



Université Constantine 2
جامعة قسنطينة 2

Web Analytics and Natural Language Processing (WANLP)

Chapitre 01 | Cours 01

Introduction à l'IA, au NLP et au Web Analytics

Professeur BOURAMOUL Abdelkrim

Département IFA, Faculté NTIC

abdelkrim.bouramoul@univ-constantine2.dz

www.bouramoul.com

Etudiants concernés

Faculté/Institut	Département	Niveau	Spécialité
NTIC	IFA	Master 1	SDIA

Plan du Cours

Section 1

- **Aperçu de l'Intelligence Artificielle (IA)**

Section 2

- **Bases du Traitement Automatique du Langage (NLP)**

Section 3

- **Bases du Web Analytics (WA)**

Section 4

- **Techniques, Outils, et Bibliothèques pour l'NLP et le WA**

Section 5

- **Conclusion**

Section 1 :

Aperçu de l'Intelligence Artificielle (IA)

Définition

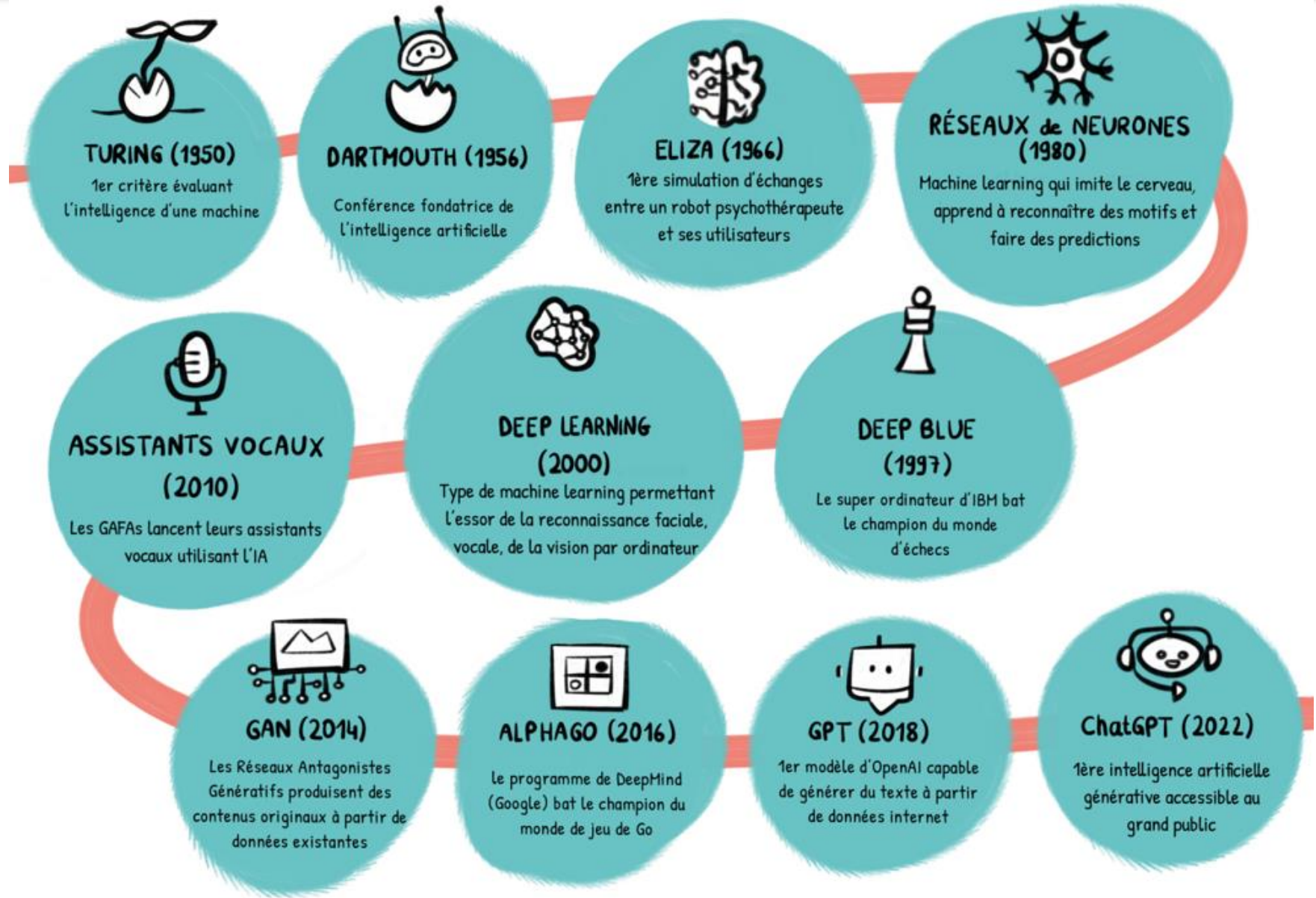
- Il s'agit de créer des systèmes informatiques capables de simuler des processus cognitifs humains et d'exécuter des tâches qui, autrement, nécessiteraient l'intelligence humaine
- Elle comprend la résolution de problèmes la reconnaissance de formes, l'apprentissage automatique (ML, DL, FL, SL), le traitement du langage naturel, la compréhension de la parole, la création de chatbots intelligents, la conception de programmes d'IA générative, entre autres...
- Le but ultime est de créer des systèmes intelligents qui peuvent raisonner, apprendre, planifier, percevoir et manipuler naturellement leur environnement.

