

Remigiusz Kinas

Tytuł:

Opis:

Bielik

Co nowego w Bielikach 2.5 i 3.0

W prezentacji omówię nowe możliwości polskich modeli językowych Bielik v2.5 i v3, które znacząco poszerzają ich zastosowania, czyniąc je jeszcze lepszymi partnerami dla polskiego biznesu. Skupię się na takich zagadnieniach jak dedykowany tokenizer APT4, zaprojektowany z myślą o języku polskim, mechanizm function calling umożliwiający integrację z systemami zewnętrznymi, generowanie strukturyzowanych danych w formatach takich jak JSON oraz jak powstał pierwszy polski model reasoningowy. Oprócz przeglądu nowych funkcji, opowiem również, jak modele te były trenowane w ramach projektu SpeakLeash — jakie dane zostały przygotowane, jakie zastosowano techniki uczenia i co pozwoliło Bielikom osiągnąć obecny poziom.

Agnieszka Karlińska**Marek Kozłowski****Anna Kołos**

Tytuł:

Opis:

PLLuM / HIVE

HIVE AI: wdrażanie modeli PLLuM w administracji publicznej

W wystąpieniu przedstawimy wnioski z projektu PLLuM, w ramach którego powstała rodzina 18 polskich dużych modeli językowych, udostępnionych w lutym 2025 roku. Omówimy, jak zdobyte doświadczenia wpłynęły na prace prowadzone w projekcie HIVE AI – inicjatywie mającej na celu rozwój i wdrażanie modeli PLLuM w polskiej administracji publicznej. Skupimy się na kwestiach związanych z pozyskiwaniem i opracowywaniem danych uczących, dostrajaniem modeli do różnorodnych zadań oraz ich zabezpieczaniem i dostosowywaniem do oczekiwań użytkowników i użytkowników. Omówimy także wdrożeniowe aspekty projektu HIVE AI, w tym integrację modeli PLLuM z aplikacją mObywatel oraz inne planowane rozwiązania wspierające cyfrową transformację administracji publicznej.

Bartosz Naskręcki

Tytuł:

Opis:

WMI UAM

Rozumowanie z AI - fakty i mity

Gwałtowny rozwój dużych modeli języka przez ostatnich kilka lat spowodował, że zaczęliśmy eksperymentować z przeprowadzaniem rozumowań i zauważać potencjalną użyteczność modeli w nauce. W ostatnich dwóch latach pojawiły się modele, które sukcesywnie pokonują kolejne bariery zaawansowanych dedukcji i potrafią do pewnego stopnia emulować pracę naukowców. Na wykładzie opowiem o najnowszych benchmarkach weryfikujących te dokonania. Opowiem czy faktycznie jesteśmy już blisko pokonania przez AI matematyków i co z tego wynika w praktyce naukowej. Zobaczymy jak różne nowe techniki, w tym modele z tokenami rozumowań oraz wieloagentowe osiągają wysoką skuteczność w niektórych typach rozumowań, ale nadal kompletnie zawodzą w innych. Spróbuję postawić tezę dlaczego tak się dzieje i czy czeka nas kres instytucji naukowych.

Paweł Gołąb

Tytuł:

Opis:

Million Monkeys Software

Jak znaleźć mordercę przy pomocy grafów, czyli co Sherlock robił z czerwoną nitką?

Kluczową umiejętnością w pracy śledczego jest zdolność łączenia faktów i znajdowania w zasobie danych powiązań, analogii i odbiegających od normy szczegółów. Tak się składa, że technologią odpowiadającą na wszystkie te potrzeby są grafy. Nie bez powodu potocznie mówimy o "łączeniu kropek". Na filmach detektywistycznych często widzimy dokumenty i zdjęcia połączone przez detektywów czerwoną włóczką (tzw. murder board). Formalnie nazwalibyśmy te rekwizyty grafami wiedzy, czyli typem sieci dostosowanym do reprezentacji powiązań między faktami.

Zastosowanie czerwonej włóczki do łączenia różnego typu dokumentów pokazuje inną ogromną zaletę grafów: naturalne łączenie multimodalnych danych z heterogenicznych źródeł w spójny system analityczny.

Oczywiście, w rzeczywistości śledczy wykorzystują systemy informatyczne, jak popularny i2 Analyst's Notebook, jednak systemy te ograniczają się zazwyczaj jedynie do manualnego modelowania wiedzy.

