

The background of the slide is decorated with numerous paint splashes in various colors, including yellow, orange, green, and blue, creating a vibrant and artistic effect.

ANALYSE ET VISUALISATION DES DONNÉES

Alexandra-Maria DOBRESCU

| Cours 12 |

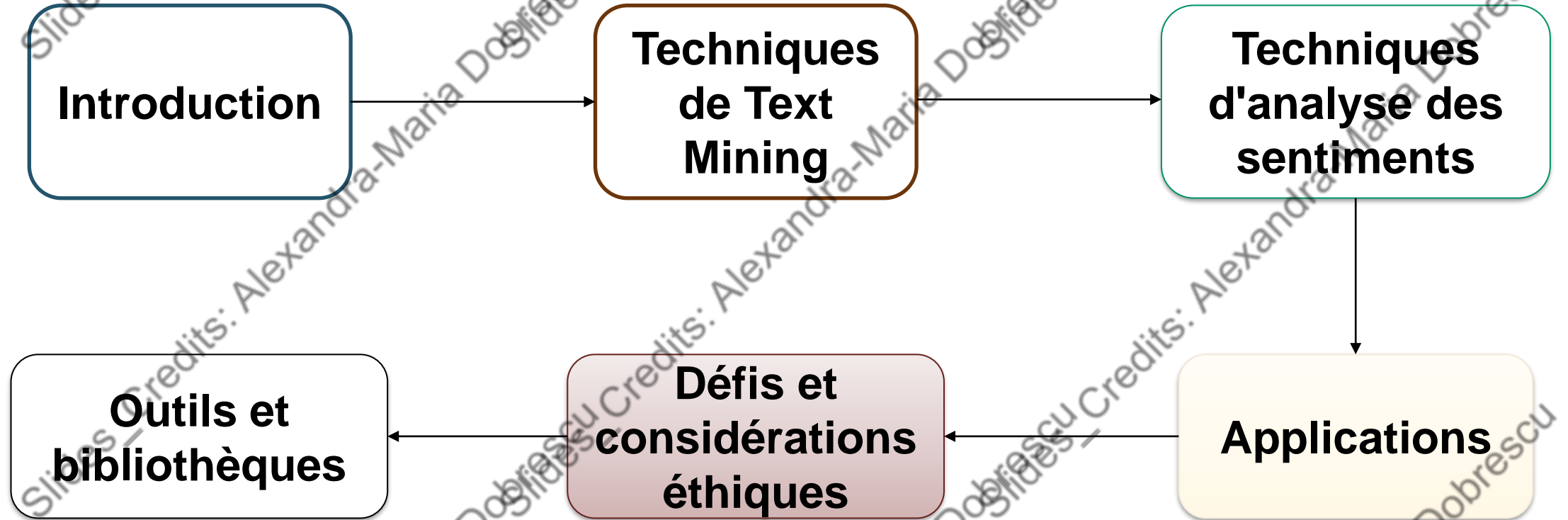
Sujets avancés ou émergents 2ème partie

AVD | Cours 12 | AMD

● **TEXT MINING ET
ANALYSE DE SENTIMENTS**

● **VISUALISATIONS INTERACTIVES
DE DONNÉES**

Text Mining et Analyse de Sentiments



Text Mining et Analyse de Sentiments

Applications

1. Récupération d'informations et moteurs de recherche

Exemple: Google Search

Cas d'utilisation: Google utilise des techniques de text mining pour analyser et comprendre le contenu des pages web. Cela permet au moteur de recherche de fournir aux utilisateurs des résultats de recherche très pertinents en fonction de l'intention qui sous-tend leurs requêtes.

2. Analyse des sentiments dans les médias sociaux

Exemple: Analyse des sentiments sur Twitter pour la surveillance des marques

Cas d'utilisation: Des entreprises comme Nike ou Starbucks utilisent l'analyse des sentiments sur Twitter pour suivre en temps réel l'opinion des gens sur leur marque. L'analyse des tweets les aide à évaluer le sentiment du public, à identifier les problèmes et à s'engager auprès de leur public.

