

Laboratorio Sentenze

Un esperimento di Intelligenza Artificiale

La ricerca dell'algoritmo

Michaela Elisa Jackson

Data Specialist

CSI-Piemonte

ARCHITETTURE E TECNOLOGIE INNOVATIVE PER LA P.A.

DATA STRATEGY, ANALYTICS E GEO SERVICES

Data Governance e Valorizzazione Dati

Torino, 19/10/2023



Giustizia predittiva

Cos'è?

Sicuramente non è questo...

Invece.

Prediction in ambito informatico ha un **significato tecnico specifico**: consiste nel **determinare il valore** di una **variabile** (non necessariamente legata al futuro) in un set di dati, tramite un **modello**.



Questa foto di Autore sconosciuto è concesso in
licenza da [CC BY-SA](#)



Questa foto di Autore sconosciuto è concesso in
licenza da [CC BY-SA](#)

Cos'è la giustizia predittiva?



Questa foto di Autore sconosciuto è concesso in licenza da CC BY-ND

Lo abbiamo chiesto a un'IA

«Se si volesse dare una definizione di giustizia predittiva, che non si focalizzi unicamente sul fatto di prevedere l'esito di una controversia, quale potrebbe essere?»

Risposta:

«La giustizia predittiva è l'utilizzo di **strumenti tecnologici**, in particolare dell'intelligenza artificiale, **per fornire supporto alle decisioni giudiziarie.**»

...

«Ecco alcuni esempi di come la giustizia predittiva può essere utilizzata per fornire supporto...:

- Ricerca di informazioni:...
- Valutazione delle prove:...
- Redazione delle motivazioni:...

...»

Intelligenza Artificiale: un po' di storia

«...what we want is a machine that can learn from experience.»

1950: Alan M. Turing – Computing Machinery and Intelligence

Anni '80: reti neurali **shallow**; le deep presentano ancora problemi legati alle risorse computazionali e al metodo di apprendimento

2016: AlphaGo di DeepMind (Google) batte il campione mondiale di Go: $2,08 \times 10^{170}$ diverse posizioni possibili!

~2019: modelli di rappresentazione del linguaggio per l'italiano, basati su transformer!

oggi: era dell' **Algoritmo**, era dei **LLM** e delle **IA** (OpenAI ChatGPT, Google BARD, Meta Llama,...)

Algoritmi e modelli

L'algoritmo è una procedura, il modello è una rappresentazione.
Alcuni algoritmi possono produrre modelli, altri usano modelli, ecc...

Algoritmo

- Treccani «La parola **algoritmo** deriva dal nome del matematico arabo Muhammad Ibn Musa al-Khuwarizmi ... e **indica una successione di istruzioni per risolvere un problema ...**»
 - *telai tessili → elettrodomestici → robot...*
- La macchina che più di tutte utilizza gli algoritmi è il **computer**: le operazioni che un computer esegue sono frutto dell'**esecuzione di un programma**, ossia la **traduzione di un algoritmo** in un elenco di **istruzioni machine-readable**.
- Oggi gli **algoritmi di machine learning** eseguono una **programmazione automatica** basata su dati di training e i **modelli** ne sono il **risultato**.

Modello

- Treccani «Nel linguaggio scientifico, il termine viene riferito a un'ampia classe di ipotesi e costruzioni complesse e articolate, ..., con cui viene **rappresentato** tutto o in parte l'oggetto di una ricerca che si propone di **organizzare dati e conoscenze**,...»
 - *modelli matematici → di machine learning → del linguaggio...*
- I **modelli di machine learning** permettono di effettuare **classificazioni**, analisi di **regressione**, in generale **predizioni**.
- Oggi i **modelli del linguaggio più evoluti** sono modelli di **deep learning**, basati su transformer: si tratta di **rappresentazioni del linguaggio naturale** che permettono di trasformare un testo in una **forma matematica equivalente** (Generative AI exists because of the transformer).

Valorizzazione dati: vecchi strumenti, nuovi strumenti



In CSI-Piemonte

Ci occupiamo da sempre di **massimizzare l'estrazione dell'informazione** contenuta nel patrimonio di dati della PA, e di **governare i dati** a tale scopo.