W prezentacji przedstawiamy Networks Notebook: nasz autorski system wspierający analizę wielomodalnych danych przy pomocy grafów wiedzy. Rozwiązanie w zautomatyzowany sposób przetwarza dane na graf, a następnie udostępnia kilka metod jego analizy.

Sercem systemu jest oparty o graf wiedzy RAG. W odróżnieniu od tradycyjnych architektur opartych na wektorach, które spłaszczają relacje i gubią kontekst informacji, systemy oparte na grafach organizują wiedzę w sieci powiązanych ze sobą elementów, odzwierciedlających relacyjny charakter rzeczywistości. Znacząco zwiększa to jakość pozyskiwanych informacji, ale daje też możliwość analizy grafu z wykorzystaniem języka naturalnego.

Następnie, system udostępnia grafowy odpowiednik analizy drill down: wychodząc z interesującego fragmentu grafu użytkownik może rozwijać stopniowo jego kontekst, dochodząc "po nitce do kłębka" do interesujących odkryć.

Ostatnim punktem są automatyczne analizy oparte o zapytania do grafu wiedzy, których wyniki są dalej interpretowane przez model językowy. W ten sposób system jest w stanie wykrywać między innymi ścieżki przebiegu informacji, luki informacyjne, czy sprzeczne relacje świadków.

Do wyróżniających cech Networks Notebook należą wysoka jakość analizy, możliwość działania w trybie offline dzięki zastosowaniu lokalnych modeli oraz obsługa bardzo dużych sieci bezpośrednio w przeglądarce internetowej.

Nasza prezentacja łączy głęboką wiedzę techniczną z praktycznym case-study. Przedstawiamy technologie, na których zbudowaliśmy system i zasady ich działania: od grafów wiedzy do grafowych systemów RAG. Prezentujemy też działanie systemu w praktyce, na przykładzie analizy znanej powieści kryminalnej, co pozwoli w przystępny sposób zademonstrować jego możliwości.

System oczywiście nie ogranicza się do analiz fikcyjnych wydarzeń: ma zastosowanie w pracy śledczych, audycie, prawie, czy medycynie.

Bartek Roszak

Tytuł:

Opis:

STX Next

Projektu DeepNext – tworzenie multiagentowego, autonomicznego software developera

Podzielę się historią projektu DeepNext – systemu multiagentowego, którego celem jest stworzenie autonomicznego software developera. Opowiem, jak wyglądały początki, jakie wyzwania napotkaliśmy po drodze oraz na jakim etapie rozwoju znajduje się obecnie nasz system. Prezentacja będzie praktycznym spojrzeniem na budowę złożonych agentów AI, integrację narzędzi developerskich i wyciągnięte wnioski.

Piotr Jabłoński Adam Janczewski Tytuł: Opis:	DomData / CSI RAG już mamy – co dalej? Na przykładzie Ferryt AI Hub opowiemy o owocach dotychczasowej współpracy badawczo-rozwojowej pomiędzy DomData a Centrum Sztucznej Inteligencji Uniwersytetu im. Adama Mickiewicza w Poznaniu. W ramach wspólnych działań opracowaliśmy szynę integracyjną AI, która umożliwia szybkie, skalowalne i bezpieczne wdrażanie rozwiązań opartych na architekturze RAG, otwartych modelach językowych (LLM) oraz modelach multimodalnych. Pokażemy, jak Ferryt AI Hub wykorzystuje systemy agentowe, mechanizmy dynamicznego generowania kodu i formularzy, oraz dlaczego samo wykorzystanie LLM nie wystarcza. Zamiast traktować go wyłącznie jako narzędzie do chatowania, wdrażamy system, w którym LLM pełni rolę asystenta LowCode dewelopera, wspierając projektowanie procesów i automatyzację działań. Przedstawimy naszą koncepcję dalszej rozbudowy architektury i kierunki rozwoju platformy, w pełni opartej na rozwiązaniach open source.
Marcin Rzetecki Tytuł: Opis:	BSV Blockchain Blockchain jako fundament dla otwartych danych w erze agentów AI oraz modeli językowych Prezentacja będzie poruszać następujące elementy: 1. Wprowadzenie do blockchaina i jego kluczowych właściwości, takich jak decentralizacja, bezpieczeństwo i niezmienność danych. 2. Znaczenie otwartych danych w kontekście rozwoju technologii AI i modeli językowych oraz korzyści płynące z ich dostępności. 3. Rola blockchaina w zapewnianiu transparentności i zaufania w procesach zarządzania danymi, które są kluczowe dla agentów AI. 4. Przykłady zastosowań blockchaina w usprawnianiu przechowywania i udostępniania otwartych danych w systemach AI. 5. Jak blockchain może wspierać rozwój modeli językowych, zapewniając dostęp do kompletnych i zabezpieczonych zestawów danych. 6. Case study pokazujące integrację blockchaina z AI i modelami językowymi w rzeczywistych zastosowaniach. Tutaj będę chciał pokazać pracę inżynierską, która prowadzi z Waszymi studentami. 7. Wyzwania i ograniczenia związane z wdrażaniem blockchaina jako fundamentu dla otwartych danych, szczególnie w środowiskach AI.
Łukasz Zbytniewski Tytuł: Opis:	Allegro Sztuczna inteligencja bez przygotowanych danych i nadzoru: jak absolwent bez doświadczenia rynkowego W swojej prezentacji pokażę jak technologia wspiera zespół analityków w pracy z danymi w Allegro - od wdrożenia po modelowanie i analizę. Zaprezentuję, jak AI tłumaczy język naturalny na SQL, odciąża zespół w tworzeniu analiz i raportów oraz omówię jego ograniczenia. Na zakończenie zaprezentuję case study, w którym interesariusz - z pomocą AI - weryfikuje hipotezy, identyfikuje przyczyny i znajduje rozwiązania problemu, z jakim się zmierzył.