Historique de l'IA

Historique de l'IA

- L'histoire de l'IA est riche et remonte à plusieurs décennies, marquée par des avancées significatives.
- Dans les années 1950, des chercheurs pionniers tels qu'Alan Turing et John McCarthy ont posé les bases de l'IA en développant des concepts et des théories fondamentaux.
- Exemple : Alan Turing a introduit le célèbre test de Turing, visant à évaluer la capacité d'une machine à exhiber un comportement intelligent.
- L'IA a traversé différentes phases, caractérisées par des périodes de progrès et de stagnation, souvent influencées par les avancées technologiques et les défis théoriques.
- Ces dernières années, l'IA a connu un essor considérable grâce à l'explosion des données et à l'amélioration des algorithmes d'apprentissage automatique.
- Exemple : Les progrès dans le domaine de l'apprentissage profond ont permis des réalisations impressionnantes, comme la reconnaissance d'images, la traduction automatique et la conduite autonome.

Les dates clés de l'IA : Du Turing à ChatGPT



Applications de l'IA

L'IA est omniprésente dans notre vie quotidienne.

- La reconnaissance faciale dans les smartphones et les systèmes de sécurité.
- Les chatbots utilisés par les entreprises pour le service client et l'assistance en ligne.

Exemples d'applications courantes :

- Les assistants vocaux tels que Siri, Alexa et Google Assistant simplifient nos tâches quotidiennes en répondant à nos questions et en exécutant des commandes vocales.
- Les systèmes de recommandation de Netflix, Amazon et Spotify utilisent l'IA pour personnaliser les recommandations de contenu en fonction des préférences de l'utilisateur.
- Les véhicules autonomes développés par Tesla, Waymo et d'autres entreprises utilisent des technologies d'IA avancées pour naviguer de manière autonome sur les routes, améliorant ainsi la sécurité et l'efficacité des transports.

