

All. 1 – Piano delle attività

- Progetto di ricerca correlato all'incarico di ricerca:

In un contesto mondiale dove i modelli di linguaggio sono ormai onnipresenti, un problema ancora in fase di ricerca è quello della proprietà individuale dei dati usati per l'allenamento, il fine-tuning e l'aggiornamento dei modelli di linguaggio. Inoltre, una volta che questi modelli producono testi in output, essi possono creare confusione con testi scritti da umani. In questo contesto, il progetto di ricerca verterà sull'analisi, lo studio e la creazione di sistemi per il watermarking e l'authorship attribution correlati ai modelli di linguaggio. Si investigherà l'applicazione di tecniche teoriche e pratiche di embedding nei testi, embedding nei pesi dei modelli, authorship attribution con lo scopo di creare nuove tecniche e pubblicare paper scientifici e avanzare la ricerca in questi ambiti.

- Progetto di ricerca correlato all'incarico di ricerca in inglese:

In a global context in which language models have become ubiquitous, an issue that remains under active investigation concerns the individual ownership of the data used for the training, fine-tuning, and updating of language models. Moreover, once these models generate textual outputs, such outputs may be difficult to distinguish from texts written by humans. Against this background, the research project will focus on the analysis, study, and development of systems for watermarking and authorship attribution in relation to language models. It will investigate the application of theoretical and practical techniques for embedding information in texts, embedding information in model weights, and authorship attribution, to publish scientific papers and advance research in these areas.

- Descrizione delle attività di assistenza alla ricerca che saranno affidate al titolare dell'incarico:

Il titolare dell'incarico svolgerà attività di assistenza alla ricerca nell'ambito dello studio e dello sviluppo di metodologie per il watermarking e l'authorship attribution applicate ai modelli di linguaggio. In particolare, contribuirà all'analisi dello stato dell'arte, alla definizione del quadro metodologico, alla progettazione e sperimentazione di tecniche per l'inserimento e il rilevamento di watermark nei testi generati e nei pesi dei modelli, nonché alla valutazione di metodi per la distinzione tra testi prodotti da esseri umani e testi generati automaticamente. Le attività comprenderanno inoltre la costruzione e l'analisi di dataset sperimentali, la validazione dei risultati ottenuti e il supporto alla preparazione di pubblicazioni scientifiche.

- Descrizione delle attività di assistenza alla ricerca che saranno affidate al titolare dell'incarico in inglese:

The holder of the appointment will carry out research assistance activities in the study and development of methodologies for watermarking and authorship attribution applied to language models. In particular, they will contribute to the analysis of the state of the art, the definition of the methodological framework, the design and experimentation of techniques for the insertion and detection of watermarks in generated texts and in model weights, as well as the evaluation of methods for distinguishing between texts produced by human authors and texts generated automatically. The activities will also include the construction and analysis of experimental datasets, the validation of the results obtained, and support for the preparation of scientific publications.

- Piano delle attività:

1. Studio dello stato dell'arte
2. Definizione del quadro metodologico
3. Sviluppo e analisi di tecniche di watermarking testuale
4. Studio del watermarking nei pesi dei modelli
5. Authorship attribution e distinzione tra testi umani e generati
6. Sperimentazione e validazione

• Piano delle attività in inglese:

1. Review of the State of the Art
2. Definition of the Methodological Framework
3. Development and Analysis of Textual Watermarking Techniques
4. Study of Watermarking in Model Weights
5. Authorship Attribution and the Distinction between Human-Authored and Generated Texts
6. Experimentation and Validation.