Dawid Głownia	Carlsberg
Tytuł:	Setting up AI Centre of Excellence
Opis:	W prezentacji przedstawię strategię oraz praktyczne wdrożenia sztucznej inteligencji i automatyzacji w Carlsberg Group, która od ponad 140 lat łączy tradycję z innowacją. Opowiem o utworzeniu Centrum Doskonałości AI i Automatyzacji oraz o u podejściu firmy, które łączy identyfikację wartościowych zastosowań AI z odpowiedzialnym wdrażaniem innowacji. Omówimy metody selekcji projektów AI według kryteriów wykonalności, skalowalności i pożądanego przez użytkowników, a także sposoby na szybkie i bezpieczne skalowanie rozwiązań w całej organizacji, wspierając przy tym kulturę wzrostu i transformacji cyfrowej. Zaprezentuję konkretne przykłady wdrożeń na bazie systemów Line Coach oraz Sales Coach -inteligentnych narzędzi wspierających efektywność produkcji oraz sprzedaży, które wykorzystują generatywną AI do monitorowania i optymalizacji procesów w czasie rzeczywistym.
Marek Grzenkowicz	Roche
Tytuł:	Produkty GenAI w Roche - 30 miliardów tokenów później
Opis:	W prezentacji omówię moje obserwacje dotyczące rozwoju produktów GenAI w dużym przedsiębiorstwie. Przechodząc między rolami dewelopera i jednocześnie użytkownika narzędzi GenAI, a także dostawcy i konsumenta platformy GenAI, wskażę różnorodne wyzwania, przed którymi stoi każda z tych grup. Zagłębimy się w szczegóły rozwoju GenAI i porównamy go z tradycyjnymi projektami IT i Data Science. Analiza tych różnic pomoże zdefiniować zakres umiejętności potrzeby podczas budowania produktów GenAI.
Maria Jędrzejczak	UAM, STX NEXT, ARAAI
Paweł Dudko	
Tytuł:	AI Act a duże modele językowe
Opis:	Już niedługo, bo od 2 sierpnia 2025 r., nastąpi rozpoczęcie stosowania przepisów AI Act dotyczących modeli AI ogólnego przeznaczenia. Jakie obowiązki spoczywać będą na dostawcach modeli AI ogólnego przeznaczenia? Czy model AI ogólnego przeznaczenia to to samo co duży model językowy? Na te i inne pytania spróbujemy odpowiedzieć w ramach analizy konkretnych stanów faktycznych. Studium przypadków obejmować będzie takie zagadnienia jak: zakres obowiązywania AI Act ze skutkiem eksterytorialnym (tzn. poza UE), odpowiedzialność twórców i użytkowników LLMów (AI Act i inne), wykorzystywanie LLMów dla biznesowych zastosowań – konsekwencje i obowiązki.