Come ci può aiutare l'**intelligenza artificiale** in questo?

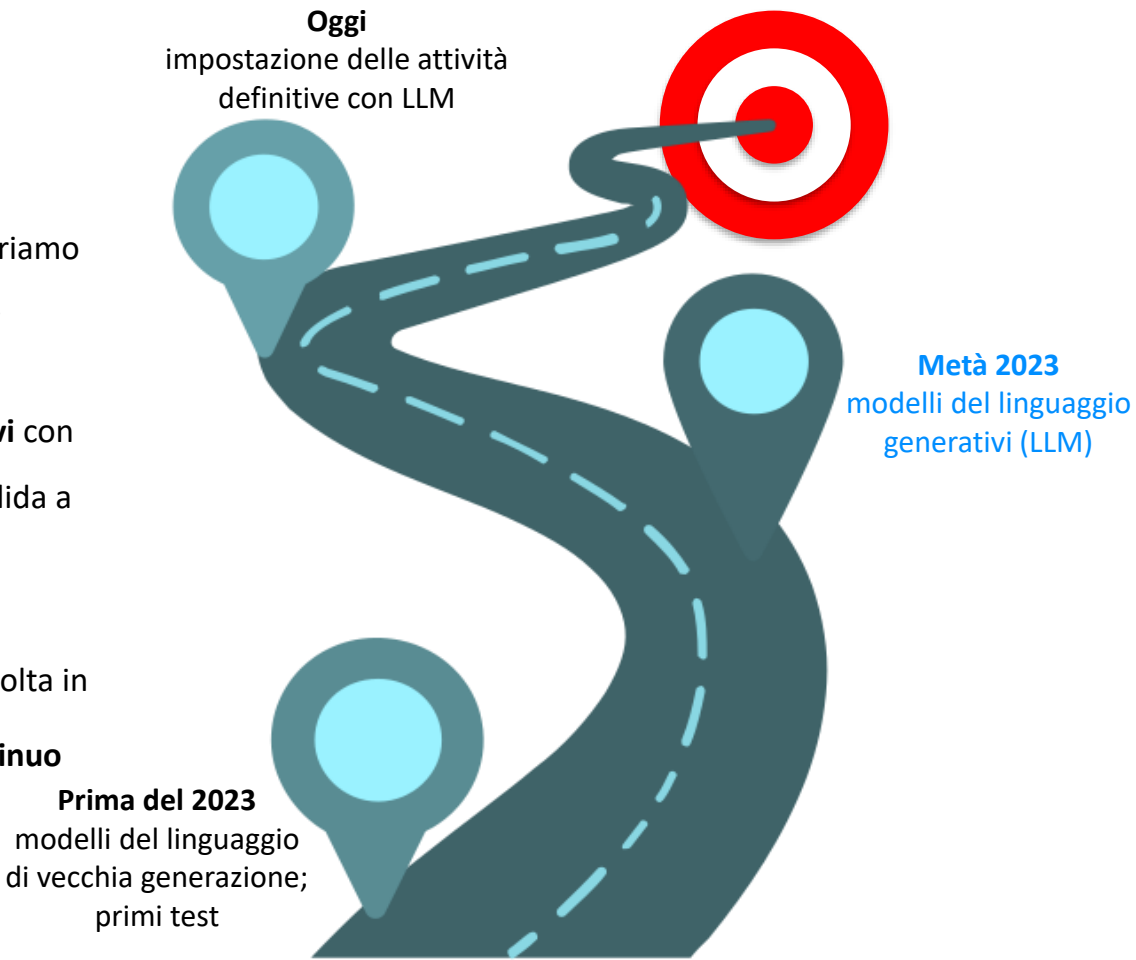
rule-based ---> **data-driven**

Si passa dall'**essere umano** che progetta **a priori** un **algoritmo di regole**, all'**IA**, la quale affianca il professionista col ruolo di **assistente**, essendo in grado di "studiare" i dati e individuarne i migliori **pattern di rappresentazione**.

