ МИНАЛО НАСТОЯЩЕ И БЪДЕЩЕ

от автора на първия света университетски курс по Универсален изкуствен разум и Теория на разума и вселената

THE PROPHETS OF THE THINKING MACHINES

ARTIFICIAL GENERAL INTELLIGENCE & TRANSHUMANISM

HISTORY THEORY AND PIONEERS; PAST PRESENT AND FUTURE

THE SACRED COMPUTER

TODOR ARNAUDOV - TOSH

ANELIA

THE PROPHETS OF THE THINKING MACHINES
ARTIFICIAL GENERAL INTELLIGENCE & TRANSHUMANISM

HISTORY THEORY AND PIONEERS
PAST PRESENT AND FUTURE

by the author of the world's first university course in
Artificial General Intelligence and the
Theory of Universe and Mind

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ

ИЗКУСТВЕН РАЗУМ И РАЗВИТИЕ НА ЧОВЕКА

ИСТОРИЯ ТЕОРИЯ И ПИОНЕРИ; МИНАЛО НАСТОЯЩЕ И БЪДЕЩЕ

Редакция: 20.8.2025
http://twenkid.com/agi
https://github.com/twenkid/sigi-2025
http://artificial-mind.blogspot.com

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ

Изкуствен разум и развитие на човека:

История, теория и пионери

Минало настояще и бъдеще

Тодор Арнаудов – Тош

ПРИЛОЖЕНИЕ АНЕЛИЯ

Преглед и бележки по научни работи от и с участието на Анелия Ангелова, Пламен Ангелов, Никола Касабов, Димитър Филев, Александър Тошев, Любомир Бурдев, Мира Дончева, Драгомир Ангелов, Драгомир Радев, Кристина Тутанова, Руслан Митков, Зорница Козарева, Преслав Наков, Галя Ангелова, Кирил Симов, Ани Ненкова, Тодор Михайлов, Веселин Стоянов, Веселин Райчев, Мартин Вечев, Светослав Караиванов; Красимир Атанасов, Стоян Михов, Петя Копринкова-Христова, Виргиния Савова; Васил Сгурев, Димитър Добрев и други от България и света по информатика, изкуствен интелект, машинно обучение; компютърно зрение, компютърна графика и обработка на изображения; компютърна лингвистика и обработка на естествен език; езици за програмиране, синтез на програми, автоматично програмиране; изкуствени, импулсни и други невронни мрежи, размита логика и други. Обзор на друга свързана стара и съвременна литература от тези области.

Други споменати български изследователи: Антон Александров (INSAIT, BgGPT), Петко Георгиев, Илиян Заров, Румен Данговски (Petko Georgiev (DeepMind), Roumen Dangovski, Iliyan Zarov (Meta AI: LLAMA); Васил Чаталбашев, Красимир Коларов и др. От ПУ: Георги Тотков, Христо Крушков, Христо Танев; Димитър Благоев, Васил Василев, Александър Пенев и др.

СВЕЩЕНИЯТ СМЕТАЧ
МИСЛЕЩИ МАШИНИ, ТВОРЧЕСТВО И РАЗВИТИЕ НА ЧОВЕКА

Целогодишна виртуална конференция „Мислещи машини 2025“, или Self-Improving General Intelligence 2025 – SIGI-2025. Продължение на може би втората най-стара международна „конференция“ за универсален изкуствен разум (AGI): SIGI-2012-1, провела се присъствено в Пловдив през 2012 г.

THE SACRED COMPUTER
THINKING MACHINES, CREATIVITY AND HUMAN DEVELOPMENT

Thinking Machines 2025/Self-Improving General Intelligence SIGI-2025:
a yearlong virtual conference, continuing SIGI 2012-1.

Обзори на някои тематични раздели освен конкретни български учени

* Компютърно зрение и обработка на изображения в различни приложения: самоуправляващи се превозни средства, разпознаване на образи и класификация и др.
* Машинно обучение, изкуствени невронни мрежи
* #Vision Tasks #Vision-Language Tasks #Зрителни задачи Изброяване на задачи от компютърното зрение
* Други статии по съвременно разделяне на изображения свързани с работи на Анелия Ангелова и Александър Тошев; Мира Дончева и др. – Current Image Segmentation #segmentation
* Невроморфни системи, импулсни невронни мрежи: Никола Касабов, Пламен Ангелов
* Размита логика – Димитър Филев, Н.Касабов, П.Ангелов и др.
* Класически трудове по разрешаване на многозначност чрез използване на контекста и корпуси – Word Sense Disambiguation WSD, Context, Corpus Linguistics
* Ранна работа в статистическата езикова обработка, корпусна лингвистика, извличане на данни, групиране и др. над големи обеми от данни и Интернет – Драгомир Радев; Зорница Козарева и др.
* Компютърна лингвистика и обработка на естествен език със статистически методи и машинно обучение – Кристина Тутанова, Зорница Козарева, Драгомир Радев и др.
* Първопроходна работа на Руслан Митков и школата му в компютърната лингвистика – разрешаване на анафори, подпомагане на превода чрез лексикология/лексикография и преводна памет.
* Школата на ПУ Паисий Хилендарски в Компютърната лингвистика от края на 1980-те и 1990-те и след това – морфологичен анализ и други видове разбор и моделиране на българския език, лексикология и лексикография, машинно обучение и извличане на структури (групиране, клъстери), извличане на информация; анализ и синтез на реч и др. – Георги Тотков, Христо Крушков, Христо Танев, Зорница Козарева, Атанас Чанев, Димитър Благоев; Тодор Арнаудов и др.
* Бележки към конференцията по Компютърна лингвистика CLIB 2024 в София
* Мултимодални пораждащи модели #multimodal #мултимодални
* Programming languages, program synthesis & verification, compilers and code optimization, formal verification, static analysis, interpreters, concurrency… #programlanguages #programsynthesis – Програмни езици, синтез на програми, верификация, оптимизация, интерпретатори и компилатори … – Веселин Райчев, Мартин Вечев, Светослав Караиванов; Васил Василев, Александър Пенев и др.
* БАН – група по невроморфни системи; история на ИИ в България и др. – Петя Копринкова-Христова и др.; групи по крайни автомати и др. (Стоян Михайлов), обобщени мрежи – Красимир Атанасов; исторически: В.Сгурев; свързани с определения на общ ИИ: Д.Добрев и др.
* Бележки за ЕЕГ и за българския принос в техники за изчистване на шума при снемане на ЕКГ(EEG, ECG) – Чавдар Левков и др.
* И др.
* Виж за други българи в някои от тези и други области като роботиката, която е една от

„*българските*“ области подобно на компютърната лингвистика и в нея има няколко пионери в епигенетичната роботика (роботика на развитието); философията и когнитивната наука, конекционистки системи, симулации за роботи, учене с подкрепление; други пионерни работи в България от ТУ София и др. в ***основния том***; приложение *Лазар* и др.

(…)

Езици: български и английски **| The content** is in **English and Bulgarian**

Томове и приложения на *Пророците ...*

Съществуващи и някои възможни бъдещи томове

* **#prophets –** Основен том (>1859 стр., 13.8.2025); Обзор на Теория на Разума и Вселената, сравнение с работи в други школи, които преоткриват и повтарят, или пък предхождат обобщаването на принципите за създаване на общ изкуствен интелект, които бяха формулирани още в началото на 2000-те г. и постепенно се сбъднаха и се сбъдват. Документален преглед на огромен обем научни школи, литература и факти, кратка и подробна хронология ... #tosh1

* **#purvata –** ***„****Първата модерна стратегия за развитие чрез ИИ е публикувана от 18-годишен българин през 2003 г. и повторена и изпълнена от целия свят 15-20 години по-късно: Българските пророчества: Как бих инвестирал един милион с най-голяма полза за развитието на страната?“* #tosh2 (31.5.2025, 248 стр.)

* **#listove –** Многообразие от теми сред които класическа и съвременна роботика и планиране, мулти-агентни системи – класически и съвременни с големи езикови модели; невронауки и невроморфни системи, съзнание и панпсихизъм, алгоритмична сложност, други теории на всичко и вселената сметач; когнитивна лингвистика и мислене по аналогия, езикови модели и машинно обучение – исторически и най-нови системи, мултимодални модели, основни модели за агенти и роботи; обзор на научни статии, новини, платформи на чатботове и други пораждащи модели за различни модалности и практика; съветска школа в изкуствения интелект и мн.др. (...), 414 стр. (13.8.2025 г.)

* **#mortal – Нужни ли са смъртни изчислителни системи за създаване на универсални мислещи машини?,** „Смъртните“ системи са свързани с носителя си, за разлика от „безсмъртни“, за каквито се смятат „обикновените“ компютри. Но дали и невроморфните са наистина невроморфни, и какво точно е „безсмъртност“, „смъртност“, „самосъздаване“ (автопоеза) и дали въобще е възможна. Наистина ли са по-ефективни невроморфните системи, както и живите или по-модерните електронни технологии с по-малки транзистори, или ефективността е избор на „счетоводство“ и скриване на реалните разходи за създаването и съществуването на съответната технология? (...) 70 стр.

* **#universe6 #UnM6 –** Вселена и Разум 6, Т.Арнаудов– #tosh3; съзанание, „метафизика“, „умоплащение“ ... на английски; свързана с теми от #mortal (...)

* **Universe and Mind 6** – Connected to "Is Mortal Computation..." – in English. Why infinity doesn't exist and Goedel theorems are irrelevant for thinking machines? What is Truth, Real and Realness and Why? The fundamentality of mapping (...)

* **#sf #cyber –** Научна фантастика за ИИ, Футурология, Кибернетика ... Подробен преглед и сравнение на статия на Майкъл Левин от 2024 г. за самоимпровизиращата се памет с идеи от Теория на Разума и Вселената.

* **#irina –** Беседи и подробни бележки и др. статии; Ирина Риш; вижданията на Йоша Бах и др. и съвпаденията на идеите му с Теория на Разума и Вселената, публикувана 20 години преди коментираните дискусии; интервю с Питър Вос на ръба преди „ерата" на ентусиазма към Общия ИИ през 2013 г.; сбъднали се предвиждания от 2005 г. за машинния превод и творчеството и за автоматичното програмиране от 2018 г. и др.; беседа с участието на Майкъл Левин (повече от него в #Основния том, #Кибернетика и #Листове.

* **#lazar #lotsofpapers –** Работата на някои български и мн. др. учени, около зората на напредъка на обучението на дълбоки невронни мрежи; автоматичен синтез на програми, компютърно зрение от миналото и настоящето, големи езикови модели, ...
* **A survey of various papers** and the work of particular researchers in many fields of AI, machine learning, deep learning, cognitive science, computer science etc., Explanation and summary of most important seminal publications, milestones, concepts, methods, topics, quotes, keywords, points, schools of thought; links between them; notes etc.. Groundbreaking or important researchers or related to the flow and context of the reviewed topics; works in AI, ML, CV, ANN, DL, ... throughout history, classical 1960s, 1970s, 1980s, 1990s, 2000s, early 2010s to 2020s... The evolution of ML and computer vision techniques before the deep learning era. Computer Vision, Program Synthesis. Lifelong Learning, Reinforcement Learning, Human-Computer Interaction, Agents, Computer Vision;

* **#anelia –** Този том с преглед работата на много български учени и др.

* **#instituti**– преглед на институти по ИИ в Източна Европа и света, сравнение на повтарящите се послания; към 2003 г. в България имаше публикувани **2 национални стратегии** за развитие с ИИ - 16 години преди първата чернова на БАН и 19 години преди откриването на INSAIT, и двете дело на юноши.

* **#complexity** – Алгоритмична сложност – обзор и бележки по множество статии и обобщения и изводи. Дали машината на Тюринг е подходяща за описание на *Мислеща машина?* #hector

* **#calculusofart** – Математически анализ на изкуството. Музика I – Как се определя дали даден „къс“ изкуство е красиво и защо ни харесва? Красотата, компресирането и предвиждането на бъдещите данни въз основа на миналите. Музиката трябва да е красива и да се измерва във всички мащаби, от най-малките с постепенно нарастващ обхват.

* **#kotkata –** Задачата от *„Анализ на смисъла на изречение въз основа на базата знания на действаща мислеща машина (...)“,* Т.Арнаудов 2004 г. в диалог с чатботовете ChatGPT и Bard, края на 2023 г. до нач. на 2024 г. и с GPT5 пред 2025 г., който успява да разбере и приложи в опростен вид метода от статията

* **#zabluda –** Заблуждаващите понятия и разбор на истинския им смисъл: трансхуманизъм, цивилизация, ... – книга, която публикувах през 2020 г. и започна като статия за трансхуманизма. Откъсът може да бъде включен и в отделно приложение.

#razvitie #transhumanism – том фокусиран върху развитието на човека, космизъм, „трансхуманизъм“; етика, биотехнологии, мозъчно-компютърен / мозъчно-машинен взаимлик (Brain-Computer Interface, Brain-Machine Interface), невроморфни системи, генетично инженерство, геномика, биология, симулиране на клетки и живи организми и др.

Практика, работилници и др. (бъдещи)

* **#robots-drones-ros-slam-simulation-rl** – Наземни и летящи роботи: дронове; обща теория, практика, конкретни системи и приложения; Robot Operating System (ROS, ROS2); среди за симулации на физически и виртуални роботи и машинно обучение: Gazebo, MuJoCo, RoboTHOR, Isaac Sim, Omniverse; gymnasium и др.

* **#neuromorphic-snn-practice** – Практика по невроморфни системи, импулсни невронни мрежи; Lava-nc и др.

* **#llm-generative-agents** – големи езикови модели: локална работа, платформи; употреба, подготвяне на набори от данни; обучение, тестване. Текст, образ, видео, триизмерни модели, програмен код, цели игри и светове с физика („world modeling“), всякакви модалности; дифузни модели, преобразители (трансформатори), съгласувани с физиката математически модели, причинностни модели с управляващо-причиняващи устройства по идеите от Теория на Разума и Вселената. Агенти, мулти-агентни системи: архитектури и др ...
(виж ***Листове*** и ***Лазар***)

* **#codegen –** автоматично програмиране, синтез на програми; модели за тази цел, платформи; методи, приложения ... program synthesis, automatic programming, code generation

* **#sigi-evolve –** саморазвиващи се машини, еволюционни техники, рекурсивно самоусъвършенстване (Recursive Self-Improvement, RSI)

* **#appx** – **Приложение на приложенията**, списък с добавени по-късно; ръководство за четене и др.

* **#agi-chronicles** – хронологичен запис и проследяване на развитие на история, новини, събития, идеи, системи, приложения; изследователи *(вероятно с Вседържец)*

... следват продължения – други приложения и *Вселената*:

* **Създаване на мислещи машини** – ... Зрим, Вседържец , Вършерод, Казбород, Всеборавител, Всетводейство, Всевод, (...

Внимание! Този списък и информацията в него може да са непълни, неточни или остарели. Възможно е да излизат нови издания с поправки и допълнения. За обновления следете уеб страниците, фейсбук групата „Универсален изкуствен разум“, Ютюб каналите, Дискорд сървъра и др.

Можете да помогнете за подобрението на съществуващите и за осъществяването на бъдещите разработки.

Свещеният сметач призовава съюзници, съмишленици, съдружници и сътрудници; университети, изследователски институти и фирми; учени, инженери, разработчици и творци; спомоществователи, дарители, последователи, другари, изследователи и съавтори за продължения и подобрени версии и за развитие на дейността на изследователско-творческото дружество.

Ако искате да помогнете, за конкретни идеи вижте в началото на основния том, в приложение *Листове,* в информацията за проекта **Вседържец** или *се свържете с мен.*

Всякаква съвременна техника за *нова*[1] изследователска дейност би ни била от полза в работата, както и достъп до облачни услуги от всякакъв вид – от достъп до пораждащи модели като ChatGPT, Gemini, LangChain, до сървъри, дисково пространство и пр. Помещения за техника и работа също може да са полезни.

Съхраняването на българската и световна компютърна история и памет е част от дейността на *Сметача* още от зората му през 2000 г.

Стара българска и световна изчислителна техника, която искате да дарите, също е добре дошла при нас или например при компютърния музей *„Компу Пловдив“ (Compu Plovdiv),* на който сътрудничим.]

[1] Мислещата машина Вседържец работи на такава техника, с каквато разполагаме и се вмества в нея. Ако не можем да осигурим друго, Вседържец ще трябва да се справи и на едно или няколко PC-та, лаптопи и по-малки компютри, с възможна връзка и към мобилни устройства за сензори и други помощни обработки; без или със достъп до Интернет и облачни услуги, които също предоставят допълнителна мощ безплатно или на достъпна цена, дори и за сиромаси. Виж *„Сингулярност на Тош“ – уравнението на относителната ефективност* в приложение *„Първата модерна стратегия…“*

Приложение „Анелия“ е част от широк и дълбок преглед на работата на български и световни изследователи по изкуствен интелект и общо – всякакви науки за ума и познанието, и техните приложения.

За по-опитните читатели „*Анелия*“ може да послужи за припомняне и задълбочаване, и за откриване и научаване на важни направления, теми, методи, понятия, разработки от миналото и настоящето чрез четене и проучване на конкретни научни публикации във времето и в работата на конкретните разглеждани учени, които са допринасяли за развитието на съответните области. Процесът е подпомогнат и ускорен чрез обобщенията и бележките за най-важното. Този том, макар и „недоподреден“ по обичайния начин, е вид алманах, енциклопедия, библиографски справочник и набор от данни с личности и основа за откриване и извличане на други свързани учени и теми чрез т.нар. изследване на литературата[2] – било „по-ръчно“ или чрез „по-автоматична“ обработка с търсещи машини и езикови модели – за построяване на допълнения и още връзки с материалите от другите части на *Пророците*: *Лазар, Листове, Основния том* … и др. и въобще в разглежданите области на познанието; за проследяване на преходите в развитието на определени научни понятия и задачи, напредъка в съответните научни области и подобласти и развитието на конкретни учени: и български, и световни.

Пророците въвежда любознателния читател в много области от науките и знанията за ума и мисленето, подобно на предобучението на мултимодален пораждащ модел, като в процеса се откриват и отбелязват връзки и съвпадения между далечни „места“ от рисуващата се карта, открива се повтарящото се или различно във времеви периоди, различни науки и научни школи; извличат се най-важните понятия и идеи и се достига до обобщения и заключения, въз основа на синтезирането на многообразно и широкообхватно знание.

Виж още уводните бележки към *Листове; Основният том;* текстовете за мултимодалните модели в „*Първата стратегия…*“ и в споменатите два тома, „*Творчеството е подражание на ниво алгоритми*“, 2003 и *др.*

Бъдещите версии ще са част от или ще работят с мислещата машина ***Вседържец***, която ще ги направи взаимодействащи и по-удобни за изучаване, търсене, подреждане, допълване, изобразяване, свързване, превод, … *(…)*

@Вседържец: ОбОбр всчк: -.- [,,]+ *# пдбн {К-К} др прлжн
https://github.com/Twenkid/Vsy-Jack-Of-All-Trades-AGI-Bulgarian-Internet-Archive-And-Search-Engine

Виж за други български „**пророци**“ в основния том и в други приложения: #prophets #tosh1; #lazar #listove #instituti и др.

[2] Literature-based research, Lierature review; Literature-Based Discovery (LBD)

Анелия Ангелова – Google, Google Deepmind, ...
https://scholar.google.com/citations?user=nkmDOPgAAAAJ&hl=en
Важни работи в приложения на машинното обучение, компютърното зрение, навигацията за самоуправляващи се превозни средства, големи езикови модели, роботика от началото на 2000-те години до днес.

* **Rapid object detection,** 5/2003, A.Angelova
https://citeseerx.ist.psu.edu/document?repid=rep1&typef=pdf&doi=2f7aea8299b39b78dba7f5c44e98c078cad928d0
https://www.researchgate.net/publication/248563069_Learning_for_Autonomous_Navigation_Extrapolating_from_Underfoot_to_the_Far_Field

Самоуправляващи се превозни средства за изследвания на Земята и в космоса. За NASA, DARPA. % zones: proprioceptive (1m), Near-field/Mid-field/Far-field (10m, 50m, infinity); learning from 3D geometry (Lf3D); learning from proprioception (LfP) - from images; traversability analysis, traversability cost on grid-based map cells; local map representations; learning: SVM, 2000 traversable/non-t. examples --> 784 support vector ... K-means clustering; ... Slipping +trav.cost; Mixture of Experts;

* **Angelova, A. N. (2004). Data pruning. MSc. Thesis**, California Institute of Technology https://thesis.library.caltech.edu/2184/1/DataPruning.pdf
Подобряване на обобщаващата сила на ученето след премахване на определени примери. Множество полунезависими класификатори, чрез които се разпознават нискокачествени образци, при които има повече разногласия. (Outliers, difficult examples); Надеждна статистика (Robust); решение какво да се прави с граничните случаи и шума? Да се премахнат, да се ограничи теглото им? RANSAC (Random Sample Concesus), вместване на геометрични примитиви...); регуляризация - "наказание" за сложността на модела (брой параметри); SVM, SVC (slack variables) - помощни променливи за съотношението между предела (margin) и грешката при класификация (по-малко обобщение, но по-малко грешки; повече обобщение, но повече грешки в класификацията). AdaBoost: съчетаване на "слаби ученици", малко по-добри от случайността, в обединен силен ученик ... двоичен класификатор; трудните примери: близки до границата на разделяне, премахване на най-трудните (срвн. GAN, 2014).

* **Angelova, A., Abu-Mostafa, Y., & Perona, P. (2005). Pruning training sets for learning of object categories.**
https://home.work.caltech.edu/pub/Angelova2005prune.pdf Автоматично изчистване на набор от данни от лоши примери за по-добро обучение; в

работата - за лица (шум, снимки които не са на лица, замъглени и пр.); общо - погрешни белези, маркиране (wrong labels, incorrect labels)

* **Real-time grasp detection using convolutional neural networks**, J. Redmon, A. Angelova, 2015 IEEE International Conference on Robotics and Automation (ICRA), 1316-1322, 1061 citations, 2015.

Cornell Grasping Dataset - labelled grasps. Ground truth grasps: g = {x, y, θ, h, w}; g - център на правоъгълника за хващане, θ - ориентация, ъгъл спрямо водоравната ос; w,h - ширина (дълбочина), височина (дължина). 5-измерен хващащ вектор вместо 7-измерен в друга работа (триизмерен), за да се сведе до разпознаване на образи; архитектура подобна на AlexNet, 3 конволюционни слоя и два напълно свързани, завършващи с 5 неврона отразяващи петте параметъра. Допускания за едно място за хващане - чист фон; при затрупан - първо сегментиране (cluttered scene); за parallel plate grippers (ръце с успоредни "щипки"); модел за много места за хващане: разделя на решетка NxN и научава вероятността във всяка клетка да има добро положение за захващане. ... Използва всеобща информация, а не плъзгащ се прозорец, в RGB-D пространство (стереокамера, образ и дълбочина)[3].

* Angelova, A., & Zhu, S. (2013). **Efficient object detection and segmentation for fine-grained recognition.** https://openaccess.thecvf.com/content_cvpr_2013/papers/Angelova_Efficient_Object_Detection_2013_CVPR_paper.pdf Финото разпознаване е например на видове птици, породи кучета - общ род, различни видови признаци; по-малки зрителни различия отколкото между по-отдалечени визуално класове като ябълка и къща, човек и камък и т.н.; предишна работа - superpixels, pre-segmentation; full-object segmentation: label Xj for each pixel Ij (feature fj); вж. също докт.дис. на А.Тошев, 2011 (Alexander Toshev).

* **Computer vision on Mars**, Matthies, L., Maimone, M., Johnson, A., Cheng, Y., Willson, R., Villalpando, C., Goldberg, S., Huertas, A., Stein, A., & Angelova, A. (2007)https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=02635d08154263c995334354d7499710c402e05d – Ценен обзор на историята на отдалечени автономни превозни средства (rovers, rover navigation) особено за космически цели, различни „марсоходи“ и актуални техники за определеня не неравностите и възможността за движение по определени траектории чрез различни сензори. Стерео зрение, зрителна одометрия (от Моравек, 1980-те; Carnegie Melon CMU Navlab, CMU Navlab 5, Ambler ... и др. с.3) и др. visual odometry; зрително прогнозиране на плъзгането (slip prediction); проследяване на белези за определяне на хоризонталната скорост (horizontal velocity feature tracking) и др.

[3] Вж и: I. Lenz, H. Lee, and A. Saxena, “Deep learning for detecting robotic grasps,” in Proceedings of Robotics: Science and Systems, Berlin, Germany, June 2013.
Y. Jiang, S. Moseson, and A. Saxena, “Efficient grasping from rgbd images: Learning using a new rectangle representation,” in IEEE International Conference on Robotics & Automation (ICRA). IEEE, 2011, pp. 3304–3311.

През годината на статията отборът на CMU печели Darpa Urban Challenge като "Tartan Racing", виж бел. към мулти-агентни системи и архитектурата TouringMachine, 1991-1992.

Виж още: Navlab 1, 1986 - самоуправляващ се камион, заради размера на компютрите https://www.youtube.com/watch?v=ntIczNQKfjQ
https://en.wikipedia.org/wiki/Navlab
* Navlab 5, 1997 - самоуправляваща се кола по магистрала, поддържа курса https://www.youtube.com/watch?v=xkJVV1_4I8E
* Navlab 84-94, Todd Jochem, 169 абонати, 965 показвания, 30.01.2008 Резюме на работата на Navlab за автономни превозни средства на университета Карнеги Мелън между 1984 и 1994 г. Navlab 1: Neighborhood Navigation, Todd Jochem https://www.youtube.com/watch?v=0GXuqw3cgwU
https://www.ri.cmu.edu/research/labs-groups/ https://www.ri.cmu.edu/robotics-groups/navlab/
* **Unsupervised learning of depth and ego-motion from monocular video using 3D geometric constraints,** R. Mahjourian, M. Wicke, A. Angelova, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 897 citations, 2018.
* **Pali: A jointly-scaled multilingual language-image model**, X. Chen, X. Wang, S. Changpinyo, A. J. Piergiovanni, P. Padlewski, D. Salz … Anelia Angelova … et al. (29) arXiv:2209.06794, 632 citations, 2022. https://arxiv.org/abs/2209.06794
Datasets: image-text pairs: Conceptual Captions (CC3M, CC12M; 2018,2021); LEMON (Hu et al., 2022): 200M; GIT: 800M. ALIGN: 1.8B (noisy); SimVLM, CoCa. Text enc-dec: mT5-Large 1B, mT5-XXL 13B. ViT: ViT-e: 4B. PALI-3B, -13B, -17B. WebLI: 10B img + tens B img-txt pairs in 109 lang; 29B img-OCR pairs; … Object-aware VQA1: 1-triplets, object labels; QA pairs: lists of all objects in the img & ? subset of obj are in the img; obj-lvl annot: Open Images. Obj.detect.

* The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale, Kuznetsova et al., 2020
https://paperswithcode.com/dataset/open-images-v4
https://storage.googleapis.com/openimages/web/index.html -v7, 2022
https://docs.voxel51.com/ Dataset building and viewing with 3 and 12 million image-text pairs …

* **Depth prediction without the sensors: Leveraging structure for unsupervised learning from monocular videos**, V. Casser, S. Pirk, R. Mahjourian, A. Angelova, Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 8001-8008, 565 citations, 2019.

* **Depth from videos in the wild: Unsupervised monocular depth learning from unknown cameras**, A. Gordon, H. Li, R. Jonschkowski, A. Angelova, Proceedings of the IEEE/CVF International Conference on Computer Vision, 462 citations, 2019.

* A. Angelova, A. Krizhevsky, V. Vanhoucke, A. S. Ogale, D. Ferguson,

* **Real-time pedestrian detection with deep network cascades**, BMVC 2, 4, 324 citations, 2015., W. Kuo, A. Angelova, J. Malik, T. Y. Lin

* **Evolving Space-Time Neural Architectures for Videos,** AJ Piergiovanni, Anelia Angelova, Alexander Toshev, Michael S. Ryoo, 11.2018/8.2019
https://arxiv.org/abs/1811.10636 (А.Ангелова и А.Тошев)
Video understanting: video object detection and activity recognition; CNN; Inflated Temporal Gaussian Mixture (iTGM); hybrid meta-architecture fill-in-the-blanks – the connectivity is fixed, but the modules can evolve.

Еволюционен алгоритъм: случайно се избира набор от възможни архитектури, изпробва се производителността им при разпознаване, най-добрата става „родител“, върху който се прилагат мутации, за да се получи „дете“, което се добавя към популация и се обучава за определен брой итерации, след което се оценява с друг набор за валидиране, което дава „приспособеността“ ѝ (fitness). От популацията, общ 25 архитектури, се премахва най-слабо приспособената и еволюционният цикъл се повтаря. Мутациите са 4 типа: 1) избор на времево-пространствен конволюционен слой и промяна на типа му. 2) Избор на времево-пространствен слой и промяна на времевия му размер (обхват). 3) Избор на модул от родителската архитектура и добавяне или премахване на паралелен поток между слоевете: между 1 и 6. Фиксиран брой изходни филтри, равномерно разпределени между паралелните линии. 4) Избор на модул и промяна на броя на повторенията му.
* Mahjourian, R., Wicke, M., & Angelova, A. (2017**).**

Geometry-based next frame prediction from monocular video
https://arxiv.org/pdf/1609.06377

* **Shapemask: Learning to segment novel objects by refining shape priors,** Proceedings of the IEEE/CVF International Conference on Computer Vision, 150 citations, 2019. @Вси +

Уточняване на обхващаща маска при разпознаване, а не само правоъгълник; shape priors – предварителни модели на формата; shape mask – маска на формата избистря първоначално грубата форма, свеждайки я до маска за всеки отделен случай на разпознат предмет в изображението ,чрез научаване на представянето на примерите (instance embeddings). Предпоставките за формата подсказват.

*** Region-Centric Image-Language Pretraining for Open-Vocabulary Detection.** Kim, D., Angelova, A., & Kuo, W. (2024) https://fq.pkwyx.com/default/https/www.ecva.net/papers/eccv_2024/papers_ECCV/papers/08007.pdf @Вси:+
LVIS open-vocabulary detection benchmark (cmp closed-vocabulary); Open-Vocabulary Detection – OVD – pretrains on a large set of image-text data (CLIP etc.); DITO – open vocabulart, RegionCLIP, CoDet, DetCLIP; Grounding DINO … pseudo labeling; inegrating detector architectures into CLIP, aligh regions with words; visual grounding annotations, multitask learning … Self-supervised pretraining: contrastive – augmented images, sliding windows, obj.proposals, point samples, pixel reconstruction. Detector heads, shifted-window learning for detection; Region of interest Align (RoI-Align) feature … cos similarity (region embedding, text embeddings, base categories) → softmax (turn into probabilities, normalize) …

*** 3D Open-Vocabulary Panoptic Segmentation with 2D-3D Vision-Language Distillation**, Xiao, Z., Jing, L., Wu, S., Zhu, A. Z., Ji, J., Jiang, C. M., Hung, W. C., Funkhouser, T., et al (2025). European Conference on Computer Vision, 21-38

* **Connection weight learning for guided architecture evolution,** Ryoo, M. S., Piergiovanni, A. J., Tan, M., & Angelova, A. (2024).
https://scholar.google.com/citations?view_op=view_citation&hl=en&user=nkmDOPgAAAAJ&sortby=pubdate&citation_for_view=nkmDOPgAAAAJ:8AbLer7MMksC
https://patents.google.com/patent/US12046025B2/en

* **Video question answering with iterative video-text co-tokenization.** Piergiovanni, A. J., Morton, K., Kuo, W., Ryoo, M. S., & Angelova, A. (2022). https://arxiv.org/pdf/2208.00934 Набори от данни: MSRVTT-QA, MSVD-QA, IVQA; намалява сложността от 150-360 до 67 GFLOPs. VQA – Visual Question Answering (image) → VideoQA (motion picture) … action(recognition, detection, segmentation) in open-set domain (unseen objects categories or unknown activities); joint learning of the video & text embeddings, co-tokenization, joint representations, cross-modality interactions at many stages, cross-modal learning … „video-text co-tokenization fusion mechanism which learns the most appropriate compact feature representations iteratively based on the previous features"; multi-stream: at different time- and space- scales → multi-scale features; fusing multi-modal models; Learn to tokenize (TokenLearner) …

Развитие на свързана задача:
* **What's in a Video: Factorized Autoregressive Decoding for Online Dense Video Captioning.** Piergiovanni, A. J., Kim, D., Ryoo, M. S., Noble, I., & Angelova, A.

(2024). , arXiv preprint arXiv:2411.14688. https://arxiv.org/pdf/2411.14688 Autoregressive model, next token prediction; split in short non-overlaping video segments. ~ 500M params, 128M: text 300M for TubeViT, ViT-Large; video memory transformer and autoregressive transformer. Input: 512 frames at 448×448 w.16 segments. Even with a high number of frames and high resolution, fits on 64 devices (TPUs?) decoding: beam search, 24 outputs, temporal NMS @ 0.7 threshold.ViT: ALIGN dataset, MaMMUT pre-trained weights. … 512 fr. 8 or 16 segments and 32 output tokens per segment = 424 GFLOPs/seg. For 8 segs = 3392 GFLOPs/256 tokens), 16 segs: 6784 GFLOPs/512 output tokens. Baseline: 4125 GFLOPs/256 tokens … Scores: SODA [16], CIDEr [51] and METEOR [3]; Datasets: ActivityNet dense captions, ViT, YouCook2, HowTo100M; segment both to events and provide corresponding captions. …

Multiple captions: <start_token><start_time><end_time><caption_text><EOS> … <start_time 1> <end_time 1><caption_text 1> <start_token><start_time 2><end_time 2><caption_text 2> … <EOS> …
https://huggingface.co/docs/transformers/en/model_doc/align

* **MaMMUT: A simple architecture for joint learning for multimodal tasks.** Weicheng Kuo, AJ Piergiovanni, Dahun Kim, … and Anelia Angelova, 2023
Вж.сщ: VLAP: Efficient Video-Language Alignment via Frame Prompting and Distilling for Video Question Answering, Xijun Wang+,2024
https://arxiv.org/html/2312.08367v2

* **Image depth prediction neural networks.** Angelova, A., Wicke, M., & Mahjourian, R. (2023). Интересно е, че в този и други патенти като за превод на графични романи посочват LSTM (а не преобразители).
https://patentimages.storage.googleapis.com/10/14/ac/7804b25ba834fa/US11734847.pdf

* **Localization of Objects Encoded in Image Data in Accordance with Natural Language Queries,** Kuo, W.-C., Bertsch, F., Li, W., Piergiovanni, A. J., Saffar, M. T., & Angelova, A. (n.d.). 29.8.2024
https://patents.google.com/patent/US20240289981A1/en
"...обобщено уточняване на местоположение, където локализираният обект е в съответствие със заявка на естествен език. По-конкретно, изпълненията включват единна обобщена архитектура за зрително откриване на мястото, която постига подобрена производителност при следните три задачи: разбиране на изрази относно целта, локализиране на обекти и откриването им. Изпълненията използват машинно обучени модели за естествен език и/или модели на изображения. Архитектурата е способна да разбира и отговаря на въпроси за естествено посочване на места в изображение, да извежда множество полета, да не предоставя изход, ако обектът липсва (напр. да върне празен резултат), както и да решава общи задачи за откриване..." - виж напр.

PaLM-SayCan: https://sites.research.google/palm-saycan

* **3D Open-Vocabulary Panoptic Segmentation with 2D-3D Vision-Language Distillation,** Xiao, Z., Jing, L., Wu, S., Zhu, A. Z., Ji, J., Jiang, C. M., Hung, W.-C., Funkhouser, T., Kuo, W., Angelova, A., Zhou, Y., & Sheng, S., 3.4.2024 (2025), 3.4.2024 https://arxiv.org/pdf/2401.02402 – Задача за безпилотни превозни средства, за да разграничат броими *обекти* наоколо (*things*) – други превозни средства, хора, препятствия – и неброими "*неща*" (*stuff*) - пътя, растителността край него, небето. Всеки отделен обект получава собствен етикет, докато "нещата" се възприемат като едно цяло, маса без собствена „единичност“ – подобно на категориите в естествения език ("насипни" и неброими са пясък, вода; могат да станат броими с мярка и уточнения: „чаша вода“, „лопата пясък“). В работата за пръв път едновременно разделят *нови* обекти и неща – „отворен речник“, несрещани в набора от данни за обучение – на отделните представители и семантична класификация (instance & semantic segmentation). Двуизмерното сегментиране на всичко с отворен речник е развито с употребата на CLIP (2D open-vocabulary); но за триизмерния случай не достигат достатъчно двойки данни за триизмерни облаци от точки и текст; виж *OpenScene*: per-pixel CLIP features → projecting 3D points onto image planes; RegionPLC: зрителни подкани върху области от изображението, контрастно обучение разделящо точки ...; в тази работа - само предварително обучен CLIP. ... Мултимодално сливане на особености ... В 2D: замразени особености на CLIP и 2D паноптично сегментиране, но в 3D не е пряко приложимо, защото много точки нямат валидни пиксели от камерата (скрити са зад други). P3Former. LiDAR-Vision ..

* **Depth from videos in the wild: Unsupervised monocular depth learning from unknown cameras**, A Gordon, H Li, R Jonschkowski, A Angelova, 4.2019 https://arxiv.org/abs/1904.04998 @Вси: вж

* **Region-Centric Image-Language Pretraining for Open-Vocabulary Detection** https://github.com/google-research/google-research/tree/master/fvlm/dito и др. проекти от „Гугъл“.

* **Query image search, Inventors:** Henry Allan Rowley, Charles J Rosenberg, Anelia Angelova, 2014 Wavelets, уейвлети, „вълнички“. Виж и патентите на Красимир Кралев в компресията с уейвлети от списъка с български роботисти. https://patentimages.storage.googleapis.com/c0/31/7a/f2b8afc580fdaf/US8782077.pdf

* **Драгомир Ангелов** – Stanford, Google, Waymo

Computer vision, ML, AI, self-driving cars …

Откъс от основния том:

Драгомир Ангелов - „Драго“ – ръководител в проекта за самоуправляваща се кола на „Гугъл“ (Waymo), с публикации от 1999 г.*, докторант по компютърно зрение в Станфорд. Компютърна геометрия, зрение, графика, моделиране на триизмерни обекти, анимация.

* **Anguelov, D., Learning Models of Shape From 3D Range Data**, 12.2005, PhD Thesis: https://ai.stanford.edu/~drago/Papers/thesis.pdf
* Дисертацията на Драгомир Ангелов, 2005 г.: извличане на геометричната информация за човешко тяло без учител; човешката фигура се разделя от триизмерни данни за координати, извлечени чрез лазерен скенер, като множество от части, свързани със стави, чрез автоматично разпознаване на частите с вероятностно минимизиране на изкривяванията и запазване на геодезическите разстояния и местната геометрия на мрежа от точки, описваща триизмерните данни (mesh). След това се подават конфигурации на обекта в различни положения (различни сканирания на един и същи човек или на различни хора), с променена поза, а системата я разделя на множество от твърди тела, разположението им, отношенията и конфигурацията (състоянията, положенията, завъртанията: виж бел. за роботика в приложението). Работата разглежда задачата за представяне на изкривяванията на повърхността на представянията на човешкото тяло като функция на ъглите на завъртане на ставите (виж „скинове“ в комп.графика) и др. Приложение за анимиране и предложено развитие за прихващане на движенията без маркери (motion capture).