Applications de l'IA

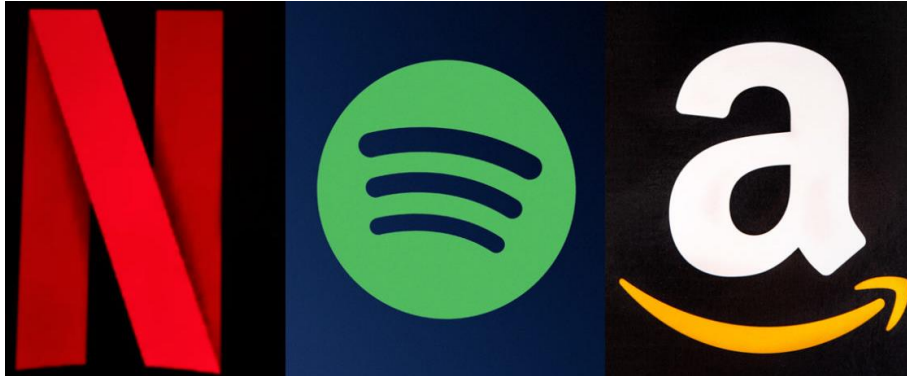
L'IA est omniprésente dans notre vie quotidienne.

- La reconnaissance faciale dans les smartphones et les systèmes de sécurité.
- Les chatbots utilisés par les entreprises pour le service client et l'assistance en ligne.

Exemples d'applications courantes :

- Les assistants vocaux tels que Siri, Alexa et Google Assistant simplifient nos tâches quotidiennes en répondant à nos questions et en exécutant des commandes vocales.
- Les systèmes de recommandation de Netflix, Amazon et Spotify utilisent l'IA pour personnaliser les recommandations de contenu en fonction des préférences de l'utilisateur.
- Les véhicules autonomes développés par Tesla, Waymo et d'autres entreprises utilisent des technologies d'IA avancées pour naviguer de manière autonome sur les routes, améliorant ainsi la sécurité et l'efficacité des transports.

Applications de l'IA



L'essor de l'IA soulève des questions éthiques et sociétales

- Comment garantir la protection des données personnelles dans un monde dominé par l'IA ?
- Comment éviter que les algorithmes ne reproduisent les biais et les préjugés existants dans la société ?
- Quels sont les risques de dépendance excessive à l'IA et comment les atténuer ?
- Quels critères éthiques devraient guider la collecte et l'utilisation des données pour l'entraînement des modèles d'IA ?
- Quelles sont les implications éthiques de l'utilisation de l'IA dans des domaines sensibles tels que la santé et la justice ?
- Quels sont les droits et les protections nécessaires pour les individus dont les données sont utilisées par des systèmes d'IA ?
- Comment prévenir les abus potentiels de l'IA, tels que la manipulation de l'opinion publique ou la surveillance de masse, et protéger les libertés individuelles ?

Section 2 :

Bases du Traitement Automatique du Langage (NLP)

Définition du TALN (NLP)

Le Traitement Automatique du Langage Naturel

- Un domaine de l'IA dédié à l'interaction entre les ordinateurs et le langage humain pour comprendre et analyser le langage naturel.
- Le TALN vise à permettre aux ordinateurs de comprendre, d'analyser et de générer le langage humain de manière naturelle.

Tâches de base :

- Tokenisation : segmentation du texte en unités lexicales (mots, phrases).
- Analyse morphosyntaxique : analyse de la structure grammaticale et des formes des mots.
- Reconnaissance d'entités nommées : identification des noms propres (personnes, lieux, organisations, etc.).
- Génération de langage : création de texte cohérent et compréhensible par les humains.

Tâches de base de l'NLP et Exemple

Tokenisation :

- Texte : "Le chat dort."
- Tokens : ['Le', 'chat', 'dort.']