Milestone della sperimentazione

Considerazioni:

- usiamo **modelli del linguaggio** perché lavoriamo su **testi** (sentenze);
- a **inizio 2023 non avevamo la possibilità** di accedere a modelli del linguaggio di ultima generazione;
- sono stati svolti alcuni **primi test esplorativi** con modelli di vecchia generazione ed è stata impostata una **metodologia** che rimane valida a tutt'oggi;
- a **metà 2023**, a fronte di nuovi strumenti utilizzabili, lo scenario si amplia: **modelli generativi = sintesi**;
- la metodologia viene **integrata** di volta in volta in relazione ai nuovi task realizzabili;
- la percezione è quella di un ambito in **continuo cambiamento**.



Sperimentazione: primi test

The screenshot shows the official website of the Italian Administrative Justice system. The header is blue with the text 'Giustizia Amministrativa' and a search bar. Below the header, there are navigation tabs: 'Giustizia Amministrativa', 'Consiglio di Presidenza', 'Consiglio di Stato', 'CGA Sicilia', 'TAR', 'Studi e approfondimenti', and 'Decisioni e pareri'. The main search area has a text input field with the placeholder 'Che contenga la seguente frase:' and the value 'servizi pubblici'. Below this, there are filters for 'Risultati per pagina:' (set to 20), 'Tipo Provvedimento:' (set to Sentenza), and 'Sede:' (set to Consiglio di Stato). There is also a field for 'Anno e numero provvedimento:' with a dropdown arrow. To the right of these filters are two buttons: 'Annulla' and 'Cerca'. Below the filters, it says 'Trovati 3408 risultati'. On the left side, there are two search results listed with their respective links and details. On the right side, there is a 'Filtra per:' section with a 'DATA' filter showing a list of years and the number of results for each year.

Giustizia Amministrativa

Cerca nel sito

Giustizia Amministrativa Consiglio di Presidenza Consiglio di Stato CGA Sicilia TAR Studi e approfondimenti Decisioni e pareri

Che contenga la seguente frase: servizi pubblici

Risultati per pagina: 20

Tipo Provvedimento: Sentenza

Sede: Consiglio di Stato

Anno e numero provvedimento:

Annulla Cerca

Trovati 3408 risultati

202308433 (CONSIGLIO DI STATO, SEZIONE 6) html
SENTENZA sede di CONSIGLIO DI STATO, sezione SEZIONE 6, numero provv.: 202308433
Numero ricorso: 202004739

202308377 (CONSIGLIO DI STATO, SEZIONE 5) html
SENTENZA sede di CONSIGLIO DI STATO, sezione SEZIONE 5, numero provv.: 202308377
Numero ricorso: 201900104

Filtra per:

DATA

- ☐ 2023 (166)
- ☐ 2022 (192)
- ☐ 2021 (171)
- ☐ 2020 (164)

Elementi

Dati: ~1000 sentenze del Consiglio di Stato, contenenti la frase "**servizi pubblici**", dal 2016 al 2023.

Task: predizione dell'esito delle controversie.

Strumenti: modelli del linguaggio di vecchia generazione.

Motivazioni: per sondare quanto effettivamente arduo sia il task, per prendere confidenza con i dati e per stabilire una metodologia generale.

Sperimentazione: primi risultati

L'**accuracy** misura la capacità di avvicinarsi al risultato specifico. Può essere a scapito della precision: per esempio, si predicono correttamente tutti i respinti, ma a scapito di molti falsi respinti.

Prove di predizione (classificazione binaria dei ricorsi in accolti e respinti)