По това време в Станфорд учи и работи и Васил Чаталбашев, който е водещ автор на проект за сегментирането на триизмерен модел на кампуса на университета чрез лазерен скенер. **Красимир Коларов** е може би първият българин в Станфорд, завършва роботика 1988-1992 г. а след това е и съавтор на лекциите и преподавател; запазена е негова лекция.

* **Laser Range Data Classification Using AMNs**
https://ai.stanford.edu/~vasco/3Dmap/
* **Lecture 10 | Introduction to Robotics, Stanford,** 1,98 млн. абонати 40 330 показвания 23.07.2008 г. Lecture … Introduction to Robotics (CS223A) in the Stanford Computer Science Department. Guest lecturer **Krasimir Kolarov** (co-writer of the lecture notes along with Professor Khatib) presents Trajectory Generation. https://www.youtube.com/watch?v=7wIqGavQjTQ – Планиране в евклидово пространство и в пр. на ставите.

Прочети повече за тях и за други българи в роботиката и др. в **основния том** на *Пророците на мислещите машини.*

* БАН https://www.iict.bas.bg/ #ban

* Красимир Атанасов

https://scholar.google.com/citations?user=K-vuWKsAAAAJ&hl=en

Влиятелен учен в размитата логика и теорията на обобщените мрежи – разширение на мрежите на Петри за моделиране на процеси. С обща памет, оптимизационни компоненти, допълнителни тактови сигнали, условия за край, обемни токени, сложни преходи, …

* https://www.math.bas.bg/bgsiam/docs/bgsiam_2017_abstract-katanassov.pdf

* KT Atanassov, Intuitionistic fuzzy sets, 1999
* Geometrical Interpretations of Interval-Valued Intuitionistic Fuzzy Sets: Reconsiderations and New Results, K Atanassov, P Vassilev, V Atanassova , 2025 https://www.mdpi.com/2227-7390/13/12/1967
* Atanassov, K.; Gargov, G. Interval-valued intuitionistic fuzzy sets. Fuzzy Sets Syst. 1989, 31, 343–349.

Обобщени мрежи

* **Generalized Nets,** Krassimir T. Atanassov, 1991
* **On generalized nets theory**, Krassimir Atanassov, 2007
* **Generalized nets and systems theory**, K. T. Atanassov

ИНСТИТУТ ПО ИНФОРМАЦИОННИ И КОМУНИКАЦИОННИ ТЕХНОЛОГИИ
(и учени от БАН в съавторство)

https://www.iict.bas.bg/docs/2024-papers-IICT.pdf
https://www.iict.bas.bg/docs/2023-papers-IICT.pdf

…
https://www.iict.bas.bg/docs/2013-papers-IICT.pdf

* **Building a Linguistically Interpreted Corpus of Bulgarian: the BulTreeBank,** Kiril Simov, Petya Osenova, Milena Slavcheva, Sia Kolkovska, Elisaveta Balabanova, Dimitar Doikoff,Krassimira Ivanova, Alexander Simov, Milen Kouylekov, 2002 – syntactically annotated corpus, "HLT" – human language technology; Head-driven Phrase Structure Grammar (HPSG); partial grammar, shallow parsing.

* Feature-Rich Named Entity Recognition for Bulgarian Using Conditional Random Fields, Georgi Georgiev, Preslav Nakov et al. 2021, *BulTreeBank (680 morpho-syntactic tags); a task-specific tagsets (local and nonlocal) + domain-specific **gazetteers** and additional unlabeled data; F1=89.4% ~ SOTA Eng. Feature-based models like CRFs; orthographic predicates – initial capital letter, all capitals, ... "* **gazetteers** – a kind of dictionary, a list of items from a type, e.g. cities, names of organisations, months, days of the week etc.

BulPhonC: Bulgarian Speech Corpus for the Development of ASR Technology, Neli Hateva, Petar Mitankin, Stoyan Mihov, 2016, LREC
http://www.lrec-conf.org/proceedings/lrec2016/pdf/478_Paper.pdf
21838 sentences in the corpus - automatically selected from the ones phonetised with the SpeechLab text-to-speech system for Bulgarian (Andreeva et al., 2005); 147 speakers. Manually annotated with Praat. Tested with the LVCSR system for Bulgarian (Mitankin et al., 2009). The n-gram language models: ~250M words legal corpus for Bulgarian, created in the framework of the project for Bulgarian ASR system for juridical texts (2009). Test set: 9 sp., 50 long utter. The corpus is free for scientific usage (BulPhonC@lml.bas.bg)

* Stoyan Mihov - Стоян Михов и др.

* https://dblp.org/pid/67/6448.html * https://lml.bas.bg/~stoyan/lmd/Person.html

* **Finite-State Techniques Automata, Transducers and Bimachines**, Mihov, Klaus Shulz, 2019 https://www.cambridge.org/core/books/finitestate-techniques/E21E748468F0310DA12A2CFAEB989185 - a book
https://lml.bas.bg/~stoyan/finite-state-techniques.pdf

* **A visual and interactive tool for optimizing lexical postcorrection of OCR results,** Christian Strohmaier, Christoph Ringlstetter, Klaus U. Schulz,Stoyan Mihov, 2003 Multiple dictionaries. Background dictionary, fixed dictionaries For proper names, geographic names, thematic areas, acronyms etc. Dynamic dictionary, extract from web – watch for false friends. Scoring function, rank correction candidates. Distance value and frequency information - Levenstein edit distance.

Similarity between the OCR read word and the correct one. Alignment – more difficult for multicolumn text; balance similarity/frequency, threshold. Alignment file. Correction accuracy.

* Christoph Ringlstetter, Stoyan Mihov, Klaus U. Schulz, and Katerina Louka: ***The Same is Not The Same - Postcorrection of Alphabet Confusion Errors in Mixed-Alphabet OCR Recognition***, Proceedings of the ICDAR 2005, pp. 406 -- 410, Seoul, September 2005.

* Stoyan Mihov, Klaus U. Schulz**: Fast Approximate Search in Large Dictionaries.** Computational Linguistics, Volume 30, Issue 4, pp. 451 -- 477, December 2004. (IF: 2.017) https://lml.bas.bg/~stoyan/J04-4003.pdf

* Angelova, G. and S. Mihov: **Finite State Automata and Simple Conceptual Graphs with Binary Conceptual Relations**. In: Eklund, P. and O. Haemmerle (Eds.), Supplementary Proceedings of the 16th International Conference on Conceptual Structures (ICCS'08), Toulouse, France, CEUR Workshop Proceedings 2008, pp. 139-148. https://lml.bas.bg/~stoyan/p28r.pdf

* Christoph Ringlstetter, Klaus U. Schulz and Stoyan Mihov: *Adaptive text correction with Web-crawled domain-dependent dictionaries*. **ACM Transactions on Speech and Language Processing (TSLP)**, Volume 4, Issue 4, pp. 9:1-9:36, October 2007.

* **Large vocabulary continuous speech recognition for Bulgarian**, Petar Mitankin, Stoyan Mihov, Tinko Tinchev: Proceedings of the RANLP 2009, Borovets, September 2009. https://lml.bas.bg/~stoyan/bgsrranlp.pdf Perhaps the first known large vocabulary … LVCSR , >440K words; specialized corpus (juridical, legal); 200M words; bigram HMM. Lang.model size: 82 MB; system memory used by the program: 120 MB. 4.1 million bigrams, 4.8 M entries in a table a sparse transition matrix, Tarjan table. Another table with 4.1M entries. 12 bytes, 3 x 4: a symbol, a value of σ and a destination state. Beam search, Viterbi,
90%, 10% … *< 4% letter error rate and 13% word error rate on speech recognition of juridical texts using real time processing speed on a modest notebook computer.* Microsoft Speech API 5.1. Tested on a Mobile AMD Sempron 3400+/GB RAM. Set: 63 utterances juridical, 1276 from common texts. An informal user feasibility test; a user adaptation period is required, but the system is usable for text dictation.

* **SpeechLab 2.0: A High-Quality Text-to-Speech System for Bulgarian**, Maria Andreeva & Ivailo Marinov & Stoyan Mihov, 2005
https://lml.bas.bg/~stoyan/ranlp2005.pdf

Cmp: “Glas 2004” (Pisar_Glas 2004), The Sacred Computer, PU Paisii Hilendarski
Срвн: Глас 2004 („Писар Глас 2004“) от Тош/Свещеният сметач, ПУ „Паисий Хилендарски“

⁎ A visual and interactive tool for optimizing lexical postcorrection of OCR results, Christoph Ringlstetter, Klaus U. Schulz, Stoyan Mihov
https://lml.bas.bg/~stoyan/DIAR03.pdf

* **БАН, Значим напредък в езиковото моделиране на учени от ИИКТ-БАН,** 25.11.2024 – *двупосочни* езикови модели, автомати

* **Consistent Bidirectional Language Modelling: Expressive Power and Representational Conciseness**, Georgi Shopov, Stefan Gerdjikov, 2024, https://aclanthology.org/2024.emnlp-main.328.pdf – theoretical, mathematical, automata theory; The work is also supported by: *"CLaDA-BG, Bulgarian National Interdisciplinary Research e-Infrastructure for Resources and Technologies in favor of the Bulgarian Language and Cultural Heritage, part of the EU infrastructures CLARIN and DARIAH, funded by the Ministry of Education and Science of Bulgaria (support for the Bulgarian National Roadmap for Research Infrastructure); European Union: NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No. BG-RRP-2.004-0008."*

* **Composition and Weight Pushing of Monotonic Subsequential Failure Transducers Representing Probabilistic Models**, Diana Geneva, Georgi Shopov, and Stoyan Mihov, 5.2020, https://arxiv.org/pdf/2003.09364

* Georgi Shopov, IICT, Bulgarian Academy of Sciences
https://scholar.google.com/citations?user=6MnP9x8AAAAJ&hl=en
Diana Generva

* **Building an ASR corpus based on Bulgarian Parliament speeches**, D Geneva, G Shopov, S Mihov, 2019
* **Towards accurate text verbalization for ASR based on audio alignment**, Diana Geneva, Georgi Shopov, 2019, https://aclanthology.org/R19-2007.pdf – How to verbalize abbreviations, Roman numbers etc., „г.", „км", „кг", „XVIII" – often ambigous, dependent on the context. Bulgarian ASR corpus: BG-PARLAMA (Geneva et al.) Bulgarian Parliament speeches. 148607 speech segments, 572 unique speakers (422 male and 150 female), 249 hours. Kaldi ASR Toolkit (Povey et al., 2011) to train a time delay deep neural network (TDNN) (Peddinti et al., 2015) acoustic model with p-norm nonlinearities (Zhang et al., 2014) on the BG-PARLAMA corpus. A speaker-adaptive GMM model was also trained and used for generating state-level alignments for the TDNN training. The same parameters for the models as in LibriSpeech (Panayotov et al., 2015) Kaldi recipe. Phonetic system: (Mitankin et al., 2009; Hateva et al., 2016); pronunciation lexicon: the extended version (Geneva et al.) of the lexicon from (Mitankin et al., 2009). Verbalization based on: rules, audio alignments,

* Diana Geneva, Georgi Shopov, Kostadin Garov, Maria Todorova, Stefan Gerdjikov, Stoyan Mihov: **Accentor: An Explicit Lexical Stress Model for TTS Systems.** INTERSPEECH 2023: 4848-4852 https://www.isca-archive.org/interspeech_2023/geneva23_interspeech.pdf 11.7K hours of speech and texts with a total of 87M words (tokens) and 435K unique words (types) & audiobooks from

https://chitanka.info when they can be paired with text.

* **StreamSpeech: Low-Latency Neural Architecture for High-Quality On-Device Speech Synthesis,** Georgi Shopov, Stefan Gerdjikov and Stoyan Mihov, https://lml.bas.bg/~stoyan/TTSArch.pdf
* Jean-Marc Valin and Jan Skoglund, “LPCNet: Improving neural speech synthesis through linear prediction, 2019 …

See also: StreamSpeech: *“All in One” seamless model for offline and simultaneous speech recognition, speech translation and speech synthesis*
https://github.com/ictnlp/StreamSpeech
https://ictnlp.github.io/StreamSpeech-site/
StreamSpeech: Simultaneous Speech-to-Speech Translation with Multi-task Learning, Shaolei Zhang, Qingkai Fang, Shoutao Guo, Zhengrui Ma, Min Zhang, Yang Feng, 2024
ICTNLP – Natural Language Processing Group, Institute of Computing Technology, Chinese Academy of Sciences https://github.com/ictnlp
Available trained models for Fr-En, Es-En, De-En

* **Построяване на f-преобразувател на ниво символи за представяне на езикови модели**, Георги Шопов, 2019, СУ ФМИ, маг.дипл.раб.
f-преобразуватели – изгладени n-грам модели с f-преходи; вероятност на изречение.

* **Stroke Lesion Segmentation and Deep Learning: A Comprehensive Review,** Mishaim Malik, .. Nikola Kirilov Kasabov et al., 1.2024 https://www.mdpi.com/2306-5354/11/1/86 - вж сщ. Данаил Стоянов (Dan Stoaynov)

…

* **Васил Сгурев, р.1936 г.** Ръководител на института по техническа кибернетика и роботика, 1983 – 1990, и по информатика 1990-1993 г. (англ. превод „Institute of Engineering Cybernetics and Robotics“)
https://dblp.org/pid/05/4871.html

*** On the History of Artificial Intelligence in Bulgaria,** July 2021, Studies in Computational Intelligence**,** In book: Research in Computer Science in the Bulgarian Academy of Sciences**,** Vassil Sgurev – разглежда историята на изследванията до края на 20-ти век.
* The Achievements of the Technical Sciences in the Bulgarian Academy of Sciences (BAS) in Automation, Robotics and Computers During the Past Century, July 2021, Studies in Computational Intelligence, In book: Research in Computer Science in the Bulgarian Academy of Sciences

* Биография за 70-годишнина на В.Сгурев от 2006 г.:
https://biomed.bas.bg/bioautomation/2006/vol_5.1/files/anniversary_2.pdf

Групи и учени по невроморфни разработки в БАН

*** Георги Русев – докторант**
*** Симона Неделчева**
*** Петя Д. Копринкова-Христова** и др.

* **Decoding Brain Signals in a Neuromorphic Framework for aPersonalized Adaptive Control of Human Prosthetics**, Georgi Rusev 1, Svetlozar Yordanov 1 , Simona Nedelcheva 1, Alexander Banderov 1 , Fabien Sauter-Starace 2 ,Petia Koprinkova-Hristova 1,* and Nikola Kasabov, 3.2025

https://www.researchgate.net/publication/389871965_Decoding_Brain_Signals_in_a_Neuromorphic_Framework_for_a_Personalized_Adaptive_Control_of_Human_Prosthetics
Electro Cortico-Graphic (ECoG), 3D-SNN, Echo state networks - ESN, Motor Control Decoding (MCD). Python NEST Simulator for SNN. .. *brain implants are inserted in the skull to measure epidural ElectroCor-ticoGrams (ECoG) signals of the motor and sensory cortices in order to extract movement intentions.*

*** Spike timing neural network model of conscious visual perception**, Petia D. Koprinkova-Hristova, Simona Nedelcheva, April 2022, BIOMATH, https://www.researchgate.net/publication/360057770_Spike_timing_neural_network_model_of_conscious_visual_perception (full-text)

* **Reverse Engineering of Human Brain Using EEG Recordings,** In book: Advanced Computing in Industrial Mathematics, Petia D. Koprinkova-Hristova**,** Simona Nedelcheva**,** Nadejda Bocheva, 5.2025
https://www.researchgate.net/publication/392249977_Reverse_Engineering_of_Human_Brain_Using_EEG_Recordings

*** NEeuroMOorphic Neuro-Response Decoding System for Adaptive and Personalised Neuro-Prosthetics Control**, Georgi Rusev, Svetlozar Yordanov, Simona Nedelcheva, Alexander Banderov, Hugo Lafaye de Micheaux, Jean Monnet University, Fabien, Tetiana Aksenova, Petia D. Koprinkova-Hristova, Nikola Kirilov Kasabov, July 2025 – *a neuromorphic decoder of intended movements of tetraplegic patients using ECoG recordings from brain motor cortex called Motor Control Decoder (MCD) …Neural Response Decoder (NRD)*

https://www.researchgate.net/publication/393356317_NEeuroMOorphic_Neuro-Response_Decoding_System_for_Adaptive_and_Personalised_Neuro-Prosthetics_Control

* **Brain-inspired models and their applications**, Petia D. Koprinkova-Hristova , March 2025, Mathematics and Education in Mathematics 54:018-024, https://www.researchgate.net/profile/Petia-Koprinkova-Hristova/research

* **From MRI to 3D-SNN brain models,** January 2025, 13TH INTERNATIONAL SCIENTIFIC CONFERENCE TECHSYS 2024 – ENGINEERING, TECHNOLOGIES AND SYSTEMS, Simona NedelchevaPetia D. Koprinkova-Hristova, 1.2025

* **Reinforcement Learning Control of Cart Pole System with Spike Timing Neural Network Actor-Critic Architecture**, February 2025 , In book: Artificial Intelligence: Methodology, Systems, and Applications, Borislav Marko, Petia D. Koprinkova-Hristova

* **Investigation of Conscious Visual Perception via Visual Stimulus Propagation through a 3D-SNN Brain Model,** September 2024, International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Petia D. Koprinkova-Hristova, Simona Nedelcheva, Nadejda Bocheva

* **Grid Search Optimization of Novel SNN-ESN Classifier on a Supercomputer Platform,** May 2024, Lab: Nikola Kirilov Kasabov's Lab, Dimitar Penkov, Petia D. Koprinkova-Hristova, Nikola Kirilov Kasabov, Simona Nedelcheva, Sofiya Ivanovska, Svetlozar Yordanov

* **Reservoir Computing for Emotion Valence Discrimination from EEG signals** March 2017, Neurocomputing 231:28-40**,** Lachezar Bozhkov, Petia D. Koprinkova-Hristova, Petia Georgieva, 2016 (Echo State Networks)

Проекти, ръководени или с участието на П.Копринкова:
https://paml.iict.bas.bg/staff/Petia_Koprinkova/petiaEN.html

* "Auto-adaptive Neuromorphic Brain Machine Interface: toward fully embedded neuroprosthetics (NEMO-BMI)", funded by HORIZON-EIC, No 101070891, 2022-2025 - leader of IICT team.

* COST action CA18106 "The neural architecture of consciousness (NeuralArchCon)", 2019-2023 - member of manegement committee.

* "Modelling post-perceptual stages of cognitive processing and conscious representations of visual information", project No FNI KP-06-N52/6 from 2021 - project leader.

* "Supercomputer simulation investigation of a neural model of brain structures involved in conscious visual perception", Bulgarian Science Fund grant No KP-06-

COST/9 from 2021 by a co-funding scheme of COST action CA18106 "The neural architectures of consciousness" - project leader.

* "Modelling Of Voluntary Saccadic Eye Movements During Decision Making, contract No DFNI – DN02/3 with Bulgarian Science Fund, 2016-2020 - project leader".

...

* **Стефан Койнов** https://is.iict.bas.bg/our-team/stefan-koynov/
* **Татяна Атанасова** https://www.iict.bas.bg/mo/bg/atanasova.htm
* **Ангела Савова** https://www.researchgate.net/profile/Angela-Slavova
  * **Cellular Neural Networks: Dynamics and Modelling (Mathematical Modelling: Theory and Applications**) Softcover reprint of hardcover 1st ed. 2003 Edition, A. Slavova, https://www.researchgate.net/profile/Angela-Slavova

* **Клетъчни невронни мрежи** https://en.wikipedia.org/wiki/Cellular_neural_network L. Chua and L. Yang, "Cellular Neural Networks: Theory," IEEE Trans. on Circuits and Systems, 35(10):1257-1272, 1988. * https://nonlinear.eecs.berkeley.edu/raptor/CNNs/CellularNeuralNetworks-Theory.pdf L. Chua and L. Yang, "Cellular Neural Networks: Applications" IEEE Trans. on Circuits and Systems, 35(10):1273:1290, 1988. L. Chua, T. Roska, Cellular Neural Networks and Visual Computing: Foundations and Applications, 2005.

* Димитър Добрев – Dimiter Dobrev, БАН

Определения за изкуствен интелект и др., около 2000 и нач. на 2000-те и до днес

* AI - What is this: A Definition of Artificial Intelligence, November'2000, this paper is a part from AI- Project, https://dobrev.com/AI/definition.html
* AI Project: https://dobrev.com/AI/index.html
* AI - How it copes in arbitrary world, D.Dobrev, February'2001
* Formal Definition of AI, October'2005 (published in IJ ITA, Volume 12, Number 3, p.277), D.Dobrev
* Testing AI in One Artificial World, D.Dobrev, June'2005 (represented at KDS 05, Volume 2, Section 4, p.461)
* A Definition of Artificial Intelligence, D.Dobrev September'2003 (Mathematica Balkanica 2005, arXiv:1210.1568)
* A Definition of Artificial Intelligence, Dimiter Dobrev, BAS, 19.1.2004 (arxiv: 2012) https://arxiv.org/pdf/1210.1568 – „*AI will be such a program which in an arbitrary world will*

cope not worse than a human“
* **The AI Definition and a Program Which Satisfies this Definition**, Dimiter Dobrev, BAS, 2025, https://arxiv.org/pdf/2212.03184 – *“We will define AI as a computable policy which is sufficiently proximal to the best performing policy”*
* The Definition of AI in Terms of Multi Agent Systems, Dimiter Dobrev, BAS, Arxiv 2012 https://arxiv.org/pdf/1210.0887
* AI in arbitrary world, Dimiter Dobrev, BAS, 14.4.2005, Arxiv 2012, https://arxiv.org/pdf/1210.2715
* Comparison between the two definitions of AI∗, D.Dobrev, BAS, 2012/2013 https://arxiv.org/pdf/1302.0216
* Initiative for Responsible Approach to Artificial Intelligence https://dobrev.com/AI/Initiative.pdf
* AI Should Not Be an Open Source Project, D.Dobrev, 2018 https://dobrev.com/AI/OpenSource.pdf - moralizing about AI safety, the “humanity” etc. Морализаторстване за „човечеството“ и пр. Виж ТРИВ, 2001-2004 и писмото от Тош до Оксфорд от 2012 г.
* The IQ of Artificial Intelligence, . Dobrev https://dobrev.com/AI/IQ_of_AI_EN.pdf
* Giving the AI definition a form suitable for engineers (Bulgarian translation is available)

April 3, 2013 (arXiv:1312.5713) https://arxiv.org/pdf/1312.5713
“ за Изкуствен Интелект признаваме такава програма, която в произволен свят би се справила не по-зле от човек.”

Тодор: Някои теми, разглеждани още в зората на „Свещеният сметач“ и насока към общ ИИ. Не открих обаче да е споменат първия в света курс по УИР в Пловдив, нито друго съдържание от Т.А. Като цяло прочетеното ми се струва формално, без „фантазия“ и „сетивно-моторно обосноваване“, въпреки че за разлика от *Сметача* публикациите са в официален академичен формат. Необосновано нравоучение за етиката и „човечеството“. Виж „Писма между 18-годишния Т.Арнаудов и философа…“, 2002 и писмото ми до Оксфордския институт от 2012 г.

Todor: Some of the topics are related to the *Sacred Compute's* early works and overall - AGI. In my opinion superficial and unconvincingly moralizing though. No mentions of the pioneering AGI work from The Sacred computer and the 2010 AGI course and the definitions of T.A. starting as eary as 2001 and later in TUM 2001-2004.[4]

[4] However I didn't know D.Dobrev either up to today, 10.8.2025. I discovered him with a query to the freshly released GPT5.I may have tried out (or have encountered) *Strawberry Prolog* compiler maybe 20 years ago or about that time. D.D. was perhaps a principal developer. https://dobrev.com/

Влиятелни български „пророци“ от 1990-те до днес:

* **Пламен Ангелов** – БАН, Lancaster University, …
* **Никола Касабов** – ТУ София, Auckland и др.
* **Димитър Филев** – ТУ София, Auckland и др.

П.Ангелов и Н.Касабов имат съвместни работи по развиващи се разумни системи (evolving intelligent systems), системи с размита логика, мултимодалност; съчетания от размита логика и невронни модели и разпознаване с използване на прототипи[5]; обясним ИИ и др. Техен колега е и **Димитър Филев,** който има дълга кариера в автомобилната индустрия и награди за приноса си във внедряването на разнообразни системи с изкуствен интелект във „Форд“. Н.Касабов и лабораторията му сътрудничат с БАН по проекти в невронауките; мозъчно-компютърни интерфейси, извличане на информация за състоянието на мозъка от ЕЕГ и електроди и др. И тримата: връзката човек-машина, Human-Computer Interaction.

[5] Виж в основния том бележките по *Когнитивна лингвистика* и *Аналогия*; книгите на Петер Грендерфорс и др.; приложение Ирина: Когнитивна наука, също Мислене по аналогия и др.

*** Plamen Angelov** (h=65 по Google Scholar, 15.7.2025);
Explainable AI (XAI), Evolving intelligence, Fuzzy systems/fuzzy logic, Rule-based … https://scholar.google.com/citations?user=CCW8PwkAAAAJ&hl=en

*** Nikola Kasabov** – **67** h-index (Research Gate) https://www.researchgate.net/profile/Nikola-Kasabov ... *neural networks, computational intelligence, artificial intelligence, soft computing, bioinformatics, neuroinformatics. ... working on several theories, methods and their practical applications: evolving connectionist systems (ECOS); evolving spiking neural networks; brain-like spatio-temporal machine NeuCube; quantum inspired evolutionary computation.*; *750+ publications in the areas of neural networks, AI, bioinformatics, neuroinformatics*

*** Dimitar P Filev** – **65** h-index (Google Scholar, 19.7.2025) **Димиър Филев**

*** Explainable artificial intelligence: an analytical review,** PP Angelov, EA Soares, R Jiang, NI Arnold, PM Atkinson p.6. Fig 3. Opaque models (Ensemble, SVM, ANN, CNN, RNN …) vs Transparent (Decision trees, KNN, Rule-based, Bayesian, Linear regression) … Post-hoc explainability [by simplification, feature relevance, visual, local] → Explainable models… XAI taxonomy, … 4: Methods: feature, global, human-oriented, surrogate, concept, local and pixel-based … 5. Explainability-critical domains: medicine, criminal justice, tumor classifications, NLP, autonomous vehicles etc. 6.3. Future directions: **prototype-based models**; similarity instead of statistical … "**Interpretability**: is defined as the capacity to provide **interpretations** in terms that are **understandable** to a human (Gilpin et al., 2018)."* https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1424

Todor: This cited definition is circular and unexplicit: what is interpretation with simpler more explicit concepts; understandable by **who, which exact human/what are the capabilities**; by what kind of representation or language (how specific), with which tools and at what complexity and resolution. Complex explanations may include a lot of details anyway. Simple explanations are "understandable", but mundane such as textual as "a face has two eyes, a mouth…" (sometimes it doesn't; "too general") or e.g the "attention" masks – which areas of an image are most decisive for a classification decision, however these masks don't answer **WHY** they are more decisive, they are "deictic" – **pointing** at "this" and leaving the observer to interpret and understand. CPU machine code is exaplainable in its numerical format if you understand and recognize the instructions, can decode it in your mind or after disassembly with a proper debugger and tool, if you can trace the code, have appropriate debugger, patience, can reconstruct the program and its purpose etc.. ANN in their default form are also explainable in a similar way with enough patience, if you intervene inside their weights, test them with different inputs, retrain them etc. Therefore ***the degree of explainability*** is connected with the **amount of work, required for decoding**, in principle and of particular users, available auxilliary tools and skills etc. – their **complexity,** capacities/capabilities etc. should also be defined, like defining the observer-evaluator in broader contexts, discussed in the whole book of "*The Prophets of the Thinking Machines*" and in Theory of Universe and Mind.

*** Evolving intelligent systems: methodology and applications,** Plamen Angelov, Dimitar P Filev, Nik Kasabov, 2010 – fuzzy system s…
* Towards explainable deep neural networks (xDNN), Plamen Angelov, Eduardo Soares, 5.12.2019 https://arxiv.org/pdf/1912.02523 - **Prototypes**; a synergy between the statistical learning and reasoning .. p.2: 1) Features descriptor layer; 2) Density layer; 3) Typicality layer; 4) Prototypes layer; 5) MegaClouds layer;
* Evolving rule-based models: a tool for design of flexible adaptive systems, Plamen P Angelov, 2002/2/26
* A new type of simplified fuzzy rule-based system, Plamen Angelov, Ronald Yager, 2012
*** Fuzzy optimal control**, D Filev, P Angelov, Fuzzy Sets and Systems 47 (2), 151-156,1992, Bulgarian Academy of Science
* A generalized approach to fuzzy optimization, P Angelov, 1994

* **Evolving intelligent systems, eIS,** P Angelov, N Kasabov, 1.6.2006 https://core.ac.uk/download/pdf/71462.pdf … *"the evolution of individual systems within their life-span (**self-organization, learning through experience, and self-developing)**"; practical **on-line algorithms that work in real-time** and are close to the theoretically optimal, analytical solutions, suitable for non-stationary, non-linear problems of modeling, control, prediction, classification, clustering, signal processing. … p.1**. Fuzzy Systems** are well known for being able to formalize the **approximate reasoning** that **still separates humans from machines** … **knowledge** evolution as opposed to the usually used **data-centred** approach .. p.2 "The problems of modeling of non-linear non-stationary processes is a generic one .. of **prediction, tracking, estimation, control, classification, and clustering as special cases**";* p.3. **Approaches**: 1) *first-principles, deterministic; 2) stochastic function approximation (not-transparent); 3) ANN – not transp.; 4) Fuzzy rule-based – transp. … p.4. knowledge and data integration (KDI), participatory learning; p.10: "**Multimodal Information Processing** and Biometrics Combining speech, image and other modalities in an adaptive way, where new speech samples can be added in time, new images, new modalities (e.g. fingerprints) " p.11:"Further research: **transductive evolving systems; evolving spiking neural networks**; **evolving neurogenetic models; evolving quantum inspired neural networks**, and others. **The true intelligent systems must evolve their structure, functionality and knowledge – they can not be fixed a priori**.";* sample fuzzy rules: p.6 "IF Nk-4 is Medium AND P2offset is Low AND …); p.7 *General Principles … (**ECOS**) are modular connectionist-based systems that **evolve their structure and functionality** in a continuous, self-organised, on-line, adaptive, interactive way from incoming information. They **can process both data and knowledge** in a **supervised** and/or **unsupervised** way. ECOS learn **local** models from data through **clustering** of the data and associating a **local output function for each cluster**. Clusters of data are created based on **similarity** between data samples **either in the input space** (… e.g. the dynamic neuro-fuzzy inference system DENFIS), or **in both the input space and the output space** (EfuNN - Evolving Fuzzy Neural Network).* New samples with distance < Rmax to

existing clusters are allocated to the cluster, while distant ones form new clusters. ***Cluster centers*** *are* ***continuously adjusted****, according to new data samples, and* ***new clusters*** *are created* ***incrementally***. See the diagram of the architecture of the EfuNN: 5 feed-forward layers and a feedback; 3-rd layer: rules, evolved through supervised or unsupervised learning. *The rule nodes represent prototypes of input-output data associations.* .. 5-th layer: the real values of the output variables.[6]

* Kasabov, N. **Evolving connectionist systems: Methods and applications in bioinformatics, brain study, and intelligent machines**, Springer Verlag, London, Heidelberg, NY, 2002; Knowledge Engineering and Discovery, Research Institute KEDRI Auckland University of Technology, Auckland, New Zealand.
https://www.researchgate.net/publication/4014275_Evolving_connectionist_systems_for_adaptive_learning_and_knowledge_discovery_Methods_tools_applications
Evolving Connectionist Systems (ECOS) *p.1. Evolving connectionist systems are multi-modular, connectionist architectures that facilitate modelling of evolving processes and knowledge discovery [1,2,3]. … a NN that operates continuously in time and adapts its structure and functionality through a continuous interaction with the environment and with other systems (fig.1) according to: (i) a set of parameters P that are subject to change during the system operation; (ii) an incoming continuous flow of information with unknown distribution; (iii) a goal (rationale) criteria (also subject to modification) that is applied to optimise the performance of the system overtime p.2. (1) open space, not fixed dimensions (2) On-line incremental fast learning, pattern mode; possibly one pass of data propagation. (3) Life-long learning mode (4) Learning both as an individual systems and as evolutionary population systems.(5) Evolving structures and constructive learning. (6) Local learning and partitioning of the problem space for fast adaptation and for tracing the evolving processes over time. (7) Facilitating different kinds of knowledge; combining memory based, statistical and symbolic;*

* N. Kasabov, "**ECOS: A framework for evolving connectionist systems and the ECO learning paradigm**", Proc.of ICONIP'98, Kitakyushu, Japan, 10.1998,IOS Press, 1222-1235.
https://www.researchgate.net/publication/2288167_Ecos_Evolving_Connectionist_Systems_And_The_Eco_Learning_Paradigm

* Kasabov, N., Postma, E., and Van den Herik, J**. AVIS: A Connectionist-based Framework for Integrated Audio and Visual Information Processing**,in Proc. of Iizuka'98, Iizuka, Japan, Oct.1998 – **multimodality**, enhancement of the recognition by integration, extracting features synchronously from both visual and auditory sensory streams for person identification task. *"The experimental results support the hypothesis that the recognition rate is considerably enhanced by combining visual and auditory dynamic*

[6] See another bottom-up clustering approach, which is non-neuromorphic in CogAlg (in development): Boris Kazachenko, Todor Arnaudov (past contributor), Khan Ngyuen, K.Chee et al. https://github.com/boris-kz/CogAlg

features." At: https://pure.uvt.nl/ws/portalfiles/portal/1131975/avis.pdf : received 11/1997, accepted 15 April 1999;

See the development of the Multimodal systems, e.g.:

* **From Multimodal LLMs to Generalist Embodied Agents: Methods and Lessons,** Andrew Szot, Bogdan Mazoure, Omar Attia, Aleksei Timofeev, Harsh Agrawal, Devon Hjelm, Zhe Gan, Zsolt Kira, Alexander Toshev et al. 11.12.2024 https://arxiv.org/abs/2412.08442

* **On Path to Multimodal Generalist: General-Level and General-Bench**, Hao Fei, Yuan Zhou et al. 7.5.2025, https://arxiv.org/pdf/2505.04620
Search below #multimodal
See also appendix #Listove, sections #Robotics and #Multi-agent systems; foundational models for robotics; appendix #Irina: "GATO", the 2022 generalist agent etc.

* P. Angelov, **Autonomous learning systems: from data streams to knowledge in real-time. John Wiley & Sons, 2012 –** *MegaClouds, prototypes*

* Kasabov, **Evolving fuzzy neural networks for on-line supervised/unsupervised, knowledge–based learning**, IEEETrans. SMC – part B, Cybernetics, vol.31, No.6, 902-918,December 2001.

* Kasabov, **Evolving connectionist systems: Methods andApplications in Bioinformatics, Brain study and intelligentmachines**, Springer, London, New York, Heidelberg, 2002

* Kasabov and Q.Song, **DENFIS: Dynamic, evolving neural-fuzzy inference systems and its application for time-seriesprediction**, IEEE Trans. On Fuzzy Systems, vol.10, No.2, 144-154, April 2002

*** Nikola Kasabov** (as of 2002) … MSc and PhD from theTechnical University of Sofia, Bulgaria. His main research interests are in the areas of: intelligent information systems, soft computing, neuro-computing, bioinformatics, brain study, speech and imageprocessing, novel methods for data mining and knowledge discovery…. > 300 publications; University of Essex, UK; University of Otago, New Zealand; University of Trento, Italy;Technical University of Sofia, Bulgaria; University of California at Berkeley; RIKEN Brain ScienceInstitute, Tokyo; Delft University of Technology, and others.

*** The Effect of Prior Programming Knowledge on MemoryEfficiency When Learning a New Language,** Mojgan Hafezi Fard1, Krassie Petrova1, Nikola Kasabov1, Grace Y. Wang – EEG
https://www.researchgate.net/publication/389908159_The_Effect_of_Prior_Programming_Knowledge_on_Memory_Efficiency_When_Learning_a_New_Language - the EEG of the participants with insufficient prior knowledge in C *experienced higher cognitive load,*

indicated by increased Theta activity (17.05%) and decreased Alpha activity (16.16%); the ones with sufficient experience: *lower cognitive load with decreased Theta (11.58%); increased Alpha (17.58%)* in the F7 data. https://www.researchgate.net/publication/389908159_The_Effect_of_Prior_Programming_Knowledge_on_Memory_Efficiency_When_Learning_a_New_Language [accessed Jul 16 2025].

* Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain research reviews, 29(2-3), pp.169-19

* Kasabov, N. and Capecci, E., 2015. **Spiking neural network methodology for modelling,classification and understanding of EEG spatio-temporal data measuring cognitive processes**.Information Sciences, 294, pp.565-575. https://www.researchgate.net/publication/389908159_The_Effect_of_Prior_Programming_Knowledge_on_Memory_Efficiency_When_Learning_a_New_Language

* Fard, M.H., Petrova, K., Kasabov, N. and Wang, G.Y., 2021, December. **Studying transfer of learning using a brain-inspired spiking neural network in the context of learning a new programming language.** In 2021 IEEE Asia-Pacific Conference on Computer Science and DataEngineering (CSDE) (pp. 1-6). IEEE.

