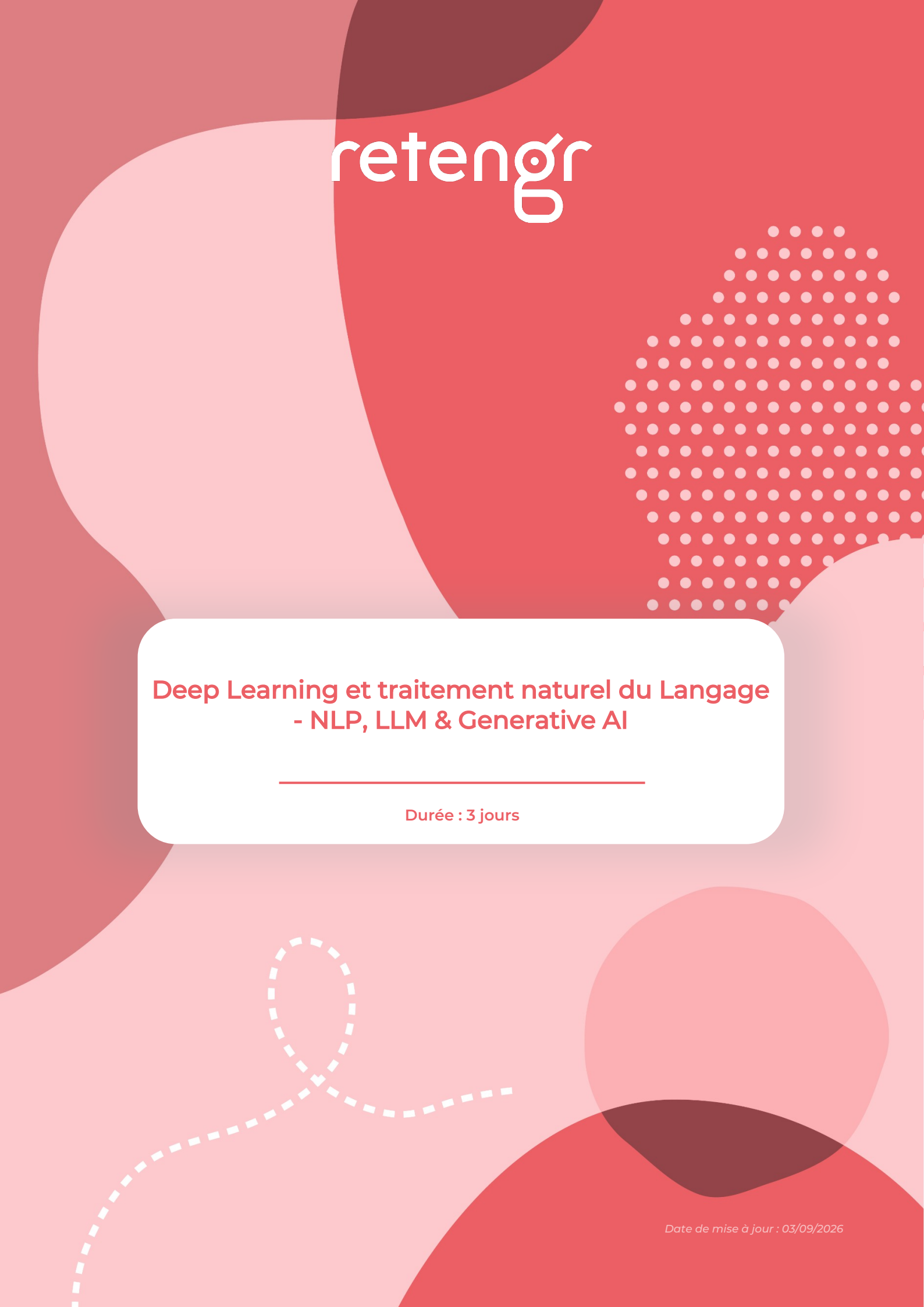




retengr

Deep Learning et traitement naturel du Langage - NLP, LLM & Generative AI

Durée : 3 jours



Méthode pédagogique

Composée à 70% de pratique, cette formation utilise des exercices illustrés et didactiques.

Une évaluation quotidienne de l'acquisition des connaissances de la veille est effectuée.

Une synthèse est proposée en fin de formation.

Une évaluation à chaud sera proposée au stagiaire à la fin du cours.

Un support de cours sera remis à chaque participant comprenant les slides sur la théorie, les exercices.

Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

Un suivi et un échange avec les participants seront mis en place quelques jours après la formation.

Cette formation ne nécessite pas d'avoir recours à une évaluation formalisée des acquis.

Présentation

Le traitement naturel du langage est un domaine à part entière de l'intelligence artificielle à l'intersection de l'informatique, des mathématiques, de la linguistique et des sciences cognitives. L'objectif est de construire des applications capables d'analyser, de modéliser, de comprendre et d'imiter le langage humain.