All. 2 - Criteri di valutazione

a) coerenza del titolo di studio per l'accesso alla selezione con il progetto oggetto del bando (<i>consistency of additional qualifications with the project covered by the call for applications</i>) fino ad un massimo di 25 punti	<i>La commissione valuta l'attinenza del titolo di laurea magistrale o a ciclo unico conseguito da non più di sei anni, con la seguente gradualità di punteggio:</i> - 20-25 punti: pienamente attinente - 15-20 punti: parzialmente attinente - 10-15 punti: non attinente <i>The Committee shall assess the relevance of the Master's degree or single-cycle degree obtained no more than six years prior, according to the following scoring scale:</i> - 20–25 points: fully relevant - 15–20 points: partially relevant - 10–15 points: not relevant
b) coerenza di ulteriori titoli di studio con il progetto oggetto del bando (<i>consistency of additional qualifications with the project covered by the call for applications</i>) fino a un massimo di 10 punti	<i>La commissione ritiene valutabili master (max 2 punti) e corso di perfezionamento (max 2 punti), con la seguente gradualità di punteggio:</i> - 2 punti: pienamente attinente - 1 punto: parzialmente attinente - 0 punti: non attinente <i>The Committee considers Master's programmes (maximum 2 points) and advanced training courses (maximum 2 points) to be assessable, according to the following scoring scale:</i> - 2 points: fully relevant - 1 point: partially relevant - 0 points: not relevant
c) le pubblicazioni presentate in termini di originalità, innovatività, rigore metodologico, rilevanza e congruenza con il progetto oggetto del bando. La Commissione prende in considerazione esclusivamente pubblicazioni o testi accettati per la pubblicazione secondo le norme vigenti, nonché	<i>La commissione valuta nel suo insieme i criteri riferiti ad ogni singola pubblicazione presentata, specificando la seguente attribuzione dei punti:</i> - 0.5 punti: originalità - 0.5 punti: rigore metodologico - 0.5 punti: rilevanza

saggi inseriti in opere collettanee ed articoli editi su riviste in formato cartaceo o digitale con l'esclusione di note interne o rapporti dipartimentali (<i>the publications submitted in terms of originality, innovativeness, methodological rigour, relevance and congruence with the project covered by the call for applications. The Committee will only take into consideration publications or texts accepted for publication in accordance with the applicable regulations, as well as any essays included in collective works and articles published in journals in paper or digital format with the exclusion of internal notes or departmental reports</i>) fino a un massimo di 10 punti	- 0.5 punti: congruenti con il tema di ricerca dell'incarico <i>The Committee shall assess, as a whole, the criteria relating to each individual publication submitted, specifying the following allocation of points:</i> - 0.5 points: originality - 0.5 points: methodological rigour - 0.5 points: relevance - 0.5 points: consistency with the research topic of the appointment
d) altri titoli collegati ad attività precedentemente svolte (es: borse di studio, partecipazione in qualità di relatore a convegni e congressi, conseguimento di premi e riconoscimenti per attività di ricerca, tirocini e stage formativi, ecc.) debitamente attestati (<i>other qualifications related to previous activities (e.g. scholarships, participation as a speaker at conferences and congresses, receipt of prizes and awards for research activity, internships and training periods, etc.) duly certified</i>) fino a un massimo di 15 punti	<i>La commissione valuta le seguenti tipologie di attività precedentemente svolte:</i> - partecipazione in qualità di relatore a convegni e congressi - attività di tutorato in lauree triennali o magistrali in informatica - periodi di ricerca superiori a 4 mesi presso qualificati atenei e istituti di ricerca esteri <i>con la seguente gradualità di punteggio per ogni voce:</i> - 5 punti: pienamente attinente - 2 punti: parzialmente attinente - 0 punti: non attinente <i>The Committee shall assess the following types of previously undertaken activities:</i> - participation as a speaker at conferences and congresses; - tutoring activities in Bachelor's or Master's degree programmes in Computer Science; - research periods of more than four months at recognised foreign universities and research institutes. <i>The following scoring scale shall apply to each item:</i> - 5 points: fully relevant - 2 points: partially relevant - 0 points: not relevant

Criteri di valutazione per il colloquio, che terrà conto anche degli aspetti motivazionali e attitudinali, nonché della capacità di esposizione e di sintesi

(Criteria for the interview, which will take also into account the candidate's motivation and aptitude, as well as their ability to present and summarize information):

- *maturità scientifica: la maturità scientifica sarà valutata sulla base dell'esposizione del proprio percorso formativo e professionale.
(scientific maturity: scientific maturity shall be assessed based on the presentation of the candidate's educational and professional background).*
- *preparazione dei candidati: la preparazione dei candidati sarà valutata sulla base di quesiti, anche in merito al proprio percorso formativo e professionale, volti ad indagare la conoscenza della materia oggetto del bando nonché la capacità di creare collegamenti fra argomenti inerenti.
(preparation of the applicants: candidates' preparation shall be evaluated through questions, including questions relating to their educational and professional background, aimed at assessing their knowledge of the subject matter of the call, as well as their ability to establish connections between relevant topics).*