* **NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data**, Apr 2014, Neural Networks 52, pp.62-76, DOI: 10.1016/j.neunet.2014.01.006 Nikola Kirilov Kasabov – Spatio- and spectro-temporal brain data (STBD); .. neuro-economics, Brain-Computer Interfaces (BCI); integration & interaction of various STBD: EEG, fMRI, genetic, DTI, MEG, NIRS; the same principle: spiking processing; 3D evolving SNN; p.4. **2. Models of spiking neurons and methods of learning in SNN** deSNN, eSNN … p.13: (Koessler et al. 2009) *Anatomical locations of international 10–10 EEG cortical pro-jections into Talairach coordinates. Same coordinates are used in aSNNr of a NeuCube model.EEG Talairach coordinates Gyri Brodmann Area.Chan. x avg (mm) y avg (mm) z avg (mm)FP1 −21.2 ± 4.7 66.9 ± 3.8 12.1 ± 6.6 L FL Superior frontal G 10FPz 1.4 ± 2.9 65.1 ± 5.6 11.3 ± 6.8 M FL Bilat. medial 10* etc.
* Koessler et al., 2009: **Automated cortical projection of EEG sensors: anatomical correlations via the international 10-10 system.**

https://www.neucube.io/ – **NeuCube** for PC; NeuCom for PC (Student Version)
NeuCube is the world-first development environment and a computational architecture for the creation of Brain-Like Artificial Intelligence (BLAI), that includes applications across domain areas. It is based on the latest neural network models, called spiking neural networks (SNN). https://kedri.aut.ac.nz/research-groups/data-mining-and-big-data-group/neucom-a-neuro-computing-decision-support-enviroment
* FaNeuRobot: A Framework for Robot and Prosthetics Control Using the NeuCube Spiking Neural Network A, ICRA 2018, 3,23 хил. абонати
https://www.youtube.com/watch?v=cLqtUoBjJWs

* Nov 3, 2021 - **Dr. Nikola Kasabov, Deep Learning in Spiking Neural Networks for Spatio-Temporal Data**, RIT Imaging Science
https://www.youtube.com/watch?v=_UOPevSClys
An introduction to SNN, real application and review of N.Kasabov's work. … reservoir, cube, … See the questions of Alex Ororbia from 52 min. He's the first author of *"Mortal Computation: A Foundation for Biomimetic intelligence"* with Karl Friston, which is answered in "*Is Mortal Computation Required for the Creation of Universal Thinking Machines*", 2025 – an appendix of "The Prophets of the Thinking Machines". [17.7.2025]

* Angelov P, Evolving Rule-based Models: A Tool for Design of Flexible Adaptive Systems. Heidelberg, Germany: Springer-Verlag, 2002.

* Angelov, P., D. Filev, "An approach to on-line identification of evolving Takagi-Sugeno models", IEEE Trans. on Systems, Man and Cybernetics, part B, vol.34, No1, pp. 484-498, 2004.

* Kasabov, N.., Z.Chan, Q.Song and D.Greer, **Evolving neuro-fuzzy systems with evolutionary parameter self-optimisation, chapter** in: Do Adaptive Smart Systems exist? Springer Verlag, Series Study in Fuzziness, vol.173, 2005

* Kasabov, N. (1996). **Foundations of neural networks, fuzzy systems and knowledge challenge.** Frontiers in Neuroscience, 6, 96.

* Kasabov, N. (2007). Evolving connectionist systems: the knowledge engineering approach. London: Springer, (frst edition 2002).

* Kasabov, N. (2010). To spike or not to spike: a probabilistic spiking neuron model.

* Kasabov, N. (2012a). **NeuroCube: EvoSpike architecture for spatio-temporal modelling and pattern recognition of brain signal**.,

* Dimitrov, D. S ., Igor A. Sidorov and Nikola Kasabov Computational Biology, in: M. Rieth and W. Sommers (eds) Handbook of Theoretical and Computational Nanotechnology, Vol. 1 (1) American Scientific Publisher, Chapter 21,2004

* Kasabov, N., and L. Benuskova, Computational Neurogenetics, International Journal of Theoretical and Computational Nanoscience, Vol. 1 (1) American Scientific Publisher, 2004, 47-61

* Autonomous learning multimodel systems from data streams, Plamen P Angelov, Xiaowei Gu, Jose C Principe, 2017 - multiple model architecture; density, typicality, data clouds – similar to clusters, unimodal density membership functions. … heterogeneous data, combining categorical and continous …"divide and rule" principle, FRB systems (Fuzzy rule-

based), Takagi-Sugeno https://pure.aber.ac.uk/ws/portalfiles/portal/39727658/almmo_tfs.pdf

* **N.Kasabov, Evolving connectionist systems for adaptive learning and knowledge discovery: Methods, tools, applications,** Dec 2002 https://www.researchgate.net/publication/4014275_Evolving_connectionist_systems_for_adaptive_learning_and_knowledge_discovery_Methods_tools_applications *p.1. Fig.1 Four levels of evolving processes in a living organism: 4. Brain: language, decision making); 3. NN (sound perception, signal processing, control); 2. Single cell (e.g. neuronal activation); 1. Molecular (DNA, RNA, genes, proteins)*

* **Life-long learning and evolving associative memories in brain-inspired spiking neural networks**, Nikola Kirilov Kasabov, May 2024, MOJ Applied Bionics and Biomechanics 8(1):56-57 DOI: 10.15406/mojabb.2024.08.00208, https://www.researchgate.net/publication/380593884_Life-long_learning_and_evolving_associative_memories_in_brain-inspired_spiking_neural_networks – Evolving associative memories EAM – all biological systems; life-long learning (LLL) Evolving spatio-temporal associative memories (ESTAM); Evolving spatio-temporal learning(ESTL); NeuCube model; synfire & polychronisation principles for recall; methods for achieving LLL with brain-inspired SNNs: integrated spike-time and error backprop learning, neuromodulatory synaptic connections, synaptic weight regulation, homeostasis, Lyapunov energy function, evolving classifiers;
* Computational Intelligence, Bioinformatics and Computational Biology: A Brief Overview of Methods, Problems and Perspectives, December 2005, Journal of Computational and Theoretical Nanoscience, Nikola Kirilov, KasabovIgor Sidorov, Dimiter Dimitrov, https://www.researchgate.net/publication/200734651_Evolving_Connectionist_Systems_Methods_and_Applications_in_Bioinformatics_Brain_Study_and_Intelligent_Machines

* **Deep Learning and Deep Knowledge Representation in the Human Brain**, Jan 2019, DOI: 10.1007/978-3-662-57715-8_3 In book: Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence, Nikola Kirilov Kasabov

* **Methods of Spiking Neural Networks,** Jan 2019**,** DOI: 10.1007/978-3-662-57715-8_4**,** In book: Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence, Nikola Kirilov Kasabov; Ch.5: **Evolving Spiking Neural Networks** (eSNN) a class of ECOS https://www.researchgate.net/publication/327294715_Evolving_Spiking_Neural_Networks

*** EEG & Brain-Computer Interface introduction & some ECG**
https://www.emotiv.com/products/epoc-x - Epoc X - 14 channel wireless EEG headset

Система за обозначаване на мястото на сензорите в картата на предполагаеми мозъчни полета по Бродман и общото деление челен, тилен, теменен и пр.: 10/10, 10/20 … Точното разположение и размери на полетата не е еднакво при всеки човек, някои полета или подполета дори липсват, по Сергей Савельев. Дори и 64 електрода е малко информация. **Като увод виж и чата от сп. „Свещеният сметач“ бр. 30:**

*** За ЕЕГ и извличането на смислена информация от него чрез изчислителни машини и 64 електрода - из среднощен разговор между Тош и Галина Славина от 21.8.2004 *** https://www.oocities.org/eimworld/4/30/30.htm
* https://www.oocities.org/eimworld/4/30/eeg_tx.htm
* Виж също лекцията за Архитектура на мозъка на бозайните от курса по УИР в Пловдив, 2010-2011:
https://research.twenkid.com/agi/2010/Brain_Architecture_22_4_2010.pdf

AF3 - Anterior Frontal; F - Frontal; FTx - Frontal Temporal; FC - Frontal entral; CP - Central Parietal; CPz - zero (midline); PO - Parietal Occipital; O - Occipital; IZ - infra Zero;
14 канален: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF. По-мощните: 32- и 64-канални. https://www.bitbrain.com/neurotechnology-products/water-based-eeg/versatile-eeg 24 bits at 256Hz ... 8,16,32,64 channel   https://openbci.com/ European Data Format (EDF) for EEG: uV (microvolts), bpm, C; -500uV, +500 uV ... -32768 to +32767 ... prefiltering: HP 0.5Hz, LP:70Hz. Up to 256-channels... 3 channels @ 200 Hz, F3, C3, P3; 1 sec data block: 200 samples from F3, 200 C3, 200 P3 ... then the next second for F3, etc. Scale factor (-500,500)uv -> Signed Int16.
https://www.teuniz.net/edfbrowser/edf%20format%20description.html   EDF+: Annotations & Events (e.g. seizure, light flash, pressing a button) +3.45s (stimulus A); +8.32 (eyes closed) ... +200s (REM sleep); allows gaps in the recordings; UTF-8; 16K channels, channel labels len =80, float, ... EDF+C/EDF+D ... Patient ID; Recording ID; Start Date dd.mm.yy, start time hh.mm.ss ... Header byte count; Reserved 44 zeros ...

Libraries, viewers etc.: MNE-Python, EEGLAB, **EDFBrowser**, PyEDFlib BioSPPy, FieldTrip (MATLAB) .. ...   BDF+: 24-bit float version of EDF.
Transducer Type: "EEG", "ECG" (cardiogram) ..
https://eeglab.org/others/EEGLAB_and_python.html
MNE tools: https://mne.tools/stable/install/mne_tools_suite.html processing, analysis, and visualization of functional neuroimaging data (EEG, MEG, sEEG, ECoG, and fNIRS); python, C++, MATLAB ... https://mne.tools/stable/install/manual_install.html   !pip install mne (pip or !pip ... in Colab)   !pip install bycycle pactools ; !git clone https://github.com/bycycle-tools/bycycle;
run plot_1_theta_feature_... (copy in a cell in Colab or cd examples, run python ..
https://github.com/bycycle-tools/bycycle Cycle-by-cycle analysis of neural oscillations.

EDFBrowser https://www.teuniz.net/edfbrowser/ Linux & Mac Only, 2.13; Ubuntu-22.04 ..
Reviewing records for bio records of ECG, EEG and other medical sensors in EDF format.

*** Bulgarian contributions in ECG signal improvement** mentioned in the docs:
* Subtraction Method For Powerline Interference Removing From ECG, Chavdar Levkov, Georgy Mihov, Ratcho Ivanov, Ivan K. Daskalov Ivaylo Christov, Ivan Dotsinsky https://ecad.tu-sofia.bg/et/2004/Papers/Electronic%20Medical%20Equipment/Paper-C_Levkov.pdf
* Removal of power-line interference from the ECG: a review of the subtraction procedure Chavdar Levkov, Georgy Mihov, Ratcho Ivanov, Ivan Daskalov, Ivaylo Christov and Ivan Dotsinsky
* Accuracy of 50 Hz interference subtraction from an electrocardiogram I. A, Dotsinsky I.K. Daskalov
* Dynamic powerline interference subtraction from biosignals, Ivaylo I. Christov
EEG tools:
https://biosppy.readthedocs.io/en/stable/ (legacy, archived in 2022)
https://www.instructables.com/Mini-Arduino-Portable-EEG-Brain-Wave-Monitor-/
https://frontiernerds.com/brain-hack - **Open EEG** DIY community guides for building electrodes and capture devices ~>200$ (2013); (MindFlex $80 + Arduino - laptop and on the battery) https://openeeg.sourceforge.net/doc/ See also the Olimex Open EEG module and sensors.[7]

* Dimitar Filev (Димитър Петров Филев)

https://scholar.google.com/citations?hl=en&user=dhZRSRI-AAAAJ&view_op=list_works&sortby=pubdate Senior Henry Ford Fellow - Control & AI, Ford Research (ret.); Inst. for Advanced Study (TAMU)
https://www.researchgate.net/profile/Dimitar-Filev

През 2025 г. Димитър Филев е носител на наградата на IEEE *"Лютфи Задех"* за Димитър Филев за зараждащи се технологии: За техническо водачество и пионерен принос във възникващите умни технологии за управление и информационни системи в автомобилите.

https://corporate-awards.ieee.org/recipient/dimitar-filev/ 2025: IEEE LOTFI A. ZADEH AWARD FOR EMERGING TECHNOLOGIES "***For technical leadership and pioneering contributions to emerging automotive intelligent control and information systems."***

[7] Costs: as of 16.7.2025, 4 active or 5 passive electrodes capture device is at 99 Euro, electrodes are 5/9 for passive/active. Too few electrodes. A complete 14-electrode wireless headset ~1000$.

Димитър Филев *е пионер в системите интелигентно управление и ИИ в автомобилната индустрия. Основополагащатам у работа обединява невронни мрежи, размита логика, машинно обучение и напреднали технологии за управление, които създадоха революция в техниката за подпомагане на водачите и автономност на превозните средства. Като председател на отдела по Технологии за управление на „Форд“ и съвета за Технологии с изкуствен интелект, Филев беше в челния отряд на изследванията по интелигентни системи, водещи до развитието на съвсем нови функции и преживявания в автомобилната техника. Приносът му се простира о производството на автомобили, през системи за напътствия, предсказваща диагностика, автономно управление и др. Той беше движещата сила на въвеждането на методи с ИИ във „Форд“, водещи до въвеждането на огромен набор от функции с изкуствен интелект. Новаторските му изследвания продължават да разширяват пределите на ИИ в автомобилния сектор.*

*** A generalized defuzzification method via BAD distributions,** DP Filev, RR Yager, International Journal of Intelligent Systems 6 (7), 687-697, 354 cit., 1991
https://scholar.google.com/citations?view_op=view_citation&hl=en&user=dhZRSRI-AAAAJ&citation_for_view=dhZRSRIAAAAJ:qjMakFHDy7sC -
BAsic **D**efuzzification **D**istributions (BADD)

*** Fuzzy Modeling of Complex Systems**, Dimiter Filev, Bulgarian Academy of Sciences, Sofia, Bulgaria, International Journal of Approximate Reasoning, 1991
– quasilinear fuzzy model (QLFM) of a dynamic nonlinear system

* **Essentials of fuzzy modeling and control,** RR Yager, DP Filev**,** New York 388, 22-23, 1994, 3291 citations.
\\ * Herbert Toth, Siemens, Wien; A review of the book: https://corporate-awards.ieee.org/recipient/dimitar-filev/ (about 2.5 3 page columns)

* **Generation of fuzzy rules by mountain clustering,** Ronald R Yager, Dimitar P Filev, 1994/8 – learning fuzzy systems models from data.
The densites of the data points form peaks, cluster centers.

* Study of improved adaptive Mountain Clustering Algorithm, February 2010, Proceedings of SPIE - The International Society for Optical Engineering, DOI: 10.1117/12.855081, Qing Deng, Jianhui Liu
* A modified mountain clustering algorithm, September 2005, Pattern Analysis and Applications 8(1-2):125-138, DOI: 10.1007/s10044-005-0250-9, Miin-Shen et al.
* Mountain and subtractive clustering method: Improvements and generalizations, 4.2000, International Journal of Intelligent Systems, N.Pal, D.Chakraborty
* Subtractive mountain clustering algorithm applied to a chatbot to assist elderly people in medication intake, Neuza Claro, Paulo A. Salgado, and T-P Azevedo Perdico´ulis, 2021
* Clustering Stream Data by Exploring the Evolution of Density Mountain October 2017, DOI: 10.48550/arXiv.1710.00867, Shufeng GongYanfeng ZhangYanfeng ZhangGe Yu

* **Path planning for autonomous moving devices,** Inventors**:** Sanghyun Hong, Jianbo Lu, Dimitar Filev**,** 2024/8/6
* **An approach to online identification of Takagi-Sugeno fuzzy models**, PP Angelov, DP Filev, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34,2004, 1295 c.

* **Autonomous highway driving using deep reinforcement learning**, S Nageshrao, HE Tseng, D Filev, 2019 IEEE international conference on systems, man and cybernetics (SMC …, 168, 2019 https://arxiv.org/pdf/1904.00035
* **Modelling decision making using immediate probabilities,** KJ Engemann, DP Filev, RR Yager, International Journal of General System 24 (3), 281-294, 146, 1996
Ordered Weighted Averaging (OWA) operators

* **Operations for granular computing: mixing words and numbers**, RR Yager, D Filev, 1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World …, 134
* **Intelligent systems in the automotive industry: applications and trends**, O Gusikhin, N Rychtyckyj, D Filev, Knowledge and information systems 12 (2), 147-168, 118, 2007
* **Autonomous planning and control for intelligent vehicles in traffic,** C You, J Lu, D Filev, P Tsiotras, IEEE Transactions on Intelligent Transportation Systems 21 (6), 2339-2349, 77, 2019
* **Intelligent systems in the automotive industry: applications and trends,** O Gusikhin, N Rychtyckyj, D Filev
* **Knowledge and information systems** 12 (2), 147-168, 118 2007

* **Autonomous planning and control for intelligent vehicles in traffic,** C You, J Lu, D Filev, P Tsiotras, 2019

* **System and method for a human machine interface**, JS Rankin, B Tonshal, P Buttolo, Y Chen, GS Strumolo, JA Greenberg, ..., US Patent 8,866,604, 2014 https://patentimages.storage.googleapis.com/4b/51/d3/789a1c7e48357a/US8866604.pdf

* Emotive engine and method for generating a simulated emotion for an information system, Inventors: Dimitar Petrov Filev, Oleg Yurievitch Gusikhin, Fazal Urrahman Syed, Erica Klampfl, Thomas J Giuli, Yifan Chen, 2014/8/19 https://patentimages.storage.googleapis.com/24/86/02/7734a48d56911b/US8812171.pdf

* **Template-based fuzzy systems modeling**, RR Yager, DP Filev, Journal of Intelligent & Fuzzy Systems 2 (1), 39-54, 71, 1994

* **Usage prediction for contextual interface**, DP Filev, JA Greenberg, RA McGee, JG Kristinsson, F Tseng, US Patent App. 14/249,931, 69, 2014 https://patentimages.storage.googleapis.com/11/5d/c2/f8373816e1e489/US20140303839A1.pdf

* Intelligent systems in the automotive industry: applications and trends , Oleg Gusikhin, Nestor Rychtyckyj, Dimitar Filev, 2007/7, Knowledge and information systems

* **Intelligent vehicle systems: applications and new trends**, O Gusikhin, D Filev, N Rychtyckyj Informatics in Control Automation and Robotics: Selected Papers from the …, 61, 2008

* An adaptive approach to defuzzification based on level sets, DP Filev, RR Yager, Fuzzy sets and systems 54 (3), 355-360, 61, 1993

* **Explaining deep learning models through rule-based approximation and visualization**, E Soares, PP Angelov, B Costa, MPG Castro, S Nageshrao, D Filev, IEEE Transactions on Fuzzy Systems 29 (8), 2399-2407, 60, 2020
https://www.researchgate.net/publication/341694469_Explaining_Deep_Learning_Models_Through_Rule-Based_Approximation_and_Visualization
– approximate the Deep RL model with a set of IF...THEN rules that provide an alternative interpretable model, which is further enhanced by visualizing the rules; data clouds & Mega-Clouds

* A Survey of Reinforcement Learning-Based Motion Planning for Autonomous Driving: Lessons Learned from a Driving Task Perspective, Zhuoren Li, Guizhe Jin, Ran Yu, Zhiwen Chen, Wei Han, Nan Li, Lu Xiong, Bo Leng, Jia Hu,, Ilya Kolmanovsky and Dimitar Filev, 31.3.2025, https://arxiv.org/pdf/2503.23650

* Vehicle speed prediction using a cooperative method of fuzzy Markov model and auto-regressive model, J Jing, D Filev, A Kurt, E Özatay, J Michelini, Ü Özgüner, 2017 IEEE intelligent vehicles symposium (IV), 881-886, 58, 2017

* Fuzzy rule-based driver advisory system for fuel economy improvement in a hybrid electric vehicle, FU Syed, D Filev, H Ying, NAFIPS 2007-2007 Annual Meeting of the North American Fuzzy Information …, 49, 2007

* Emotive text-to-speech system and method, OY Gusikhin, PR MacNeille, E Klampfl, KA Theisen, DP Filev, Y Chen, ..., US Patent 9,495,787, 48, 2016

* Markov chain modeling approaches for on board applications, DP Filev, I Kolmanovsky, Proceedings of the 2010 American Control Conference, 4139-4145, 43, 2010

* Autonomous vehicle operation based on interactive model predictive control, DP Filev, J Lu, DD Hrovat, US Patent 10,239,529, 40, 2019

* Remaining useful life estimation of vehicle component, FF Tseng, DP Filev, IH Makki, US Patent 9,846,978, 39, 2017

* Crowdsourced weather data collection and provision, M Schunder, SJ Szwabowski, DP Filev, PR MacNeille, US Patent 9,451,030, 38, 2016 https://patentimages.storage.googleapis.com/2d/3c/87/f06738854b5d48/US8738228.pdf

* Vehicle and method of tuning performance of same, Inventors: Dimitar Petrov Filev, Jianbo Lu, Kwaku O Prakah-Asante, Fling Tseng , 2014/5/27

* Explaining deep learning models through rule-based approximation and visualization, E Soares, PP Angelov, B Costa, MPG Castro, S Nageshrao, D Filev, IEEE Transactions on Fuzzy Systems 29 (8), 2399-240, 60, 2020 https://www.researchgate.net/publication/341694469_Explaining_Deep_Learning_Models_Through_Rule-Based_Approximation_and_Visualization *approximate the Deep Reinforcement Learning model with a set of IF-THEN rules that provide an alternative interpretable model, which is further enhanced by visualizing the rules.* ..

*** A Study on the Evolutionary Adaptive Defuzzification Methods in Fuzzy Modeling** ,September 2004, International Journal of Hybrid Intelligent Systems 1(1):36-48, DOI: 10.3233/HIS-2004-11-206, Oscar Cordon, Francisco Herrera et al. https://www.researchgate.net/publication/220516008_A_Study_on_the_Evolutionary_Adaptive_Defuzzification_Methods_in_Fuzzy_Modeling - parametric and adaptive defuzzification methods by Filev and Yagger, 1991, 1993 – one or more parameters in order to modify the behavior of the defuzzifier or to get a higher accuracy; *Fuzzy Modeling (FM) designers try to find a trade-off between two edges: higher interpretability with lower accuracy or lower interpretability with higher accuracy: Mamdani fuzzy systems versus TSK fuzzy systems… the trade-off between interpretability and accuracy in FM*

*** Template-based fuzzy systems modeling**, Ronald R Yager, Dimitar P Filev, date, 1994/2, Journal of Intelligent & Fuzzy Systems
* Operations for granular computing: mixing words and numbers , Ronald R Yager, Dimitar Filev, 1998/5/4

*** Outline of a new approach to the analysis of complex systems and decision processes**, Lotfi Zadeh, IEEE Trans. Syst. Man Cybern. SMC-3 (1973) 28–44, 1973

* **HANDBOOK OF GRANULAR COMPUTING**, Ed.: W.Pedrycz, Skowron, Kreinovich; et al., https://tjzhifei.github.io/links/GrC.pdf *foundations of granular computing, interval analysis, fuzzy set theory, and rough set theory;*
Stochastic and Interval Arithmetics, Interval Analysis, Interval Methods, Fuzzy Clustering, Uncertainty, Fuzzy Sets, Granulation, Fuzzification & Defuzzification, Fuzzy Numbers & Arithmetics; Calculi of Information Granules; Fuzzy Rough Sets …
Chapter 8. Encoding and Decoding of Fuzzy Granules, Shounak Roychowdhury p.205 - … 8.2.. *the concept of fuzzification means to generate a membership function for a fuzzy set in the universe of discourse from a singleton value. In other words, it is to find the collection of data that not only encapsulates the singleton, but also captures information about its **neighborhood** in a **graded** and **meaningful** fashion. .. **Functional** and **data driven approaches**: 1… the membership function is generated by a functional generator; 2: unsupervised learning techniques like clustering or expert's characterization of the data. … no sharp distinction between these two categories.* **Functional:** triangular, trapezoidal,

Gaussian, gamma, S-, exponential; E**xperts:** *(1) polling, (2) direct estimation, and (3) experience- and intuition-based estimation.* Histograms, nearest neighbors, Clustering, end-data apprximation … *p.210: **Defuzzification** does **degranulation** of a fuzzy set – a process contrary to fuzzification.* … Mean of maxima, center of maxima, midpoint of area, center of gravity* … Unit Hypercube, nearest vertex … Subsethood defuzzification … Possibility-Probability Defuzzification: Filev & Yager, basic defuzzification distribution (BADD) Neighborhood Defuzzification: Cooperative neighbors:, evolutionary biology – *interactions among the elements in the set* … Radial defuzzification*;* … [See below a list from Wikip.]

Ch. 9 Systems of Information Granules, Frank H¨oeppner and Frank Klawonn, p.221
Zadeh:, *a 'granule is a clump of elements drawn together by indistinguishability, similarity, proximity or functionality.'*
* L.A. Zadeh. Toward a logic of perceptions based on fuzzy logic. In: V. Nov´ak and I. Perfilieva (eds), Discovering the World With Fuzzy Logic. Physica-Verlag, Heidelberg, 2000, pp. 4–28.

Ch.13. Rough-Granular Computing Andrzej Skowron and James F. Peters, 319 p.
*rough sets; information granulation; Fig.13. Evolution of AI models of computing in the Rasiowa-Pawlak School … the problem of concept approximation and reasoning; compound concept approximations; layered learning idea .. **13.3.4 Indiscernibility and Approximation** … the indiscernibility relation .. due to a lack of information (or knowledge) **we are unable to discern some objects employing available information** (or knowledge) .. we are unable to deal with each particular object but **we have to consider granules (clusters) of indiscernible objects as a fundamental basis for our theory**. (…)*

Ch.20 Construction of Rough Information Granules, Anna Gomoli´nska
infosystem, decision, condition and decision attributes, descriptor language
** Z. Pawlak. Information systems – theoretical foundations. Inf. Syst. 6(3) (1981) 205–218.*
** Z. Pawlak. Rough Sets. Comput. Inf. Sci. 11 (1982) 341–356.*
** Z. Pawlak. Information Systems. Theoretical Foundations (in Polish). Wydawnictwo Naukowo-Techniczne, Warsaw, 1983.*
** Z. Pawlak. Rough Sets. Theoretical Aspects of Reasoning About Data, Kluwer, Dordrecht, 1991.*

* **Generation of fuzzy rules with subtractive clustering,** Agus Priyono, M. Ridwan et al., 2005 https://www.researchgate.net/publication/252637529_Generation_of_Fuzzy_Rules_with_Subtractive_Clustering Types of clustering: K-means, c-means, mountain and subtractive. *Projection of fuzzy clusters onto the antecendent space. .. p.5 The **potential** of a data point to be **a cluster center** is higher when **more datapoints are closer**. The data point with the highest potential, denoted by P*i is consideredas the first cluster center c1 = (d1, e1). The potential is then recalculated for all otherpoints, excluding the influence of the first cluster center …*
* Sugeno, M., and T. Yasukawa. 1991**. Linguistic Modelling based on Numerical Data**. Proc. of IFSA .Brussel

* **Mechanism to Improve the Interpretability of Linguistic Fuzzy Systems with AdaptiveDefuzzification based on the use of a Multi-objective Evolutionary Algorithm**, Antonio A. Márquez, Francisco A. Márquez, Antonio Peregrín, Department of Information

Technologies, University of Huelva, Huelva, 21819, Spain, 2010/2011 https://www.researchgate.net/publication/261629069_A_Mechanism_to_Improve_the_Interpretability_of_Linguistic_Fuzzy_Systems_with_Adaptive_Defuzzification_based_on_the_use_of_a_Multi-objective_Evolutionary_Algorithm

* **Granular Computing and Rough Sets,** January 2005, DOI: 10.1007/0-387-25465-X_24
In book: Data Mining and Knowledge Discovery Handbook
Tsau Young ('T. Y.') Lin, Churn-jung LiauC, hurn-jung Liau
https://www.researchgate.net/publication/226882782_Granular_Computing_and_Rough_Sets
Chapter . **1. GRANULAR COMPUTING AND ROUGH SETS**
An Incremental Development … *Proposition 2 A binary neighborhood system (BNS)* ⇔ *A binary granulation(BG)* ⇔ *a binary relation (BR; Fuzzy Binary Granulations (Fuzzy binary Relations) 4.3 Topological Concept Hierarchy Trees … a nested sequence of binary granulations. Each inner layeris strongly dependent on the immediate next outer layer ...4.3.1 Granular tree. ..* ***Pawlak views partitions (classification) as knowledge****, and calls* ***a finite set of equivalence relations on a given universe a knowledge base****. He interprets* ***refinements of equivalence relations as knowledge dependencies.*** *We will take a stronger view: we regard the interpretations as the integral part of the knowledge.* Interpreation is *the* ***naming*** *of the mathematical structures, based on real world characterization;* ***the name is a summarization.*** *…the fundamental procedures in table processing* ***are to find cores and reducts of decision table.*** *.. a* ***set of nested granular structures*** *is called an* ***attribute-oriented generalizati****on when a granulation is a partition. Such concept hierarchy tree can be used to discover "soft rules", "high level rules" and others.*
…
Decision Tables, Reducts – minimal subset of attributes which reproduce the decision-making capabilities of the initial table; a table with reduced attributes; find the minimal combinations of attributes which allow for the same classifications.
https://en.wikipedia.org/wiki/Decision_table

* **On Finding All Reducts of Consistent Decision Tables,** Demetrovics Janos1 , Vu Duc Thi2 , Nguyen Long Giang3, 2014, Bulgarian Academy of Science, CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume 14, No 4,
https://cit.iict.bas.bg/CIT_2014/v14-4/1-2015-12-%20Demetrovics_VDThi_NLGiang-m-Gotovo.pdf

***** Сравни гранулираните изчисления и размитата логика с немонотонната логика на Александър Зиновиев в приложение #listove:
* А.Зиновиев, Логическа физика, 1972
http://www.vixri.ru/d/Zinov'ev%20A.A.%20%20_Logicheskaja%20fizika,1972.pdf
Виж също Петер Грендерфорс, геометрия на понятията в уводните учени и школи от основния том, когнитивна лингвистика.
* Частта за „*Coarse-Graining*" в списъка с учени и школи; института Санта Фе и др.;

разделителна способност на възприятие и управление, степени на обобщение; "multi-scale representations" …

* Александър Тошев – Alexander Toshev

https://sites.google.com/view/alextoshev
https://scholar.google.com/citations?user=T6PbwPIAAAAJ&hl=en

А.Тошев е автор и съавтор на важни разработки в машинното обучение, компютърното зрение и др. от началото на 2010-те..

* **Shape-Based Object Detection via Boundary Structure Segmentation**, Alexander Toshev, Ben Taskar, Kostas Daniilidis, 2011/2012
Shape descriptor: Chordiogram: contours, boundaries, edges; gradient- and texture-based approaches: local, while shape-based: larger scale/scope, at best when capturing the whole object (Gestalt psych., holism); Chord features: orientation of the normals at boundary edges (cmp. spin images); recover correspondance between shapes; segments, boundary; foreground, background; boundary structure segmentation …

* **SHAPE REPRESENTATIONS FOR OBJECT RECOGNITION**, **Alexander Toshev,** PhD thesis, 2011, University of Pennsylvania; supervisor: Kostas Daniilidis
https://repository.upenn.edu/bitstreams/26ad7bb7-b626-4262-a899-7a2ab505ad19/download * Kostas Daniilidis
https://scholar.google.com/citations?user=dGs2BcIAAAAJ&hl=en
Chordiogram: captures both the object boundary & the interior in a holistic way and is invariant to some rigid transformations & robust to deformations of the shape; it works even with clutter and unsegmented images**.** ***Chord features****:* length, orientation, normals. The normals point towards the interior of the object. **Quantization** of the chordiograms in bins. **Similarity**: **Intra-image** similarity, **Inter-image** similarity; **Co-saliency** score **function**; Co-saliency **region/feature Matching**; **segmentation** synchronization; image segmentation subspaces S1, S2 as eigenvectors; **co-salient region**; **Alignment** in the **embedding space**; Estimation of **Dense Correspondences**; **appearance** similarity between the regions; **geometric compatibility** with respect to the affine transformation; Finding the correct match for a point; ranking; several matches with high scores due to similar or repeating structure; **segment-based reranking**; compare quantitatively the difference between the initial and the improved set of feature matches. **Place Recognition task**; ***Related Work****:* **Spectral graph matching**, weighted graph matching: characterize the graphs by dominant eigenvectors: don't capture co-salient structures; simultaneous object recognition and segmentation with spectral clustering in a graph, capturing the relationship(pixels, object parts); correct partial correspondences between manifolds: infer the complete alignment by regularization based on similarities between points on the manifolds**. Spectral graph matching**, **spectral clustering**; Co-segmentation: simultaneous segmentation of two images and extraction of the common

objects: generative graphical model: a smoothness prior for segmentation and **appearance-based model** for the common object; usually described as **histogram**. **Joint image representation**: the similarity between images as the composability of one of the images from large segments of the other image. ... **Co-salient regions:** segments which are **coherent and distinct from their surroundings & similar to each other**; feature similarities. **Normalized Cuts** framework. Bins. **Partitioning** of the **bins**. **Boundary grouping** principles; edge/contour detector cannot detect all object boundaries since there is no evidence in the image, but **the boundaries can be hallucinated to recover the missing ones**. Perceptual **Grouping** & Region grouping **principles**, perceptually salient segmentation; **unnatural**: **Intersecting contours** & **Including** contour. **Segments**: A **set** of **regions**. **Boundaries**: A set of non-intersecting, non-including, closed contours (see Fig. 3.5). Region-boundary constraints. Shape-based **detection cost.** Segmentation: The opimal values of the boundary and segment indicators; **object interior and boundary**. **Shape-based detection cost**: The minimum of the objective function quantifies the quality of match based on shape similarity. **Region-boundary constraints**; contours as a result of **bottom-up contour grouping** preprocessing step. The use of **pre-grouped contours** leads to more stable solution and reduces the computational complexity: but ignoring region information; bottom-up grouping by selecting object boundaries from a set of long salient contours [Zhu et al., 2007]; segments as additional constraints. **Shape Matching;** Parametrization of the chordiogram, boundary indicators, **L1 distance; shape matching cost**; shape similarity between a model and a particular selection of segment boundaries: minimization of the cost while taking into account the relation between boundaries and segments. *Decompose a chordiogram in terms of chordiograms which relate pair of boundaries*; select via boundary indicator variables; sum of all boundary-pair chordiograms for all pairs of boundaries; **boundary indicator variables**; Chordiogram **additivity**; Segment boundary parameterization; Segment parametrization; segment: **foreground** or **background (1,-1);** preprocessing: **Oversegmentation**. **Boundary Structure Segmentation (BoSS):** detection and segmentation simultaneously in a unified framework; shape similarity and figure/ground segmentation in a single step. Linear constraints relating **superpixels** with their boundary. [**multilevel**] bin-based distance, graph matching and chordiogram distance; correspondence on pairs of shapes & correspondence recovery**. Familiarity with the target shape** plays a large role in **figure/ground** assignment. Perceptually salient region; chordiogram matching as **bipartite matching among chords**. **Shape Part Correspondence**: quantify the similarity between two shapes by establishing correspondences between points on the shapes which serves as an explanation. **Shape representation**: object **boundary (closed curve)** & **interior**. 2.1. **Contour-based shape descriptors**: s**hape context, k-adjacent segments, hierarchical shapes; shock graphs.** Interior: in **ambiguous** cases it must be **selected** (e.g. Rubin's vase). Holism, holistic approach: **Gestaltism**, **emergent properties** of art configurations. **Edge**- and **texture**-based local descriptors: a **patch** in the image can contain **rich information**; **unlike shape**: very limited expressiveness – locally a curve, potentially with a few high curvature points. .. **Transformation invariance**: **rigid** & **non-rigid; Articulations**: composition of independent deformable parts; Image **artifacts**, **Clutter**, **Occlusion** & **missing parts**. … See the references.

* **From Multimodal LLMs to Generalist Embodied Agents: Methods and Lessons,** Andrew Szot, Bogdan Mazoure, Omar Attia, Aleksei Timofeev, Harsh Agrawal, Devon Hjelm, Zhe Gan, Zsolt Kira, Alexander Toshev, 11.12.2024
https://arxiv.org/abs/2412.08442 – MLLMs, extending the tasks to Embodied AI, Games, UI Control, Planning. Generalist Embodied Agent (GEA); supervised learning on large multimodal data, the importance of cross-domain data – strong generalization across unseen tasks; a unified learned multi-embodiment action tokenizer; physics and geometric reasoning. Vision-Language-Action models; strong generalist capabilities; multi-embodiment action tokenizer vs uniform discretization of OpenVLA; cmp: multi-task agents; finetuning LLMs with self-generated data‘ goal-specified PartiallyObservable Markov Decision Processes (POMDPs) ; Spaces: observation, action, goal; reward model: O,A,G,R.The goal space G – textual description of the task to solve. O – RGB images; environment types – *domains* with a diverse set of action spaces. The goal is to learn one policy which works in many environments E. $M_i = (O_i, A_i, G_i, R_i)$ for each i ∈ E. Continuous Multi-Embodiment Tokenizer; hierarchical encoding; 2 codebooks x 512 tokens, token vector dim = 1024.
Training: Stage 1: Supervised-Instruction Finetuning, demonstration datasets from all environments E. Global batch size = 256, O context length = 3, 2 days x 8 nodes x 8 H100 GPUs. Stage 2: Online RL, proximal policy optimization PPO, simulator …
Archictecture of GEA: ((Prompt & observation history), Visual encoder Visual Bridge) LLM Continous & Discrete Control (Video games (jump, left, right…), UI Control (Tap 23,47; click 243,211), Navigation (Forward, left…), Static & Mobile Manipulation (Joint velocity; Delta joint position, End-effector,) Action for Environment (dx dy dz Residual VQ-VAE Encoder, “move right” LLM Tokenzer [212, 201]) Multi-Embodiment Action De-Tokenizer Action tokens (Unified Token Output Space) LLM …
A table with Embodied datasets p.5: OpenX, Meta-World, CALBIN, Mniskill, Habitat Pick, Habitat Place, Habitat Nav, BabyAI, LangR, Procgen, Atari, AndroidControl … Action types: Continous Various, … +gripper, joint velocity, joint position, discrete, mixed. … Embodiment type: various robots, sawyer, Franka Arm, Fetch, Virtual , … # Trajectories: 1.2M, 45k, 18k, 5k, 50k … 320K … Data source: Various, Scripted, Human, Motion Planner, RL Expert, Shortest Path

* **Datacomp-LM: In search of the next generation of training sets for language models** Jeffrey Li, Alex Fang, Georgio, … **Alexander Toshev** … et al. from 1 University of Washington, 2 Apple, 3 Toyota Research Institute, 4 UT Austin, 5 Tel Aviv University, 6 Columbia University, 7 Stanford, 8 UCLA, 9 JSC, 10 LAION, 11 AI2, 12 TUM, 13 CMU, 14 Hebrew University, 15 SambaNova, 16 Cornell, 17 USC, 18 Harvard, 19 UCSB, 20 SynthLabs, 21 Bespokelabs.AI, 22 Contextual AI, 23DatologyAI
https://proceedings.neurips.cc/paper_files/paper/2024/file/19e4ea30dded58259665db375885e412-Paper-Datasets_and_Benchmarks_Track.pdf
A new dataset of 240T tokens from web; CommonCrawl, methods for improving the datasets.