ESPERIMENTO	DATI DI INPUT	ACCURACY	CONFUSION MATRIX	NOTE
NUMERO 1 - Pipeline tf-idf su FATTO intero + classifier	modello allenato su 1076 sentenze del Consiglio di Stato, contenenti la frase "servizi pubblici", dal 2016 al 2023	0,69	TP=44, FP=18, TN=39, FN=19	benchmark con tecnica tradizionale, nessun modello del linguaggio
NUMERO 2 - Pipeline chunk 200 sovrapposti + sentence transformer distil + classifier	modello allenato su 1076 sentenze del Consiglio di Stato, contenenti la frase "servizi pubblici", dal 2016 al 2023	0,64	TP=39, FP=23, TN=38, FN=20	quantità di dati sufficiente per questa pipeline? incertezza nell'individuazione dell'esito anche solo per il dataset di training
NUMERO 3 - Pipeline frasi della class "popolo" + sentence transformer distil + classifier	modello allenato su 1076 sentenze del Consiglio di Stato, contenenti la frase "servizi pubblici", dal 2016 al 2023	0,71	TP=44, FP=18, TN=42, FN=16	quantità di dati sufficiente per questa pipeline? incertezza nell'individuazione dell'esito anche solo per il dataset di training
NUMERO 4 - Pipeline testo intero con LSG16k + CLS + classifier	modello allenato su 1076 sentenze del Consiglio di Stato, contenenti la frase "servizi pubblici", dal 2016 al 2023	0,65	TP=39, FP=23, TN=40, FN=18	i dati sembrano essere troppo pochi per questa pipeline con modello LSG16k incertezza nell'individuazione dell'esito anche solo per il dataset di training
NUMERO 5 - Pipeline testo intero con LSG16k + AutoClass	modello allenato su poco più di 4000 sentenze del Consiglio di Stato, senza filtro sul contenuto	0,48		il dataset di training non è significativo i dati sembrano essere troppo pochi per questa pipeline con modello LSG16k incertezza nell'individuazione dell'esito anche solo per il dataset di training

Cambio di passo: LLM e attività definitive

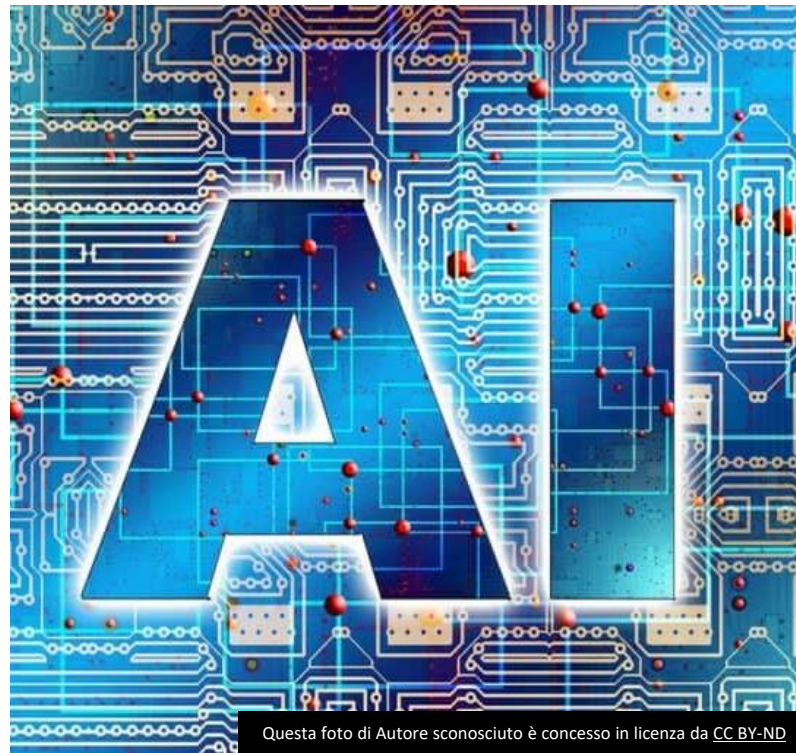
LLM

Parliamo di **IA effettiva** perché, oltre ad essere in grado di rappresentare il linguaggio, i **LLM** sono oggi il **core** di **agenti digitali** capaci di **generare** testo, **imitando** l'uomo.

I **modelli generativi** ci permettono di "aggredire" la questione della **sintesi**.

