

Découvrez SEMIA : le projet d'intelligence artificielle générative de l'ONERA !



L'intelligence artificielle (IA) prend une place croissante dans les activités de recherche mais également au sein des services à l'ONERA. L'initiative SEMIA, mise en œuvre en interne, illustre parfaitement cette dynamique. Pour mieux comprendre les enjeux de cette expérimentation, nous avons rencontré Nicolas Trouvé, chef de l'unité Simulation Environnement Modélisation au DEMR et acteur clé du déploiement de cette intelligence artificielle.

30 avril 2025

Pouvez-vous nous présenter l'intelligence artificielle SEMIA et son origine ?

SEMIA a été mis en place dans le contexte du Labo IA notamment avec Aurélien Plyer et Adrien Chan Hon Tong afin de répondre rapidement aux enjeux de l'IA dite « générative », c'est-à-dire des modèles d'IA capables de créer de la donnée qui peut être de différente nature (texte, images, vidéos...). L'utilisation professionnelle de ces outils pouvant amener à manipuler des données internes comme des rapports, documents, codes sources, résultats de mesures il était important de rapidement disposer d'une solution locale qui ne dépende pas des GAFAs afin d'évaluer les différentes applications pour l'ONERA.

Quels sont les principaux objectifs de SEMIA ?

La plateforme poursuit trois objectifs majeurs :

- Permettre au personnel de l'ONERA de se familiariser et de se former à l'utilisation des modèles d'IA génératifs dans un environnement sécurisé.
- Evaluer les cas d'usages de l'IA possibles pour l'ONERA, que ce soit pour les études mais également au sein des services.
- Soutenir les activités d'études et de recherche en intelligence artificielle générative, notamment appliqué au développement informatique mais également en génération ou analyse d'images radar.

L'utilisation de SEMIA est possible pour tous les salariés de l'ONERA. Quelles sont ses principales fonctionnalités et comment peut-on y accéder ?

SEMIA est accessible par tout le personnel de l'ONERA depuis le réseau interne dans un navigateur internet à l'adresse <http://semia:8080/>. Il reproduit une interface d'utilisation très similaire à celles des outils du commerce. Depuis 2024 des formations sont organisées par la DRH est un manuel utilisateur accessible sur Gedeon2.

Il est principalement utilisé pour traiter des données textuelles. Les modèles d'IA proposés permettent la mise en forme ou la correction de textes, la traduction, la réalisation de synthèses ou la recherche d'information dans des grands volumes de texte. Ils peuvent également assister à la traduction, et à la recherche documentaire. Enfin pour les usages scientifiques il s'agit d'un outil extrêmement puissant en support à la programmation et à la maintenance du code informatique.

Cette IA peut être utilisée comme un assistant interactif mais également depuis peu en mode « API » c'est-à-dire utilisable comme un moteur au sein d'environnement de développement ou pour des opérations massives automatisées sur des grands ensembles de documents.

Enfin SEMIA dispose d'un outil de génération d'image à partir de description textuelle.

Quels ont été les principaux défis rencontrés dans le déploiement de cette plateforme ?

Notre premier chantier a été d'évaluer les différents modèles d'IA ouverts disponibles, leurs paramétrages possibles et leur adéquation avec le matériel disponible et les besoins des utilisateurs. Cette problématique reste encore d'actualité aujourd'hui compte tenu de la publication régulière de nouveaux modèles d'IA par la communauté scientifique.

Le second enjeu concerne la planification de l'activité des moyens, en effet l'utilisation des modèles génératifs se fait très majoritairement aux heures ouvrées. Les utilisations de SEMIA pour les besoins de recherche doivent donc s'insérer en dehors des horaires de bureau la nuit et le week-end ce qui a nécessité la mise en place d'une organisation particulière.

Quelles perspectives pour SEMIA ?

Les évolutions à venir concernent principalement l'enrichissement des modèles d'IA avec des bases de données documentaires internes à l'ONERA afin de proposer des assistants « sur mesure ». Par exemple un agent conversationnel pourrait être développé pour répondre aux questions sur l'organisation de l'ONERA, l'accès aux différentes procédures ou formulaires etc.

Nous avons également reçu de nombreuses demandes d'utilisateurs qui souhaitent bénéficier de solutions de transcription audio afin de faciliter la prise de notes lors des réunions et la rédaction des compte rendus.

Enfin nous allons relancer un cycle de formation cette année en proposant probablement plusieurs thèmes pour nous adapter aux différents niveaux d'expérience et objectifs d'utilisation.

En conclusion, quel message souhaitez-vous faire passer ?

J'invite l'ensemble du personnel de l'ONERA à découvrir et à se familiariser avec ces outils. Bien que leur utilisation semble aisée au premier regard, une certaine expérience est nécessaire pour les exploiter pleinement : il faut pratiquer !

SEMIA offre un environnement sécurisé idéal pour tester sans contraintes et se former à leur utilisation. N'hésitez pas à vous y essayer : ces outils génèrent des gains de productivité significatifs, et ce principalement sur les tâches les plus fastidieuses, libérant du temps pour les activités de fond.

Je suis convaincu qu'à terme, ces outils seront intégrés au socle numérique de base, devenant ainsi des ressources essentielles pour toutes les personnes travaillant sur un poste informatique au sein de l'ONERA.