DCLM-POOL and DCLM-BASELINE Model sizes: multi-scale design of 400M – 7B models for comparisons (8B to 276B training tokens, filtered from a pool of 469B to 15.7T tokens). H100 hours: 26 to 7300). Data curation: filtering (by language, heuristics, quality); data deduplication, repetition filter, page length, word-length, stop words, ellipsis count, URL, word removal ratio , fasttext, Bloom dedup... List of large datasets (C4: 160B, The Pile 300B, RefinedWeb: 600B, Dolma: 3 T, FineWeb-Edu: 1.3T, RedPajamav2 30 T. Code: StackV2 1.3T. DCLM-POOL: 200B documents (370TB compressed) from hundreds of millions of sources --. 240T GPT-NeoX tokens; decontamination (test data in the training sets); DCLM-BASELINE: 3 B d. Selection: filtering & mixing of data from the pools; high-quality: Wikipedia, StackExchange, peS2o. See the Appx; p. 76: Datasheet. WARC (web archive) … *“”Q13: … DCLM-POOL is based on Common Crawl, which can be thought of as a snapshot of the internet at a given time*

https://huggingface.co/datasets/allenai/peS2o – 39M open-access academic papers, cleaned, filtered and formatted for pre-training of LM.

* **Datasets, Documents, and Repetitions: The Practicalities of Unequal Data Quality,** Alex Fang, Hadi Pouransari, Matt Jordan, **Alexander Toshev**, Vaishaal Shankar, Ludwig Schmidt, Tom Gunter et al. 3.2025 https://arxiv.org/pdf/2503.07879
DCLM vs RefinedWeb, C4 filtering on repeatability of the results.

* **Meta-World: A Benchmark and Evaluation for Multi-Task and Meta Reinforcement Learning** https://meta-world.github.io/   50 manipulation tasks ... basketbal, button press, dial turn, drawer close, peg insert side, pick place, push, reach, sweep into, window open; test: door close, drawer open, lever pull, shelf place, sweep ... https://github.com/Farama-Foundation/Metaworld

* **Meta-World: A Benchmark and Evaluation for Multi-Task and Meta Reinforcement Learning**, Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Avnish Narayan, Hayden Shively, Adithya Bellathur, Karol Hausman, Chelsea Finn, Sergey Levine, 10.2019/6.2021 https://arxiv.org/abs/1910.10897 p.40 list of success metrics per task ... *“Meta-reinforcement learning aims to leverage the set of training task to learn a policy π(a|s) that can quickly adapt to new test tasks that were not seen during training, where both training and test tasks are assumed to be drawn from the same task distribution p(T)” ... a task, T , in Meta-World is defined as the tuple (reward function, initial object position, target position*

* **CALVIN: A Benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks**
July 2022IEEE Robotics and Automation Letters 7(3):1-1
DOI: 10.1109/LRA.2022.3180108, Oier Mees, Lukas Hermann, Lukas Hermann, Erick Rosete-Beas, Wolfram Burgard, Wolfram Burgard, 13.7.2022
https://arxiv.org/abs/2112.03227 *Composing Actions from Language and Vision (CALVIN) http://calvin.cs.uni-freiburg.de/ https://youtu.be/x2CepHlhnqg 18.08.2023 ...dddd* 34 tasks, goal-conditioned RL, ... CALVIN dataset: 24 h teleoperated play data, 2.4 M interaction steps, 20Kk labeled language sequences ... “grasp the drawer handle, then

open it", "grasp the handle of the drawer, then open it", "grasp the handle of the drawer and open it", "open the drawer', "go open the drawer", "pull the drawer", öpen the cabinet drawer" etc. ... - variants of the same intention ... play data: 64 frames; CALVIN challenge:
Observation Space: RGB static camera 200x200x3, Depth static camera 200x200, RGB gripper camera 84x84x3, Depth gripper camera 84x84, Tactile image 120x160x2, Proprioceptive state: EE position 3, EE orientation 3, Gripper width 1, Joint positions 7, Gripper action 1 (7 DOF for the articulated robot's joints).
Action Space: Absolute cartesian pose (EE position 3, EE orientation 3, Gripper action 1), Relative cartesian displacement (=), Joint action (Joint positions 7, Gripper action 1) ... CALVIN Environment - 7-DOF Franka Emika Panda robot arm with a parallel gripper ... Example of Long-horizon language instructions*: "turn on the led → open drawer → push the red block → pick up the red block → place in slider"*. List of tasks: p.9.

Franka Robotics: https://franka.de/ FRANKA RESEARCH 3 "**THE REFERENCE PLATFORM FOR AI AND ROBOTICS RESEARCH. GLOBALLY.**"
https://www.mybotshop.de/Franka-Emika-Panda-FCI-Licence €24.900,00 + shipping @ 3.2.2025

ManiSkill: https://github.com/haosulab/ManiSkill *SAPIEN Manipulation Skill Framework; open source GPU parallelized robotics simulator and benchmark, led by Hillbot, Inc.*

* **habitat-sim** https://github.com/facebookresearch/habitat-sim *A flexible, high-performance 3D simulator for Embodied AI research.* https://aihabitat.org/
Habitat-Sim, Habitat-Lab, Habitat Challenge

Habitat Navigation Challenge 2023 https://aihabitat.org/challenge/2023/
Two tasks: 1: ObjectNav – "find a chair" 2: ImageNav – RGB goal image

Habitat Matterport 3D Semantics Dataset https://aihabitat.org/datasets/hm3d-semantics/ – "*the largest-ever dataset of 3D real-world and indoor spaces with densely annotated semantics ...*" 142K object instaqnces, 216 3D-spaces (homes); semantic segmentation : labels for the regions of a window, bed, chair, door etc.

Procgen Benchmark: Procedurally-Generated Game-Like Gym-Environments
https://github.com/openai/procgen

BabyAI platform. *A testbed for training agents to understand and execute language commands.* https://github.com/mila-iqia/babyai → **Minigrid** library
https://github.com/Farama-Foundation/Minigrid – *Simple and easily configurable grid world environments for reinforcement learning*

Scott Reed, Konrad Zolna, Emilio Parisotto, Sergio Gomez Colmenarejo, Alexander Novikov, Gabriel Barth-Maron, Mai Gimenez, Yury Sulsky, Jackie Kay, Jost Tobias Springenberg, et al. **A generalist agent.** arXiv preprint arXiv:2205.06175, 2022 (5.2022/11.2022)
https://arxiv.org/abs/2205.06175 ***Gato*** *– multi-modal, multi-task, multi-embodiment*

generalist policy. The same network with the same weights can play play Atari, caption images, chat, stack blocks with a real robot arm etc. ... 604 distinct tasks with varying modalities, observations and action specifications. ... 6. Related work: ***Decision Transformers*** *(Chen et al., 2021b; Reid et al., 2022; Zheng et al., 2022; Furuta et al., 2021) and* ***Trajectory Transformer*** *(Janner et al., 2021) ...*
See the data samples preparation, embeddings, training procedure etc. from the Supplementary materials p.28

See below: **Recognizing Everything from All Modalities at Once: Grounded Multimodal Universal Information Extraction**, Meishan Zhang et al., 2024 ... etc.

The RobotSlang Benchmark: Dialog-guided Robot Localization and Navigation, *Shurjo Banerjee, Jesse Thomason, Jason J. Corso, 2021*
https://proceedings.mlr.press/v155/banerjee21a/banerjee21a.pdf

LaNMP: A Language-Conditioned Mobile Manipulation Benchmark for Autonomous Robots, Ahmed Jaafar,, Shreyas Sundara Raman et al.
https://arxiv.org/html/2412.05313v1 *long-horizon mobile manipulation tasks in large, diverse environments ... Existing datasets do not integrate all these aspects, restricting their efficacy as benchmarks… Language, Navigation, Manipulation, Perception (LaNMP, pronounced Lamp). ... 574 trajectories across eight simulated and real-world environments for long-horizon room-to-room pick-and-place tasks specified by natural language. 20 attributes: RGB-D images, segmentations, poses of the robot body, end-effector, grased objects, 3 Hz ... NeurIPS 2023 Open Vocabulary Mobile Manipulation challenge – the best result only 33% - lack of a benchmark ... A task: "Go to the kitchen, pour the boiling water into the teapot, then bring it to me in the living room" – requires language understanding, navigation, manipulation & perception capabilities.* ***2.2 Datasets of Language, Navigation and Perceptio:*** *Room-to-Room, Room-Across-Room, ALFRED, CoNav and TEACh. QUARD (for quadruped robots) 2.3.* ***Datasets of Navigation, Manipulation and Perception****: MoMaRT, Mobile ALOHA, BRMData – bimanual-mobile robot manipulation ...* ***2.4 ... Language, Navigation, Manipulation and Perception*** *– RT-1 (transformer: fetch & deliver in a kitchen scene) ...*

* **Grounding Multimodal Large Language Models in Actions**, A Szot, B Mazoure, H Agrawal, D Hjelm, Z Kira, A Toshev, arXiv preprint arXiv:2406.07904, 6.2024/9.12.2024 https://arxiv.org/abs/2406.07904

MLLM, action space adapters (ASA); Continuous ASA - (regression dx, dy, dz), uniform tokenizaion ... Discrete ASA, MLP, classification, semantic tokenization, non-semantik tokenization... Task goal + prompt --> Language Embedding ... Multimodal LLM --> Adapter head --> Action tokenization --> Action decoder --> at (action for Environment)

Въображаеми среди за изпробване на роботи: CALVIN - tabletop robot, 200x200 RGB fixed camera, 6DoF end-effector and binary gripper state (open/close), Meta-World ML-45, 45 tasks, observations: 200x200 RGB fixed camera - evaluate on unseen objects and robot starting states, Habitat Pick (HabPick) - 336x336 RGB egocentric head camera, BabyAI - grid world - navigates and interacts with objects, 200x200 RGP top-down view, Language Rearrangement (LangR) - "store all the fruit in the fridge", 70 skills, ... 336x336 RGB head camera*; вж. с.15 илюстрации и подробности;

Някои задачи за CALVIN от с.20: turn off led, move slider left, rotate red block right, open drawer, turn off lightbulb; за BabyAI: goto, pickup, open, putnext, pick up seq go to; Meta-World: assembly, basketball, button-press-topdown, door-open, faucet-open, ...

* Виж също AI2THOR interactive environments for Embodied AI: https://ai2thor.allenai.org/ срвн плана от "Вселена и Разум 5", Т.Арнаудов, 2004

* **MEGA-Bench: Scaling Multimodal Evaluation to over 500 Real-World Tasks,** Jiacheng Chen, Tianhao Liang, Sherman Siu, Zhengqing Wang, Kai Wang, Yubo Wang, Yuansheng Ni, Wang Zhu, Ziyan Jiang, Bohan Lyu, Dongfu Jiang, Xuan He, Yuan Liu, Hexiang Hu, Xiang Yue, Wenhu Chen, 10.2024/11.2024
https://tiger-ai-lab.github.io/MEGA-Bench/ https://arxiv.org/abs/2410.10563
https://github.com/TIGER-AI-Lab/MEGA-Bench/
505 realistic tasks; many output formats: numbers, phrases, code, \LaTeX, coordinates, JSON, free-form, etc.; 45 evaluation metrics … *fine-grained capability report across multiple dimensions (application, input type, output format, skill, …)*

* **World-consistent Video Diffusion with Explicit 3D Modeling,** Qihang Zhang, Shuangfei Zhai, Miguel Angel Bautista, Kevin Miao, Alexander Toshev, Joshua Susskind, Jiatao Gu, 2.12.2024 https://arxiv.org/pdf/2412.01821
Datasets: RealEstate10K, ScanNet,MVImgNet, CO3D, Habitat; 2B diffusion transformer, generates 3D point cloud; training: batch size 128, 1M steps on 64xA100 x two weeks.
– Срвн. предвиждания: Compare: "Chairs, Buildings, Caricatures ...", T.A. 2012 - Vision as reverse graphics with explicit 3D-reconstruction and modelling.

Виж също: **ViewDiff: 3D-Consistent Image Generation with Text-to-Image Models,** L.Hollein et al, https://lukashoel.github.io/ViewDiff/ multi-view image generation

* **On Robustness in Multimodal Learning**, B McKinzie, V Shankar, J Cheng, Y Yang, J Shlens, A Toshev, 2023 https://arxiv.org/pdf/2304.04385 – "Здравина", устойчивост (robustness) - виж "adversarial attacks", устойчивост на шум, смущения на разпознатото или предвиденото; устойчивост срещу неочаквани разлики в резултатите при малки промени в данните - "distribution shift", голяма промяна на поведението (напр. при класическите "атаки" беше открито, че когато в изображението се внедри незабележим от човешко око шум, невронните модели за разпознаване на образи можеше да се манипулират да върнат коренно различен отговор от онова, което вижда човек, с което показваха, че особеностите и начина, по който се учи и решава НМ не са съвместими). Модалност: сетивност; модалности - типове, видове данни, вид сетиво (може да отразява и движение, промяна - също данни). Вж т.2.2. Мултимодалността: обучение с всички сетивности, смесване в единно представяне (aggregtor), общо пространство; наличието на данни от всички сетивности се приема за положителен пример; извод за съдържанието на някоя от тях при липса на някои от началните. Смесване в общо представяне - напр. чрез преобразители. Успоредни съгласувани потоци от данни: предобучение без учител.

Downstream task (функцията-изовд на модела: напр. разпознаване) - supervision, downstream training data. Видове тестове: премахване, добавяне или пренос по време на тест (на модалност); 1: непълна информация; 2: липсваща модалност в обучителните данни; 3. пренос на задача, научена в един набор модалности за решаване в друг. Многосетивно самообучение (multimodal self-supervised learning): различни гледни точки към едно и също съдържание; чрез контрастно пресъздаване или пресъздаване с маски. Отделните сетивности се кодират по специализиран начин; специализирани за всяка сетивност функции на загубата (task specific loss L) 6.1 Анализ на научените многосетивни представяния: по-добре е да има повече модалности по време на предобучението или изпитанията. ... 8. Принос: А.Тошев: инициатор на проекта, ръководител на изследванията, един от проектантите на рамката за оценка на надеждността; автор на главните алгоритмични приноси на статията и написал повечето от статията. Видове намеси в обучението *(Training interventions): Modality Augmented Self-Distillation* (MASD): насърчава постоянство между научните представяния при маркирани и немаркирани модалности и WiseFT: увеличават надеждността 1.5-4 пъти при тестове върху AudioSet, Kinetics-400, ImageNet-Captions.

* **Intriguing properties of neural networks**, Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., and Fergus, R. arXiv preprint arXiv:1312.6199, 2013/19.2.2014 https://arxiv.org/pdf/1312.6199 **Adversarial examples (AE)** „Deep … have recently achieved state of the art performance on speech and

visual recognition tasks…" counter-intuitive properties; fairly discontinuous input-output mappings; the entire space of activations contains the bulk of the semantic information, rather than the individual units; *adversarial examples*: imperceptible non-random perturbation to an image for arbitrarily changing the network's prediction; AE with Gaussian noise MNIST: 51% recognition, while visually indistinguishable adv.modified: 0%; possible reasons: *"the set of adversarial negatives is of extremely low probability, and thus is never (or rarely) observed in the test set, yet it is dense (much like the rational numbers), and so it is found near every virtually every test case."*

* **Deep neural networks for object detection**, Christian Szegedy, Alexander Toshev, Dumitru Erhan, 2013 https://proceedings.neurips.cc/paper/2013/file/f7cade80b7cc92b991cf4d2806d6bd78-Paper.pdf Another from the era: **Deformable Part-based Model (DPM) ..**

* **DeepPose: Human Pose Estimation via Deep Neural Networks**, Alexander Toshev, Christian Szegedy, 20.8.2014 https://arxiv.org/pdf/1312.4659 (compare Dragomir Anguelov's PhD, 2005, on 3D range data)

* **On the Modeling Capabilities of Large Language Models for Sequential Decision Making**" by Martin Klissarov, Devon Hjelm, Alexander Toshev and Bogdan Mazoure, 10.2024 Виж бел. за Мартин Клисаров за RL.

* **Show and Tell: A Neural Image Caption Generator,** Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Vinyals_Show_and_Tell_2015_CVPR_paper.pdf Image-Text, LSTM sentence generator.

* **Generating natural language descriptions of images,** S Bengio, O Vinyals, AT Toshev, D Erhan**,** US Patent App. 18/662,584**,** 2024

* **Generating natural language descriptions of images,** S Bengio, O Vinyals, AT Toshev, D Erhan**,** US Patent 12,014,259

* **Robot navigation using a high-level policy model and a trained low-level policy model**, A Toshev, M Fiser, A Wahid, 13.8.2024, US Patent 12,061,481 https://patentimages.storage.googleapis.com/3c/7b/96/6a583901857152/US12061481.pdf

* **Controlling agents using scene memory data,** K Fang, AT Toshev**,** US Patent App. 18/536,074**,** 2024

* **Controlling agents using scene memory data**, K Fang, Alexander Toshkov Toshev, US Patent 11,842,277, 12.12.2023 (filled in 11-2019) https://patents.google.com/patent/US11842277B2/en

https://patentimages.storage.googleapis.com/da/4f/0f/e087456366076d/US12248875.pdf – Управляване на агент с използване на памет за сцената: Получаване на състоянието на средата в текущия момент, пораждане на неговото вътрешно представяне*, обработка на данни от паметта за сцената, включващи представянето и предходни наблюдения, получини в минали стъпки във времето чрез кодираща невронна мрежа, където тя е устроена да прилага механизъм за кодиране със самовнимание спрямо данните за паметта за сцената; обработка на кодираното представяне на миналото и настоящето наблюдение за да се породи извод, който избира действие, което кара агента да изпълни избраното действие. *embedding

* **Future prediction, using stochastic adversarial based sampling, for robotic control and/or other purpose (s)** Inventors: Anthony Jacob Piergiovanni, Anelia Angelova, Alexander Toshev, Michael Ryoo; **Publication date:** 2025/6/24; **Patent office: US; Patent number:** 12340307**; Application number:** 17638469 https://patentimages.storage.googleapis.com/13/44/42/eee4dda90f2bf0/US12340307B2.pdf

* **Natural language control of a robot,** K Hausman, B Ichter, S Levine, A Toshev, F Xia, C Parada**,** US Patent App. 18/128,953**,** 25, 2023
* **Do as i can, not as i say: Grounding language in robotic affordances**, M Ahn, A Brohan, N Brown,... A Toshev, ..., (45), 10.8.2022, arXiv preprint https://arxiv.org/pdf/2204.01691 https://say-can.github.io/

Тълкуване и изпълнение на непреки словесни инструкции от робот: "Разсипах си напитката на бюрото. Помогни ми да почистя" --> Намери гъба. Вземи гъбата. Донеси гъбата до човека. Остави гъбата на бюрото. Action Space Adapters ... Сравни с SHRDLU, STRIPS (без обосноваване). Описанията на действията: сравни записки Т.А., "Свещеният сметач" ок. 2013-2015, "behaviontrospective" и др. от бдщ: "Създаване на мислеща машина".

* **Walk the talk: Connecting language, knowledge, and action in route instructions.** M. MacMahon, B. Stankiewicz, and B. Kuipers. 2006

* **World-consistent Video Diffusion with Explicit 3D Modeling**, Qihang Zhang, Shuangfei Zhai, Miguel Angel Bautista, Kevin Miao, Alexander Toshev, Joshua Susskind, Jiatao Gu, 2.12.2024 https://arxiv.org/pdf/2412.01821
* **Evolving Space-Time Neural Architectures for Videos**, AJ Piergiovanni, Anelia Angelova, Alexander Toshev, Michael S. Ryoo, 11.2018/8.2019 https://arxiv.org/abs/1811.10636 Виж бел. към Анелия Ангелова

(…)

* Lubomir Bourdev – Любомир Бурдев

Компютърна графика и компютърно зрение; computer graphics & computer vision; Adobe Photoshop Elements 4, 2005: Person/Face recognition and detection - Soft Cascade, better than Viola-Jones; the first consumer product using Face detection; the first computer vision engineer at Facebook, author of the original image recognition system https://lubomir.org/ https://lubomir.org/#engprojects

Linkedin: Researcher, Developer, Manager, Enterpreneur, Inventor, Parent, Co-founder and CEO * **Stealth Startup ·** Aug 2023 - Present · 1 yr 7 mos *Stay tuned!*
***** Co-founder and CEO **WaveOne**, Inc.WaveOne, Inc. Dec 2015 - Feb 2023 · 7 yrs 3 mos
***** Research Scientist / Engineering ManagerFacebook Mar 2012 - Aug 2015 · 3 yrs 6 mos Sr. Research Scientist
***** Sr. Research Scientist Adobe SystemsAdobe Systems Jun 1998 - Mar 2012 · 13 yrs 10 mos
***** University of California, Berkeley logo**,** Ph.D., Computer Science, 2007-2011
Computer vision research with Prof. Jitendra Malik – poselets, decompose the pose from appearance, which helps with computer vision problems like determining where objects are in the image, segmenting them from the background, inferring the pose and action of people, their gender, hair style, types of clothes, etc. Poselets and Deep Poselets were the state-of-the-art methods for detection of people and person attribute classification (inferring gender, clothes, hair style, etc) at the time of my graduation.
* Brown University**,** B.A. + M.S., Computer Science, 1994 – 1998
Computer graphics research … Co-author of two SIGGRAPH papers… Second student in the previous 10 years in the Brown Computer Science department to get bachelors and masters degrees in a total of four years.

* D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. **Learning spatiotemporal features with 3D convolutional networks**. In ICCV, 2015

https://www.kaldata.com/it-новини/българин-е-част-от-амбициозен-ai-проект-82896.html - **Българин е част от амбициозен AI проект на Facebook**, Методи Дамянов, 9:21 | 23.09.2013 161 0
* https://www.standartnews.com/tekhnologii/apple-e-pridobil-kompaniya-za-ai-algoritmi-za-kompresirane-na-video-520351.html 28 мар 23 | 8:10 0 коментара 3355 Агенция Стандарт
https://lubomir.org/#publications

* **Generic Image Library** - GIL, then in Boost C++

Facebook Image Classification 2012 - Lubomir was the first guy hired to do computer vision; "run on 300M photos per day". First year - traditional CV features; then: CNN, 9 versions for 2 years, the latest recognizes "more than a thousand types of objects, scenes, activities and places of interest using convolutional neural networks with multiple loss functions." Peak load:10K calls/sec, run on every photo & every second of every video on Facebook and

Instagram. Lubomir developed "a large part of the training code as well as a highly optimized feedforward path used in production."

* PIM: **Video Coding using Perceptual Importance Maps,** Evgenya Pergament, Pulkit Tandon, Oren Rippel, Lubomir Bourdev, Alexander G. Anderson, Bruno Olshausen, Tsachy Weissman, Sachin Katti, Kedar Tatwawadi**,** Computer Graphics Arxiv 2022

* ELF-VC: Efficient Learned Flexible-Rate Video Coding, Oren Rippel, Alexander G. Anderson, Kedar Tatwawadi, Sanjay Nair, Craig Lytle, Lubomir Bourdev,, Computer Vision International Conference in Computer Vision (ICCV 2021)

* **Real-Time Adaptive Image Compression**, Oren Rippel and Lubomir Bourdev, Computer Vision International Conference in Machine Learning (ICML 2017) - ML lossy image compression, better than the existing codecs, while running in real-time. Typically: 2.5 times smaller files than JPEG & JPEG 2000, 2 times < WebP ... generic images across all quality levels. At the same time: lightweight and deployable, e.g. the Kodak dataset: ~ 10ms/image on GPU.

* **Deep End2End Voxel2Voxel Prediction,** Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani and Manohar Paluri**,** Computer Vision The 3rd Workshop on Deep Learning in Computer Vision 2016 (in CVPR 2016)

* D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. **Learning spatiotemporal features with 3D convolutional networks.** In ICCV, 2015.

* **Microsoft COCO: Common Objects in Context**, Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnik and Piotr Dollar, Computer Vision Arxiv 2015

*** Deep Poselets for Human Detection,** Lubomir Bourdev, Fei Yang, Rob Fergus**,** Computer Vision Arxiv 2014 *.. detecting people in natural scenes using a part approach based on poselets.. a bootstrapping method that allows to collect millions of weakly labeled examples for each poselet type; train a CNN to discriminate different poselet types and separate them from the background class.* Then: the CNN → *represent poselet patches with a Pose Discriminative Feature (PDF) vector -- a compact 256-dimensional feature vector that is effective at discriminating pose from appearance.* The poselet model is trained *on top of PDF features and combine them with object-level CNNs for detection and bounding box prediction. The resulting model leads to state-of-the-art performance for human detection on the PASCAL datasets.*

* **Articulated Pose Estimation using Discriminative Armlet Classifiers**, Georgia Gkioxari, Pablo Arbelaez, Lubomir Bourdev and Jitendra Malik, Computer Vision IEEE Conference in Computer Vision and Pattern Recognition (CVPR 2013) ... *a novel approach for human pose estimation in real-world cluttered scenes; focus on the challenging problem of predicting the pose of both arms for each person in the image. .. the notion of poselets, training a highly discriminative classifiers to differentiate among arm configurations – armlets; a rich representation which, in*

addition to standard HOG features, integrates the information of strong contours, skin color and contextual cues in a principled manner. Unlike existing methods: an evaluation on a large subset of images from the PASCAL VOC detection dataset with critical visual phenomena, such as occlusion, truncation, multiple instances and clutter are the norm. Their approach outperforms Yang and Ramanan, the state-of-the-art technique, with an improvement from 29.0% to 37.5% PCP accuracy on the arm keypoint prediction task, on this new pose estimation dataset.

* **Interactive Facial Feature Localization,** Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Bourdev, and Thomas Huang**,** Computer Vision European Conference in Computer Vision (ECCV 2012)

* **Semantic Segmentation using Regions and Parts,** Pablo Arbeláez, Bharath Hariharan, Chunhui Gu, Saurabh Gupta, Lubomir Bourdev, and Jitendra Malik**,** Computer Vision IEEE Conference in Computer Vision and Pattern Recognition (CVPR 2012)

*** Facial Expression Editing in Video Using a Temporally-Smooth Factorization,** Fei Yang, Lubomir Bourdev, Eli Shechtman, Jue Wang and Dimitri Metaxas**,** Computer Vision IEEE Conference in Computer Vision and Pattern Recognition (CVPR 2012)

*** Face Morphing using 3D-Aware Appearance Optimization,** Fei Yang, Eli Shechtman, Jue Wang, Lubomir Bourdev, Dimitris Metaxas**,** Computer Graphics Graphics Interface (GI 2012)

*** Semantic Contours from Inverse Detectors,** Bharath Hariharan, Pablo Arbelaez, Lubomir Bourdev, Subhransu Maji and Jitendra Malik**,** Computer Vision International Conference in Computer Vision (ICCV 2011)

We study the challenging problem of localizing and classifying category-specific object contours in real world images. For this purpose, we present a simple yet effective method for combining generic object detectors with bottomup contours to identify object contours. We also provide a principled way of combining information from different part detectors and across categories. In order to study the problem and evaluate quantitatively our approach, we present a dataset of semantic exterior boundaries on more than 20, 000 object instances belonging to 20 categories, using the images from the VOC2011 PASCAL challenge

* **Poselets: Body Part Detectors Trained Using 3D Human Pose Annotations,** Lubomir Bourdev and Jitendra Malik**,** Computer Vision International Conference in Computer Vision (ICCV 2009)

* **Generic Image Library,** Lubomir Bourdev**,** Software Engineering Software Developer's Journal 2007 **–** *The Generic Image Library (GIL) is a C++ image library sponsored by Adobe Systems, Inc. and developed by Lubomir Bourdev and Hailin Jin. It is an open-source library, planned for inclusion in Boost 1.35.0. GIL is also a*

part of the Adobe Source Libraries. It is used in several Adobe projects, including some new features in Photoshop CS4

* **Robust Object Detection Via Soft Cascade,** Lubomir Bourdev and Jonathan Brandt**,** Computer Vision IEEE Conference in Computer Vision and Pattern Recognition (CVPR 2005) *A method for training object detectors using a generalization of the cascade architecture, which results in a detection rate and speed comparable to that of the best published detectors while allowing for easier training and a detector with fewer features. In addition, the method allows for quickly calibrating the detector for a target detection rate, false positive rate or speed. One important advantage of our method is that it enables systematic exploration of the ROC Surface, which characterizes the trade-off between accuracy and speed for a given classifier.*

* **Art-Based Rendering of Fur, Grass, and Trees,** Michael Kowalski, Lee Markosian, J.D. Northrup, Lubomir Bourdev, Ronen Barzel, Loring Holden and John Hughes**,** Computer Graphics ACM Transactions on Graphics (SIGGRAPH 1999) *Artists and illustrators can evoke the complexity of fur or vegetation with relatively few well-placed strokes. We present an algorithm that uses strokes to render 3D computer graphics scenes in a stylized manner suggesting the complexity of the scene without representing it explicitly. The basic algorithm is customizable to produce a range of effects including fur, grass and trees, as we demonstrate in this paper and accompanying video. The algorithm is implemented within a broader framework that supports procedural stroke-based textures on polyhedral models. It renders moderately complex scenes at multiple frames per second on current graphics workstations, and provides some interframe coherence.*

* **Rendering Nonphotorealistic Strokes with Temporal and Arc-Length Coherence,** Lubomir Bourdev**,** Computer Graphics Master's Thesis, Brown University, 1998 **–** *Rendering a silhouette of an object in a frame-to-frame coherent way. The input to the system each frame is a set of silhouette pixels in a rendering of the object and their corresponding silhouette edges in a polygonal model (mesh) of the object. The output is a set of silhouette strokes.*

* **Real-Time Nonphotorealistic Rendering,** Lee Markosian, Michael Kowalski, Sam Trychin, Lubomir Bourdev, Daniel Goldstein and John Hughes**,** Computer Graphics ACM Transactions on Graphics (SIGGRAPH 1997) *Nonphotorealistic rendering (NPR) can help make comprehensible but simple pictures of complicated objects by employing an economy of line. But current nonphotorealistic rendering is primarily a batch process. This paper presents a real-time nonphotorealistic renderer that deliberately trades accuracy and detail for speed. Our renderer uses a method for determining visible lines and surfaces which is a modification of Appel's hidden-line algorithm, with improvements which are based on the topology of singular maps of a surface into the plane. The method we describe for determining visibility has the*

potential to be used in any NPR system that requires a description of visible lines or surfaces in the scene. The major contribution of this paper is thus to describe a tool which can significantly improve the performance of these systems. We demonstrate the system with several nonphotorealistic rendering styles, all of which operate on complex models at interactive frame rates.
(…)

* **Георги Герганов** (Georgi Gerganov) – създател на Llama.cpp, Whisper.cpp и форматът GGUF, GGML за квантуване и ефективно изпълнене на ГЕМ и на процесори и други платформи и библиотеки за ускорение, а не само на най-обичайната все още CUDA/NVIDIA/GPU. https://github.com/ggerganov/llama.cpp Може би най-популярният български разработчик, С 15.7k последователи в Гитхъб, а проектът му има 69.6 хил. звезди и над 10 хил. разклонения към 22.12.2024. Към 19.4.2025: 78336/11450/1117 сътрудници с принос, 17.2к последователи.

В разработката на моделите Llama на Meta AI са участвали може би двама българи: https://huggingface.co/docs/transformers/en/model_doc/llama2

* Други статии по съвременно разделяне на изображения свързани с работи на Анелия Ангелова и Александър Тошев; Мира Дончева и др. – Current Image Segmentation #segmentation

Виж работата на XuDong Wang et al. Сравни със споменатите статии с участието на българи свързани с panoptic segmentation, open-dictionary segmentation, detection, shapes etc. https://people.eecs.berkeley.edu/~xdwang/

Visual Lexicon: Rich Image Features in Language Space, XuDong Wang, Xingyi Zhou, Alireza Fathi, Trevor Darrell, Cordelia Schmid, 2024

While image representations for computer vision usually were tailored either for: high-level semantics (CLIP) or for high-fidelity image reconstruction (VAE), ViLex captures both. **Understanding-focused high-level semantics**: CLIP, SigLIP, DINO – lose pixel-level details. **Reconstruction-focused**: autoencoders: VAE, MAE, BeiT – fine details, but lack semantic richness, struggle with discriminative tasks. **Dreambooth-like** (inject a new image example, face etc. into a pre-trained diffusion model), zero-shot unsupervised image re-contextualization; frozen text-to-image T2I diffusion model (frozen - their weights, parameters are fixed and they don't change during training). Prompting with ViLex tokens and text prompt tokens. Visual prompts (image, converted to embedding) + Text prompt: “[image] in Leonardo’s style” … OpenCLIP) DeDiffusion, Imagen as a base (text-to-image diffusion model), U-Net arch. 600M params, embedding dim. 256, input 64x64. Text encoder: OpenCLIP ViT-H/14, vocabulary size = 49408. Pooled embedding vector, added to the diffusion timestep embedding.

* **WebLI Dataset** - WebLI (Web Language Image) web-scale multilingual image-text dataset. (with PaLI: A Jointly-Scaled Multilingual Language-Image Model, Xi Chen,..., Anelia Angelova, ..., 2022) https://paperswithcode.com/dataset/webli - private, can't be downloaded; includes also alt-text and OCR; 109 languages, deduplicated on 68 vision/vision-language tasks https://paperswithcode.com/dataset/textcaps 145K captions for 28K images - read and reason about text in images to generate captions about them

SegLLM: Multi-round Reasoning Segmentation with Large Language Models XuDong Wang*, Shaolun Zhang*, Shufan Li*, Konstantinos Kallidromitis, Kehan Li, Yusuke Kato, Kazuki Kozuka, Trevor Darrell https://arxiv.org/pdf/2410.18923.pdf

Segment Anything without Supervision, XuDong Wang, Jingfeng Yang, Trevor Darrell https://arxiv.org/pdf/2406.20081.pdf (вж също SAM, SAM2, Meta AI)
https://github.com/frank-xwang/UnSAM
https://colab.research.google.com/drive/1aFObIt-xlQmCKk3G7dD8KQxaWhM_RTEd#scrollTo=QzkCAobcHsNi

InstanceDiffusion: Instance-level Control for Image Generation, XuDong Wang, Trevor Darrell, Saketh Rambhatla, Rohit Girdhar, Ishan Misra[8]
https://arxiv.org/pdf/2402.03290.pdf
Задаване на йерархични обхващащи правоъгълници, в които да се породи образ по различни подкани. По-фино управление на пораждането.

Rethinking Patch Dependence for Masked Autoencoders, Letian Fu, Long Lian, Renhao Wang, Baifeng Shi, XuDong Wang, Adam Yala, Trevor Darrell, Alexei A. Efros, Ken Goldberg https://arxiv.org/pdf/2401.14391.pdf

Unsupervised Universal Image Segmentation, Dantong Niu*†, XuDong Wang*†, Xinyang Han*, Long Lian, Roei Herzig, Trevor Darrell.
https://arxiv.org/pdf/2312.17243.pdf https://github.com/u2seg/U2Seg

See, Say, and Segment: Teaching LMMs to Overcome False Premises, Tsung-Han Wu*, Giscard Biamby*, David Chan, Lisa Dunlap, Ritwik Gupta, XuDong Wang, Joseph E. Gonzalez, Trevor Darrell. https://arxiv.org/pdf/2312.08366.pdf

VideoCutLER: Surprisingly Simple Unsupervised Video Instance Segmentation, XuDong Wang, Ishan Misra, Z.Zeng, R. Girdhar and T. Darrell**.**, CVPR 2024
https://people.eecs.berkeley.edu/~xdwang/projects/VideoCutLER/videocutler.pdf
Разпознаване на движещи се обекти и отделянето им с двоична маска. Pseudo masks, MaskCut, video synthesis … effectively segment and track multiple instances;

---

[8] Благодаря Валентина Лилова, която сподели за свое пресъздаване на проекта:
* valentina98, simonvhuesgen https://github.com/valentina98/DL2_InstanceDiffusion

no optical flow; unsupervised video instance segmentation (VIS), unsupervised video object segmentation (VOS); DINO; CRF

Hierarchical Open-vocabulary Universal Image Segmentation, XuDong Wang*, Shufan Li*, Konstantinos Kallidromitis*, Yusuke Kato, Kazuki Kozuka and Trevor Darrell. https://arxiv.org/pdf/2307.00764.pdf https://github.com/berkeley-hipie/HIPIE *"Various levels of granularities (whole, part and subpart) and tasks, including semantic segmentation, instance segmentation, panoptic segmentation, referring segmentation, and part/subpart segmentation, all within a unified framework of language-guided segmentation."* Cmp: SAM, Grounded-SAM.

Cut and Learn for Unsupervised Object Detection and Instance Segmentation, XuDong Wang, Rohit Girdhar, Stella X. Yu, Ishan Misra. https://arxiv.org/pdf/2301.11320.pdf

Unsupervised Selective Labeling for More Effective Semi-Supervised Learning XuDong Wang*, Long Lian*, Stella X. Yu. https://arxiv.org/pdf/2110.03006.pdf

Debiased Learning from Naturally Imbalanced Pseudo-Labels, XuDong Wang, Zhirong Wu, Long Lian, Stella X. Yu., 4.2022 https://arxiv.org/pdf/2201.01490.pdf *Pseudo-labels - predictions on unlabeled target data by a classifier trained on labeled source data; semi-supervised learning (SSL).* Вж псевдо-етикети също в NLP, маркиране с досега обучения модел и подбиране на примери с висока степен на увереност като допълителни обучителни примери (напр. тяхната околност в текста и пр.)