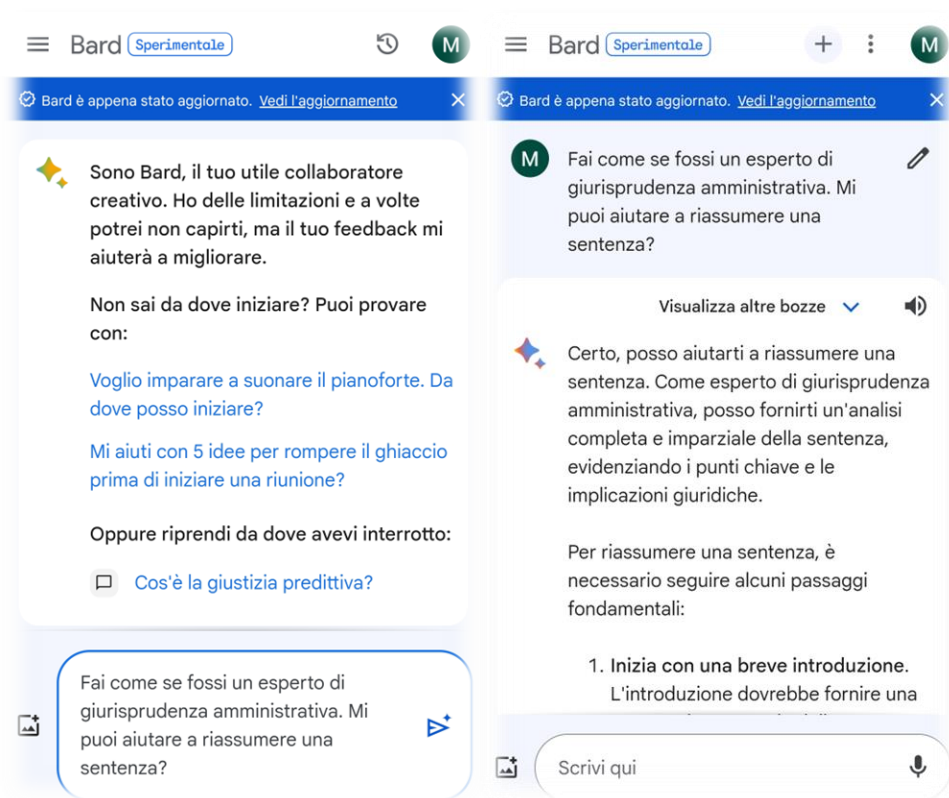
Attività definitive

1. **sintesi automatizzata** di sentenze/fatti
2. **similarità automatizzata tra fatti**
3. **predizione automatizzata** dell'esito delle controversie



Questa foto di Autore sconosciuto è concesso in licenza da [CC BY-ND](#)

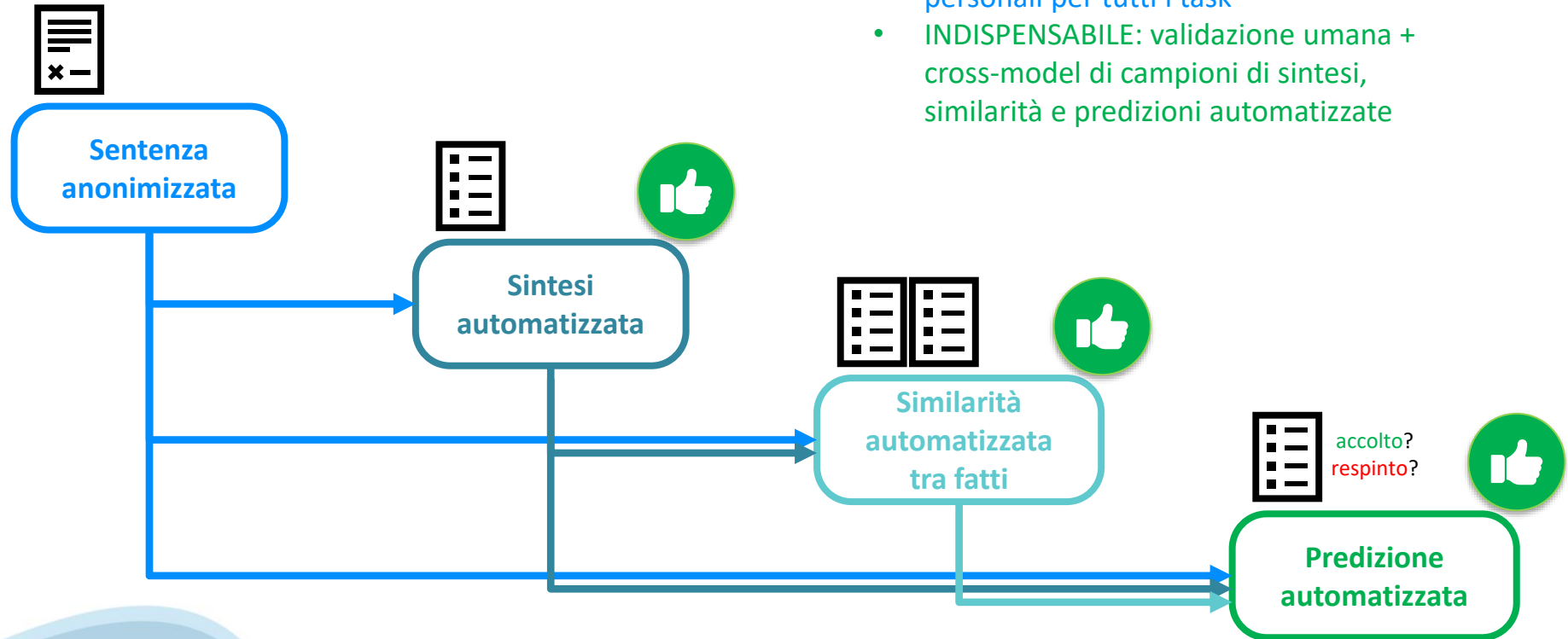
Copilot!