Text Mining et Analyse de Sentiments

Applications

3. Témoignages et évaluations des clients

Exemple: Avis sur les produits Amazon

Cas d'utilisation: Amazon utilise l'exploration de texte pour analyser les commentaires sur les produits soumis par les clients. Cette analyse permet de comprendre la satisfaction des clients, d'identifier les caractéristiques les plus appréciées et de prendre des décisions fondées sur des données pour améliorer les produits.

4. Exploration de textes médicaux et cliniques

Exemple: IBM Watson pour l'oncologie

Cas d'utilisation: IBM Watson pour l'oncologie analyse la littérature médicale, les données des essais cliniques et les dossiers des patients afin d'aider les oncologues à faire des recommandations de traitement personnalisées. Il améliore la prise de décision clinique dans les soins du cancer.

Text Mining et Analyse de Sentiments

Applications

5. Analyse de l'actualité financière

Exemple: Bloomberg Terminal

Cas d'utilisation: Les professionnels de la finance utilisent Bloomberg Terminal, qui intègre des techniques de text mining, pour analyser les nouvelles financières, les rapports de marché et les médias sociaux. Cela les aide à se tenir informés des sentiments du marché et à prendre des décisions d'investissement.

6. Analyse de documents juridiques

Exemple: Lex Machina

Cas d'utilisation: Lex Machina utilise le text mining pour analyser les documents juridiques, les dossiers des tribunaux et les données relatives aux litiges. Il fournit des informations sur les stratégies d'affaires, le comportement des juges et les résultats juridiques, soutenant ainsi les professionnels du droit dans leur stratégie de litige.



Set a Winning Patent Litigation Strategy with Legal Analytics

Lex Machina[®] legal analytics provide accurate, valuable data that combines the knowledge of legal experts with machine learning to predict the behavior of judges, attorneys and parties in U.S. courts.

[REQUEST A DEMO](#)

Text Mining et Analyse de Sentiments

Applications

7. L'administration de la preuve électronique dans les affaires juridiques

Exemple: Relativity

Cas d'utilisation: Relativity est une plateforme d'*e-discovery* qui utilise le text mining pour aider les équipes juridiques à rechercher, examiner et analyser de grands volumes de documents électroniques pendant la phase de découverte des affaires juridiques.

8. Détection de la fraude dans le secteur bancaire

Exemple: SAS Fraud Analytics

Cas d'utilisation: Des banques comme HSBC utilisent SAS Fraud Analytics, qui emploie des techniques d'exploration de texte, pour analyser les données textuelles dans les transactions et les communications. Elle permet de détecter des schémas indiquant des activités frauduleuses.

Engineered to Empower You



Artificial Intelligence

We're building AI solutions that help you better manage e-discovery and do what humans do best: tell the story.

[Learn About Our AI](#)



Simply powerful experience

Powerful doesn't mean complicated. RelativityOne has the speed, scale, and automation you need for faster, easier results.

[Learn More](#)



Proactive security

We build our products secure from the ground up and layer on threat intelligence and 24/7 monitoring. Security isn't part of the program. It is the program.

[Explore Relativity Security](#)



Legendary support

We'll give you a whole team dedicated to your success, plus 24/7/365 tech support. No extra charge, no hidden fees. It's why customers love us.

[Meet the Team](#)

Leading Organizations Rely on Relativity

 HonelliErede

 INNOVATIVE LITIGATION

 bsi.

 LinkedIn

 MCA

 icourts

 BDO

 FRONTEO

 FOLEY
FOLEY & LARDNER LLP



 CDS
COMPLETE DISCOVERY SOLUTIONS

 MCS GROUP

 EimerStahl

 FUJITSU

 Ballard Spahr

 McGUIREWOODS

 Parcels

 THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Text Mining et Analyse de Sentiments

Applications

9. Recherche universitaire


Exemple: JSTOR

Cas d'utilisation: JSTOR, une bibliothèque numérique de revues universitaires, utilise le text mining pour aider les chercheurs à explorer et à analyser une vaste collection d'articles savants. Les chercheurs peuvent identifier les tendances, extraire des informations et découvrir la littérature pertinente.