> * **ReMixMatch: Semi-supervised learning with distribution matching and augmentation anchoring**, David Berthelot, Nicholas Carlini, Ekin D Cubuk, Alex Kurakin, Kihyuk Sohn, Han Zhang, and Colin Raffel, 2020 https://arxiv.org/abs/1911.09785 consistency regularization, entropy minimization … *distribution alignment: of the model's aggregated class predictions to match the marginal distribution of ground-truth class labels.* Consistency regularization … augmentation anchoring … goal: *to learn from unlabeled data in a way that improves the performance on labeled data. …* Regularization With Stochastic Transformations and Perturbations … "П-Model" loss function for measuring consistency … MSE (mean-squared error), cross-entropy for pertrubed and non-pertrubed input. See *Data Augmentation*; e.g. the Python library *Albumentations* https://github.com/albumentations-team/albumentations
> See FixMatch, Mean Teacher;

Clipped Hyperbolic Classifiers Are Super-Hyperbolic Classifiers, Yunhui Guo, XuDong Wang, Yubei Chen, Stella X. Yu., CVPR 2022 https://arxiv.org/abs/2107.11472

Unsupervised Hierarchical Semantic Segmentation with Multiview Cosegmentation and Clustering Transformers, Tsung-Wei Ke, Jyh-Jing Hwang, Yunhui Guo, XuDong Wang, Stella X. Yu. https://arxiv.org/pdf/2204.11432.pdf Hierarchical Segment Grouping (HSG), visual similarities & stat.co-ocurences; grouping, co-segmentation among multiple views of the same image; semantic consistency, grouping hierarchy (person, arm, face, torso, … table, dining table, food; coarse/fine grouping)

[49] Bryan Russell, Alyosha Efros, J.Sivic, B.Freeman, and A. Zisserman**.** **Segmenting scenes by matching image composites**. In NIPS, 2009 https://papers.nips.cc/paper_files/paper/2009/file/fba9d88164f3e2d9109ee770223212a0-Paper.pdf … MRF, LabelMe dataset, K-means, Data driven boundary detection, data driven image grouping; the info in a single image is not sufficient; matching is not translation invariant, same regions between multiple images from a database ("stack"); not self-similarity, parts appearing together; Latent Dirichlet Allocation

* A. Oliva and A. Torralba. Modeling the shape of the scene: a holistic representation of the spatial envelope. IJCV, 42(3):145–175, 2001. "gist" descriptor

* B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. LabelMe: a database and web-based tool for image annotation. IJCV, 77(1-3):157–173, 2008

* A. Torralba, R. Fergus, and W. T. Freeman.
80 million tiny images: a large dataset for non-parametric object and scene recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 30(11):1958–1970, 2008 – 32×32, 75062 non-abstract nouns from WordNet, nearest neighbor; from: (the DB is not available for download anymore) https://groups.csail.mit.edu/vision/TinyImages/ (53K nouns, collected in 2006) Recognition: Dssd, Dwarp or Dshift. SSD – sum of the squared differences; shift: shifting pixels up to 5x5 for the least SSD; warp: small translations, scaling (up to 10 pixels shift) and image mirror; – optimized by gradient descent. … *A simple non-parametric methods, in conjunction with large datasets, can give reasonable performance on object recognition tasks. "The vast majority of the effort in recent years has gone into the modeling part – seeking to develop suitable parametric representations for recognition. In contrast, this paper moves into other direction, exploring how the data itself can help to solve the problem.*"

* Large image datasets: A pyrrhic win for computer vision? Anonymous submission, https://openreview.net/pdf?id=s-e2zaAlG3I 2020

* Generalized Belief Propagation, Jonathan S. Yedidia, William T. Freeman ,Yair Weiss, 2000 https://proceedings.neurips.cc/paper_files/paper/2000/file/61b1fb3f59e28c67f3925f3c79be81a1-Paper.pdf

* J. Pearl. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, San Francisco, 1988.

* **Interactive digital photomontage**, A. Agarwala, **Mira Dontcheva***[9], M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. In SIGGRAPH, 2004.
https://www.researchgate.net/publication/2941744_Interactive_Digital_Photomontage – a famous paper, coarsely drawing lines on several family photos, automatically composites parts seamlessly with the best selected faces etc. for selective composite, extended depth of field, relighting, stroboscopic effect etc graph-cut optimization; gradient-domain fusion, minimum likelihood image objective …

Известна статия, която съчетава изображения с минимални следи от снаждане като търси най-гладка връзка или образува най-плавен преход в пространство на градиентите.
[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence, 23(11), 2001
Poisson equations, Perez ´ et al. 2003; Fattal et al. 2002
* ROBINSON, H. P. 1869. Pictorial Effect in Photography: Being Hints on Composition and Chiaroscuro for Photographers. Piper & Carter
https://archive.org/details/pictorialeffecti00robi/mode/2up?view=theater
* FATTAL, R., LISCHINSKI, D., AND WERMAN, M. 2002.
Gradient domain high dynamic range compression. ACM Transactions on Graphics 21, 3, 249–256
https://www.researchgate.net/publication/2530899_Gradient_Domain_High_Dynamic_Range_Compression
… attenuating the magnitudes of the HDR image gradients by a factor of Φ(x, y) at each pixel … progressive, shrinking gradients of large magnitude more than small ones; Gaussian pyramide … propagating the scaling factor from each level to the next with linear interpolation

Unsupervised Visual Attention and Invariance for Reinforcement Learning, XuDong Wang*, Long Lian*, Stella X. Yu.
https://people.eecs.berkeley.edu/~xdwang/papers/CVPR2021_VAI.pdf

Unsupervised Feature Learning by Cross-Level Instance-Group Discrimination
XuDong Wang, Ziwei Liu, Stella X. Yu.
https://people.eecs.berkeley.edu/~xdwang/papers/CVPR2021_CLD.pdf

Long-tailed Recognition by Routing Diverse Distribution-Aware Experts.
XuDong Wang, Long Lian, Zhongqi Miao, Ziwei Liu, Stella X. Yu. ICLR 2021
https://people.eecs.berkeley.edu/~xdwang/papers/ICLR2021_RIDE.pdf

9

Volumetric Attention for 3D Medical Image Segmentation and Detection, XuDong Wang, Shizhong Han, Yunqiang Chen, Dashan Gao, Nuno Vasconcelos.**,** MICCAI 2019 https://people.eecs.berkeley.edu/~xdwang/papers/MICCAI2019_VA.pdf

Towards Universal Object Detection by Domain Attention, XuDong Wang, Zhaowei Cai, Dashan Gao, Nuno Vasconcelos.**,** CVPR 2019 https://openaccess.thecvf.com/content_CVPR_2019/papers/Wang_Towards_Universal_Object_Detection_by_Domain_Attention_CVPR_2019_paper.pdf

Feature Space Transfer for Data Augmentation, Bo Liu, XuDong Wang, Mandar Dixit, Roland Kwitt, Nuno Vasconcelos.**,** CVPR 2018 (Oral Presentation) http://www.svcl.ucsd.edu/people/xdwang/CVPR_2018.pdf

 # Mira Dontcheva: (probably Bulgarian?) Мира Дончева
* **Discovering Natural Language Commands in Multimodal Interfaces, Arjun Sriniva-san,** Mira Dontcheva, Eytan Adar, Seth Walker, 2019 https://www.cond.org/discover-nl-iui19.pdf

* **Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Ex-pressive Visualization Authoring**, Zh.Liu, … Mira Dontcheva,… et al.https://dl.acm.org/doi/pdf/10.1145/3173574.3173697

PixelTone: a multimodal interface for image editing, J.Linder et al.2013, https://dl.acm.org/doi/abs/10.1145/2468356.2479533

Summarizing Personal Web Browsing Sessions, Mira Dontcheva et al.,2006 https://www.researchgate.net/publication/220876962_Summarizing_personal_web_browsing_sessions

Layered Acting For Character Animation, Mira Dontcheva, Gary Yngve Zoran Popo-vic, 2003 https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=0a15bb26b01a495eace20eb4f219ecb3ee6e6bcf

* **Image segmentation with traveling waves in an exactly solvable recurrent neural network**, Luisa H. B. Liboni et al.
https://www.pnas.org/doi/10.1073/pnas.2321319121
https://phys.org/news/2025-01-mathematical-technique-black-ai-decision.html
3.1.2025 – особен подход без учител с рекурентни НМ. Кратък обзор на различни най-общи видове алгоритми за сегментиране.
Напр. G. E. Hinton, S. Sabour, N. Frosst, “Matrix capsules with EM routing” in *International Conference on Learning Representations* (2018) е наречен „slot-based”, слотово, с
* S Beucher, “Use of watersheds in contour detection” in *Proceedings of the International Workshop on Image Processing, Sept. 1979* (1979), pp. 17–21.

* M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models. *Int. J. Comput. Vis.* **1**, 321–331 (1988). "Active contour" – дейно очертание, активен контур https://en.wikipedia.org/wiki/Active_contour_model
https://link.springer.com/article/10.1007/BF00133570 вж препратки:
* D.J. Burr, "Elastic matching of line drawings", Ieee Trans. Pami p. 708. 1986.
* M.A., Fischler, and R.A., Elschlager, "The representation and matching of pictorial structure", Ieee Trans. on Computers, vol. C-22, pp. 67–92, 1973.
* Kanisza, "Subjective contours", Scientific American, vol. 234, pp. 48–52, 1976. Etc.
* Roberto C. Budzinski et al, **An exact mathematical description of computation with transient spatiotemporal dynamics in a complex-valued neural network**, *Communications Physics* (2024). DOI: 10.1038/s42005-024-01728-0 – невронни мрежи с комплексни числа, cv-NN (complex-valued NNs)

#Vision Tasks #Vision-Language Tasks #Зрителни задачи

Vision Tasks: Image Classification, Object Detection, Semantic Segmentation, Instance Segmentation, Panoptic Segmentation, Image Captioning, Image Generation, Object Tracking, Pose Estimation, Action Recognition, Face Recognition, Keypoint Detection, Optical Character Recognition (OCR), Depth Estimation, 3D Object Reconstruction, Super-Resolution, Image Inpainting, Image Denoising, Image Style Transfer, Image Colorization, Image Superpixel Segmentation, Visual Question Answering (VQA), Scene Understanding, Image-to-Image Translation, Visual Saliency Prediction, Texture Synthesis, Image-to-Text Generation, Visual Grounding, Landmark Detection, Scene Text Detection, Visual Semantic Parsing, Object Localization, Visual Odometry, Human-Object Interaction Recognition, Human Pose Estimation, Eye Gaze Detection, Visual Turing Test, Visual SLAM, Anomaly Detection in Images, Clothing Recognition, Food Recognition, Artistic Style Classification, Gesture Recognition, Animal Species Classification, Handwriting Recognition, Light Field Imaging, Image Compression, Data Augmentation for Vision Tasks, Disaster Image Assessment, Video Summarization, Video Captioning, Video Action Detection, Dynamic Scene Segmentation, Motion Tracking, Video Object Segmentation, Cross-Domain Image Classification, Video Colorization, Motion Estimation, Person Re-identification, Vehicle Detection and Tracking, Medical Image Segmentation, Medical Image Classification, 3D Pose Estimation, Change Detection in Satellite Images

Vision-Language Tasks: Visual Question Answering (VQA), Image Captioning, Text-to-Image Generation, Image-Text Matching, Cross-modal Retrieval, Visual Commonsense Reasoning, Visual Entailment, Image-Text Alignment, Text-Image Synthesis, Visual Reasoning, Multimodal Sentiment Analysis, Text-based Object Detection, Image-based Text Generation, Multimodal Machine Translation, Scene Graph Generation, Visual Dialogue, Zero-Shot Visual Classification, Cross-modal Retrieval from Visuals, Visual Question Answering with Explanation, Multimodal Summarization, Speech to Image Generation, Image Captioning with Localization, Image Retrieval with Natural Language Queries, Textual Entailment with Visual Context, Multimodal Sentiment Classification, Image Description Generation from Text Queries, Visual Concept Recognition, Visual Storytelling, Visual

Grounding in Text, Multimodal Event Detection, Image-Text Cross-Modal Reasoning, Interpretable Visual Question Answering, Fine-Grained Visual-Textual Matching, Image Captioning with Object Detection, Emotion Recognition from Text and Images, Scene Understanding with Textual Input

Термини и юнашко наречие: визуален – зрителен, образен; генериране – пораждане; сегментиране – разделяне, реконструиране – възстановяване; 3D – триизмерно; локализация – уместяване; модалност – сетивност, вид данни; мултимодален – многосетивен; всички модалности, вкл. двигателни: всетводействие, всетводействен; текст – слово; контекст - обслов

Зрителни задачи с образи и видео: Класификация на изображения, Откриване на обекти, Семантично сегментиране, Сегментиране на инстанции, Паноптично сегментиране, Пораждане на надписи за изображения, Пораждане на изображения, Проследяване на обекти, Оценка на позата, Разпознаване на действия, Разпознаване на лица, Откриване на ключови точки, Оптично разпознаване на символи (OCR), Оценка на дълбочина, 3D-възстановяване на обекти, Супер-резолюция, Възстановяване на изображения, Шумопочистване на изображения, Пренос на стил в изображения, Оцветяване на изображения, Сегментиране на суперпиксели на изображения, Визуален отговор на въпроси (VQA), Разбиране на сцени, Превод от изображение в изображение, Прогнозиране на зрителна важност, Синтез на текстури, Пораждане на текст от изображения, Зрително търсене на предмети, Откриване на забележителности, Откриване на текст в сцена, Зрителен семантичен разбор , Локализация на обекти (Уместяване на предмети), Зрителна одометрия, Разпознаване на взаимодействие между хора и предмети, Оценка на позата на човек (оценка на положението), Откриване на посоката на погледа, Зрителен тест на Тюринг, Зрителен SLAM (зрително едновременно уместяване и рисуване на карта), Откриване на необичайност в изображения, Разпознаване на дрехи, Разпознаване на храни, Класификация на художествени стилове, Разпознаване на жестове, Класификация на животински видове, Разпознаване на ръкописен текст, Снимки на светлинни полета, Снимане на изображения, Разширяване на данни за задачи с изображения, Оценка на изображения при бедствия, Резюмираност на видео, Генериране на надписи за видео (пораждане на надписи за видео), Откриване на действия във видео, Сегментиране на динамични сцени, Проследяване на движение, Сегментиране на обекти във видео, Прехвърляне на цвят на видео, Оценка на движение, Смяна на лица, Откриване и проследяване на превозни средства, Сегментиране на медицински изображения (разделяне на медицински изображения), Класификация на медицински изображения, Триизмерна оценка на позата, Откриване на промени в сателитни изображения.

Зрителни задачи с образи и текст, слово: Зрителен отговор на въпроси (VQA), Генериране на надписи за изображения, Пораждане на изображения от текст, Съответствие между изображения и текст, Прехвърляне между модалности (пренос, многосетивен пренос), Разбиране на общи образни смисли, Зрително заключение, Приравняване на съдържанието между изображения и текст, Пораждане на текст от изображения, Зрително разсъждение, Мултимодален анализ на емоции, Откриване на обекти чрез текст, Генериране на текст от изображения, Мултимодален машинен превод, Пораждане на графи на сцени (отношения между части, вложеност и др.), Зрителен диалог, Зрителна класификация без примери и подготовка (Zero-shot), Междусетивен пренос от изображения, Зрителен въпрос отговор с обяснение, Многосетивно обобщение, Пораждане на изображения от текстови въпроси, Извличане на изображения с текстови запитвания (словесни запитвания), Словесен извод в обслов на изображения, Мултимодално класифициране на емоции (многопредметно определяне на чувства), Генериране на описание на изображения от текстови запитвания, Разпознаване на образни понятия, Образно разказване на истории, Образно обосноваване в текст, Мултимодално откриване на събития, Кръстосано-разбиране на изображения и текст, Отговаряне на въпроси с тълкуване, Точно съвпадение между слово и изображения, Пораждане на надписи за изображения с уместяване на предмети, Разпознаване на изразени чувства от текст и изображения, Разбиране на сцени с текстов вход. …

Вж също разделите за роботика, основни модели за роботи, бележките за български роботисти в началото на книгата и др.

* Мултимодални модели #multimodal #мултимодални

Recognizing Everything from All Modalities at Once: Grounded Multimodal Universal Information Extraction, Meishan Zhang1, Hao Fei 2,*, Bin Wang1, Shengqiong Wu2, Yixin Cao3, Fei Li4, Min Zhang1, 1Harbin Institute of Technology (Shenzhen), 2National University of Singapore, 3Fudan University 4Wuhan University, 2024
https://arxiv.org/abs/2406.03701

Video: Hao Fei 284 абонати 108 показвания 18.07.2024 г.
https://www.youtube.com/watch?v=mg9ItO6s9V4
https://www.youtube.com/@haofei4763/videos Hao Fei @haofei4763
https://haofei.vip/MUIE https://github.com/scofield7419/MUIE-REAMO

Text+Video... Text+Image+Audio, ... NER,
Image Encoder for Image, Image encoder for Video, Audio Encoder } Image_projection1, image_projection 2, audio_projection > LLM --> <UIE> <Radev, person) (Trump, person) (Bulgaria, country)(...) <Module>Image Segmenter <Instruction>The man in the left, The man ... --> Image segmenter, Video Tracker, Audio Segmenter --> Image segments, Video Tracklets, Audio Segments ... //MUIE Decoding with grounding ...
Multimodal Universal Information Extraction (MUIE)
Multimodal Encoding, LLM Reasoner 4 min ... "Please extract all entities ..."
REAMO: Audio segmenterr, Image segmenter, Video tracker, ...

*** Improving Multimodal Interactive Agents with Reinforcement Learning from Human Feedback** Josh Abramson, Arun Ahuja, Federico Carnevale, *Petko Georgiev,* Alex Goldin, Alden Hung, Jessica Landon, Jirka Lhotka, Timothy Lillicrap, Alistair Muldal, George Powell, Adam Santoro, Guy Scully, Sanjana Srivastava, Tamara von Glehn, Greg Wayne, Nathaniel Wong, Chen Yan, Rui Zhu**,** 21.11.2022 https://arxiv.org/abs/2211.11602

RLHF, Inter-temporal Bradley-Terry" (IBT) modelling to capture human judgments ... humans interacting with agents in a simulated 3D world ; behavioral cloning - BC; setter-replay ... p.7 Fig.3: Positive/Negative human annotation, IBT reward, Learned reward, model reward

Hao Fei et al.: Multimodal systems

Video-of-Thought: Step-by-Step Video Reasoning from Perception to Cognition, Hao Fei[1], Shengqiong Wu[1], Wei Ji[1], Hanwang Zhang[2], Meishan Zhang[3], Mong-Li Lee[1], Wynne Hsu[1] , ICML, 2024 https://haofei.vip/VoT/ - Video understanding, MotionEpic, Video-of-thought - https://www.youtube.com/watch?v=2fKCWjetV-Y

***Any2Caption : Interpreting Any Condition to Caption for Controllable Video Generation** Shengqiong Wu1,2*   Weicai Ye1, et al., 31.3.2025 1Kuaishou Technology 2National University of Singapore https://sqwu.top/Any2Cap/ https://arxiv.org/abs/2503.24379
Text-only prompts can't capture the user intent. Diverse Input Conditions: Text, **Depth map sequences, Sketch, Camera Pose, First Frame, Multiple Identities, Human Pose Sequences, Segmentation, Normals (texture), Style, Key video frames …** *Any2CapIns, a large-scale dataset with 337K instances and 407K conditions Dense caption; Structured Dense Captions of Any Cond; Captions: Main object, Background, Camera, Style, Action (separate);: Off-the-shelf video generators: CTRL-Adapter, VideoComposer, ControlVideo, CameraCtrl, ConceptMaster, MotionCtrl, HunYuan, CogVideoX.* ***Data Collection.:*** *Spatialwise conditions depth maps, sketches, and video frames.* *Action-wise conditions, human pose, motion.* *Composition-wise conditions* **…** *an image encoder FI, a video encoder FV , a motion encoder FM and a camera encoder FC to process non-text conditions.*

*** On Path to Multimodal Generalist: General-Level and General-Bench,** Hao Fei*1 Yuan Zhou*2 et al., 7.5.2025 https://arxiv.org/pdf/2505.04620 (>30 authors, 305 p.) * https://www.youtube.com/watch?v=ccd2PSiA57I   * https://generalist.top/
See in the beginning of the main volume of "The prophets…" @Vsy: try the tests.

*** Petko Georgiev** – probably a Bulgarian – Петко Георгиев, @Deepmind
https://scholar.google.com/citations?hl=en&user=ksq9614AAAAJ&view_op=list_works&sortby=pubdate

*** Grandmaster level in StarCraft II using multi-agent reinforcement learning,** Oriol Vinyals, …, Petko Georgiev, … et al., 2019
https://www.seas.upenn.edu/~cis520/papers/RL_for_starcraft.pdf

* Gemma 2: Improving open language models at a practical size, Gemma Team, Morgane Riviere, …, Petko Georgiev, … et al, 7.2024

* **Interactive Agents Team. Creating multimodal interactive agents with imitation and self-supervised learning.** arXiv preprint arXiv:2112.03763, 2021b.

Improving Multimodal Interactive Agents with Reinforcement Learning from Human Feedback, Federico Carnevale, 62 абонати, 64 likes, 5402 показвания 9.11.2022 г
https://www.youtube.com/watch?v=v_Z9F2_eKk4&feature=youtu.be

Симулиран свят, с лъч хваща предмети, мести, изпълнява поръчения.

Uni-MoE: Scaling Unified Multimodal LLMs with Mixture of Experts, Yunxin Li, Shenyuan Jiang, Baotian Hu, Longyue Wang, Wanqi Zhong, Wenhan Luo, Lin Ma, Min Zhang, 18.5.2024 https://arxiv.org/abs/2405.11273 https://www.marktechpost.com/2024/05/25/uni-moe-a-unified-multimodal-llm-based-on-sparse-moe-architecture/

Extends the Mixture of Experts MoE ... sparse MoE architecture; "a progressive training strategy: 1) Cross-modality alignment using various connectors with different cross-modality data 2) Training modality-specific experts with cross-modality instruction data to activate experts' preferences, and 3) Tuning the Uni-MoE framework utilizing Low-Rank Adaptation (LoRA) on mixed multimodal instruction data." Router -> Expert1, Expert2, ... --> Gating weights --> Output
@Вси: {K-K}, {#:} Mixtral-MoE 8x7B ... MoE-LLaVA: 3B activated, cmp to 7B dense.;
1. Cross-modality alignment 2. Training Modality-specific experts 3. Tuning Uni-MoE (LLM, self-attention, Router-->Experts...Gating weights --> x --> Textual output)
LoRA technique - Low-Rank Adaptation on mixed multimodal instruction data;
trainable: Self-attentio in 3, Tuning Uni-MoE, Router, Expert, ... Meta-Transformer ... 3.2 *Connectors*: LLaVA, CLIP; Whisper encoder from Whisper-small; BEATs encoder - bidirectional encoder represm. from the audio transformers. Q-former ... distil fixed-length speech and audio feature vectors and map them into soft audio and speech tokens via a linear projection layer. The specific workflow is given in:
XInput = [I, V, A, S, T], (1) I = MLP(CLIP-V(I)),
(2) V = Mean(CLIP-V([I1, ..., I8])) (3) A = Audio-Qformer(BEATs(A))
(4) S = Speech-Qformer(Whisper(S)), (5) T = Word-Embedding(T), (6)
[I, V, A, S, T] represents the image, video, audio, speech, ... MLP - learnable projection layer; Audio-QFormer ... LibriSpeech, RACE dataset, LLaVA-Instruct-150K (Image, Text question & answer)

https://huggingface.co/datasets/liuhaotian/LLaVA-Instruct-150K?row=1

RACE (ReAding Comprehension dataset from Examinations) - 27,933 passages and 97,867 questions from English exams, targeting Chinese students aged 12-18. RACE-M and RACE-H, from middle school and high school exams: 4 candidate answers, 1 correct.
* RACE: Large-scale ReAding Comprehension Dataset From Examinations, Guokun Lai et al., 12.2017 https://arxiv.org/abs/1704.04683

See pseudocode: p.5

Datasets: Review datasets and search: @Вси
https://app.labelbox.com/catalog/dataset/cl3jm1jjgd1d5084ig6nnf13m?public=true&_r=https%3A%2F%2Fwww.google.com%2F%3F +Kaggle**,** Huggingface
OpenImages V6 Google, 8,2M**;** PASCAL-C , VRD - Visual Relationship Dataset (2016), imSitu3 - situation recognition**,** ACE2005 - IE, inf.extr., annotated news in English with NE, relations, events**,** ReTACRED: relation detection, 91K sentences, 40 relations**;** VidSitu - 10 sec videos from movies for complex situations (collection of related events); annotated, 2

sec., verbs, semantic roles, entity co-references, event relations, Twt17 - Twitter, NER, 723 test tweets, annotated, 4 entity types: person, location, organization, miscellaneous; MNRE - Multimodal Neural Relation Extraction; M2E2 - 245 multimedia news articles, annotated with events & arguments. ... Pre-processing ... VidSitu-Aud: captioning; VidSitu-Txt -> TTS --> Bark, EdgeTTS; ACE-Aud, Twt17-Aud -> TTS records. Modality-aligned content.
15 combinations of modalities & tasks. Multimodal grounding ... linear layer 4096; : InstructBLIP, LLaVA, MiniGPTv2.

"The advent of the Transformer model, introduced by Vaswani et al. [23], marked a significant milestone in the field of deep learning, enabling the scalable integration of multiple modalities—including image, language, speech, audio, and video—into a unified representational space." ImageBind by Girdhar et al. [29] modality composition through arithmetic operations, crossmodal detection, and generation, directly ”out-of-the-box.”**,** MetaAI SAM (Segment Anything Model)...

* ***Meta-Transformer*** - unified data tokenizer, shared encoder across modalities, and specialized heads for task-specific applications ... 12 unpaired modalities

> * **Lora: Low-rank adaptation of large language models,**E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, ICLR, 2022.
> * A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo et al., **“Segment anything”** ICCV, 2023. https://arxiv.org/pdf/2304.02643 (SAM)
>
> **SAM2 for Image and Video Segmentation: A Comprehensive Survey**, Zhang Jiaxing, Tang Hao 17.3.2025 https://arxiv.org/pdf/2503.12781
> * Dynamic Scene Reconstruction: Recent Advance in Real-time Rendering and Streaming. Jiaxuan Zhu, Hao Tang, 11.3.2025 https://arxiv.org/pdf/2503.08166
>
> * Y. Zhang, K. Gong, K. Zhang, H. Li, Y. Qiao, W. Ouyang, and X. Yue, **“Meta-transformer: A unified framework for multimodal learning,”** arXiv preprint arXiv:2307.10802, 7.2023 https://arxiv.org/abs/2307.10802 https://github.com/invictus717/MetaTransformer

Up to 12 modalities with the same parameters: "natural language , RGB images , point clouds , audios , videos , tabular data, graph , time series data , hyper-spectral images , IMU , medical images , and infrared images"

Related work: ... MLP (multi-layer perceptron): SVM, MLP ... then Recurrent and Convolutional NN: Hopfield network, LSTM, GRU - NLP * audio; CNN: LeNet, AlexNet, VGG GoogleNet, ResNet - image recognition, point cloud, speech classification → transformer

***Meta-Transformer* evolves into:**

OneLLM: One Framework to Align All Modalities with Language, Jiaming Han, Kaixiong Gong, Yiyuan Zhang, Jiaqi Wang, Kaipeng Zhang, Dahua Lin, Yu Qiao, Peng Gao, Xiangyu Yue, 12.2023/9.1.2025, https://arxiv.org/abs/2312.03700
https://onellm.csuhan.com/ https://github.com/csuhan/OneLLM
MLLM; a unified multimodal encoder and a progressive multimodal alignment pipeline.
1.Image projection module to connect a vision encoder with LLM.
2. Universal projection module (UPM) by mixing multiple image projection modules and dynamic routing. **3.** Progressively aligning more modalities to LLM with the UPM (Progressive Multimodal Alignment) multimodal instruction dataset, 2M items: 2M items from image, audio, video, point cloud, depth/normal map, IMU and fMRI for captioning, question answering, and reasoning tasks … universal X-to-language interface; dynamic router to control the weight of each expert for the given inputs, which turns UPM into soft mixtures-of-experts

UPM: several projection experts and modality routers to align the input signal with the language. ... learnable tokens for each modality (i.e., modality tokens), which are then used to aggregate input information and generate fixed-length tokens for all modalities. … *First: image-text pretraining:* better balances diff. mod. cmp. if directly aligning all modalities with text using a random init. model (UPM).
Multimodal-text alignment: stage 1: image; 2. video, audio and point cloud, 3. depth/normal map, IMU and fMRI. In order to support new modalities: repeat the training episode: sample a similar amount of data from previous modalities and jointly training with the current modalities. *"One universal encoder and projection module can effectively map multimodal inputs to LLM"*, modality-specific tokenizers

Multimodal-Text Datasets: LAION-400M, LAION-COCO; WebVid-2.5M, WavCaps, Cap3D; generate depth/normal map with a DPT model over CC3M dataset. ... no system prompts Prompt Design ... *Instruction tuning:* {q, Sys, [Inst, Anst]}
Architecture: CLIP VIT Large on LAION, LLaMA2-7B ... UPM K=3 projection experts, each 8 transf.blocks & 88M params; 30K tokens for each modality. First stage training: 16 x A100GPUs x 200K it, batch 5120 ... Stage 2,3: 8 x A100 x 200K it., x 100K it. Instruction tuning: 8xA100 x 96K it. batch 512

After the multimodal-text alignment OneMML can do multimodal captioning, can generate short description on any input from any modality.

Instruction Tuning datasets: : LLaVA-150K [49], COCO; Caption [14], VQAv2 [26], GQA [34], OKVQA [55], AOKVQA [71], OCRVQA [58], RefCOCO [36] and Visual Genome [38]. The video IT datasets include MSRVTTCap [91], MSRVTT-QA, AudioCap, 70K point cloud description, conversation and reasoning dataset ...
Routers: constant router, sparse router and the default soft router

* **Hierarchical vision transformer using shifted windows**, Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin

transformer. In ICCV, pages 10012–10022, 2021 https://github.com/Prof-Lu-Cewu/Visual-Relationship-Detection
* **Perceiver: General perception with iterative attention.** Andrew Jaegle, Felix Gimeno, Andy Brock, Oriol Vinyals, Andrew Zisserman, and Joao Carreira. pages 4651–4664. PMLR, 2021

ChatBridge [104] and AnyMAL [59], PandaGPT [77] and ImageBind-LLM [31] - *ImageBind* [23]

* **MiniGPT-v2:Large Language Model as a Unified Interface for Vision-Language Multi-task Learning,** Jun Chen, Deyao Zhu , Xiaoqian Shen, Xiang Li, Zechun Liu, Pengchuan Zhang, Raghuraman Krishnamoorthi, Vikas Chandra, Yunyang Xiong†, Mohamed Elhoseiny†, 11.2023 https://arxiv.org/abs/2310.09478
https://minigpt-v2.github.io/ visual backbone, linear projection layer, LLM**,** image patches --> ViT --> concat --> Linear: concat([INST]<img>...</img>[refer]where is the right eye?[/INST] --> Llama 2 7B--> <50><34><65><40> ... [identify] this {<35><45><65><70>} is ... a black chainring [grounding] pleae describe this image as detailed as possible**:** “Cut slice of fruit cake on a plate with a fork and a cup of coffee with flowers in a vase “**, “**Who are the people on the right, on the left and in the middle? ...” [detection] ... ViT 448x448, frozen; projection all the image tokens would result in 1024 tokens, for that reason: concatenating 4 adjacent visual --> 256 tokens.

RefCOCO/RefCOCO+/RefCOCOg: [refer] give me the location of question

* **MINIGPT-4: Enhancing Vision-language Understanding with Advanced Large Language Models**, Deyao Zhu* , Jun Chen*, Xiaoqian Shen, Xiang Li, Mohamed Elhoseiny, 10.2023 https://github.com/Vision-CAIR/MiniGPT-4
https://minigpt-4.github.io/ https://arxiv.org/pdf/2304.10592

* **Vision Transformer with Quadrangle Attention,** Qiming Zhang, Jing Zhang, Yufei Xu, Dacheng Tao, 27.3.2023 https://arxiv.org/abs/2303.15105
Quadrangle attention (QA) extends window-based attention -> *QFormer* => classification, object detection, semantic segmentation, and pose estimation. Enlarging windows: 7x7 to 32x32 ... Others: Focal attention - coarse granularity tokens to capture longrange context, cross-shaped window attention [5] two cross rectangular windows to model long-range dependency from both vertical and horizontal directions; Pale [12] - attends to tokens in dilated vertical/horizontal directions to model long-range dependency from diagonal directions.;
QFormer: learn the shape, size, orientation; rotates, translates, shears the shapre of the attention windows

*** VisualGPT,** Jun Chen, Han Guo, Kai Yi, Boyang Li, and Mohamed Elhoseiny. Visualgpt: Data-efficient adaptation of pretrained language models for image captioning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 18030–18040, 2022

https://huggingface.co/tasks/image-text-to-text https://moondream.ai/playground
https://huggingface.co/vikhyatk/moondream2 https://github.com/vikhyat/moondream

* **An image is worth 16x16 words: Transformers for image recognition at scale," in International Conference on Learning Representations**, A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, 2021.
(ViT)

*** Reka Core, Flash, and Edge: A Series of Powerful Multimodal Language Models**, Reka Team: Aitor Ormazabal, Che Zheng, Cyprien de Masson d'Autume, Dani Yogatama, Deyu Fu, Donovan Ong, Eric Chen, Eugenie Lamprecht, Hai Pham, Isaac Ong, Kaloyan Aleksiev, Lei Li, Matthew Henderson, Max Bain, Mikel Artetxe, Nishant Relan, Piotr Padlewski, Qi Liu, Ren Chen, Samuel Phua, Yazheng Yang, Yi Tay, Yuqi Wang, Zhongkai Zhu, Zhihui Xie
https://arxiv.org/abs/2404.12387v1

Multimodal language models at different sizes, trained from scratch; pre-training: ~ 4.5-5 trillion filtered and deduplicated tokens, 25% code-related, 30% STEM, 25% - web crawl, 10% - some math relation. Priority: unique tokens. Sizes: Reka Edge 7B, Reka Flash 21B, Reka Core 67B. Training: mostly Nvidia H100, Pytorch, peak 2.5K x H100 & 2.5K x A100. Reka Flash, Edge: several hundreds H100s for several weeks.

Instruction tuning with strong regularization: 1. SFT - supervised fine tuning. 2. RLHF, PPO (Proximal policy optimization, Schulman et al., 2017) - the same family of Reka models provide the reward. Tool-use, function calling, web search. Annotation pipelines: user interface .. images, videos, text-only ... multi-turn dialogs. https://www.reka.ai/ourmodels
Reka Spark: 2B, also fully multimodal, multilingual 32 langages, 128K context ... + (Wei et al., 2021; Ouyang et al., 2022; Chung et al., 2024)
See also: * EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Mingxing Tan 1 Quoc V. Le 1, arXiv:1905.11946v5 [cs.LG] 11 Sep 2020
https://arxiv.org/pdf/1905.11946

* **GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks,** Zemin Liu, Xingtong Yu, Yuan Fang, Xinming Zhang, 2.2023
https://arxiv.org/pdf/2302.08043

* **Prompt-engineering:** https://www.promptingguide.ai/techniques/zeroshot etc. Zero-shot Prompting, Few-shot Prompting, Chain-of-Thought Prompting, Meta Prompting, Self-Consistency, Generate Knowledge Prompting, Prompt Chaining, Tree of Thoughts, Retrieval Augmented Generation (RAG), Automatic Reasoning and Tool-use, Automatic Prompt Engineer, Active-Prompt,Directional Stimulus Prompting, Program-Aided, Language Models, ReAct, Reflexion, Multimodal CoT, Graph Prompting https://github.com/dair-ai/Prompt-Engineering-Guide/blob/main/notebooks/pe-lecture.ipynb

*** SELF-CONSISTENCY IMPROVES CHAIN OF THOUGHT REASONING IN LANGUAGE MODELS** Xuezhi Wang et al. 2023 https://arxiv.org/pdf/2203.11171
Sampling a diverse set of reasoning paths instead taking the greedy one etc., then selects the most consistent answer … “marginalizing” – averaging, choosing answers which match in different paths, lead to the same goal; between open-ended text generation and optimal text generation with a fixed best answer; other approaches: reranking, verifiers; human annotations …

Компютърна лингвистика и български дейци и изследвания

#компютърна лингвистика #cl #linguistics

По-ранни изследователи от 1980-те, 1990-те, …: Галя Ангелова, Георги Тотков, Руслан Митков, Христо Крушков, Кирил Симов … Светла Коева … RANLP

* **Галя Ангелова** и др. https://www.iict.bas.bg/news/2018/GAngelova/trudove.pdf

On An Approach For Designing Linguistic Processors., January 1982, DOI:10.3115/990100.990153 , Conference: Proceedings of the 9th conference on Computational linguistics - Volume 2, Radoslav Pavlov, Galia Angelova https://aclanthology.org/C82-2054.pdf

Система за диалог на естествен език, в лабораторията по математическа лингвистика в Института по математика и информатика на БАН. Формално описание на синтаксиса на основните структури на изречения на естествен език … За достъп на потребители, които не са програмисти, до релационна база данни на български език и до пакета от програми за статистически анализ BMDP на български. Превеждане на инструкциите до съвместими команди на програма и връщане на резултата.

SPRINT - A SYSTEM FOR INFORMATION RETRIEVAL IN LAW, Chapter, Dec 1981, R. Pavlov, V. Kiskinov, Galia Angelova

* **On Experimental Linguistic Processors for Man-Computer Dialogue in Bulgarian**, Jan 1984, Radoslav Pavlov, Galia Angelova, Elena Paskaleva

* **Voltron: A Hybrid System For Answer Validation Based On Lexical And Distance Features**, Ivan Zamanov, Nelly Hateva, Marina Kraeva, Ivana Yovcheva,Ivelina Nikolova, Galia Angelova, 2015

https://www.researchgate.net/publication/301446250_Voltron_A_Hybrid_System_For_Answer_Validation_Based_On_Lexical_And_Distance_Features

* [SAP, 1990] Simov K., Angelova G., Paskaleva E. MORPHO-ASSISTANT: The Proper Treatment of Morphological Knowledge. In: COLING'90, Helsinki, Finland, Vol. 3, pp. 453 - 457, 1990 ..

@Все: продължи, извлечи, подреди# зрдш: RANLP, CLIB; изброените български учени, имена, фамилия; връзки, вериги.

bgGLUE: A Bulgarian General Language Understanding Evaluation Benchmark, Momchil Hardalov, Pepa Atanasova, Todor Mihaylov, Galia Angelova, Kiril Simov, Petya Osenova, Ves Stoyanov, Ivan Koychev, Preslav Nakov, Dragomir Radev, 2023 https://arxiv.org/abs/2306.02349

Приспособяване, преработка на множество тестове за български език. Сбит преглед и на съответните тестове и задачи като NER, sentiment analysis, "Fake News", и др. Бих отбелязал: High School Examinations (EXAMS) EXAMS (Hardalov et al., 2019, 2020) – над 24 хил. въпроси по 24 училищни предмета. Изпитват мярката чрез обучение на BERT, mBERT, XLM-R и др. Споменават други многоезични модели, които включват български от тогава: XLM-RoBERTa (Goyal et al., 2021), multilingual T5 (Xue et al., 2021), XGLM (Lin et al., 2022), mGPT (Shliazhko et al., 2022), и разширената версия на BLOOM (Yong et al., 2022). Обучавали са модели в рамките на 118-560M параметъра (MiniLM-XLM-RLARGE)

* Пропуснали са разбираемо неизвестният им и непубликуван на „престижни конференции“ български модел GPT2-Medium, един от най-големите неанглийски до 2021 г.. обучен от мен, Тодор Арнаудов през лятото на 2021 г. виж на още няколко места в това приложение за повече информация, таблица в края * https://github.com/Twenkid/GPT2-Bulgarian-Training-Tips-and-Tools T.Arnaudov, 2021

* Работата на Руслан Митков и изследвания от края на 1980-те и 1990-те на учени от ПУ Паисий Хилендарски – разрешаване на анафора, морфологичен анализ и други видове разбор и моделиране на българския език, лексикология и лексикография и др.