L'IA è uno strumento:

- può essere **integrato** in algoritmi;
- è in grado di ispezionare in poco tempo una **quantità di dati vastissima**;
- è a **supporto** del professionista e **non ne sostituisce** il lavoro;
- è straordinariamente utile come **copilot**, sempre sotto la **supervisione umana**;
- trattandosi di un approccio **statistico**, il **controllo umano** a posteriori è **sempre necessario**.

Flusso attività



- PUNTO DI ATTENZIONE: anonimizzazione all'origine dei dati personali per tutti i task
- INDISPENSABILE: validazione umana + cross-model di campioni di sintesi, similarità e predizioni automatizzate

Risultati: sintesi automatizzata

Richiesta di generazione di sintesi (sentenza N. 00007/2016 del C.d.S.) **tre volte con le stesse istruzioni.**
Occorre **affinare le istruzioni** e lavorare sui **parametri** che influiscono sul **comportamento generativo.**
La **ripetibilità** della risposta aumenta agendo su alcuni parametri: la **temperatura** misura la **creatività.**

temperatura 0,7	«La sentenza impugnata dichiara il difetto di giurisdizione del giudice amministrativo sul ricorso riguardante l'assegnazione di ore di sostegno a un alunno disabile. L'appello contesta tale decisione e chiede l'annullamento del provvedimento impugnato. La decisione finale è ancora in attesa.»	«La sentenza afferma che il giudice ordinario ha giurisdizione sul ricorso riguardante il piano educativo individualizzato (PEI) di un alunno con disabilità. L'appello contro la decisione è stato trattenuto in decisione.»	«La sentenza impugnata dichiara il difetto di giurisdizione del giudice amministrativo sul ricorso riguardante l'assegnazione delle ore di sostegno a un alunno disabile. L'appello è trattenuto in decisione.»
temperatura 0,2	«La sentenza impugnata dichiara il difetto di giurisdizione amministrativa sul ricorso riguardante l'assegnazione di ore di sostegno a un alunno disabile. L'appello contesta questa decisione e chiede l'annullamento del provvedimento. La decisione finale è ancora in attesa.»	«La sentenza impugnata dichiara il difetto di giurisdizione amministrativa sul ricorso riguardante l'assegnazione di sostegno a un alunno disabile. L'appello contesta questa decisione e chiede l'annullamento del provvedimento impugnato. La decisione finale è ancora in attesa.»	«La sentenza impugnata dichiara il difetto di giurisdizione amministrativa sul ricorso riguardante l'assegnazione di ore di sostegno a un alunno disabile. L'appello contesta questa decisione e chiede l'annullamento del provvedimento impugnato. La decisione finale è ancora in attesa.»