Les transformers introduits en 2017 et intégrant des capacités d'attention ont révolutionné ce domaine en offrant une nouvelle méthode pour la compréhension du langage naturel. D'une part, nous pouvons désormais réutiliser de nombreux modèles pré-entraînés extrêmement performants et d'autre part, ils ont contribué à l'IA générative permettant de générer du contenu original à partir de données existantes.



Vous connaissez les principes fondamentaux de l'apprentissage automatique et vous avez déjà mis en œuvre différentes architectures de réseaux de neurones, mais vous souhaitez en savoir plus sur les opportunités proposées par le Machine et le Deep Learning dans le domaine du traitement naturel du langage, cette formation est faite pour vous.

Vous serez capables de pré-traiter des données textuelles et serez en mesure d'implémenter des modèles permettant de les valoriser en TensorFlow ou PyTorch : classification de texte, recherche d'informations, analyse de sentiments, extraction d'informations, traduction, résumé de texte, etc.

Objectifs

- Mettre en place un prétraitement efficace d'un jeu de données textuelles
- Maîtriser les architectures de réseaux de neurones récurrents et des transformers
- Comprendre les bases théoriques et pratiques des LLM
- Réutiliser des modèles existants avec du transfer learning
- Comprendre le principe du RAG
- Mesurer la pertinence des modèles mis en œuvre
- Mettre en œuvre des cas concrets pour chaque architecture/solution

Audience

Data Analyst, Data scientist, Développeur



Le formateur

Le formateur est un expert du domaine qui intervient sur le sujet depuis plusieurs années en formation mais aussi en conseil.

Doté d'une grande qualité d'écoute, sa pédagogie et sa compétence technique vous permettront d'acquérir les compétences visées.

Il saura alterner entre théorie, pratique, et retours d'expérience.

Pré-requis

Connaissance de Python

Connaissance des librairies scientifiques (scikit-learn, pandas, numpy)

Connaissance du Deep Learning ou avoir suivi la formation "maîtriser le deep learning" (maîtrise des concepts du machine learning et avoir mis en œuvre différentes architectures de réseaux de neurones : convolutionnels, récurrents, ...)

Connaissance d'un framework Deep Learning TensorFlow ou PyTorch.

Programme

Introduction [2h]

- Les bases de la linguistique
- Introduction au Text Mining / Fouille de texte (Data Mining pour le texte)
- Extraction d'informations
- Recherche d'informations
- Catégorisation du texte
- Résumé du texte



Text Normalisation / Normalisation du texte [3h]

- n-grams
- Tokenization
- Stop Word
- Stemming
- Part-Of-Speech (POS) tagging
- Lemmatization

Vectorisation du texte [3h]

- Analyse de la fréquence des termes (Counter, TF-IDF, Word vectors)
- Bag of word
- Word Embedding : Word2vec, GloVe, FastText, etc.
- Création d'une pipeline de preprocessing

Les réseaux de neurones récurrents [3h]

- Principes fondamentaux des RNN
- LSTM et GRU
- Les approches Encodeur-Decodeur
- Utilisation de connexions résiduelles (skip connections)

L'architecture Transformers [3h]

- Seq2Seq
- Attention is All You Need
- Architecture des transformers
- Mécanisme d'attention
- Tokens & Embeddings

Les LLM [3h]

- Les familles de modèles : encoder, decoder & encoder-decoder
- Panorama de quelques modèles
- Méthodes d'apprentissage et d'optimisation
- Closed model APIs vs OpenModel weights vs Fully open model
- Coût et empreinte écologique des LLM
- Des LLM vers les SLM



Hugging Face et Keras Hub [2h]

- HuggingFace et ses pipelines
- Keras Hub et Kaggle

RAG - Retrieval-Augmented Generation [2h]

- Fondements du Retrieval-Augmented Generation (RAG)
- Les bases vectorielles
- Premier pas avec LangChain
- Du RAG avancé vers l'Agentic AI

Modalités et délais d'accès à la formation

Les inscriptions sont possibles jusqu'à 48 heures ouvrées avant le début de la formation, en interentreprises, dans la limite des places disponibles.