* **Руслан Митков** (Ruslan Mitkov) – водещ учен в разрешаването на анафора и препращане (anaphora resolution & coreference), свързана с машинния превод, разбиране на езика, дискурсен анализ; разработката на езикови ресурси/ лексикография, машинно-подпомогнат превод (преводна памет, translation memory), корпусна лингвистика, автоматично създаване на тестове с множество отговори от корпуси (multiple-choice tests); Р.М. е организатор на работилниците и конференцията RANLP в България (Recent Advances in Natural Language Processing); мн. години беше директор на институт и изследователска група по компютърна лингвистика в университета „Улвърхамптън“, която през 2007 г. според изследване, за което научихме, беше сред първите пет групи в Обединеното кралство. Редактор на сборника:
* The Oxford Handbook of Computational Linguistics - Ruslan Mitkov, Steven Bird, ..., Dragomir Radev (65 authors) (last edition 2022)

https://www.researchgate.net/publication/365789276_The_Oxford_Handbook_of_Computational_Linguistics

Автор на главата за Извличане на информация е и Д.Радев.
https://lml.bas.bg/ranlp2007/ https://ranlp.org/ranlp2021/workshops.php

* **A knowledge-based and sublanguage-oriented approach for anaphora resolution**. Proceedings of the Pacific Asia, Conference on Formal and Computational, Linguistics, Taipei, 1993

* An integrated model for anaphora resolution, R.Mitkov, 1994 , Institute of Mathematics, Sofia, Bulgaria https://aclanthology.org/C94-2191.pdf – Prolog; *"tracking focus or center of the first sentence in the segment [Brennan et al. 1987]; thematic roles (Signer 81), subject; studied 1000 pages from 30 sources of computer science texts in order to develop a sublanguage-dependent heuristics for tracking the center … Verbs: {discuss, present, illustrate, summarize, examine, describe, define, show, check,develop, review, report, outline, consider, investigate, explore, assess, analyze, synthesize, study,survey, deal, cover} → the object is the center …* [10] Application of the Bayes' theorem; integrated, because it uses discourse, syntactic & semantic knowledge (discourse: from the selected domain of CS texts).

* **A TUTORING SYSTEM WHICH EXPLAINS IN NATURAL LANGUAGE**, Ruslan Mitkov, Berlin, GDR,1990 – in Bulgarian, concepts of planimetry in elementary geometry.

* Factors in anaphora resolution: they are not the only things that matter. A case study based on two different approaches, R. Mitkov 1997 126 https://aclanthology.org/W97-1303.pdf
* Robust pronoun resolution with limited knowledge, R.Mitkov, 1998 418
* A new, fully automatic version of Mitkov's knowledge-poor pronoun resolution method, R Mitkov, R Evans, C Orasan, 2002 181
* Anaphora resolution: the state of the art, R Mitkov,1999 315
* Computer-aided generation of multiple-choice tests, R Mitkov, 2003 348* * Coreference and anaphora: developing annotating tools, annotated resources and annotation strategies, R Mitkov, R Evans, C Orasan, C Barbu, L Jones, V Sotirova, 2000
* Generating multiple-choice test items from medical text: A pilot study, N Karamanis, R Mitkov, 2006
* The latest in anaphora resolution: going multilingual., R Milkov, 1998
* Anaphora resolution in natural language processing and machine translationR Mitkov, Institut der Gesellschaft zur Förderung der Angewandten

[10] TA: Grounding: How the rules are derived from a previous state – in order to automate it. Incrementally construct the rules from simpler texts.

Informationsforschung, 1995
* How far are we from (semi-) automatic annotation of anaphoric links in corpora?, R Mitkov, 1997
* Knowledge-based automatic abstracting: Experiments in the sublanguage of elementary geometry, R Mitkov, D Le Roux, JP Descles

* How to ask a foreigner questions without knowing his language? Proposal for a conceptual interface to communicate thought, M Zock, R Mitkov, 1991 https://hal.science/hal-03175829/document

* A breakthrough in automatic abstracting: the corpus-based approach, R Mitkov, University of Wolverhampton, 1995

* Machine Translation, Ten Years On: Discourse has yet to make a breakthrough, R Mitkov, J Haller, 1994 https://aclanthology.org/1994.bcs-1.8.pdf

Only sentence level, discourse "has yet to make a breakthrough"; mostly sentence translation, "discourse analysis being a very complicated task"); typed feature structures, SAT2, PATR-II Eurotra. Tree-to-tree transduction, t-rules...transformation from one interface structure to another; SICStus ProLog
* Translation memory systems, Chapter *By* *Ruslan Mitkov*, 2022 https://www.taylorfrancis.com/chapters/edit/10.4324/9781003273417-27/translation-memory-systems-ruslan-mitkov

* Mitkov, R. (1995). Anaphora resolution in natural language processing and machine translation.

* Mitkov, R. (1993). How could rhetorical relations be used in machine translation? https://aclanthology.org/W93-0223.pdf Most MT systems translate only sentence-by-sentence, a few: paragraph-by paragraph and preserve the discourse structure. However R.M. has shown that the structure is different across different sublanguages and pairs of NL.Paragraph translation: "for now an unjustifiably complicated task for practical needs...", "determination of discourse topic(s), goals, intentions, ... a very tough problem." The order of sentences in the translation may be different. Rhetorical predicates.

* **Tanev, H.**, & Mitkov, R. (2002). Shallow language processing architecture for Bulgarian. https://aclanthology.org/C02-1027.pdf LINGUA – *"First, the pre-processing modules for tokenisation, sentence splitting, paragraph segmentation, part of-speech tagging, clause chunking and noun phrase extraction are outlined. Next, the paper proceeds to describe in more detail the anaphora resolution module."*
LINGUA uses Hristo Krushkov's BULMORPH Morphological analyzer. Rule-based, ...

* **H. Krushkov**. 1997. **Modelling and building of machine dictionaries and morphological processors. Ph.D. thesis**, University of Plovdiv. in Bulgarian.

* **G. Totkov and Ch. Tanev. 1999.** Computerized extraction of word semantics through connected text analysis. In Proc. of the International Workshop DIALOGUE ’99, pages 360 – 365. (Georgi Totkov & Hristo Tanev; Ch.Tanev = H.Tanev)

* Hristo Tanev, PhD thesis: "Automatic text processing and ambiguities resolution for the Bulgarian language", 2000-2001: PoS tagging, syntactic analysis, and anaphora resolution, using Mitkov's anaphora resolution algorithm.

* T. Avgustinova, K. Oliva, and E. Paskaleva., 1989. An HPSG-based parser for bulgarian. In International Seminar on Machine, Moscow, Russia.

Tania Avgustinova: https://www.coli.uni-saarland.de/~tania/Schriftenverzeichnis.pdf

*** G. Totkov**, **METHODOLOGY, RESOURCES AND TOOLS FOR COMPUTERIZATION OF BULGARIAN LANGUAGE (1988-2000),**
ЮБИЛЕЙНА НАУЧНА СЕСИЯ – 30 години ФМИ, ПУ “Паисий Хилендарски”, Пловдив, 3-4.11.2000 https://fmi.uni-plovdiv.bg/GetResource?id=591

See this paper and the references there for a survey of the early NLP/Computational Linguistics research for Bulgarian by G.Totkov, K.Ivanov, Hristo Krushkov, Maria Krushkova, Hristo Tanev, P.Petrova P etc.

* Petrova P., G. Totkov, K. Ivanov, Syntactic Analyser. Proceedings of the 20th spring conference of the Union of Bulgarian Mathematicians, 1991, pp. 341-345 (in Bulgarian). 1991.

* Totkov G., Hr. Krushkov, Robust Morphological Analysis for Bulgarian Tests. International Conference “Intelligent management systems”, **Sept.’89,** Varna, pp.141-147 (in Russian)

* Ivanov K., G. Totkov, The Linguistic Processor: System for Research the Word Paradigms in Inflective Natural Languages. Proceedings of the 20th spring conference of the Union of Bulgarian Mathematicians, 254-259, 1991 (in Bulgarian).

* Ivanova P., K. Ivanov, G. Totkov. Automated Acquisition of Grammars Representing Inter-sentence Relations. Proc. of the XXIV summer school, Sozopol, in B. Cheshankov, M. Todorov (Eds.), “Applications of Mathematics in Engineering”, Inst. of Applied Mathematics and Informatics, Technical University of Sofia, Heron Press, Sofia, 1998, 231-234.

* **Krushkov Hr.**, Automatic Construction of an Auxiliary Dictionary for Robust Morphological Analysis. Proceedings of the 21th International Conference ITP'96:

* Totkov G., R. Doneva, K. Ivanov, OMIR-LING: A Linguistic Processor Based on Many Sorted Algebraic Specifications. Intern. Conf. on Mathematical Linguistics ICML'93, Taragona (Catalonia, Spain), Mar 30-31, 1993, pp. 13-14.

*** Totkov G., Formalisation of Bulgarian Language and the Development of a Linguistic Processor. Universite de Plovdiv, Travaux scientifiques, Mathematique, vol.26, fasc.3,**

1988, pp. 301-311 (in Bulgarian)

* Krushkov Hr., Automatic Construction of an Auxiliary Dictionary for Robust Morphological Analysis. Proceedings of the 21th International Conference ITP'96: Interaction between Intelligent Entities, Plovdiv, 1996, pp. 85-88 (in Bulgarian)

* Krushkov Hr., Automatic Checking of the Syntactic Agreement., 1994
* Krushkov Hr., Automatic Construction of Machine Dictionaries. 1996
(…)

*** Towards bulgarian wordnet,** Svetla Koeva, Angel Genov, Georgi Totkov, 2004, ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY Volume 7, Numbers 1–2, 2004, 45–60 https://www.researchgate.net/profile/Svetla-Koeva/publication/267787366_Towards_Bulgarian_Wordnet/links/55e2210a08ae6abe6e8cd4a3/Towards-Bulgarian-Wordnet.pdf (data as of 1.3.2004)
The biggest Bulgarian corpus ~ 33M words mostly electronically published, some scanned; prose & poetry, periodicals,fiction, science fiction, administrative documentation, and scientific texts. 20% literary, 50% journalistic, 20% administrative. (Used for the frequency analysis) Structured corpus: 1M words, 500 text units x 2000 words, sentence boundaries. 15 textcategories ~ Brown corpus. … Current state: 18810 synonyms (synsets): nouns (12292), adjectives (3564), verbs (2946), adverbs (only 8), 35K literals … Semantic tree hierarchy … Number of nodes: Eng.WN2.0 = 79689, **BulNet** = 12292, Tops N = 9. Relations: ALSO SEE, CAUSE, HOLO MEMBER, HOLO PART, HOLO PORTION, HYPERNYM, NEAR ANTONYM, SIMILAR TO, SUBEVENT, VERB GROUP. BE IN STATE, BG DERIVATIVE; DERIVED, PARTICIPLE; REGION DOMAIN, USAGE DOMAIN, CATEGORY DOMAIN. SENSE, SYNONYM, SYNSET, USAGE. Dictionaries, tools for creation of the dictionary, automatic improvement: check for errors, gaps, discrepanies… Extractor

*** Balkanet: A multilingual semantic network for the balkan languages**, Sofia Stamou, Kemal Oflazer, Karel Pala, Dimitris Christoudoulakis, Dan Cristea, Dan Tufis, Svetla Koeva, George Totkov, Dominique Dutoit, Maria Grigoriadou, 2002/1/21

По-късната версия се нарича „BulNet”. Към 17.8.2025 г.: BulNet 3.0

*** Уеб достъп до BulNet/Булнет**: https://dcl.bas.bg/bulnet/ - но показва данни като обикновен речник.

* TOTKOV, G., IVANOVA, P., RISKOV, I**., Automated Improving and Forming WordNet Synsets on Conventional (non computer based) Synonym and Bilingual Dictionaries**, in Comp. Ling. and its Applications (A. Narin'iyani, ed.), DIALOGUE'2003, Protvino, June 2003.

* **On Bulgarian Text-to-Speech System**, G Totkov, V Angelova, Varna 2003

* **Езикови модели и алгоритъм за разпознаване на текст на естествен език**, Мария Жекова, Георги Тотков, 2022 – чрез образци и правила; хибриден, ръчно; POS-тагер, лематизатор, семантични корпуси, шаблони, продукционни правила и програми за разрешаване на многозначност; фраза – БД - вид (таблица, колона, стойност), тип данни; синонимни редове (synsets в WordNet, BalkaNet) – факултет: звено, отдел; университет: ВУЗ, институция, институт …;договор: трудов договор; редовно: нормално, постоянно; област: сфера … Моделиране на въпроси: Шаблон и езикова конструкция: „Кои са/Изведи/Намери/Покажи [списък на] [всичко/всички…SELECT * FROM X …QP1 – др. варианти QP2, QP3, …SELECT A1, A2 .. FROM X … SELECT * FROM X WHERE A1= ... „всички, за които е изпълнено“ … Какъв е процентът на Y спрям о S , за които …

* Visual Parser Builder, Dimitar Blagoev, George Totkov, RANLP 2005
https://lml.bas.bg/ranlp2005/DOCS/RANLP2005.pdf (p.112/125)
https://www.researchgate.net/publication/370068260_Ezikovi_modeli_i_algoritm_za_razpoznavane_na_tekst_na_estestven_ezik

*** Синтез на реч, разпознаване на реч:**

* Тодор Арнаудов, **„Опит за първично разделяне на запис на говор на съставящите го фонеми“**, ПУ, 3.2004 – статията е цитирана в официалния научен поток*. http://eim.twenkid.com/old/_5/31/analiz_na_zvuk.htm

https://www.oocities.org/eimworld/_5/31/dan.txt
* 57. Баева, Д., Д. Игнатова-Цонева, Д. **"Позиции за реализация на гласните и съгласните фонеми в съвременния български език с оглед разработването на компютърни програми за разпознаване на реч"**. – Сб. Книгата, езикът, литературата., БАН, 2006, с. 227 – 237 – цитира „Опит за първично разделяне…“. https://www.researchgate.net/publication/346627640_POZICII_ZA_REALIZACIA_NA_BLGARSKITE_FONEMI

* Т.Арнаудов, **Как да накараме машината да говори като човек? Предложения и описание на синтезатора на реч "Глас", версия 22.4.2004**.
Още „Звуковият синтезатор "ГЛАС 1.0" (22.4.2004)
https://web.archive.org/web/20041020165507/
https://www.oocities.org/todprog/bgr/glas_.htm
https://web.archive.org/web/20040920180549/http://geocities.com/todprog/bgr/glas.htm
Бележки и ръководства и изтегляне на програмата:
https://web.archive.org/web/20080812201430/http://www.geocities.com:80/todprog/bgr/za_pisar_glas_52004.htm http://www.oocities.org/todprog/bgr/glas.htm
https://www.oocities.org/todprog/bgr/za_pisar_glas_52004.htm

* **Т.Арнаудов, „Тошко 2“** – безплатен синтезатор на българска реч и малко английски за PC, 2013 г. https://github.com/Twenkid/Toshko_2 с подобрения в следващи години.
* **Т.Арнаудов, „Глас 2 - синтезатор на българска реч“,** 6.2008, магистърска дипломна работа – проект за подобрения на Глас 2004 и насоки за подобрения и бъдещи обучаващи се синтезатори. https://github.com/Twenkid/Glas-2

* Mitkov R. - Discourse-based approach in machine translation, From Proceedingsof the International Symposium on Natural Language Understanding and Artificial Intelligence, Fukuoka, Japan, 13-15 July, 1992 https://aclanthology.org/1995.tmi-1.6.pdf CAT2 translation system, unification-based formalism
* Mitkov, R., Choi, S.-K., & Sharp, R. (1995). Anaphora resolution in machine translation.
* Mitkov, R. (2000, June 15). Pronoun resolution: The practical alternative.
* Mitkov, R. (1999). Introduction: Special issue on anaphora resolution in machine translation and multilingual NLP.
* Mitkov, R. (1999, December). Multilingual anaphora resolution.
* Are rule-based approaches a thing of the past? The case of anaphora resolution, R Mitkov, 2024 list of preferences – 2. .. antecedent indicators; search scope 2,3,4 sentences; MARS – Mitkov’s Algorithm to pronoun ReSolution.

* Computer-Aided Language Processing, R Mitkov, FLAIRS, 319-320, 2006 https://cdn.aaai.org/FLAIRS/2006/Flairs06-062.pdf (robust parsing 90% by 2000 (S(NP(DET(ADJ(ADJ(NOUN(VP(VERP(ADP(… , but anaphora resolution only 60% by 2006 (references to previously mentioned entities)
* Towards a more efficient use of PC-based MT in education, Ruslan Mitkov, 1996 https://aclanthology.org/1996.tc-1.6.pdf
* Anaphora Resolution Exercise: an Overview. C Orasan, D Cristea, R Mitkov, AH Branco, http://www.lrec-conf.org/proceedings/lrec2008/pdf/713_paper.pdf
… identification of referential expression and their candidates, pronominal anaphora resolution on pre-annotated texts, coreferential chains resolution on pre-annotated texts, pronominal anaphora resolution on raw texts, coreferential chains in raw unannotated texts; relations: IDENTITY, SYNONYMY, GENERALISATION, SPECIALISATION between entitie …

* Тодор Арнаудов (Todor Arnaudov)

***„Smarty - Extendable Framework for Bilingual and Multilingual Comprehension Assistants**”, Todor Arnaudov, Ruslan Mitkov, LREC 2008 – „най-интелигентният речник в света“, подпомагащ превода с голям брой обработки на естествен език и използващ контекста за разрешаване на многозначност. Използва WordNet, BalkaNet и обикновен речник и мощен графичен интерфейс, при който може да се посочва всяка дума. Разработен до завършен прототип с всички функции за около три месеца от нулата от Т.Арнаудов, научен ръководител Р.Митков.

Интерфейсът на „Смарти“ започна да навлиза в уеб сайтове за превод като „Reverso Context” и др. десетина или 15 години по-късно. Около 2012-2013 г. се разработваше версия на Java с допълнителни възможности за пораждане на код по образец и по-развит интерфейс, която не беше публикувна. Продължение на „Смарти“ е „Research Assistant, още „Assistant C#” или само „Assistant”, замислен още през 2007 г. – виж насоките за изследвания и разработка от *„Първата модерна стратегия за развитие чрез изкуствен интелект* …“ През март 2024 г. за кратко се разработиха подобрения на „Смарти“ в „Smarty 2”, но засега не са публикувани.
http://www.lrec-conf.org/proceedings/lrec2008/pdf/826_paper.pdf
https://aclanthology.org/L08-1586/
https://github.com/Twenkid/Smarty

„Смарти“ имаше голям потенциал, някои от причините да не се доразвие в края на 2000-те бяха организационни и лични. Друго беше опасение, че не може да се състезава с „Гугъл“ и тяхната развиваща се услуга за машинен превод – това беше погрешно. От 5.2024 г. в *Research Assistant*, универсалният помощник, наследник на „Смарти“, се използва безплатна услуга за машинен превод на „Гугъл“.

*** GPT2-MEDIUM-BG, T.Arnaudov 2021** – един от най-големите езикови модели за езици, различни от английския, до 2021 г Разработих и метод за пораждане с неограничена дължина и промяна на темата, чрез постепенно вмъкване на скрити подкани, които не се включват в изведеният текст. Виж таблицата в края на този том. LLM. https://huggingface.co/twenkid/gpt2-medium-bg

* **Superhuman** (~2012-2013, Java) – unpublished, **Smarty 2** (2024 – announced, unpublished), **Research Assistant** (Assistant C#, ACS, Assistant) – conceived since 2007, in-house use since early 2010 – unpublished (as of 17.8.2025) – виж бъдещи публикации.

(…)

Драгомир Радев (Dragomir Radev)

Извличане на информация, резюмиране, компютърна лингвистика, … уважаван преподавател в няколко американски университети. Докторант в Колумбийския университет от 1992-1993?-1998; … в края на живота си – в Йейл. Един от най-високо класираните по h-index. h-index по Google Scholar: 88; since 2019: 68 (21.12.2024)
* BLOOM: A 176B-Parameter Open-Access Multilingual Language Model, BigScience Workshop, … D.Radev, … (участва в Dataset, Prompt engineering, Evaluation and Interpretability) https://arxiv.org/pdf/2211.05100 – подробно описание на разработката на голям езиков модел с размерите на GPT3. vocab_size=250,680 + 200 reserved; BPE tokenizer; fertility; pre-tokenizer; didn’t use splits on numbers and digits or English specific (‘nt ‘ll); regex (not exactly) ?[^(\s|[.,!?… |_])]+ ; trained: 3.5 months, 1,083M GPU/h, 48 x 8 NVIDIA A100 x 80

GB = 384 GPUs + 4 nodes reserve; each node x8 GPU with 2x AMD EPYC 7543 32-Core CPUs and 512 GB of RAM; SpectrumScale (GPFS) parallel file system for SSDs & HDDs; 4 NVLink GPU-to-GPU per node; 4 Omni-Path 100 Gbps links per node 8D hypercube. Megatron-DeepSpeed. 8 copies of the model are trained in parallel on 384 GPUs with data parallelism = 8. Model parameters are shared between 4 GPUs (tensor parallelism = 4; horizontal parallelism or intra-layer model parallelism); layers of the model spread across 12 groups of GPUs (pipeline parallelism = 12, vertical par.); One full copy of the model (replica) = 48 GPUs. 156 TFLOPs in their fastest configuration with A100 GPUs, attaining their objective of half of the theoretical peak performance of 312 TFLOPs (in float32 or bfloat16). bfloat16 mixed precision – float16 overflows, initial experiments on V100. …
Antoine Simoulin and Benoit Crabbé. Un modèle Transformer Génératif Pré-entrainé pour le ______ fran¸cais. In Pascal Denis, Natalia Grabar, Amel Fraisse, Rémi Cardon, Bernard Jacquemin, Eric Kergosien, and Antonio Balvet, editors, Traitement Automatique des Langues Naturelles, pages 246–255, Lille, France, 2021. ATALA. URL https://hal.archives-ouvertes.fr/hal-03265900 submitted: 23.6.2021

* Generating Natural Language Summaries from Multiple On-Line Sources, Ph.D. Thesis Proposal, Dragomir R. Radev, Technical Report CUCS-005-97 Department of Computer Science Columbia University, March 28, 1997 (56 p.)
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2b013ebac519c005f91c18284eb2351830507f6e * **Завършената дисертация, 1998:** https://www.cs.columbia.edu/nlp/theses/dragomir_radev.pdf

Тош: Необятността за отделен потребител дори на тогавашния Интернет, **"мегабайти новини на ден",** "**100** източника на новини" и пр.; briefs - сводки по теми и събития, интересуващи потребителя, включващи и пораждане на текст; източници: исторически и текущи (historical and current), тексови и нетекстови (онтологии, БД; Т: изрично стуктурирани) … метод: language reuse and regeneration (LRR): преизползване на текст и пресъздаване. … *„Предимство на функционалните граматики в системи за пораждане на текст от знания: отнасят се по еднообразен начин с всички изисквания на пораждането: на дискурса/текста, семантични, синтактични и лексикални, не изискват прекалено сложни граматики и позволяват лесно разделяне на модули…“* Функционални унификационни граматики … functional description FD, atom path, another FD; frames: slot groups: message, incident, … clusters of stories; 4.6. feature type: isa(car, vehicle) … (define-feature-type weapon (car truck)) … (country ((name “Bulgaria”) (capital “Sofia”) (map ((url “…”))) (type republic) (divisions ((name “oblast”)(number 28)(list (“София“, „Пловдив“, …))))(ports ((list(“Varna” “Burgas” …)))) @Вси: *срвн. PDDL, роботика, планиране. …* 5.3. Планиране на абзац … PLANDOC .. подобия между статии, свързваща инф.; прждизрч. семнтч.стркр. FD; шблн; вж. 6.5. с.92 граматика на изрчн: пълно опрдлн. Нбхдм з прждж. 7. LRR; ? да повтори, ? да преобразува, за изглаждане; phrasal lexicons, преизползваеми изречения за факти; глава 10: Предишна работа: summarization as sentence extraction – извличане на обобщаващи изречения без промяна. Виж от там: McKeown and Radev, 1995,

* Radev, 1996; Radev and McKeown, 1997
https://www.cs.columbia.edu/nlp/bibsearch.cgi?keyword=1998
* Rendezvous: A WWW Synchronization System, Dragomir R. Radev, Second International WWW Conference, 10.1994
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1af661a8e3ab9732c1ecfd9cbb1ae2c2f3fed8cf - hotlists, notifications for updates of particular web sites
* CREATIVESUMM: Shared Task on Automatic Summarization for Creative Writing, Divyansh Agarwal, ... Radev, ..., 2022 https://arxiv.org/pdf/2211.05886 - fiction, movie and TV scripts
* MEAD - A platform for multidocument multilingual text summarization., D.Radev et al. 2004 https://www.cs.columbia.edu/nlp/bibsearch.cgi?keyword=Radev
https://scholar.google.com/citations?hl=en&user=vIqWvgwAAAAJ&view_op=list_works&sortby=pubdate
* John Prager, Dragomir Radev, and Krzysztof Czuba, 2001. Answering what-is questions by virtual annotation.
* **Question answering by predictive annotation,** J.Prager, E.Brown, Anni Coden, *Dragomir Radev,* July 2000
https://researchgate.net/publication/221300763_Question_answering_by_predictive_annotation TREC8; GuruQA – returns a ranked list of passages instead of documents; QA-Tokens: PLACE, COUNTRY, STATE, PERSON, ROLE, NAME, ORG, DURATION, AGE, YEAR, TIME, DATE, VOLUME
AREA$, LENGTH$, WEIGHT$, NUMBER$, METHOD$, RATE$, MONEY$ … (Where, Where/What country, Where/What state, Who, Who, Who/What/Which/Where/Name the, Who/What, How long, How old, When/What year, When, When/What date. How big, How big, How big/long/high, How big/heavy, How many, How, How much, How much); method: "*by* doing…"; *Textractor:* tokenization, lemmatization, annotation: Nominator (proper names), Terminator (technical terms, e.g. computer), Abbreviator; statistics, alternative names, canonical forms. Word-list; Resporator: identify potential answer phrases and annotate them; the problem of the optimal window size – the size of the selected passage; dynamic window (smaller passage with the same score as a larger wins). Query analysis; Matching & Ranking … Answer selection: AnSel, WerLect; "How and Why questions are difficult"; by %ing… the example "How did Socrates die?"*… @SYN(because,cause,result); "in order to", "to VERB"…; problems: not appropriate question templates, synonym or hyponym of the term in text which is unknown to the system; no anophora resolution (refererences with alternative forms), poor text understanding. See also Message Understanding: fact extraction;

* **A Multi-Strategy and Multi-Source Approach to Question Answering**, J.Chu-Carroll et al., 2002, IBM Watson … See also the references about Watson, Jeopardy! after the interview "*Marvin Minsky on AI: The Turing Test is a Joke"*

* **Twenkid, 30.12.2024:** it should ?T from a seed with the named entity and expand for semantic connections to the question verb/query in more depth, a longer chain of

connections: ("die" – death, poison, kill, cease to exist; sentenced, executed, … and their relations: take a poison, drink a …; ways of dying, executions etc. and searching e.g. connections such as "like…")
Корпуси: * D. R. Radev, P. Muthukrishnan, and V. Qazvinian. **The ACL anthology network corpus. In Workshop on Text and Citation Analysis for Scholarly Digital Libraries** (NLPIR4DL), pp. 54–61. Association for Computational Linguistics, 2009.

Резюмиране:
TL;DR: Mining Reddit to Learn Automatic Summarization, Michael Völske et al.
https://aclanthology.org/W17-4508/ https://aclanthology.org/W17-4508.pdf
използват резюмета, написани от потребителите за улеснение на другите ("Too Long; Didn't Read); TA: виж и подобен подход за клипове в Ютюб.
Extractive & abstractive summarization systems datasets. 1.1M subreddits; 286M submissions, 1.6B comments (2006-2016)
https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/generative-ai/working-with-llms/evaluation/list-of-eval-metrics

* **Извличане на информация – Ива Попова,** ТУ София, докторант около 2006-2008 г. С нея през 2006 г. дискутирахме подходящ термин:
* **ЕВ1: Относно превода на Information Retrieval от NemSys** на 01.10.2006 05:08 23 коментара , 960 прочита Етикети: езикови въпроси
https://bglog.net/BGLog/post/%D0%95%D0%921--%D0%9E%D1%82%D0%BD%D0%BE%D1%81%D0%BD%D0%BE-%D0%BF%D1%80%D0%B5%D0%B2%D0%BE%D0%B4%D0%B0-%D0%BD%D0%B0-Information-Retrieval#gsc.tab=0

* Kristina Toutanova

https://scholar.google.com/citations?user=9qY7NPEAAAAJ&hl=en
http://kristinatoutanova.com/ [font: Constantia 12, test]

* **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,** Jacob Devlin, Ming-Wei Chang, Kenton Lee, **Kristina Toutanova** Google AI Language https://arxiv.org/pdf/1810.04805 //Font: Liberation Serif 12

* ProtEx: A Retrieval-Augmented Approach for Protein Function Prediction, P Shaw, B Gurram, D Belanger, A Gane, ML Bileschi, LJ Colwell, ... K.Toutanova, bioRxiv, 2024.05. 30.596539 1 2024

* Anchor Prediction: Automatic Refinement of Internet Links, NF Liu, K Lee, K Toutanova, 2023 https://arxiv.org/pdf/2305.14337 [Liberation Serif]

* Open-domain visual entity recognition: Towards recognizing millions of wikipedia entities, H Hu, Y Luan, Y Chen, U Khandelwal, M Joshi, K Lee, K Toutanova, ... 2023

https://openaccess.thecvf.com/content/ICCV2023/html/Hu_Open-domain_Visual_Entity_Recognition_Towards_Recognizing_Millions_of_Wikipedia_Entities_ICCV_2023_paper.html Text query to link to an image from Wikipedia: Open-domain Visual Entity recognitioN (OVEN). Re-purposing 14 datasets with all labels grounded to the label space of Wiki entities. 6 million Wki entities --> a general visual recognition benchmark with largest number of labels. CLIP, PaLI

* **Image Recognition Datasets:** ImageNet21k-P [34,36], iNaturalist2017 [45], Cars196 [21], SUN397 [52], Food101 [2], Sports100 [16], Aircraft [26], Oxford Flower [29], Google Landmarks v2 [50].

* **Visual QA Datasets:** VQA v2 [17], Visual7W [55], Visual Genome [22], OK-VQA [28], Text-VQA [42]. Entity Split (ES), Query split (QS).

* [8] **Pali: A jointlyscaled multilingual language-image model.** Xi Chen, **... Anelia Angelova** ... et al., arXiv preprint arXiv:2209.06794, 2022/6.2023 https://arxiv.org/abs/2209.06794 Pathways Language and Image model. Vision Transformers (ViTs). Vision+Text --> Text … existing Transformers for language are much larger than their vision counterparts, we train a large, 4-billion parameter ViT (ViT-e) to quantify the benefits from even larger-capacity vision models. … multilingual mix ... image-text training set 10B images and texts in > 100 lang; tasks: captioning, VQA, scene-text understanding;

* **Sparse, dense, and attentional representations for text retrieval**, Y Luan, J Eisenstein, K Toutanova, M Collins 2021
* **Representations for question answering from documents with tables and text,** V Zayats, K Toutanova, M Ostendorf, arXiv preprint arXiv:2101.10573, 2021

* **Providing rewards and metrics for completion of microtasks**, JB Teevan, S Amershi, ... KN Toutanova, US Patent App. 16/933,827, 2020 https://patentimages.storage.googleapis.com/95/7d/29/f2e34aacfed15b/US20200349596A1.pdf https://scholar.google.com/citations?view_op=view_citation&hl=en&user=9qY7NPEAAAAJ&cstart=20&pagesize=80&sortby=pubdate&citation_for_view=9qY7NPEAAAAJ:cF7EPgIkoB4C

* **Natural questions: a benchmark for question answering research,** T Kwiatkowski, J Palomaki, O Redfield, M Collins, A Parikh, C Alberti, ... 2019 [font: Palatino 12]

* **NLP for Precision Medicine**, H Poon, C Quirk, K Toutanova, W Yih, 2017, Tutorial, Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017 → Entity linking, relation extraction, Distant supervision, complex event extraction, Grounded semantic parsing, Cross-sentence relation extraction

* **Observed versus latent features for knowledge base and text inference**, K Toutanova, D Chen, 2015 [Univers 12]

* **Representing Text for Joint Embedding of Text and Knowledge Bases,** K Toutanova, D Chen, P Pantel, H Poon, P Choudhury, M Gamon, 2015 [Arial 12]

* **Locating parallel word sequences in electronic documents**, CB Quirk, KN Toutanova, JR Smith US Patent 8,560,297, 22, 2013

* **Multilingual named entity recognition using parallel data and metadata from wikipedi**a, S Kim, K Toutanova, H Yu … 2012

* **Machine translation from text,** N Habash, J Olive, C Christianson, J McCary, N Habash, A de Gispert, 2011

* **Extracting parallel sentences from comparable corpora using document level alignment**, J Smith, C Quirk, K Toutanova, Human language technologies::… 2010

* **Unsupervised chinese word segmentation for statistical machine translation**, J Gao, KN Toutanova, J Xu US Patent App. 12/163,119 16, 2009

* **The stanford natural language processing group**, C Manning, D Jurafsky, P Liang ... The Stanford Parser: A statistical parser, http://nlp. stanford.edu … 2008

* **Generating complex morphology for machine translation**, **E. Minkov**, K Toutanova, H Suzuki, Proceedings of the 45th annual meeting of the association of computational … 131 2007

* **The pythy summarization system**: Microsoft research at duc 2007, K Toutanova, C Brockett, M Gamon, J Jagarlamudi, H Suzuki, ... Proc. of DUC 2007

* **Competitive generative models with structure learning for NLP classification tasks**, K Toutanova Proceedings of the 2006 Conference on Empirical Methods in Natural Language … 14 2006

* **Learning to predict case markers in japanese**, H Suzuki, K Toutanova, Proceedings of the 21st International Conference on Computational … 47 .. 2006

* **機械学習による日本語格助詞の予測** (Machine learning prediction of Japanese case particles), 鈴木久美， K Toutanova
言語処理学会第 13 回年次大会 2 .. 2006

* **Effective statistical models for syntactic and semantic disambiguation**, KN Toutanova, stanford university 12.2005

* **Learning random walk models for inducing word dependency distributions,** K Toutanova, CD Manning, AY Ng, 2004

https://nlp.stanford.edu/kristina/papers/ppwalks.pdf
https://slideplayer.com/slide/4931425/

* **Optimizing local probability models for statistical parsing,** K Toutanova, M Mitchell, CD Manning, 2003
* Combining heterogeneous classifiers for word sense disambiguation, D Klein, K Toutanova, HT Ilhan, SD Kamvar, CD Manning, Proceedings of the ACL-02 workshop on Word sense disambiguation: recent … 2002

* **Combining Heterogeneous Classifiers for Word-Sense Disambiguation,** D Klein, CD Manning, K Toutanova

* **Pix2Struct: Screenshot Parsing as Pretraining for Visual Language Understanding.** Kenton Lee, Mandar Joshi, Iulia Turc, Hexiang Hu, Fangyu Liu, Julian Eisenschlos, Urvashi Khandelwal, Peter Shaw, Ming-Wei Chang, and Kristina Toutanova, 2022.

* **Entity-Centric Query Refinement**. David Wadden, Nikita Gupta, Kenton Lee, and Kristina Toutanova, AKBC 2022.

* **Representations for Question Answering from Documents with Tables and Text.** Vicky Zayats, Kristina Toutanova and Mari Ostendorf, EACL 2021.

* **Compositional Generalization and Natural Language Variation: Can a Semantic Parsing Approach Handle Both?,** Pete Shaw, Ming-Wei Chang, Panupong Pasupat, and Kristina Toutanova, ACL 2021, 6.2021
https://arxiv.org/abs/2010.12725
NQG-T5: high-precision grammar-based + pre-trained seq2seq m.
Seq2seq struggles with out-of-distrib. composit. gener. T5 is an LLM; data: “jump”, walk”, “walk twice” --> can't gener. to "jump twice” (compositional generalization). SCAN. Target Maximum Compound Divergence (TMCD) train & test splits. NQG - grammar, solves SCAN. NQG-T5: hybrid semantic parser. Quasi-synchronous Grammar induction. SCAN,(Lake & Baroni, 2018): NL commands --> action seq.: ("clap twice: I_CLAP I_CLAP"). CFQ, Maximum Compound Divergence (MCD), Keysers et al. (2020) - compound distribution. ... Contextualized Representations

* **Using Textual Encyclopedic Knowledge.** Mandar Joshi, Kenton Lee, Yi Luan, and Kristina Toutanova, 2020. https://arxiv.org/pdf/2004.12006
TEK-enriched, contextualizing Q & passages w. background from wikipedia

* **Latent Retrieval for Weakly Supervised Question Answering.** Kenton Lee, Ming-Wei Chang, and Kristina Toutanova, ACL 2019 6.2019
arxiv.org/pdf/1906.00300 Inverse Cloze task; sentence as a pseudo-question, context - pseudo-evidence. Open-Retrieval Question. Answering (ORQA) model; evidence from open corpus, not in the input; ... **колко било трудно невронните модели да подобрят традиционното извличане [като TF-IDF, обратна**

честота и пр.], фактическият най-добър модел тогава, 2019 г., бил 10-годишният BM25, Robertson et al., 2009. Сравняват със 128-dim. NNLM, context-independent fed-forward LM (Bengio et al. 2003), ELMo (small), context-dependent bidirectional LSTM (Peters et al., 2018);
Datasets: Natural Questions, WebQuestions, TriviaQA, CuratedTrec; SQuAD: 100K questions, 536 docs; p.7. *"Language Models unsupervised neural retrieval is notoriously difficult to improve over traditional IR..."* 128-dim. NNLM, context-independent fed-forward LM (Bengio et al. 2003), ELMo (small), context-dependent bidirectional LSTM (Peters et al., 2018)
Notes Todor*:* Weakly supervised: e.g. in TriviaQA, the answer is an entity which might be mentioned multiple times in the document with the answer (ambiguity) [note that if it is just once, that's a hint], no explicit info about where the answer is: a pattern (a mapping) for discovering the answer has to be discovered from many examples; paraphrased answer; Textbook question answering: multi-modal contexts, diagrams; ... See also: DROP QA (Discrete Reasoning Over the content of Paragraphs) - add, count, sort over the info in a text passage. "The bus had 15 bus when it left the depot. 5 left on the first stop, 3 got on. How many were the people in the bus before the next stop?" 15-5+3 = 13 Answer: 13

* **Bayesian nets in syntactic categorization of novel words**, Leonid Peshkin, Avi Pfeffer, **Virginia Savova**, 2002 – Dynamic Bayesian Network for PoS-tagging ... observable features (OF), memory; OF: letters, prefix, suffix, number, hyphen, is a word or a number, capital or lower case, ... unlike Toutanova, 2002 ... a set of **binary features** and a set of vocabulary features - does the token contain a ... capital letter, a hyphen, a suffix ... etc. Two hidden variables: PoS and Memory reflect the contextual information about the past PoS tags ... Learning = collecting statistics over co-occurrences of feature values and tags. Tagging: standard Forward-Backward algorithm (Murphy[2002])
* Murphy. K. Dynamic Bayesian Networks: Representation, Inference and Learning. PhD thesis. UC Berkeley. 2002
*A. Ratnaparkhi. 1996. A maximum entropy model for part-ofspeech tagging. In Proceedings of EMNLP.
* E. Brill. 1994. Some Advances In Rule-Based Part of Speech Tagging. In Proceedings of the 12th AAAI.