10. Ressources humaines et feedback des employés

Exemple: Glint

Cas d'utilisation: Glint, une plateforme d'engagement des employés, utilise le text mining pour analyser les enquêtes et les commentaires des employés. Elle aide les organisations à identifier les domaines à améliorer, à évaluer le sentiment des employés et à accroître la satisfaction générale sur le lieu de travail.

**Viva Glint**





Dashboard

Focus Areas

Reporting


Feedback



My Pulses



Employee Signals

Closed on April 24, 2023

 Settings

 Switch Program 

Hi Nora, see what's next

It's easy to take these results personally, obsess over finding the root cause in the data, or immediately jump to solutions. Instead, see this dashboard as a starting point to an open conversation with your team.

1

Share feedback

Completed on April 17

2

Intrepret results

View the guide

3

Have a Team Conversation

Launch conversation

4

Commit to action

Choose a focus area

Survey Participation

See how representative the results are and how vocal your employees were

Response Rate

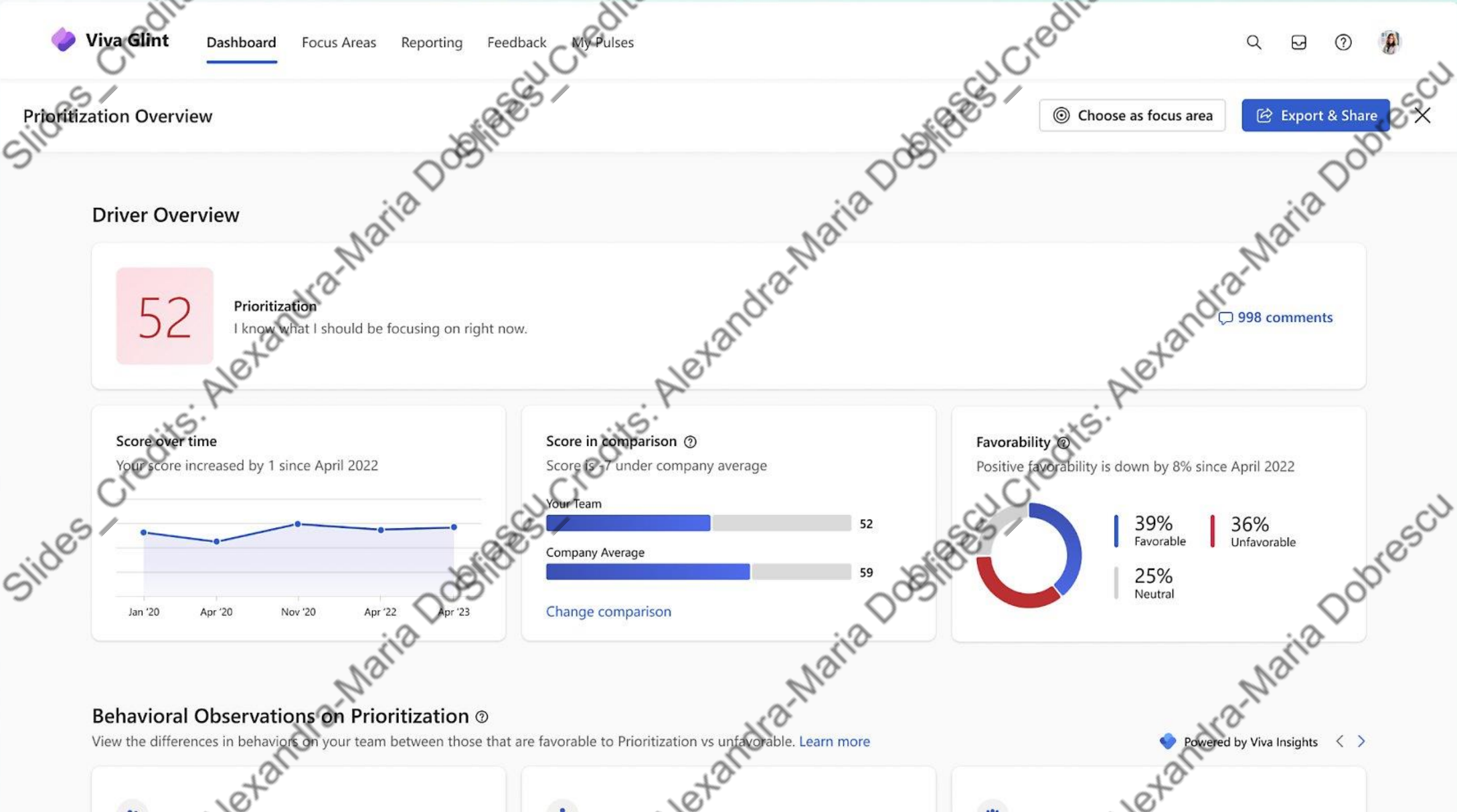
22 out of 28 of your team members responded, compared to a benchmark response rate of 79%