Risultati: sintesi automatizzata

Il fatto in 150 parole

«Con la sentenza impugnata, il TAR ha dichiarato il difetto di giurisdizione amministrativa sul ricorso presentato dalla madre di un alunno disabile, riguardante l'assegnazione delle ore di sostegno scolastico. Il TAR ha motivato la sua decisione con il rispetto di una sentenza della Corte di Cassazione che stabilisce che la giurisdizione ordinaria è competente per le controversie relative alla fase successiva all'adozione del piano educativo individualizzato (PEI). La madre ha presentato un appello contro questa decisione, contestando la sua correttezza e chiedendo l'annullamento del provvedimento impugnato. La Corte ha rimesso la questione alla Corte di Cassazione per decidere sull'estensione della giurisdizione amministrativa alla fase di esecuzione del PEI.»

Il testo generato **dipende fortemente dalle istruzioni fornite: ruolo assegnato all'IA, contesto fornito, linguaggio usato e richiesto, dettagli sul compito assegnato, limiti imposti, ecc.**

Risultati: altri elementi estratti dal fatto

Principio di diritto

"Il principio di diritto presente nel testo è il seguente: *la cognizione delle controversie relative alla fase successiva all'adozione del Piano Educativo Individualizzato (PEI) deve essere assegnata alla giurisdizione ordinaria.*"

Parole chiave

Sentenza impugnata	Ordinanza istruttoria	Piano educativo individualizzato (PEI)	Alunno affetto da disabilità	Difetto di giurisdizione	Giudice amministrativo	Giudice ordinario	Dirigente scolastico
Ore settimanali di sostegno	Stato di disabilità del minore	Overruling	Cognizione delle controversie	Fase successiva all'adozione del PEI	Giurisdizione ordinaria	Appello	Illegittimità del provvedimento impugnato
Annullamento della sentenza impugnata	Ambito della giurisdizione esclusiva amministrativa	Erogazione di pubblici servizi	Estensione della giurisdizione amministrativa	Fase di esecuzione del PEI	Sospensione dell'esecutività della sentenza impugnata	Pubblica udienza	Decisione trattenuta