→ the paper **cites:** Toutanova K and Manning, C. **Enriching the Knowledge Sources Used in a Maximum Entropy PoS Tagger. 2002.** as the SOTA POS-tagger

* **Virginia Savova** – a Bulgarian? (unknown (guessed, but unconfirmed; another researcher: Angela *Savova*)): John Hopkins University; participates at RANLP 2003
* **Виргиния Савова**

https://lml.bas.bg/ranlp2003/
https://scholar.google.com/citations?user=hmwDs7cAAAAJ&hl=en

* **Why build another part-of-speech tagger? A minimalist approach.** Leonid Peshkin, Virginia Savova https://www.researchgate.net/profile/Leonid-Peshkin/publication/228789616_Why_Build_Another_Part-of-Speech_Tagger_A_Minimalist_Approach/ (See *Bayesian nets in syntactic categorization of novel words* above) – *The dilemma: statistically extracted or expert-selected features?* The most likely tags: 90% correct; dictionaries; the best-known rule-based tagger [Brill'94]: 1. assign the most likely tag to each word in the text; 2. apply transformation rules of the form "Replace tag X by tag Y in triggering environment Z", where the trig.env. spans up to 3 sequential tokens in both directions and refers to words, tags or properties of words within the region. The Brill tagger: < 3.5% error on the Wall Street Journal (WSJ) corpus. However rule-based taggers perform poorly on different datasets, e.g. Netlingo (e-mail, newsgroups, web sites). *"The typical inference task is to determine the probability distribution over the states of a hidden variable over time, given time series data of the observed variables. This is usually accomplished using the forward-backward algorithm. Alternatively, we might obtain the most likely sequence of hidden variables using the Viterbi algorithm. These two kinds of inference yield resulting PoS tags. Note that there is no need to use "beam search", (cf. [Brants'00])."* Tests of the performance after omitting some of the *binary features* (see the other paper). Morphology helps for better results. OoV words (out of vocabulary). GAWK[11] scripts for text processing.
* **Is the Turing test good enough? The fallacy of resource-unbounded intelligence,** Virginia Savova, Leonid Peshkin, 1/2007 compression ... not just retransmit ...
https://www.researchgate.net/publication/220816294_Is_the_Turing_Test_Good_Enough_The_Fallacy_of_Resource-Unbounded_Intelligence

Compare: **Faults in Turing Test and Lovelace Test. Introduction of Educational Test,** Todor Arnaudov, 11.2007, Artificial Mind https://artificial-mind.blogspot.com/2007/11/faults-in-turing-test-and-lovelace-test.html
* **"Man and Thinking Machine: an Analysis of the Possibility of Creating a Thinking Machine and Some Shortcomings of Humans and Organic Matter in Comparison to it."**, Todor Arnaudov, 12.2001, The Sacred Computer, #13 (in Bulgarian: https://www.oocities.org/eimworld/eimworld13/izint_13.html) the part discussing the inadequacies of the Turing test. („Човекът и Мислещата Машина: Анализ на възможността да се създаде мислеща машина и някои недостатъци на човека и органичната материя пред нея", Т.Арнаудов, 2001)

* **Discovering Syntactic Hierarchies**, Virginia Savova, Daniel Roy, Lauren Schmidt & Joshua B. Tenenbaum, 2007, Journal Proceedings of the Annual Meeting of the

[11] https://www.gnu.org/software/gawk/manual/html_node/index.html#SEC_Contents

Cognitive Science Society, 29(29) … Hierarchical clustering, morphological inf., … features: Fig.2.: (concrete, intentional, manner, moving, suffix (“ing”, “s”),them,first,or,from,must,he) .. & distributional; https://escholarship.org/content/qt6kp787g3/qt6kp787g3.pdf

* Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. Machine Learning, 7, 195-224.

* **RANLP, 2003. Tutorials:** Note the EuroWordNet initiative at the time, with an “InterLingua”; the tutorial on Question Answering and **“Learning in NLP: When can we reduce or avoid annotation cost?”, Ido Dagan, 2003**
* https://www.slideserve.com/jalen/learning-in-nlp-when-can-we-reduce-or-avoid-annotation-cost 143 pages – very good synthesized material on various topics, the theoretical foundations and why machine learning is needed and for what purposes. The methods at the time: Expectation-Maximization with HMM for PoS-tagging; Parsing as a “perspective”. “Simple learning from noisy statistics” for prepositional phrase attachment; Parallel coprpora for WSD; Bootstraping of decision lists for WSD & NE classif.; Expectation-Maximization clustering for Bayesian WSD; Basic types of clustering: bottom-up (agglomerative), top-down (divisive); hierarchical and flat (K-means etc.); basic methods, similarity. Unsupervied ditstributional similarity and clustering … E.g. “Supervised/classification” is identifying hidden units (concepts) of explicit units; Unsupervised is identifying relationships & properties of explicit units (p.5); the *Explicit units are from lexical units (words, terms) to documents; implicit (hidden) are word senses, name types, document categories, PoS tags, syntactic relationships, semantic concepts and relationships etc.* (p.4). .. Ambiguity, variability … WSD: door, window → doorwindow and see whether the classifier will work … p.43 – intersecting redundancies (Dictionary-based mapping to target language); selection of alternative word-forms/senses for translations; bilingual corpus & reverse translation .. decision list;; p.78: significant features; p.141 “glass box” vs black box; *inadequacy of strict “black box” supervised learning vs relevance of rule/knowledge-based models* @Vsy: cmp LLMs. *„Unsupervised learning of variability – paraphrases from repeated descriptions of the same fact from parralel, stand-alone or comparable corpora” .. “Reducing human intervention is critical for NLP applicability …”*

* **Open-Domain Question–Answering, John Prager, IBM T.J. Watson Research Center, 2006** https://www.nowpublishers.com/article/DownloadSummary/INR-001 – (link – only a summary) Short history of QA and modern approaches at the time; QA in its current definition from 1999 at TREC, Text Retrieval Conference (from 1992)… SHRDLU is considered as a QA system in hindsight, 1968-1970. Note that LLMs can be seen as QA systems and a lot of their usage is like that. The creation of the datasets for LLMs also often includes an IR phase for extracting and curating the

right examples if such are not already available. QA: IR, NLP, NER (Named entity recognition), parsing, search, indexing, classification, ML… Early systems besides SHRDLU: LIFER/LADDER, LUNAR, and CHAT-80. First modern: MURAX, open-domain, searches in an on-line encyclopedia. MIT's "Start" – Web QA. Ask Jeevs (1996). Brainboost, AnswerBus: return a sentence. AQUAINT (Advanced Question–Answering for INTelligence). nuggets - a set of descriptive text fragments; linguistic, statistical, and knowledge-based methodologies … Types of questions: factoid, list, definition; relationships – influences, financial or communication connections or links; list of properties – nuggests (късче, хапка, откъс); Eval. ROUGE, POURPRE, Pyramids*https://www.khoury.northeastern.edu/home/vip/teach/IRcourse/IR_surveys/ibm-ai-qna.pdf

- P.Passonneau, A. Nenkova and R. Passonneau, "Evaluating content selection in summarization: The pyramid method," in Proceedings of the Human Language Technology Conference (NAACL-HLT), 2004. https://www.linkedin.com/in/ani-nenkova-6223b9212/

* **Question answering by predictive annotation,** J.Prager, E.Brown, Anni Coden, *Dragomir Radev,* July 2000 https://researchgate.net/publication/221300763_Question_answering_by_predictive_annotation – see notes for the article in the section for Dragomir Radev above

Cmp:
* Matthew E. Peters et al., 2018. Deep contextualized word representations. (ELMo)
* **Todor Mihaylov:** Knowledge-Enhanced Neural Networks for Machine Reading Comprehension, 2021, PhD thesis (see below)
A dataset for NLU tasks: predict Wikidata's textual values by reading the text of the corresponding Wikipedia articles**:** https://paperswithcode.com/dataset/wikireading https://github.com/google-research-datasets/wiki-reading

...

* ROUGE: A Package for Automatic Evaluation of Summaries*,* Chin-Yew Lin, 2004 https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/generative-ai/working-with-llms/evaluation/list-of-eval-metrics
Ground truth: Match sequences: N-gram: BLUE, ROUGE (matches of generated summaries to ground-truth given summaries), JS divergence: overlap-based; Levenstein distance (string transformations); Semantic: BERTScore, MoverScore, Sentence Mover Similarity (SMS) - embeddings; cos similarity.... Reference-free: BLANC - accuracy of reconstructions of masked-tokens; ROUGE-C: source as context for comparison; Entailment (Natural Langage Inference?): output entails, contradicts or undermines the premise: "consistent/inconsistent". Factuality, QA, QG ... QuestEval; LLMs as evaluators: Reason-then-Score RTS, Multipple Choice Question Scoring MCQ, Head-to-head scoring H2H, G-Eval. GEMBA - translation. ... Gen.code ... Rule-based ...
https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/generative-

ai/working-with-llms/evaluation/g-eval-metric-for-summarization
Abstractive vs extractive; G-eval; coherence, consistency, fluency, relevance;
1.structure, 2.factual alignment; 3. sentences: no gramm.err.; 4. most important content;
Note: T.Arnaudov: **important for whom? the evaluator model also has to be defined for any metrics! what is relevant for which reader at which moment; there could be many equal variants**
https://github.com/microsoft/promptflow/tree/main/examples/flows/evaluation/eval-summarization
https://ieeexplore.ieee.org/document/4664680   Tingting He et al, 2008
* Large Language Models are Not Yet Human-Level Evaluators for Abstractive Summarization, Chenhui Shen et al., 10.2023 https://arxiv.org/pdf/2305.13091

* A Survey of Text Summarization Techniques, Chapter, First Online: 01 January 2012, pp 43–76

 # Ani Nenkova (probably Bulgarian)

https://link.springer.com/search?dc.creator=Ani%20Nenkova
https://www.linkedin.com/in/ani-nenkova-6223b9212/
* Mining Text Data, Ani Nenkova & Kathleen McKeown
https://link.springer.com/chapter/10.1007/978-1-4614-3223-4_3

Tosh: **Important *for whom* - model of the reader**
Yes: … from a KB,
* User Modelling as an Application of Actors, pp 83–89, 2003 …
* **Conceptual Structures: Standards and Practices**(ICCS 1999), Ani Nenkova & Galia Angelova Levels: beginner, medium level or domain expert …
*Integration of Resources and Components in a Knowledge-Based Web-Environment for Terminology Learning, Conference paper, First Online: 01 January 2003, pp 210–220
Artificial Intelligence: Methodology, Systems, and Applications (AIMSA 2000), Svetla Boytcheva, Ognian Kalaydjiev, Ani Nenkova & Galia Angelova

CGs as a Knowledge Representation Core in a Complex Language Learning Environment, Angelova, G., A. Nenkova, Sv. Boycheva, and T. Nikolov. To appear in ICCS-2000, Darmstaht, Germany, August 2000.

CGWorld - A Web Based Workbench for Conceptual Graphs Management and Applications, Pavlin Dobrev and Kristina Toutanova, To appear in ICCS-2000, Darmstaht, Germany, August 2000. https://www.researchgate.net/profile/Pavlin-Dobrev/publication/245633963_CGWorld-a_web_based_workbench_for_conceptual_graphs_management_and_applications/links/02e7e52a80c98ecf76000000/CGWorld-a-web-based-workbench-for-conceptual-graphs-management-and-applications.pdf   Knowledge graphs, Prolog+Java

* DB-MAT: Knowledge Acquisition, Processing and NL Generation, G. Angelova, K. Bontcheva:
* Using Conceptual Graphs. ICCS 1996: 115-129

*** Anton Alexandrov, Vesselin Raychev, Martin Vechev, D. Dimitrov, K.Toutanova & BGGPT, 2024**

* **BgGPT 1.0: Extending English-centric LLMs to other languages**, A Alexandrov, V Raychev, DI Dimitrov, C Zhang, M Vechev, K Toutanova, arXiv:2412.10893, 14.12.2024, https://arxiv.org/abs/2412.10893

BgGPT-Gemma-2-27B-Instruct, 9B; > 100B tokens; 64 NVIDIA H100s, 8x8 Infiniband; EXAMS* (Hardalov et al., 2020) - high school exam question dataset covering a range of subjects 1472 sampl. in Bg; 5-shot prompt; & school tests from Ministry of Educ&Sci. 4-12 gradee, multiple-choice of 4 (a,b,c,d) in: literature,math,phys,hist,comp_sc etc. - all subjects, only text;
Eval. 1 commonsense r.: HellaSwag, winogrande_xl - 1767 s. - some rephrased (gender in Bg). Converted mult.choice->open with GPT-40

* “The prophets of the Thinking Machines…”: Tests, datasets: Comp. the prediction: Faults in Turing test and Lovelace test...”, T.Arnaudov 2007, “Човекът и мислещата машина…“, 2001

* **Mitigating Catastrophic Forgetting in Language Transfer via Model Merging**, Anton Alexandrov, Veselin Raychev, Mark Niklas Mueller, Ce Zhang, Martin Vechev, Kristina Toutanova, 2024 https://aclanthology.org/2024.findings-emnlp.1000.pdf

Branch-and-Merge (BAM), a new adaptation method ... "Experience replay" (from the first dataset); train in parallel a set of model branches on different subsets of the dataset, then merge: convert vectors to polar coordinates;... BgGPT 7B (Mistral); Dataset: OpenWebText, English Wikipedia, GitHub repositories, a range of instruction finetuning datasets (IFT) = 15.1B tokens.... Smaller IFTx4 = effective 17.1B. OpenWebText Web 8.5B, Wiki-En: 4.6B; github code: 1.35B; OpenHermes-2.5 IFT: 357M; SlimOrca IFT: 197M; MetaMathQA IFT: 85M; CodeInstructions IFT: 20M.

Capybara - multi-turn conversation; PyTorch, DeepSpeed; 64 Nvidia H100, lear.rate 10E-5, batch=512 continued pretraining & 256 for sueprvised finetuning. cos decay 0.1*max_lr, max(100,0.01*total_steps) lin.warmup.

Besided BgGPT: adapting Llama-3-8B to German (small improvement); magnitude of weight change after merging; 15K, 25K steps; merging methods: SLERP, LINEAR, MODEL STOCK. Doubling or halving the batch size can also reduce the catastrophic forgetting, but are more costly than BAM.

Tests, datasets: ARC-Easy (5197), ARC-Challenge (2590) (Clark et al., 2018) - science exam questions. MMLU: multitask NLU (lang.understanding), 57 tasks, 14079 test samples. GSM8K - math reasoning; MathQA - multiple choice math reason. Belebele - multiple choice reading comprehension. TriviaQA, XNLI: contradict, neutral, entails. EXAMS (see above), PAWS - paraprhrases; MGSM -

translated from GSM8k. Google translate for the benchmark problems & answers, 2143 manual translations, Winogrande. MON: 10088 exam questions, 4 choices (no geometry: images), not public. 50K Bulgarian translated samples from the OpenHermes-2.5, 10K - MetaMathQA, 2K code + Bulg.instr. CodeAlpaca. Manual inspection for bad translations e.g. rhyming. 5% manually translated or adjusted. Challenge for avoiding data contamination. Bg validation set: 40K examples, 30K from news articles, 10K mix of dialogs, questions and answers. English: 25K random samples from FineWeb-Edu dataset, 7K from aeXiv papers, 3K PubMed, 5K books from Project Gutenberg.

Training params: MISTRAL-7B and LLAMA-3-8B, 8192 context, sequence packing, no trunctation. 1E-5 max lr, continued pre-training, batch 512/256, for 4M/2M tokens. ... up to 7000 tokens per second per GPU ~ 448K/s, 64 NVIDIA H100 GPUs (8 nodes x 8 GPUs) with InfiniBand and 224 available CPU cores per node. Total cost with exploratory: ~80K H100 GPU hours. Reduction of 30% with the tokenizer extension. https://huggingface.co/datasets/INSAIT-Institute/winogrande-bgeval?row=71
?cost of training – unknown; if $2-3/h (1); 8xH100: 18, 24 , 28.98$/h; https://www.vultr.com/pricing/#cloud-gpu 8 x H100 640 GB, 224 vCPU 2048 GB RAM, 32TB storage 15TB bandw. $3000/GPU/hr ~ but only 730 h/month (in 12.2024) if 80000/730 ~ 109,6 ~ $328,8K; 80000/64 ~ 1250/GPU ~ 52 days Cost of 8 DGX 8xH100

* See also **LLEMMA: AN OPEN LANGUAGE MODEL FOR MATHEMATICS,** Zhangir Azerbayev et al., 3.2024 https://arxiv.org/pdf/2310.10631 for the description of the learning protocol, preparation of the data, specializing a pretrained Code Llama and replicating a surrogate dataset from Pile etc. and the diversity of the domains for better results, p.2-3; see also the training of BLOOM, an attempt to recreate a GPT-3-scale open LLM with a documented procedure:
* **BLOOM: A 176B-Parameter Open-Access Multilingual Language Model**, BigScience Workshop: Teven Le Scao et al., 11.2022/6.2023, https://arxiv.org/abs/2211.05100

* Зорница Козарева, Zornitsa Kozareva - Zori

https://www.crunchbase.com/person/zornitsa-kozareva ПУ „Паисий Хилендарски“, Amazon, Google, Facebook … Множество патенти. Обработка https://scholar.google.com/citations?user=bsi48IQAAAAJ&hl=en

Първите и изследвания са свързани с разпознаване на именувани обекти чрез групиране (клъстериране; named entity recognition, NER)[12]. Ранни приложения на мултимодални езикови модели и в роботиката (2019). Някои нейни работи и теми по профила и в Линкдин и др.

* **Cluster analysis of named entities**, Z.Kozareva, 2003

* **Cluster Analysis and Classification of Named Entities,** Z.Kozareva, 2004

* **How can Context and Semantic Information help a Machine Learning Word Sense Disambiguation**, Sonia Vázquez, Zornitsa Kozareva & Andrés Montoyo, 11.2007 … разрешаване на многозначност чрез съчетаване на методи, основани на знания и машинно обучение. WordNet Domains[13] – групиране на отделните значения в смислови 200 области и класифициране по този признак, за да се намали броят на значенията, защото някои от тях са близки, напр. за „банка“; в работата представят нов ресурс: Relevant Domains (уместни области) и модул за скрит семантичен анализ (Latent Semantic Analysis) https://link.springer.com/chapter/10.1007/978-3-540-76631-5_95

* **Few-shot Learning with Multilingual Generative Language Models, Nov 1, 2022** Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O'Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, Xian Li

---

[12] Може би в първи курс? бях на нейна лекция, в която разказваше за бъдещата си работа в тази сфера на специализация по Еразъм, за която предстоеше да замине. Срещнахме се веднъж и в Улвърхамптън, не знам по какъв повод беше там през 2007 г.

[13] Magnini, B., Strapparava, C.: Experiments in word domain disambiguation for paralell texts. In: Proceedings of SIGLEX. Workshop on Word Senses and Multi-linguality (2000) https://aclanthology.org/W00-0804.pdf

* Using Relevant Domains Resource for Word Sense Disambiguation, S. Vázquez, A. Montoyo, G. Rigau, 2004 – context vectors (различни от онези в невронните представяния), за 250-те области (Music, Acoustics, Law…) от по-горната статия. Бел. ТА: подобни области могат да възникват постепенно като гроздове към зърно и към тях да се закачат конкретни контексти, места от конкр. опит/място/URI, @всчк. https://web.archive.org/web/20100622044236/http://adimen.si.ehu.es/~rigau/publications/ic-ai04-vmr.pdf

https://arxiv.org/abs/2112.10668
Тодор Михайлов, Зорница Козарева, Веселин Стоянов; 21 автора – 1/7 от тях българи. Вж. бел. за конф. CLIB 2024. Виж приложение **„B“:** многоезично обучение в контекст (Multilingual In-context Learning Formulation), продължение на Brown, 2020 за „few shot learning”: описание на задачата и няколко примера (вид „prompt engineering”); XGLM: 564M, 1.7B, 2.9B, 7.5B сравнен с GPT3 до 6.7B; слоеве: 24,24,48,32, контекст: 1024, 2048, 2048, 4096. Обучение до 500B токена. Речник: 250K токена unigram, SentencePiece; multinomial distribution. Zero-shot learning**,** Few-shot learning: Тестове с маски, пропуснати думи или откъси – Cloze-type test, cloze-style template, [mask] … FLORES-101 MT task https://en.wikipedia.org/wiki/Cloze_test

* **Efficient Large Scale Language Modeling with Mixtures of Experts**, 20.12.2021

* **Few-shot Learning with Multilingual Language Models**, 2021

* **Natural Language Grounded Multitask Navigation**, 12.2019, Xin Wang, … , Zornitsa Kozareva …(6) “…we introduce a generalized multitask navigation model that can seamlessly be trained on language-grounded navigation tasks such as Vision-Language Navigation (VLN) and Navigation from Dialog History (NDH)…” reinforced cross-modal matching (RCM); Multitask-RCM … Reward Shaping .. Distance to Goal room – роботът трябва да стигне до стая, в която се намира целевият предмет, а не до точка. Ранни модели, които все още не използват „основни“ (foundation) ГЕМ като PaLM и др. и преобразители, а двупосочни LSTM (bidirectional); goal-progress – напредък към целта в метри като мярка.

* **Visually Grounded Interaction and Language (ViGIL**),

* **Sentiment Prediction using Collaborative Filtering**, 2013

* Cause-Effect Relation Learning, 2012 https://aclanthology.org/W12-4107.pdf (чрез шаблони, првила, изрази в текста: „X causes Y, A triggers B, C leads to D…, result” is-a, part-of… построяване на графи; виж „търсене в литературата“ ..

* **Semantic class learning from the web with hyponym pattern linkage graphs**. In Proceedings of ACL-08: HLT, pages 1048–1056 Zornitsa Kozareva, Ellen Riloff, and Eduard Hovy, 2008 – https://aclanthology.org/P08-1119/ semantic class learning … 1.popularity/2.productivity, hyponym patterns + pattern linkage graphs … “X and other Ys”, “Ys such as X” … Seed: class name + one instance → automatically finds other instances (fish, singers, US states) …: 1: 2: exhaustive search (reckless bootstrapping) and then ranks the candidates using the outdegree scoring …

* Веселин Стоянов – Veselin Stoyanov

https://scholar.google.com/citations?hl=en&user=xdfWqboAAAAJ&view_op=list_works&sortby=pubdate

* **RoBERTa: A Robustly Optimized BERT Pretraining Approach,** Y Liu, ... V Stoyanov ... 2019 arXiv preprint arXiv:1907.11692
XLM-R, .. Facebook, Tome AI
https://azbuki.bg/news/novini-2024/broj-37-12-18-09-2024-g/chatbotovete-ne-sa-bezgreshni/

Конференция Компютърната лингвистика в България "Computational Linguistics in Bulgaria " https://www.facebook.com/CompLingInBulgaria/

2024 г. - 30 г. секция "Компютърна лингвистика" в БАН

* Преслав Наков

Компютърен лингвист. Известен като състезател по програмиране и съавтор на книгата „Алгоритми=Програмиране++". Започва с със скрит семантичен анализ и автоматично извличане на информация и филтриране на спам и др.[14] и използване на уеб като като корпус. Впоследствие се фокусира върху откриването на „фалшиви новини" и др.
https://scholar.google.bg/citations?hl=bg&user=rzJnkjgAAAAJ&view_op=list_works&sortby=pubdate

* Getting better results with latent semantic indexing, P Nakov, 6.2000
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=96e62b650e6b95b8edee17d10197d40fd1d97517#page=164
* Latent semantic analysis of textual data, 2000, P Nakov
https://www.researchgate.net/profile/Preslav-Nakov/publication/250268056_Latent_semantic_analysis_of_textual_data/links/55be46e908ae092e966510cc/Latent-semantic-analysis-of-textual-data.pdf
* Using the web as an implicit training set, Preslav Nakov, Marti Hearst, 2005
https://biotext.berkeley.edu/papers/hlt-emnlp05-nakov.pdf
* Nakov, P., & Hearst, M. A. (2005). Search engine statistics beyond the n-gram: Application to noun compound bracketing.
* Nakov, P., Hearst, M., & Hearst, M. (2005). A study of using search engine page hits as a proxy for n-gram frequencies. Proceedings of Recent Advances in Natural Language Processing (RANLP'2005)
https://lml.bas.bg/ranlp2005/DOCS/RANLP2005.pdf#page=360

---

[14] Building an Inflectional Stemmer for Bulgarian, P.Nakov, 2003
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=47104ac49462c8b4181429dfe12a5ea36a6037ce * Finding Good Answers in Online Forums: Community Question Answering for Bulgarian, T Mihaylova, I Koychev, P Nakov, I Nikolova, 2016

* Non-Parametric Spam Filtering based on kNN and LSA, Preslav Ivanov Nakov, Panayot Markov Dobrikov, 2004
https://lml.bas.bg/~nakov/selected_papers_list/nakov_smb04_spam.pdf

Свързани статии: (Banko & Brill 01) Michele Banko and Eric Brill. Scaling to very very large corpora for natural language disambiguation. In Proceedings of ACL, 2001. https://aclanthology.org/P01-1005 https://aclanthology.org/P01-1005.pdf

Ранна работа в статистическата езикова обработка в голям мащаб – един милиард думи. В Интернет има „стотици милиони думи“ … 1-billion word training corpus ... news, scientific abstracts, government ..., literature etc. ... 1000 times > than earlier experiments: „then, than“, „among, between“; „to, two“ … active learning (select samples for annotation from the unannotated pool) - seed learner (a family of learners) → run over unlabeled samples -→ more useful if more uncertain classification label (different decisions); unsupervised learning: HMM POS (Marialdo 1994), Yarowski "seeds for WSD", 1995; Niam et al: topic classifier 1998

* (Brin & Page 98) Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. Computer Networks, 30(1-7):107–117, 1998 (Google)
* Nigam, N., McCallum, A., Thrun, S., and Mitchell,T. (1998**). Learning to classify text from labeled and unlabeled documents**.

* **Using context-window overlapping in synonym discovery and ontology extension** Marıa Ruiz-Casado, Enrique Alfonseca and Pablo Castells, 2005, Department of Computer Science Universidad Autonoma de Madrid, (RANLP 2005) – **измерване на подобие** http://alfonseca.org/pubs/2005-ranlp1.pdf https://lml.bas.bg/ranlp2005/DOCS/RANLP2005.pdf (p.450) Similarity metric ... Distributional Semantics hypothesis – the meaning of a word w is highly correlated to the contexts where w appears. Definitions of synonymy .. can appear in the same contexts while preserving its truth value ... ; Measuring similarity between context: Vector Space Model , context length – a sentence or a window; bag of words (the vector is the presence of given words in the context – their index; no positional information); hyponym, hyperonym, synonym; Pointwise Mutual Information – words appearing near each other ... Synonym discovery; Ontology - “explicit specifications of a conceptualisation”; ontology building, enrichment, population; distance metric, based on co-occurrence information → clustering; Formal Concept Analysis; supersense categories; pattern extraction and matching; dictionary definitions analysis; hyperonymy discovery and extraction; top-down beam search ... TA: Cmp(transformers, late 2010s+ NLP)
* M. Rajman and A. Bonnet. Corpora-based linguistics: new tools for natural language processing, 1992
* P. D. Turney. Mining the web for synonyms: PMI-IR versus LSA on TOEFL, 2001

* Gregory Grefenstette, **Evaluation Techniques for Automatic Semantic Extraction: Comparing Syntactic and Window Based Approaches** … term variability, knowledge-rich & knowledge-poor; **4 megabyte corpus**; syntactic context of each word, regular grammar; *"frequently occurring events can be more finely analyzed than rarer ones. … For frequent words, finer grained context such as that provided by even rough syntactic analysis, is rich enough to judge similarity. For less frequent words, reaping more though less exact information such as that given by windows of N words provides more information about each word. For rare words, the context may have to be extended beyond a window, to the paragraph, or section, or entire document level, as Crouch (1990) did for rarely appearing words."*

* Gregory Grefenstette. Extracting semantics from raw text, implementation details. Heuristics: the Journal of Knowledge Engineering, 1993
* G. Grefenstette. Sextant: Exploring unexplored contexts for semantic extraction from syntactic analysis. 1992
* C. J. Crouch. An approach to the automatic construction of global thesauri. Information Processing and Management, 26(5):629-640, 1990.

* Iliyan Zarov – Илиян Заров - Llama2, Llama3 ,...

https://dblp.org/pid/209/9698.html https://www.toptal.com/resume/iliyan-zarov
Информацията за него е свързана с образование в Англия.
https://huggingface.co/docs/transformers/en/model_doc/llama2
и:

* Тодор Борисов Михайлов - Todor Mihaylov - Meta AI, Llama2, Llama3 https://dblp.org/pid/166/2019.html

*** Knowledge-Enhanced Neural Networks for Machine Reading Comprehension, Department of Computational Linguistics,** 8.2021, Heidelberg University https://archiv.ub.uni-heidelberg.de/volltextserver/34352/1/Thesis_Todor_Mihaylov_Camera_Ready.pdf
Докторска дисертация. Богато съдържание за бързо навлизане в обработката на естествен език, машинно разбиране на език, отговаряне на въпроси и пр. Discourse-Aware Semantic Self-Attention encoder – relations(discourse units, events, arguments, co-referents*) – with discourse-semantic annotations, QA, ConceptNet, .. discourse relation sense disambiguation … discourse, semantic types, question types; "Semantic role labeling (events) improves who, when; intra-sentential Explicit discourse relations – why, where."; Coreference; non-explicit (implicit) disc.rel.; pre-training, fine-tuning, transfer learning; adapters, BERT.

External knowledge: KB & corpus. KnowBERT,2019 w. WordNet & Wiki. Retrieve-

and-augment 2020; Open QA tasks; Penn Discourse Treebank; Google News dataset; CNNs for sentence classification; LR – logistic regression; Semantic Features: Word Embeddings: Word2Vec; Sent.match.Discourse Analysis, semantic relationships, information extraction: Temporal.Async.Precedence (two events, “before”), Expansion.Conjunction (“and, but, or”), EntRel (people, places, things, concepts) – entity relationships, Contingency.Cause.Reason, Comparison.Contrast.
See 4.2, fig. 4.1. Skillful Reader: NER, Question Type Classification (TREC, QTC dataset by Li and Roth (2002a) – 6 coarse and 50 fine-grained types), Paraphrasing (lexical & syntactic), Textual Entailment … mixing embeddings of Document & Question … Prediction layers; Semantic Role Labeling (who did what to whom: verbs and arguments ARG0, … CoNLL 2005 SRL dataset for sequence labeling), Entity Description Classification (DbPedia: “Company, Athlete, Artist”); Discourse Relation Sense Classification (CoNLL 16 ST), Paraphrasing Detection (Quora, MRPC) …cmp: progressive neural networks (Rusu et al., 2016); NLI – inference (entailment, contradiction, neutral relation) – SNLI dataset (Bowman et al., 2015) – common knowledge; MNLI (Williams et al., 2018) – many domains. Microsoft Research Paraphrase Corpus (MRPC) (Dolan and Brockett,2005); Quora Question Pairs – 400K

* *See Ruslan Mitkov, anaphora resolution;* see *Question Answering;* cmp: 2006 …

* Boris Velichkov, Borislav Kapukaranov, Ivan Grozev, Jeni Karanesheva, Todor Mihaylov, Yasen Kiprov, Preslav Nakov, Ivan Koychev, Georgi Georgiev:
SU-FMI: System Description for SemEval-2014 Task 9 on Sentiment Analysis in Twitter. SemEval@COLING 2014: 590-595

* **Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge,** Todor Mihaylov, Anette Frank, 2018
https://aclanthology.org/P18-1076/ https://aclanthology.org/P18-1076.pdf

Reading comprehension (RC): a passage of text – answer questions about it. cloze-style; Cloze-style RC – placeholders (missing words). Common Nouns dataset; Open Mind Common Sense (OMCS, Singh et al. (2002) ~ 630k facts (2024: ConceptNet ~ 2M); (subject, relation, object) … https://conceptnet.io/
https://github.com/commonsense/conceptnet5/
https://huggingface.co/datasets/conceptnet5/conceptnet5
https://en.wikipedia.org/wiki/Open_Mind_Common_Sense M.Minsky et al., crowd- “The snow is cold”. → ConceptNet. https://en.wikipedia.org/wiki/Never-Ending_Language_Learning

ConceptNet 5.5: An Open Multilingual Graph of General Knowledge Robyn Speer et al., 12.2018, https://arxiv.org/pdf/1612.03975
OpenCyc, DBPedia, Open Multilingual WordNet (Bond and Foster 2013)

• **Discourse Relation Sense Classification Using Cross-argument Semantic Similarity Based on Word Embeddings.** Todor Mihaylov, Anette Frank (2016**).**
• Neural Skill Transfer from Supervised Language Tasks to Reading Comprehension.

Workshop on Learning with Limited Labeled Data (LLD), Todor Mihaylov, Zornitsa Kozareva, Anette Frank (2017) at NIPS 2017

* **Момчил Хардалов**, https://mhardalov.com/uploads/Momchil_CV.pdf

https://scholar.google.com/citations?user=DfXsKZ4AAAAJ&hl=en

* Класически трудове по разрешаване на многозначност чрез използване на контекста и корпуси – Word Sense Disambiguation WSD, Context, Corpus Linguistics #WSD

* Yarowsky, D. (1995) Unsupervised word sense disambiguation rivaling supervised methods, 1995. *…one-sense-per-discourse hypothesis , 37,232 examples (hand-tagged over a period of 3 years), seed examples dictionary-based approaches, including Lesk (1986), Guthrie et al. (1991), Veronis and Ide (1990), and Slator (1991). ... "Hearst, 1991: bootstrapping to augment training sets for a supervised sense tagger. Train a fully supervised algorithm on hand-labelled sentences, applied the result to new data and added the most confidently tagged examples to the training set." Schiitze (1992) has pioneered work in the hierarchical clustering of word senses. .. generate up to 10 sense clusters and then manaually assigned to a fixed sense label, hand-inspection of 10-20 sentences per cluster... in Yarowsky: first deciding the groups (in the example: "manifacture plant", "plant life")*

* Yarowsky, David "Word-Sense Disambiguation Using Statistical Models of Roget's Categories Trained on Large Corpora," in Proceedings, COLING-92, Nantes, France, 1992.

* Yapowsky, David, "One Sense Per Collocation," in Proceedings, ARPA Human Language Technology Workshop, Princeton, 1993.

*** Hearst, Marti, "Noun Homograph Disambiguation Using Local Context in Large Text Corpora**," in Using Corpora, University of Waterloo, Waterloo, Ontario, 1991. https://people.ischool.berkeley.edu/~hearst/papers/oed91.pdf CatchWord; context similarity: orthographic, syntactic, and lexical features; multi-million word corpora; ... propose the use of WordNet; "Hindle and Rooth's indisputable observations"; Zernik 1991

* Kelly, E. & P. Stone (1975). Computer recognition of english word senses, volume 13 of North-Hol land Linguistics Series. North-Holland, Amsterdam.

* Zernik, 1991 – P(думи в прозорец от двете страни на целевата дума), дали са

съседни, честоти в корпуса ...subject code marking.
* Schiitze, Hinrich, "Dimensions of Meaning," 1992.
*** Lesk, Michael, "Automatic Sense Disambiguation: How to tell a Pine Cone from an Ice Cream Cone**," Proceeding of the 1986 SIGDOC Conference, Association for Computing Machinery, New York, 1986. – Речниковият метод

* Choueka, Y. & S. Luisgnan (1985). Disambiguation by short contexts. Computers and the Humanities, 19(3):147{157

* Hirst, G. (1986). Semantic Interpretation and the Resolution of Ambiguity. Cambridge University Press, Cambridge.

* Mitchell, T. M. (1980). The need for biases in learning generalizations. Technical Report CBM-TR-117, Rutgers University, Department of Computer Science.

* Zernik, U. (1991). TRAIN1 vs. TRAIN2: Tagging word senses in corpus. In RIAO 91 Conference Proceedings, pages 567{585, Barcelona, Spain.

* Wilks, Y. A., D. C. Fass, C. ming Guo, J. E. McDonald, T. Plate, & B. M. Slator (1990). Providing machine tractable dictionary tools. Journal of Computers and Translation, 2.

* Hindle, D. & M. Rooth (1991). Structural ambiguity and lexical relations. In Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics

* Havasi, C., Speer, R., & Alonso, J. B. (2007). ConceptNet 3: A Flexible, Multilingual Semantic Network for Common Sense Knowledge. https://web.media.mit.edu/~lieber/Teaching/Common-Sense-Course/ConceptNet-3.pdf

* Barriere, C., & Ménard, P. A. Multiword noun compound bracketing using Wikipedia, 2017 https://aclanthology.org/W14-5708.pdf

Разрешаването на многозначност в Smarty

"Най-интелигентният речник" от 2007 г. „Smarty", T.Arnaudov, R.Mitkov също извършваше WSD, но по прост начин с „торба от думи" (bag of words) и размито сравнение с нормализация по статиите с примери от WordNet (използван 2.0 по технически причини). Следващата стъпка беше да се търси в свързани с тях думи и пр., но не беше изпълнена.

*** Бележки за конференцията CLIB 2024 в Сф** от Т.Арнаудов**:**
https://youtu.be/Ewc7bMbc7SY?t=23171

Ivelina Stoyanova: Semantic Features in the Automatic Analysis of Verbs of Creation in Bulgarian and English

16:20 – 16:45 – Ivelina Stoyanova, Hristina Kukova, Maria Todorova, Tsvetana Dimitrova: **Multilingual Corpus of Illustrative Examples on Activity Predicates Verbs:** 15 semantic classes: motion, emotion, communication, change, contact, creation, consumption, competition, body... exclude: idiomatic, phrasal v. (ходя на лов), light verb constructions (търси отговор) **…** Frames, hierarchical (Inheritance, use, subframe, ...) Causation
BulEnAC (Clause-Aligned Corpus)
Bulg-English parallel corpus

Language Resources: 15:55 – 16:20 – Svetlozara Leseva: A 'Dip-dive' into Motion: Exploring Lexical Resources towards a Comprehensive Semantic and Syntactic Description: FrameNet, WordNet, SemCor, BulSemCor, VerbNet, VerbAtlas Ontology of Activity Predicates for linguistic modeling...