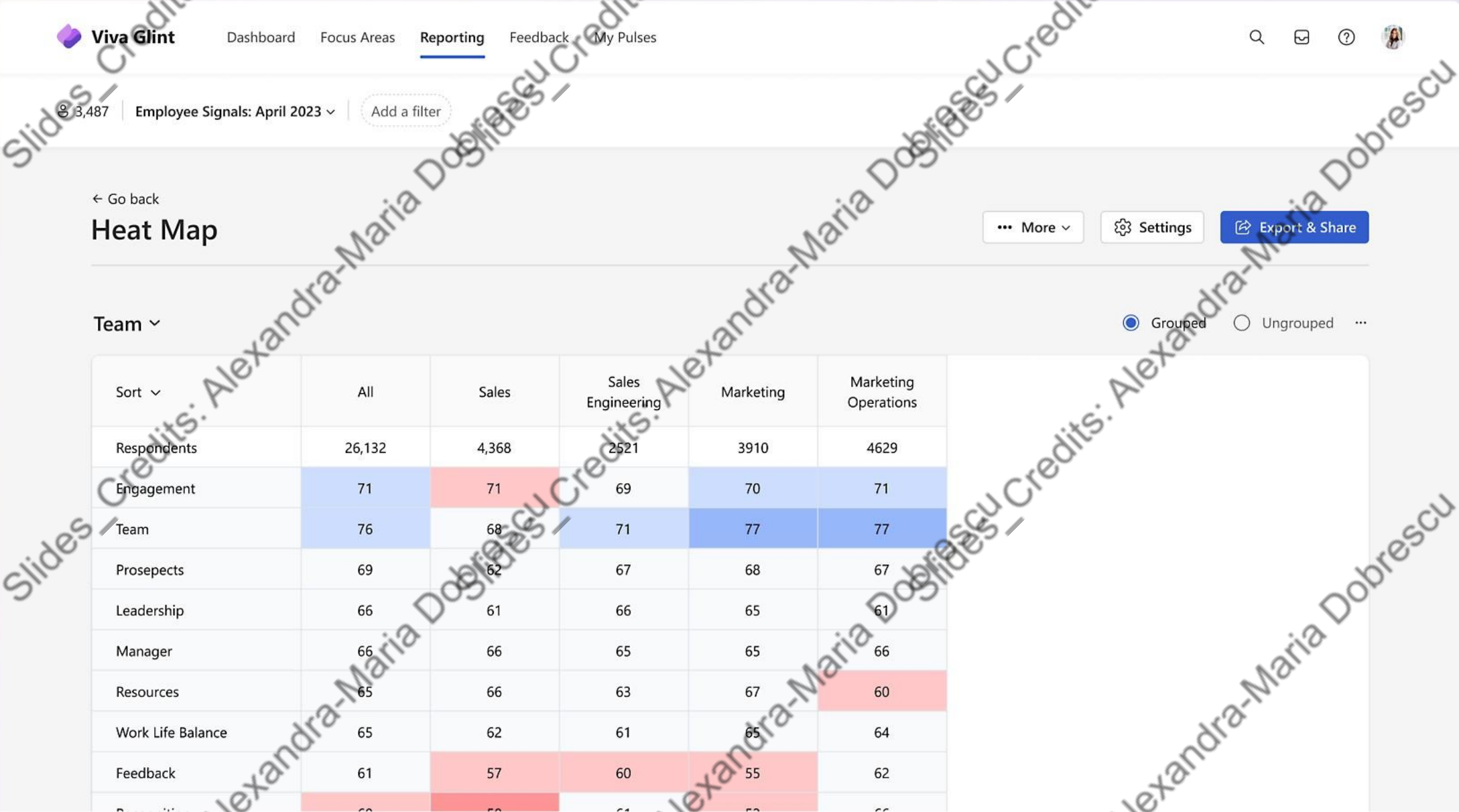
79%

Comments

Questions that received the most comments were about [Work-life Balance](#) and [Leadership](#)

3,293





Comment Keywords ②


Keywords are words and phrases that occur most frequently across comments.

Cloud ▾








A word cloud visualization of comment keywords. The words are arranged in a circular pattern, with larger words indicating higher frequency. The words are color-coded: red for 'Reorganization', 'Hours', and 'Headcount'; blue for 'Flexibility', 'Supportive', 'Leadership', and 'Training'; and black for 'Decision making', 'Delegation', 'Vacation', 'Days off', 'Energized', 'Short staffed', 'Ownership', 'Demoralize', 'Resources', and 'Flexibility'.

Demoralize Resources Supportive
Flexibility Reorganization
Hours Short staffed Energized Leadership
Ownership Headcount Days off Vacation
Decision making Delegation Training

**Viva Glint**

[Dashboard](#) [Focus Areas](#) [Reporting](#) [Feedback](#) [My Pulses](#)


   



 3,487



Employee Signals: April 2023 ▾

Add a filter

[← Go back](#)

Alerts  Settings

[Alert Summary \(108\)](#)  Elevated Attrition Risk (93)  Your Custom Predictive Alert (18)

Population	Respondents ▾	Negative Alerts		Positive Alerts	Risks
▼ Thrive Engage	26,132	4		0	 
Area	Score	Alert	Comparison	Impact	
★ Career - I have good career opportunities at Thrive Eng...	52	-14	vs Benchmark	Very High	
★ Empowerment - I feel empowered to make decisions reg...	58	-17	vs Benchmark	Very High	
★ Feedback - My manager provides me with feedback tha...	61	-14	vs Benchmark	High	

Text Mining et Analyse de Sentiments

Applications

11. Intelligence compétitive

Exemple: Kompyte

Cas d'utilisation: Kompyte utilise l'exploration de texte pour surveiller les sites Web, les médias sociaux et les stratégies de marketing de ses concurrents. Les entreprises l'utilisent pour mieux comprendre les activités de leurs concurrents, identifier les tendances du marché et prendre des décisions éclairées.