Analyse Morphosyntaxique

- Phrase : "Le chat noir dort paisiblement sur le tapis."
- Analyse morphosyntaxique :
 - "Le" : déterminant masculin singulier
 - "chat" : nom commun masculin singulier
 - "noir" : adjectif masculin singulier
 - "dort" : verbe du troisième groupe, troisième personne du singulier, indicatif présent
 - "paisiblement" : adverbe
 - "sur" : préposition
 - "le" : déterminant masculin singulier
 - "tapis" : nom commun masculin singulier

Tâches de base de l'NLP et Exemple

Reconnaissance d'Entités Nommées

- Phrase : "Belmadi a été limogé de l'équipe nationale algérienne après l'élimination de la Coupe d'Afrique en 2024."
- Reconnaissance d'Entités Nommées :
 - "Belmadi" : Personne (nom propre)
 - "équipe nationale algérienne" : Organisation
 - "Coupe d'Afrique" : Événement sportif
 - "2024" : Date

Génération de Langage

- Entrée : "Le chat"
- Génération de Langage :
 - "Le chat est un animal domestique qui aime jouer avec une balle."
 - "Le chat noir se promène dans le jardin."
 - "Le chat miaule doucement en regardant par la fenêtre."

Section 3 :

Web Analytics

Introduction au Web Analytics

Définition, Rôle et Importance:

- Analyse du comportement des utilisateurs sur les sites web et les applications.
- Comprendre et optimiser les performances en ligne.
- Guider les décisions stratégiques et améliorer l'expérience utilisateur.

Notes supplémentaires

- Le Web Analytics permet de mesurer et d'analyser diverses interactions des utilisateurs, telles que les visites de pages, les conversions et les taux de rebond.
- Il aide les entreprises à mieux comprendre leur audience et à prendre des décisions informées pour améliorer leurs stratégies en ligne.

Objectifs du Web Analytics

Principaux objectifs :

- Comprendre le comportement des utilisateurs : identifier les pages les plus visitées sur un site web et les actions les plus fréquemment effectuées par les utilisateurs.
- Améliorer l'expérience utilisateur : en analysant les données de navigation pour optimiser la convivialité du site et réduire les temps de chargement.
- Optimiser les conversions et les performances en ligne : en suivant les taux de conversion pour identifier les opportunités d'amélioration et ajuster les stratégies marketing en conséquence.

Exemples d'utilisation :

- Suivi du parcours des utilisateurs sur un site web : Analyse des pages de destination et des chemins de conversion pour identifier les points forts et les points faibles de l'expérience utilisateur.
- Évaluation de l'efficacité des campagnes marketing: Suivi des performances des campagnes publicitaires en ligne pour déterminer le retour sur investissement et ajuster les budgets en conséquence.

Métriques et Indicateurs Clés

Trafic :

- Nombre total de visites sur un site web ou une application.
- Utilisation : Mesurer l'attractivité et la popularité du site.

Taux de rebond :

- Pourcentage de visiteurs qui quittent le site après avoir consulté une seule page.
- Utilisation : Évaluer la pertinence du contenu et l'expérience utilisateur.

Durée de session :

- Temps moyen passé par un utilisateur sur le site lors d'une visite.
- Utilisation : Analyser l'engagement des utilisateurs et la qualité de l'expérience.

Pages vues :

- Nombre total de pages consultées par les utilisateurs.
- Utilisation : Évaluer la profondeur de l'interaction des utilisateurs avec le contenu.

Processus d'Analyse des Données

Processus d'analyse des données en Web Analytics :

- Collecte de données : Acquisition de données à partir de diverses sources telles que les balises de suivi, les cookies et les outils d'analyse web.
- Nettoyage et prétraitement : Élimination des données inexactes, incomplètes ou redondantes pour garantir la qualité des données.

Analyse et interprétation :

- Analyse des données : Utilisation d'outils et de techniques pour extraire des informations utiles à partir des données collectées.
- Interprétation des données : Compréhension des tendances, des modèles et des insights pour prendre des décisions éclairées.

Outils et techniques d'analyse des données :

- Utilisation de logiciels spécialisés tels que Google Analytics, Adobe Analytics, etc.
- Techniques d'analyse statistique, visualisation de données, modélisation prédictive, etc.

Méthodes et outils de collecte de données web :

- Balises de suivi : Codes insérés dans les pages web pour collecter des données sur le comportement des utilisateurs.
- Cookies : Fichiers stockés sur les appareils des utilisateurs pour suivre leur activité en ligne.