Krystian Zawistowski Tytuł: Opis:	Samsung Agenty AI do automatyzacji interfejsu użytkownika Agenty do obsługi GUI są aktywnym polem badań, jako kolejna generacja asystentów głosowych i obiecujące narzędzie do automatyzacji. Tematem wystąpienia będą rozwiązania upublicznione przez Samsung R&D Polska, m.in. agentowy model na urządzenia mobilne TinyClick, jak również zarys teorii i zastosowań tego rodzaju technologii. Powiązane artykuły: https://arxiv.org/abs/2410.11871v3 https://arxiv.org/abs/2410.11872
Marcin Borzymowski Tytuł: Opis:	RQLabs Syntetyczne dane i agenty AI, czyli jak AI generuje dane dla AI Systemy agentowe potrafią skutecznie radzić sobie z zaawansowanymi wyzwaniami. Czy możliwe jest zatem stworzenie zespołu agentów, który zastąpi ludzi przy generowaniu danych treningowych i to w języku polskim? Myślę, że tak! W trakcie wystąpienia podzielę się doświadczeniami z budowy takiego systemu oraz omówię jakie problemy i wyzwania napotkałem. Przedyskutujemy również, czy syntetyczne dane generowane przez AI i dla AI mają sens i jak je można wykorzystać w praktyce.
Joanna Dziekańska Tytuł: Opis:	Omniviser MapReduce dla LLM Wraz z rosnącą skalą i złożonością dużych modeli językowych (LLM), efektywne potoki przetwarzania danych stają się kluczowe dla obsługi ogromnych ilości tekstu wykorzystywanych do trenowania, dostrajania i wnioskowania. Czerpiąc inspirację z paradygmatów Big Data, w jaki sposób zasady MapReduce—pierwotnie zaprojektowane dla rozproszonego przetwarzania w systemach takich jak Hadoop—mogą zostać zaadaptowane do potoków LLM. Przeanalizujemy, jak operacje ""Map"" mogą być wykorzystywane do równoległego przetwarzania i tokenizacji ogromnych zbiorów danych, podczas gdy kroki ""Reduce"" służą do agregacji wyników, eliminacji duplikatów i optymalizacji przetwarzania danych przez LLMy. Tak jak MapReduce zrewolucjonizowało przetwarzanie Big Data poprzez rozproszenie obciążeń na klastry, tak potoki LLM mogą wykorzystywać podobne techniki do skalowania kuracji danych, ewaluacji modeli, a nawet samego wnioskowania, rozkładając zadania na wiele węzłów obliczeniowych. Podczas sesji omówimy rzeczywiste implementacje, wyzwania i optymalizacje związane z zastosowaniem technik rozproszonego przetwarzania w potokach LLM, łącząc świat inżynierii Big Data z nowoczesnym rozwojem sztucznej inteligencji. Uczestnicy zdobędą wiedzę na temat projektowania skalowalnych, odpornych na błędy potoków LLM, które odzwierciedlają efektywność MapReduce w systemach opartych na danych.
Bartosz Walter Tytuł: Opis:	PCSS Fabryka AI - Piast-AI - Infrastruktura to fundament - nie tylko domu, także inteligencji - Fabryka AI Piast-AI: co, kiedy, dla kogo? - Budowa ekosystemu usług AI: dla każdego coś specjalnego - Polska w drugim rzędzie: blisko czy daleko? - Poznać na tle Polski

Wojciech Darłowski

Tytuł:

Opis:

Beyond.pl

Suwerenna Fabryka AI w Beyond.pl

Wojciech Darłowski, Członek Zarządu Beyond.pl, podczas wystąpienia przybliży projekt uruchomienia na terenie kampusu data center Beyond.pl Fabryki AI wraz z ekosystemem usług: kolokacji AI, GPUaaS, AI PaaS i AlaaS oraz wspomagających usług zarządzanych. Omówione zostaną elementy związane z uruchomieniem Fabryki AI, w tym wdrożenie jednego z najsilniejszych superkomputerów w regionie NVIDIA DGX SuperPOD z kartami graficznymi NVIDIA B200 oraz współpraca technologiczna z NVIDIA i Pure Storage. Podkreślona zostanie rola partnerstw i otwartego modelu współpracy jako fundamentu stworzenia skalowalnego i suwerennego środowiska infrastrukturalnego na potrzeby projektów AI w regionie.