Risultati: similarità automatizzata

id_fatto	fatto	result1_id	result2_id	result2_score	result3_id	result3_score	result4_id	result4_score	result5_id	result5_score
324	Si controverte intorno alla delibera meglio indicata in epigrafe	324	320	0.9999938011169434	322	0.9999499320983888	326	0.9996491074562072	318	0.9996464252471924
261	§ 1. Gli appellati (originari ricorrenti in primo grado) sono geni	261	180	0.999660849571228	199	0.9996562004089355	285	0.9996554851531982	219	0.999645888053894
238	§ 1. Le appellate esercitano la responsabilità genitoriale su mi	238	247	0.9994428753852844	248	0.9993582367897034	237	0.999341368675232	236	0.999329686164856
215	§ 1. Gli appellati esercitano la responsabilità genitoriale su un	215	200	0.9999562501907348	188	0.9992653131484984	203	0.9991888403892516	186	0.9991703033447266
1000	1. Giunge all'attenzione di questo Persona giuridica_226 il con	1000	1001	0.9993115663528442	968	0.9991235733032228	967	0.9991199374198914	973	0.9991179704666138
152	1. Con il ricorso di primo grado n. 1017 del 2011 (proposto al T	152	154	0.9996235370635986	150	0.9995778799057008	151	0.9990647435188292	155	0.9989806413650512
151	1. Con il ricorso di primo grado n. 924 del 2011 (proposto al TA	151	155	0.9999117851257324	153	0.9995051622390748	152	0.9990646839141846	154	0.9989117980003356
154	1. Con il ricorso di primo grado n. 819 del 2011 (proposto al TA	154	152	0.9996235370635986	150	0.9994849562644958	153	0.9991999864578248	151	0.9989117980003356
155	1. Con il ricorso di primo grado n. 923 del 2011 (proposto al TA	155	151	0.9999117851257324	153	0.999575138092041	152	0.9989806413650512	154	0.9988144636154176
153	1. Con il ricorso di primo grado n. 818 del 2011 (proposto al TA	153	155	0.999575138092041	151	0.9995052814483644	154	0.9991999864578248	152	0.9987687468528748
372	1.La Persona giuridica_1205 Ireca s.r.l. ha proposto appello av	372	369	0.9993914365768432	373	0.9991024136543274	371	0.999043107032776	370	0.9987687468528748
713	1. Nel giudizio di primo grado la Persona giuridica_1205 qui ric	713	714	0.9999723434448242	715	0.999692976474762	716	0.999647855758667	717	0.9987321496009828
150	Con il ricorso di primo grado n. 1016 del 2011 (proposto al TA	150	152	0.9995778799057008	154	0.9994848966598512	151	0.9985729455947876	155	0.9984580874443054
328	Si controverte intorno alla delibera meglio indicata in epigrafe	328	317	0.9999271035194396	325	0.9999271035194396	320	0.9981775879859924	324	0.998171091079712
785	1. Ricorrendo dinanzi a questo Persona giuridica_428, l'Perso	785	784	0.9997551441192628	756	0.9977298378944396	757	0.9977216720581056	755	0.997719943523407
329	Si controverte intorno alla delibera meglio indicata in epigrafe	329	327	0.9975900650024414	321	0.9966809153556824	323	0.9966803789138794	318	0.9964879751205444
995	1. Giunge all'attenzione di questo Persona giuridica_226 il con	995	966	0.9952558279037476	964	0.9952455759048462	967	0.9952399730682372	973	0.9952350854873656
38	Con la sentenza n. 4270/2010 in epigrafe appellata il Persona	38	44	0.9954236745834352	42	0.9953716993331908	45	0.9952965378761292	43	0.9942506551742554
45	Con la sentenza in epigrafe appellata n. 313/2010 il Persona g	45	38	0.9952965378761292	42	0.9932066798210144	43	0.993039608001709	44	0.9921849370002748
42	Con la sentenza in epigrafe appellata n. 4278/2010 il Persona	42	38	0.9953716993331908	37	0.9953651428222656	45	0.9932066798210144	43	0.9919716715812684
37	Con la sentenza in epigrafe appellata n. 4287/2010 il Persona	37	42	0.9953651428222656	38	0.9933091402053832	45	0.991883397102356	43	0.9915088415145874
43	Con la sentenza in epigrafe appellata il Persona giuridica_8 reg	43	38	0.9942506551742554	45	0.993039608001709	42	0.9919716715812684	37	0.9915088415145874
687	Con la sentenza appellata indicata in epigrafe il Tar per il Luog	687	681	0.9962010383605956	682	0.9948563575744628	688	0.9918232560157776	680	0.99121356010437
44	Con la sentenza in epigrafe appellata n. 4274/2010 il Persona	44	38	0.9954236745834352	45	0.99218487739563	42	0.9917245507240297	43	0.99071604013443
682	1. Con la sentenza impugnata indicata in epigrafe il Tar del Luc	682	688	0.996044397354126	687	0.9948564767837524	680	0.9923378229141236	681	0.9903959035873412
222	1. Con la sentenza in epigrafe appellata n. 162/2010 il Perso	222	222	0.9999999999999999	221	0.9999999999999999	221	0.9999999999999999	221	0.9999999999999999

Punti aperti

1. Qual è il ruolo più appropriato da assegnare all'IA prima di sottoporre un'istruzione specifica? Quale linguaggio deve adottare? Di quale contesto deve tener conto? Di quali limiti?
2. Come possiamo formulare meglio le istruzioni specifiche da sottoporre all'IA per ottenere sintesi più precise e accurate? Su quali elementi della sentenza deve concentrarsi l'IA ("scena della controversia", storia del processo, parti in causa...)?
3. Quale ampiezza di parole deve avere la sintesi di un fatto di una sentenza per essere utile? E per la parte di PQM?
4. Quali tipi di parole chiave potrebbe essere sensato estrarre dal fatto?
5. E' corretto cercare il principio di diritto nel fatto?
6. ...
7. Quali sono i passi da seguire per una validazione ottimale dei risultati delle IA, dei modelli e, più in generale, degli algoritmi impiegati?
8. In che modo devono essere proposti i risultati dell'IA in modo da essere utili, fruibili e verificabili da parte del professionista in ambito giustizia?
9. In attesa dell'AI Act, tenendo conto che l'IA soffre di bias e allucinazioni, quali compiti è meglio non assegnare all'assistente digitale e quali invece sì?

Grazie

This work is licensed under CC BY-SA 4.0.

To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