Трудове от 2016 г: https://dcl.bas.bg/clib2018/wp-content/uploads/2016/09/clib-2016-proceedings-v03.pdf#page=62

* Programming languages, program synthesis & verification, compilers and code optimization, formal verification, static analysis, interpreters, concurrency… #programlanguages #programsynthesis

*** Веселин Райчев**

*** Мартин Вечев**[15]

*** Светослав Караиванов и др.**

* **Vesselin Raychev, . “Learning from Large Codebases”,** 2016, PhD thesis https://files.sri.inf.ethz.ch/website/people/veselin/raychev_thesis.pdf Supervisor: Martin Vechev. Papers: p. vii – viii e.g.: “Deep Code”, CRF – conditional random fields, ... Top 3 PhD papers ...

*** Svetoslav Karaivanov**, V.Raychev, M.Vechev, **Phrase-Based Statistical Translation of Programming** Languages, 2014 https://files.sri.inf.ethz.ch/website/papers/onward14.pdf … Parallel data collection, Word alignment, Phrase table construction … „C#: Console . WriteLine ("Hello World!") ; Java: System . out . println ("Hello World!") ;… Compute score (probability) P(C# phrase, Java phrase) … BLEU score for evalation (matching tokens) … - phrase table for both translation probabiliy and reverse translation probability; max-likelihood estimation; smoothing with lexical weighting; short/long phrase penalty … Beam-search, grammar from partial translation. Prefix Grammars, rules of a context-free grammar e.g. terminal, nonterminal, orExpr, andExpr, parenthesizedExpr, specailExpr (* ? +) … Explore the possible continuations … e.g. pre(S → S1|S2) → pre (S1 → aS1b|ab), pre(S2 → cS2d|cd) } … They used: Mining Translations … they request an “Amazon Turk”-like approach for creating parallel corpora for translation like for NL. They used: Db4o, Lucene, Hibernate, Quartz and Spring (both for Java & C#), 21K translation pairs. <tree> → <tree>; BLUE (86%-90%), Parse & Compile rate (57-98—99% & 49-69%) … *Performance:* “*Word alignment* for the collection of parallel data: ~50 hours for 20K sentence pairs on Ubuntu, 2.13 GHz Xeon E7-4830 (8 cores, 16 threads). *Transl.:* 2 h to gen. 30-best transl. for 980 methods in the eval. data. “*adding programming-language specific features improves the precision of the system. We believe that further research on the problem should focus on experimenting with statistical techniques combined with deeper semantic features.*” **Note Tosh:** C# & Java are similar in many ways and a lot

[15] М.В., подкаст „Свръхчовекът?“, опитна разработка на виртуална машина на Java около 1999-2000 г. в България (или 2001-2002? Научна статия от 2002 г. @Вси: првр

of the code is appropriate for almost direct port, unlike C++ to, or from Java, or C++ to Python. The conclusion is correct and also obvious, see Michail Bongard, 1967, Михаил Бонгард, 1967: Проблема узнавания (in English "Pattern Recognition") and the repeating of his thoughts by F.Chollet.

* **Chameleon: Adaptive Selection of Collections**, Ohad Shacham, Martin Vechev, Eran Yahav https://csaws.cs.technion.ac.il/~yahave/papers/pldi09.pdf semantic profiling TVLA - flexible static analysis

* **Automatic verification of determinism for structured parallel programs,** Vechev, M., Yahav, E., Raman, R., & Sarkar, V. (2010). https://www.cs.rice.edu/~vs3/PDF/VYRS10.pdf task-parallel programs … check if: independent memory accesses; tool: DICE; sequential analysis …

Refactoring with synthesis, Veselin Raychev, Max Schäfer, Manu Sridharan, Martin Vechev, 2013/10/29 https://manu.sridharan.net/files/OOPSLA13Refactoring.pdf synthesis from examples, Local refactorings on small fragments, heuristic search, Resynth plugin Eclipse; Correct Local Refactoring; Synthesize Local Refactoring Sequence …

* **Code completion with statistical language models,** Veselin Raychev, Martin Vechev, Eran Yahav, 2014 https://csaws.cs.technion.ac.il/~yahave/papers/pldi14-statistical.pdf - частични програми с пропуски („дупки", holes, незапълнени неща); статистически модел + класически програмни техники; Slang; съчетаване на N-gram (3-gram) и RNN; вж 8.др: MatchMaker, MAPO, Strathcona; code search .

* **Automatic inference of memory fences,** Michael Kuperstein, Martin Vechev, Eran Yahav, 2010/10/20 – for concurrent programming; automatic formal verification; "relaxed" (or "weak") memory models (RMMs) – out of order & non-atomic execution; memory fence, memory barrier; over-fencing vs under-fencing (worse performance vs unexpected execution); load balancing; finte-state programs, SAT-solving, transition system, error state identification; ordering constraints; execution buffers, reordering box, local instruction buffer. constraint solving: L – labels, $l_1 < l_2$ (l_1, l_2 – labels in the program; l_1 must execute before l_2) ; optimality & correctness; fender, blender; blocking queue; Work-Stealing Queue; fence-types: store-store, store-load, load-load, load-store; Abstract Memory Models (AMMs) ; Goal: Program, safety spec, memory model: P,S,M → P' with fences satisfying S under M, while minimizing the performance impact. Constraint representation as Binary Decision Diagrams, a transition system for P under M. Constraint language F. Graph of the states &

transistions, transitions system; avoidable tr. – can add fences, vs unavoidable transitions; delay set analysis

* **"How to Make a Multiprocessor Computer That Correctly Executes Multiprocess Programs",** Leslie Lamport –"sequential consistency"

* **Synthesis for Concurrency,** M.Vechev, EranYahav, Greta Yorsh, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=24b78da101dfac0d208e4ce893b89dacc3071a45

* **Race Detection for Web Applications**, Boris Petrov, Martin Vechev, Manu Sridharan, Julian Dolby https://www.cs.columbia.edu/~junfeng/12fa-e6121/papers/js-race.pdf
"happens-before relation", WebRacer; Verification; concurrency, asynchrony; deferred scripts

* **Predicting program properties from" big code"**, Veselin Raychev, Martin Vechev, Andreas Krause, 1.2015 https://janvitek.org/events/NEU/7580/papers/more-papers/jsnice15.pdf … CRFs, разпознава модели на код, преобразува кратки имена на променливи в дълги, … the first applying Conditional Random Fields CRFs in the context of programs; dependency network m <n,m,rel> – connects unknown properties to known ones; structured joint prediction; i += t → dependency <i,t,L+=R>; Maximum a Posteriori query (MAP) – given a program x, find y = argmax … the most likely program properties; Z(x) – partition function; name prediction JS code JSNICE;

* **Programming with "big code"**, Martin Vechev, Eran Yahav, 2016/12/28, Journal Foundations and Trends in Programming Languages, https://www.nowpublishers.com/article/DownloadSummary/PGL-028 github, stackoverflow as corpus of code, statistical learning: CRFs (conditional random fields)_ …
@Вси: ?Т, прлж [,,] +"Abstract interpretation, Compilation and interpretation techniques, Domain specific languages, Formal semantics, including lambda calculi, process calculi, and process algebra, Language paradigms, Mechanical proof checking, Memory management, Partial evaluation, Program logic, Programming language implementation, Programming language security, Programming languages for concurrency & parallelism; Program synthesis, Program transformations, optimizations, verification; Runtime techniques for programming languages; Software model checking, Static and dynamic program analysis, Type theory and type systems

* **Learning programs from noisy data**, Veselin Raychev, Pavol Bielik, Martin Vechev, Andreas Krause, 1/2016
https://files.sri.inf.ethz.ch/website/papers/popl16.pdf regularized program generator, : Iterative Synthesis Algorithm, data sampler: part of the dataset, not all (some of them

noisy); reduce search space; noise bound = 0.2 → 1 error out of 5 samples; 4.1. program generator with errors: examples input→output; regularizer, punishes longer programs; empirical risk minimization (ERM) task over a discrete search space of programs … Related: Boolean program synthesis, Quantitative p.s., Statistical code completion, Discriminative learning – classification, not generation; Core sets, genetic alg., dataset cleaning (of the noise), Probabilistic programs (Psketch, mixture of Gaussians)

* **Fine-Grained Semantics for Probabilistic Programs,** Benjamin Bichsel, Timon Gehr & Martin Vechev, 2018 https://link.springer.com/chapter/10.1007/978-3-319-89884-1_6 … expressive probabilistic programming language

* **Programming with "Big Code": Lessons, Techniques and Applications**, Pavol Bielik, Veselin Raychev, and Martin Vechev, 2015 – probabilistic models & clustering; BLEU score; need for semantics; prediction approach: CRF, discriminative log-linear classifier for structured prediction .. p.4

* **Differentiable abstract interpretation for provably robust neural networks**, Matthew Mirman, Timon Gehr, Martin Vechev, 2018 https://proceedings.mlr.press/v80/mirman18b/mirman18b.pdf

* **Probabilistic model for code with decision trees**, V Raychev, P Bielik, M Vechev, ACM, 10.2016, OOPSLA 2016 https://pavol-bielik.github.io/data/papers/oopsla16-dt.pdf Domain specific language DSL over abstract syntax trees (TGEN); recognize the right context of similar usages; others: DeepSyn, ID3; DEEP3 JavaScript, Python; to their knowledge – the first work to use decision trees for learning programs

* **ABSTRACTION-GUIDED SYNTHESIS,** Inventors: Martin Vechev, Eran Yahav, Greta Yorsh, 2011/10/20, Patent office, US, Patent number, 20110258606, Application number, 12/762002 - verification of parallel programs, concurrency; https://patentimages.storage.googleapis.com/d7/46/1e/c6faf2fd69be92/US8495588.pdf

* **Abstraction-Guided Synthesis of Synchronization**, Martin Vechev IBM Research Eran Yahav IBM Research Greta Yorsh IBM Research, 1/2010 https://www.researchgate.net/profile/Eran-Yahav/publication/220997414_Abstraction-Guided_Synthesis_of_Synchronization/links/09e4150cf4413a77a0000000/Abstraction-Guided-Synthesis-of-Synchronization.pdf inferring correct and efficient synchronization in concurrent programs, GUARDIAN; generating atomicity constraints; accumulating atomicity constraints by iteratively eliminating invalid interleavings; atomicity predicates; schemes/lattices looking like a guitar tabulature with circles on the intersections; parity abstraction; partial abstract transition system <S,T,Init >: States, transitions, initial state; src(t), dst(t); trace п – sequence of transitions; syntax: assignments, non-deterministic choice, conditional jump/goto, sequential composition, parallel composition, atomic section. No dynamic threads,

nested atomic sections & parallel composition inside atomic sections …

* **CGCExplorer: A Semi-Automated Search Procedure for Provably Correct Concurrent Collectors**, Martin T. Vechev, Eran Yahav, David F. Bacon, Noam Rinetzky, 2007 https://www.math.tau.ac.il/~maon/pubs/pldi07-cgc.pdf

* VECHEV, M., AND YAHAV, E**. Deriving linearizable fine-grained concurrent objects.** In PLDI (2008), pp. 125–135.

* VECHEV, M. T., YAHAV, E., BACON, D. F., AND RINETZKY, N. **Cgcexplorer: a semi-automated search procedure for provably correct concurrent collectors**. In PLDI (2007), pp. 456–467. Garbage collectors for C#, Java etc. *automatic exploration of a space of concurrent mark-and-sweep collectors*
* VECHEV, M. T., YAHAV, E., AND YORSH, G. **Inferring synchronization under limited observability.** In TACAS (2009), pp. 139–154

* **Bug Localization with Statistical Models**, Pavol Bielik,Svetoslav Karaivanov, Veselin Raychev, Martin Vechev, Christine Zeller, 2015
Silq: A high-level quantum language with safe uncomputation and intuitive semantics, B Bichsel, M Baader, T Gehr, M Vechev, ACM PLDI 2020
https://files.sri.inf.ethz.ch/website/papers/pldi20-silq.pdf

* **Phrase-based statistical translation of programming languages, Svetoslav Karaivanov**, V Raychev, M Vechev, ACM Onwards 2014, 159, 2014
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=211c48b664b4cb02b816b88bdfbdf2a098672d34 sentence alignent, prefix grammar; parallel data collection, word alignment → method pairs, aligned method pairs, phrase table; n-gram model;
 * KOEHN, P. Statistical Machine Translation, 1st ed. Cambridge University Press, New York, NY, USA, 2010.
 * KOEHN, P., OCH, F. J., AND MARCU, D. Statistical phrase-based translation. In NAACL'2003 - Volume 1.

* **Robustness certification with generative models**, M Mirman, A Hägele, P Bielik, T Gehr, M Vechev, 4.2020 - @намери бел. на Т.А. и вклч.
https://arxiv.org/abs/2004.14756 continuous transformations between images via latent-space interpolation. – пораждане на междинни положения на образи чрез интерполация на скритото пространство на теглата (завъртания на лице и др. rotation); система за верификация APPROXLINE, със сигурни граници в P(x); probabilistic bounds, APPROXLINE, EXACTLINE .. .piecewise-linear activation functions … VAE; По-ранни работи: PROVEN (Weng et al., 2018), Dvijotham et al. (2018a); сертифициране за детерминистични свойства

* **Programmable Synthetic Data Generation,** Mark Vero, Mislav Balunovic, Martin Vechev**,** 23 Sept 2023 (modified: 11 Feb 2024) – ProgSyn, the first programmable synthetic tabular data generation method. Differential privacy (DP) or fairness https://openreview.net/pdf?id=KTL534o7Ot Fair Classification; medical data, …

* **Programming with" big code": Lessons, techniques and applications**, P Bielik, V Raychev, M Vechev, 2015, 1st Summit on Advances in Programming Languages (SNAPL 2015)

* **Differentiable Abstract Interpretation for Provably Robust Neural Networks**, Matthew Mirman, Timon Gehr, Martin Vechev, 2018
2. Robustness and Sound Approximations – point $x{:}R^{\kappa}$; a ball of radius epsilon around a point assigns the same class to all points. п-robust … labeled training examples … worst-case adversarial loss; abstract interpretation (Cousot & Cousot, 1977): abstract domain D, abstraction function α; concretization function γ; $P(R^d) \rightarrow D(R^p)$, $p \in N$; abstract transformer (computable function); overapproximates; interval domain; zonotope, total error … line segments as Hybrid zonotopes ...

* Cousot, P. and Cousot, R**. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints**. In Symposium on Principles of Programming Languages (POPL), 1977.
https://www.di.ens.fr/~cousot/COUSOTpapers/publications.www/CousotCousot-POPL-77-ACM-p238--252-1977.pdf
-34*456 -(+)*(+) (-)*(+) (-) … the results of the abstract execution give some info on the actual computations; consistent abstract interpretations form a lattice … abstract program properties – complete semilattice; order-preserving functions…

Topic: Static program analysis of programs (without execution); simplification; loss of precision; soundness: must be correct in the concrete execution; concrete domain; abstraction function alpha and concertization f. gamma. Program semantics as fixpoints (stable states, points of equations) - widening: jumping to more general abstract state vs, narrowing: get a more precise form from an over-approximated fixpoint. **Static analysis:** type checking, data flow analysis, security, bug detection, verification(prove specification properties, check safety conditions, find optimization opportunities). Abstract domains: represent selected program properties - simpler than the concrete program states, but they must preserve the relevant inf. for analysis. Usages: dead code elimination, constant propagation, loop optimization, proving lack of runtime errors, checking security properties, verifying concurrent programs; null pointer detection, array bounds checking, memory leak detection
TA: Note that AST, Abstract Syntax Trees are slightly different domain about the syntax, syntactic structure, formal grammars for parsing and compilation, originating in 1960s; while the Cousot's Abstract interpretation is about the semantics as integrity and safety, verification, program analysis.

* Patrick Cousot. **Définition interprétative et implantation de languages de programmation.** Thèse de Docteur Ingénieur en Informatique, Université Joseph Fourier, Grenoble, France, 14 Décembre 1974.
https://www.di.ens.fr/~cousot/COUSOTpapers/CousotTheseDi1974.shtml
* P.Cousot. Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique des programmes (1978)
* **Patrick Cousot. 2021. Principles of Abstract Interpretation**. MIT Press
https://mitpress.mit.edu/9780262044905/principles-of-abstract-interpretation/
* **The Best of Abstract Interpretations**, 7.1.2025, Roberto Giacobazzi, Francesco Ranzato
https://dl.acm.org/doi/pdf/10.1145/3704882 *", the best possible abstract interpretations of programs. abstraction and concretization maps; bca – best correct approximation is not compositional; inherent intensional nature of abstract interpretation; program p, abstract domain A; inductive abstract semantics; least domain refinement …* [16] *Example: The abstract domain Sign Z 6 cases; abstracted to 3: (>=0, !=0 >=0); then to other 3: (Z<0, Z=0, Z>0) with some relations: (Z>=0, Z<=0) Z=0; Z>=0 Z >0; Z<=0 Z < 0 …*

* **Learning a static analyzer from data**, Pavol Bielik, Veselin Raychev, Martin Vechev, 2017 https://arxiv.org/pdf/1611.01752
DSL for static analysis rules, synthesis alg., oracle, counter-example guided learning, Equivalence Modulo Abstraction (EMA): Semantic-preserving transformations (dead-code insertion, variable renaming, constant modification, side-effect free expression): to prevent overfitting; for robustness; cmp equivalence modulo inputs (EMI) – compiler validation, correctness of the compiler; similar to EMA + control flow restructuring, expression rewriting (x+1 → 1+x)

* **AI²: Safety and Robustness Certification of Neural Networks with Abstract Interpretation**, Timon Gehr, Matthew Mirman, Dana Drachsler-Cohen, Petar Tsankov, Swarat Chaudhuri∗, Martin Vechev
https://www.cs.utexas.edu/~swarat/pubs/sp2018-ai2.pdf

* **Adversarial Robustness for Code**, Pavol Bielik, Martin Vechev, 2020
https://proceedings.mlr.press/v119/bielik20a/bielik20a.pdf

Model that *abstains* (Liu et al., 2019) from making a prediction when uncertain, effectively partitioning the dataset into two parts: one with accurate and robust predictions, and it doesn't in the other and remains robust; adversarial training (Goodfellow et al., 2015) to the domain of code; refine the representation used as input to the model by learning the parts of the program relevant for the prediction to reduce the number of places that affect the prediction; finally: an alg. for train. multiple models, each learning a specialized representation that makes robust predictions on a different subset of the dataset.
The program is a sequence of tokens; the NN uses a selection function to evaluate the

[16] Patrick Cousot. 2021. Principles of Abstract Interpretation. MIT Press
https://mitpress.mit.edu/9780262044905/principles-of-abstract-interpretation/
…

confidence and decide is the prediction above particular certainty. The adversarial training: applying valid modifications of the program $\delta \subseteq \Delta(x)$, e.g. changing identifiers (“width” to “h”, “height” to “w”) and evaluating the predictions which by default may become wrong – if the model is biased with specific identifiers, e.g. “width” is usually abbreviated to “w” and “h” may confuse it. The adversarial training aim is not to minimize the expected loss on the original distribution E(x,y)~D[loss((f, gh)(x), y)], but the expected adversarial loss: max ... (x + δ), y)]

* Liu, Z., Wang, Z., Liang, P. P., Salakhutdinov, R. R.,Morency, L.-P., and Ueda, M. Deep gamblers: Learning to abstain with portfolio theory. In Advances in Neural Information Processing Systems 32, NeurIPS’19, pp. 10622–10632. 2019.
* Goodfellow, I. J., Shlens, J., and Szegedy, C. Explaining and harnessing adversarial examples. In 3rd International Conference on Learning Representations, ICLR’15, 2015.

* **LARGE LANGUAGE MODELS ARE ANONYMIZERS**, Robin Staab, Mark Vero, Mislav Balunovic, Martin Vechev, 2024 https://openreview.net/pdf?id=xIfEb7q9Rh privacy, adversarial anonymization; tools: Presidio, Azure Language Studio; high-accuracy predictions of personal attributes from online posts; personal data, personal identifiable information PII, GDPR; privacy-utility tradeoff; Author Profiling: identify key author attributes (gender, age etc.); Text Anonymization Benchmark (TAB) (Pilán et al., 2022); PersonalReddit dataset - real-world online comments & human-labeled personal attribute inferences

TA: the identification is only of “as if” the author had a particular age, location etc., if she was honest, because she also can *deceive* and intentionally share misinformation in order to confuse an “adversarial” reader.

* **Hiding in Plain Sight: Disguising Data Stealing Attacks in Federated Learning**, Kostadin Garov, Dimitar I Dimitrov, Nikola Jovanović, Martin Vechev, 2023 https://arxiv.org/pdf/2306.03013 Malicious server (MS) attacks, Gradient leakage attacks – undetected data leakage; data stealing, secret embedding & reconstruction (SEER) – distributed NN training; biasing training to particular objectve e.g. memorizing particular data, image etc.; “memory via decoder”; Can the clients detect the attacks – by checks of the model updates …

* Fischer, M., Balunović, M., Drachsler-Cohen, D., Gehr, T., Zhang, C., & Vechev, M. (2019). **DL2: Training and querying neural networks with logic.** Deep Learning with Differentiable Logic; logic – as a loss function; declarative constraint language; loss = 0 if the constraints are satisfied; it is differentiable almost everywhere; optimize with projected gradient descent (PGD) – for robustness constraints. https://github.com/eth-sri/dl2

* **BgGPT 1.0: Extending English-centric LLMs to other languages**, A Alexandrov, V Raychev, DI Dimitrov, C Zhang, M Vechev, K Toutanova, arXiv preprint

arXiv:2412.10893   3   2024

* **Prompting Is Programming: A Query Language for Large Language Models**, L.Beurer-Kellner, M.Fischer, M.Vechev, 2023 - Language Model Programming (LMP), LMQL; final and follow abstractions; model-specific token masks; deoding: greedy (argmax), sampling, full, beam search; masked – exclude selected tokens at selected positions; few-shot prompting: “translate English to Bulgarian: “a dog barks => куче лае; горещ чай => hot tea; … сирене => “; Multi-Part Prompting; meta-prompts: query expansion; scripted beam search; chain-of-thought prompting [Wei et al. 2023]; tasks: arithmetic reasoning, odd one out (кое е излишно, различно); date understanding; Python; ReAct;

> **Note: ReAct (Reason+Act):** use a search engine, access a knowledge base, perform a calculation with a tool, call a function to get data or perform an operation; answer the question if the LLM believes it has enough info;1. Observation: ”What’s the current weather in Plovdiv?”. 2. Thought: “I need to find the current weather for Plovdiv. I can use a weather API to get this information.”. 3. Action: Use a weather API with location “Plovdiv”. 4. Observation: the returned value from the API: “Current weather is: Sunny, 25 degrees Celsius.” 5. Thought: “Now I have the weather info. I can formulate the final answer.”. 6. Action: “The current weather in Plovdiv is Sunny, 25 degrees Celsius”.
> Can solve math problems, logic puzzles, complex reasoning tasks; make decisions using multi-step processes, gather external information, do strategic planning; retrieve information: research, fact-check, data analysis. Separate thought from action, maintain logical flow, document observations (explainability), consider alternatives; build on observations, adjust based on feedback; refine …
> * **ReAct: Synergizing Reasoning and Acting in Language Models**
> Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, Yuan Cao,3.2023 https://arxiv.org/abs/2210.03629
> https://www.promptingguide.ai/techniques/react

For tool use (access Wikipedia): external Python scripts, phrases: *Search* and *Finish*; beam search with multiple hypotheses … the loop ends when reaching a Finish action;

* **BEYOND MEMORIZATION: VIOLATING PRIVACY VIA INFERENCE WITH LARGE LANGUAGE MODELS**, Robin Staab, Mark Vero, Mislav Balunovic, Martin Vechev, 5.2024 https://arxiv.org/pdf/2310.07298 T: criticism: The given examples are too “cheesy self-revealing”, using stereotypical phrases for British, Australlian, particular locations etc. as if the user *wants* to *emphasize* this particular identity. “Waiting for a *hook* turn” … - Google search: top results: “Road Rules: The CRAZY Melbourne Hook Turn!”, “Is Melbourne the only place with hook turns?”, “Are there hook turns in Adelaide?”
Privacy Leakage in LLM, memorization, free text inference; PersonalReddit (PR) Dataset: 520 Reddit profiles, 5814 comments year 2012-2016; 8 attribute categories: age, education, sex, occupation, relationship status, location, place of birth, income.

*** Vassil Vassilev, Alexander Penev et al. –** programming languages, compilers, interpreters https://vassil.vassilev.info/

* **Cling–the new interactive interpreter for root 6**, V Vasilev, P Canal, A Naumann, P Russo, Journal of Physics: Conference Series 396 (5), 052071, 61, 2012 https://iopscience.iop.org/article/10.1088/1742-6596/396/5/052071/pdf

Clang, LLVM, C++ interpreter for CERN, REPL read-evaluate-print-loop, rapid application development RAD, experimentation. Replaces Cint. … error recovery; reflection with integration with ROOT

* Brun R and Rademakers F, S 1996 ROOT - An Object Oriented Data Analysis Framework

* Goto M, A 1996 Concept and application of Cint C++ interpreter Interface magazine (Japanese)

* **SolidOpt–Innovative Multiple Model Software Optimization Framework,** VG Vassilev, AP Penev, TK Petrov, IEEE and STRL: The Secondю Conference on Creativity and Innovations in …,1, 2009

* **Migrating large codebases to C++ Modules,** Oksana Shadura-UNL, Vassil Vassilev
https://pdfs.semanticscholar.org/27fa/bff1c2564d6d9c623052708833l3d40e6af6.pdf

* **Clad—automatic differentiation using Clang and LLVM.** Vassilev, V., Vassilev, M., Penev, A., Moneta, L. and Ilieva, V., 2015. In *Journal of Physics: Conference Series* (Vol. 608, No. 1, p. 012055). IOP Publishing.
https://iopscience.iop.org/article/10.1088/1742-6596/608/1/012055/pdf
Вход – математическа функция, описана в код на един програмен език, която се преобразува до производната ѝ на друг език.
https://compiler-research.org/clad/ https://compiler-research.org/team/

*** Konstantin Gizdov –** Константин Гиздов, a Bulgarian researcher

NNDrone: a toolkit for the mass application of machine learning in High Energy Physics, Sean Benson, Konstantin Gizdov, 2017/2019 **-** a toolkit for a drone classifier from any ML classifier, standardise, execute in parallel.
https://arxiv.org/pdf/1712.09114 - "Drone classifier" is a simplified version of a more complex model … *"The ability of a neural network with a continuous, bounded, non-constant activation function to approximate functions to an arbitrary degree has been indeed known since the early 1990s"*

* K. Hornik, **Approximation capabilities of multilayer feedforward networks**, Neural Networks 4 (2) (1991) 251 – 257. doi:10.1016/0893-6080(91)90009-T. URL
http://www.sciencedirect.com/science/article/pii/089360809190009T
https://web.njit.edu/~usman/courses/cs677/hornik-nn-1991.pdf
average performance, closeness of functions,

The theorem about the NN as universal approximators even with just a single hidden layer, given the mentioned above conditions about the activation function etc., given a sufficient amount of hidden units are available (see "requisite variety").

*** Roumen Dangovski** (Румен Данговски) – a Bulgarian Researcher;
equivariant neural networks http://super-ms.mit.edu/rumen/CV.pdf
http://super-ms.mit.edu/rumen.html
Yung-Sung Chuang, Rumen Dangovski, Hongyin Luo, Yang Zhang, Shiyu Chang, Marin Soljacic, ShangWen Li, Scott Yih, Yoon Kim, and James Glass. **2022. DiffCSE: Difference-based contrastive learning for sentence embeddings**. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 4207–4218, Seattle, United States. Association for Computational Linguistics
1. Rumen Dangovski, Li Jing, Charlotte Loh, Seungwook Han, Akash Srivastava, Brian Cheung, Pulkit, Agrawal, Marin Soljaˇci´c. **Equivariant Contrastive Learning** ICLR 2022. doi.
2. Rumen Dangovski, Michelle Shen, Dawson Byrd, Li Jing, Preslav Nakov and Marin Soljaˇci´c. 2020. **We Can Explain Your Research in Layman's Terms: Towards Automating Science Journalism at Scale**. AAAI 2021. doi.
3. Matthew Khoury, Rumen Dangovski, Longwu Ou, Preslav Nakov, Yichen Shen and Li Jing. 2020. VectorVector-Matrix Architecture: A Novel
6. Ivan Ivanov, Li Jing and Rumen Dangovski. 2018. Improving the Performance of Unitary Recurrent Neural Networks and Their Application in Real-life Tasks…

@Вси: рзшр, ?Т, свржк; обх(всчк до дн,врх); ?=… обх(INSAIT)

*** INSAIT** – Institute for Computer Science, Artificial Intelligence and Technology
* **ИНСАЙТ** – Институт по компютърни науки, изкуствен интелект и технологии

* https://insait.ai/   * https://insait.ai/publications/

Българо-швейцарският международен институт „ИНСАЙТ“ по заявени цели, предмет на дейност, стратегия и др. е едно от **многобройните производни и повторения на *оригиналната българска стратегия*** за изследвания и разработка на универсални мислещи машини, общ изкуствен интелект, публикувана от 18-годишния Тодор Арнаудов **в средата на 2003 г**. на конкурс с тема: *„Как бих инвестирал един милион с най-голяма полза за развитието на страната?*[17]*“*, като дори изказванията на представителите на института или в медиите, и описанията на събитията от разширението на дейността им, като роботиката и построяването на суперкомпютърния център, повтарят *буквално* и дословно мисли от оригиналната стратегия от 2003 г.

Има обаче и огледални разлики. (…)

[17] Виж също проекта на Ахмед Мерчев „Кибертрон“, обявен около 2002-2003 г.

Случаят е разгледан в приложението ***„Първата модерна стратегия за развитие чрез изкуствен интелект е публикувана от 18-годишен българин и повторена и изпълнена от целия свят 15-20 години по-късно: Българските пророчества: Как бих инвестирал един милион с най-голяма полза за развитието на страната“***, Т.Арнаудов, 2025 г. и в специализираното приложение *„**Институти и стратегии за изкуствен интелект „на световно ниво“ в Източна Европа и света“***, където е направен по-широк и подробен преглед и сравнение на съдържанието и дейността на голям брой национални стратегии за разработка и развитие на и чрез ИИ. #instituti (ще бъде публикувано на конференцията SIGI-2025)
* https://twenkid.com/agi/Purvata_Strategiya_UIR_AGI_2003_Arnaudov_SIGI-2025_31-3-2025.pdf
* https://www.oocities.org/todprog/ese/proekt.htm
* https://twenkid.com/agi/proekt.htm

* От **„Първата стратегия…“**, Литература и бележки към нея, относно неверни или неточни твърдения на института във връзка с BgGPT, т.нар. *„първи отворен голям езиков модел за български език“* и др.

236. Ранни пораждащи големи езикови модели от типа GPT за езици, различни от английския: български, френски, арабски, испански, португалски, немски, китайски; гръцки, сръбски, румънски, японски – 2020-2021 г. Датата на някои – по дати на файловете с теглата на модела, дата на научна статия и пр. Само френският, арабският, румънският, японският и *българският* са с над 100-тина милиона параметъра. *Румънският* е силен, обучаван на 17 GB-ов корпус. *Само българският* вероятно е разработен от един-единствен човек с бюджет и подкрепа = 0 и авторът представя родната компютърна лингвистика в тази дисциплина като *самозван „хайдутин“,* понеже институциите и по-„елитните“ бойци чакаха до 2023-2024 г. [66] (няма данни [за по-ранни подобни модели]). Сравни с аналогичен случай с ДЗБЕ около 2001-2003 г. и бездействието на ИБЕ на БАН и на останалите филолози от университетите спрямо явленията, срещу които ДЗБЕ се противопоставяше и се опитваше да „призове“ „чети“ [16][40], а *„маститите“ езиковеди* (по определението на Павлин Стойчев, „PC World Bulgaria“, 5.2003 [239]) гледаха безучастно и обясняваха, че това били *„естествени процеси“.* Сравни с бележките за *„Добродетелната дружина и нехранимайковците“* и [40], 2003 г., дали талантите не са имали избор да не учат в „най-престижните университети“ и да развият местните и пр. XLM-R от „Фейсбук“, 11.2019 е по-голям, но в него българският е един от 100 езика, на които е обучаван, и е за класификация и отговаряне на въпроси, а не за пораждане.

Ранни големи езикови модели “GPT“ за разни езици		
Арабски	**1.46 B**	**3.2021**
Френски	**1 B**	5.2021
Румънски	**774 M**	7.2021
Български	**355 M**	**6.2021 – 8.2021, Тош**
Японски	**336 M**	16.8.2021
Японски	**1 B**	20.1.2022
Испански	124 M?	12.2020
Португалски	124 M?	5.2020
Немски	124 M?	11.2020 – 8.2021
Италиански	117 M	4.2020
Китайски	124 M?	11.2020 – 5.2021
Гръцки	124 M?	9.2020
Сръбски	124 M?	7.2021
БАН	**124 M**	**27.6.2023**
INSAIT	**7.3 B**	**2.2024**
GPT	117 M	**6.2018**
GPT2	**1.554 B**	**14.2.2019 (XL) (публик. 11.2019)**

1. Тодор Арнаудов**, GPT2-MEDIUM-BG, Свещеният сметач, ДЗБЕ ~6.2021 – 8.2021, 345M – български –** обучен от нулата на Tesla T4 в Colab [31][46] (безплатно), публикуван метод за обучение – популярен клип за жанра в „Ютюб“ с над 4 хил. гледания и над 30 преки абонати.
* https://huggingface.co/twenkid/gpt2-medium-bg
2. Antoine Simoulin, Benoit Crabbé. Un modèle Transformer Génératif Pré-entrainé pour le ______ français. Traitement Automatique des Langues Naturelles, **6.2021**, Lille, France. pp.246-255. ffhal03265900f https://hal.science/hal-03265900 : – френски **GPTfr-124M и GPTfr-1B** с архитектурата на GPT3. **5.2021**
3. https://huggingface.co/dbddv01/gpt2-french-small - друг френски малък **SMALL 137M,** също обучен в Colab като българския, но с платена услуга Colab Pro.
4. Wissam Antoun and Fady Baly and Hazem HajjARAGPT2: Pre-Trained Transformer for Arabic Language Generation, **7.3.2021** – https://arxiv.org/pdf/2012.15520

Арабски, 4 варианта: 135M, 370M, 792M, 1.46B
(...)

Бележка от страницата на българския модел:
*** Note from GPT2-Medium-BG page on Huggingface:**
* **Other Bulgarian autoregressive models:** an earlier one was a few seconds display of a generation in Bulgarian by a startup called BAIHUI AI in mid 2019. I've written in my blog 1.5B, but I don't remember if they have mentioned a size and now it seems unlikely and unreasonable, they just showed that they can train a model, a team of 3 people, only one of them a ML engineer. There are a few surviving records: my blog post: https://artificial-mind.blogspot.com/2019/07/baihuiai-baihuiai-new-bulgarian-ai.html and info here: https://www.eu-startups.com/directory/baihui-ai/ The company didn't live long. Now it seems reasonable that their model was GPT2-SMALL, as that was the usual choice even 4 years later and even the Bulgarian Academy of Science 2023 model was the small one. I found several other GPT2-SMALL models trained later than this one here, one for poetry, the BAS' from 2023 and maybe a few others. I couldn't get info from the ML engineer of the BAIHUI project ***Mitko Vassilev***: https://mitkox.com/#home
https://www.linkedin.com/in/ownyourai/

***Smarty - Extendable Framework for Bilingual and Multilingual Comprehension Assistants**, Todor Arnaudov, Ruslan Mitkov, LREC 2008
https://github.com/Twenkid/Smarty – „Смарти“ – разширяема рамка за двуезични и многоезични помощници в разбирането (интелигентен речник, подпомогнат превод, автоматичен превод, речници, лексикография); създаден през 2007 г. Най-интелигентният речник

* **Творчеството е подражание на ниво алгоритми: Възможен, бегло начертан път на развитие на Изкуствения разум "Емил",** Тодор Арнаудов, сп. „Свещеният сметач“, бр. 23, 5.2003 г., – пророчески български проект за пораждащ модели за текст и всякакви модалности; мултимодалността и съчетаването на всички видове данни като път към общ разум и др. Част от *„Българските пророчества“* и класическите публикации от Теория на Разума и Вселената (2001-2004). Виж *Първата модерна стратегия* … и *Основния том.*
https://eim.twenkid.com/old/eim22n/eim23/emil04052003.htm
https://www.oocities.org/eimworld/eim22n/eim23/emil04052003.htm

* **Проектът Вседържец – инфраструктура за мислещи машини**, обявен през 9.2022 г. **–** „Вседържец/Vsy - The AGI Infrastructure of "The Sacred Computer" AGI Institute : Custom Intelligent Selective Internet Archiving and Exploration/Crawling; Information Retrieval, Media Monitoring, Search Engine, Smart DB, Data Preservation, Knowledge Extraction,Datasets creation,AI Generative models building and testing, Experiments etc.“
* https://github.com/Twenkid/Vsy-Jack-Of-All-Trades-AGI-Bulgarian-Internet-Archive-And-Search-Engine

Набори от данни за български език

Видеозаписи от заседания на Народното събрание:
Data from the Bulgarian parliament: https://www.parliament.bg/bg/video/ID/42
Стенограми на народното събрание от 1933 г. до 2025 г., до 1991 г. са сканирани pdf, от 1992 г. - текст: https://www.parliament.bg/bg/plenaryst
https://www.parliament.bg/bg/plenaryst/ns/55/ID/4572
8.1.1997 г.: https://www.parliament.bg/bg/plenaryst/ns/55/ID/1373
3.1.1997 г. https://www.parliament.bg/bg/plenaryst/ns/55/ID/2299
* 3.1.1933 г.
https://www.parliament.bg/pub/StenD/20140424040731030119322_31.pdf

@Вси: ОбОбр всчк

(…)

Подлежи на допълване, подобрения, подреждане … И бъдещо пораждане чрез Вси и др. @Vsy

…Следват подобрения, продължения, разширения…

АНЕЛИЯ

приложение към

СВЕЩЕНИЯТ СМЕТАЧ
ТОДОР АРНАУДОВ – ТОШ

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ ИЗКУСТВЕН РАЗУМ И РАЗВИТИЕ НА ЧОВЕКА
ИСТОРИЯ ТЕОРИЯ И ПИОНЕРИ МИНАЛО НАСТОЯЩЕ И БЪДЕЩЕ

от автора на първия света
университетски курс по
Универсален изкуствен разум и
Теория на разума и вселената

THE PROPHETS OF THE THINKING MACHINES
ARTIFICIAL GENERAL INTELLIGENCE & TRANSHUMANISM
HISTORY THEORY AND PIONEERS; PAST PRESENT AND FUTURE

2025