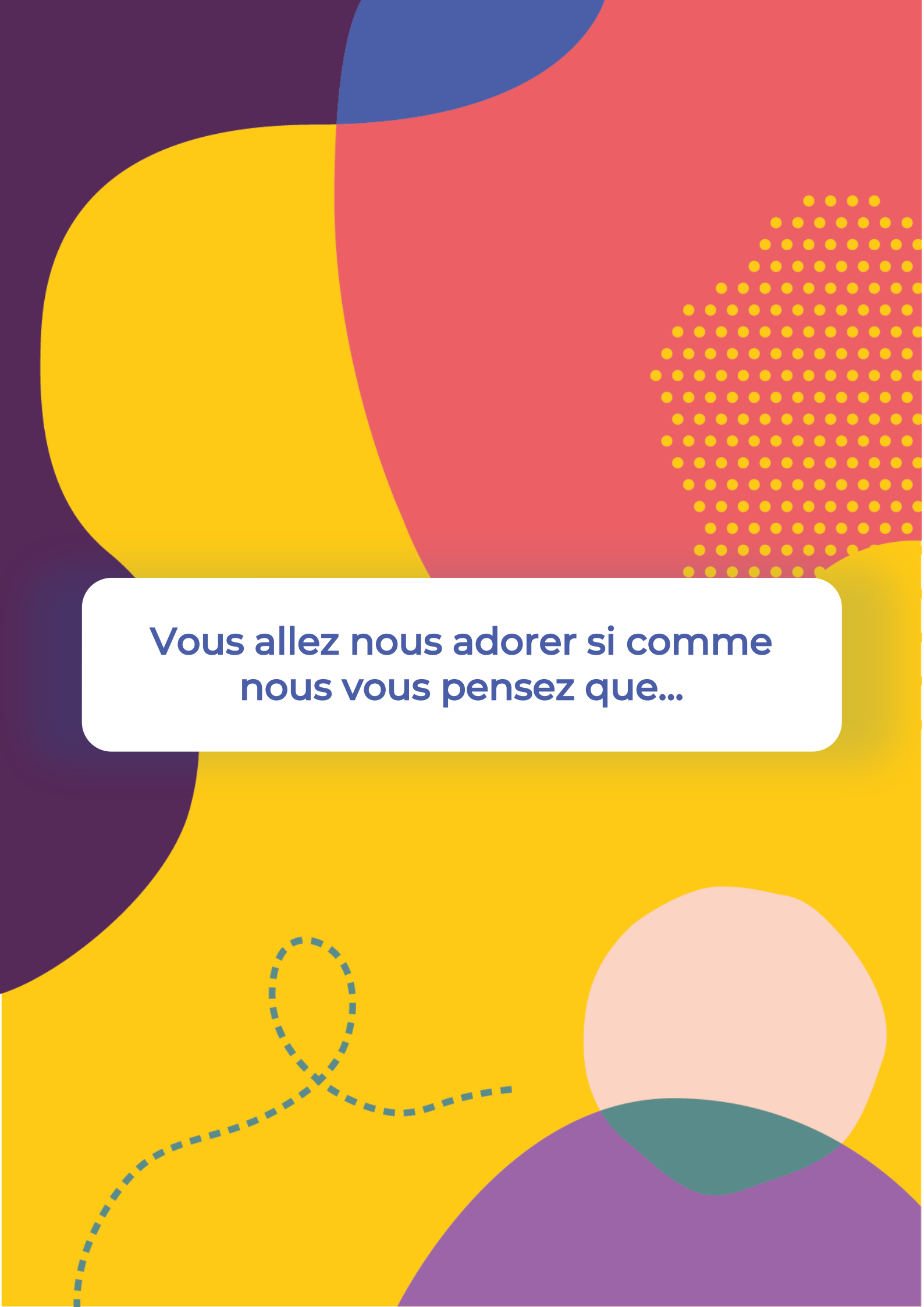
Pour les formations organisées en intra entreprise, la liste des participants peut être modifiée jusqu'à 24h ouvrés avant le début de la formation.

Accessibilité

RETENGR facilite l'accessibilité de ses formations.

Cette formation est accessible aux personnes en situation de handicap.

Si vous avez un besoin d'accès spécifique, contactez Céline BOURREIL (celine.bourreil@retengr.com) qui étudiera avec Handifiel's (notre référent handicap) votre demande et vous proposera les meilleures solutions.

The background is a vibrant abstract composition. It features large, overlapping organic shapes in shades of purple, yellow, and pink. A semi-circular area on the right side is filled with a dense pattern of small yellow dots. In the lower-left quadrant, there is a dashed teal line that forms a loop and then extends outwards. The overall aesthetic is modern and playful.

**Vous allez nous adorer si comme
nous vous pensez que...**

Une formation doit être au service de la performance du collaborateur et de l'entreprise

Ceci nécessite une quête constante d'excellence de la part de l'organisme formateur avec une adaptation systématique aux enjeux de l'entreprise, la mise à jour régulière des supports de cours et une veille technologique indispensables pour toujours être à la pointe du domaine.



L'expertise technique est aussi importante que les qualités pédagogiques



Nos formateurs sont tous des experts de leur domaine. Mais qu'ont-ils de plus que les autres ? Nous les sélectionnons en plus pour leurs qualités de pédagogue et leurs méthodes d'enseignements. Nous plaçons les qualités pédagogiques au même niveau que l'expertise afin que nos stagiaires tirent le meilleur de leurs formations.

L'excellence naît de l'excellence

Beaucoup de nos clients se classent parmi les leaders de leurs industries respectives ou parmi les start-ups les plus prometteuses. Nous savons que former les collaborateurs de telles entreprises nécessite de prêter attention à chaque détail en prodiguant un accompagnement à la hauteur de l'ambition de nos stagiaires. C'est pourquoi nous savons faire des leaders d'aujourd'hui les champions de demain !





retengr

**Faire du leader d'aujourd'hui, le
champion de demain**