Importance de la collecte de données précises et fiables :

- Les données précises et fiables sont essentielles pour des analyses efficaces et des prises de décision informées.
- Une collecte de données rigoureuse garantit la qualité des informations utilisées pour évaluer les performances et optimiser les stratégies en ligne.

Section 4 :

Techniques, Outils, et Bibliothèques

Introduction aux Outils et Techniques

Aperçu des outils et techniques :

- Outils pour le Traitement Automatique du Langage (TALN) : NLTK, spaCy, TensorFlow.
- Techniques d'analyse en Web Analytics : segmentation des utilisateurs, analyse des entonnoirs de conversion, analyse prédictive.

Importance des outils et techniques :

- Facilitent la compréhension et l'interprétation des données.
- Permettent d'identifier des tendances et des modèles significatifs.
- Favorisent la prise de décision éclairée et l'optimisation des stratégies en ligne.

Outils et bibliothèques pour l’NLP

NLTK (Natural Language Toolkit) :

- Bibliothèque Python largement utilisée pour le traitement du langage naturel, offrant une gamme complète d'outils pour le prétraitement, l'analyse et la génération de texte.
- Exemple d'utilisation : prétraitement de texte pour la suppression des mots vides et la lemmatisation des mots.

spaCy :

- Bibliothèque Python hautement performante pour le TALN, reconnue pour sa rapidité et sa précision dans le traitement de texte en langage naturel.
- Exemple d'utilisation : Analyse morphosyntaxique pour l'identification des parties du discours et des entités nommées dans un texte.

TensorFlow :

- Plateforme open-source développée par Google, offrant des outils puissants pour la création et le déploiement de modèles d'apprentissage automatique, y compris dans le domaine du TALN.
- Exemple d'utilisation : création d'un modèle de traitement du langage naturel pour la classification de texte ou la génération de texte.

Intégration d'outils et de techniques

Discussion sur l'intégration d'outils et de techniques

- Exploration des synergies entre les outils et les techniques utilisés dans le domaine du Traitement Automatique du Langage (TALN) et du Web Analytics.
- Réflexion sur la manière dont ces deux domaines peuvent se compléter et s'enrichir mutuellement dans l'analyse des données et la génération d'insights.

Exemples d'utilisation conjointe :

- Utilisation de techniques de NLP pour l'analyse de contenu web : Extraction d'entités nommées à partir de commentaires d'utilisateurs ou d'avis sur les réseaux sociaux pour mieux comprendre les préférences et les besoins des clients.
- Utilisation de techniques de Web Analytics pour optimiser les modèles de TALN : Analyse des entonnoirs de conversion pour identifier les points de friction dans l'expérience utilisateur et ajuster les modèles de classification de texte en conséquence.

Section 5 :

Conclusion

Conclusions

Récapitulation :

- Nous avons exploré les concepts fondamentaux de l'Intelligence Artificielle (IA), du Traitement Automatique du Langage (TALN) et du Web Analytics.
- Ces domaines offrent des outils et des techniques puissants pour l'analyse et l'optimisation des données textuelles et en ligne.

Importance :

- L'IA, le NLP et le Web Analytics jouent un rôle crucial dans de nombreux aspects de notre vie quotidienne et des entreprises, de la reconnaissance vocale aux stratégies de marketing en ligne.

Perspectives :

- La combinaison de ces domaines ouvre de nouvelles possibilités pour des analyses avancées et des insights plus profonds.
- L'apprentissage continu et l'exploration des tendances émergentes sont essentiels pour rester à la pointe dans ces domaines en constante évolution.

Prochaines étapes :

- Dans les prochains chapitres, nous plongerons plus en profondeur dans chaque domaine, explorant des techniques avancées et des cas d'utilisation pratiques pour approfondir notre compréhension et nos compétences.



Université Constantine 2
جامعة قسنطينة 2

Fin de Chapitre 01 | Cours 01