12. Recherche pharmaceutique

Exemple: BenevolentAI

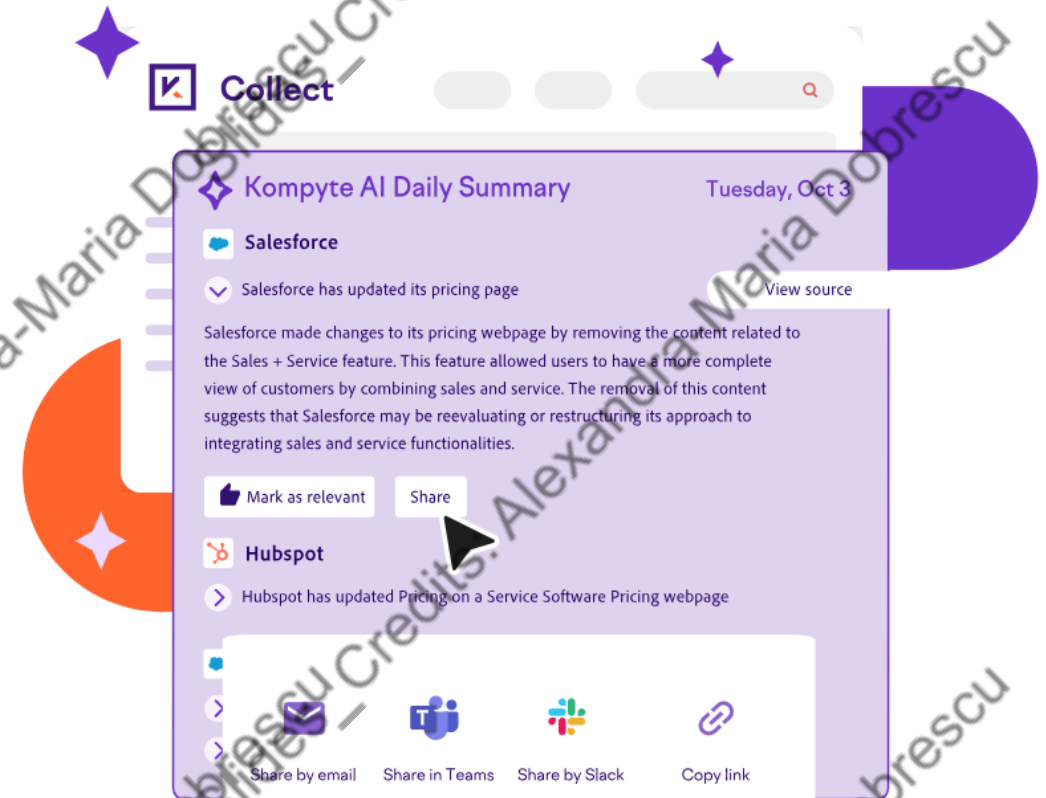
Cas d'utilisation: BenevolentAI utilise l'exploration de texte et l'intelligence artificielle pour la découverte de médicaments. Il analyse la littérature scientifique et les données cliniques pour identifier les médicaments potentiels, accélérant ainsi le processus de recherche.

Stay on Top of Competitor Activity Effortlessly—with AI-Powered Daily Summaries

Does tracking your competitors' every move leave you feeling more overwhelmed than exhilarated? Transform a time-consuming, tedious review process into a quick, efficient routine with AI Daily Summaries!

Scan everything at a glance or filter by update type or competitor(s). Easily dig in where you need to, share with a click, and train the AI with your preferences as you go.

Get early access



Text Mining et Analyse de Sentiments

Applications

13. Éducation et apprentissage en ligne

Exemple: Coursera

Cas d'utilisation: Coursera utilise le text mining pour analyser les commentaires des apprenants, les évaluations des cours et les forums de discussion. Cette analyse permet d'améliorer le contenu des cours, d'identifier les sujets les plus populaires et d'améliorer l'expérience d'apprentissage en ligne.

The Coursera logo, consisting of the word "coursera" in white lowercase letters on a blue square background.

**Qualité des
données et
bruit**

**Ambiguïté
et
polysémie**

**Manque de
compréhension
du contexte**

**Traitement de
grands
ensembles
de données**

**Nature
dynamique du
langage**

**Biais dans les
données et les
modèles**

**Exploration de
textes
multilingues**

Défis

Text Mining et Analyse de Sentiments

Défis

Qualité des données et bruit: Les données textuelles peuvent être bruyantes, contenir des erreurs, des abréviations, des fautes d'orthographe et un langage informel. Cela peut affecter la précision des modèles d'exploration de texte.

Ambiguïté et polysémie: les mots ont souvent des significations multiples en fonction du contexte, ce qui rend difficile la saisie précise du sens d'un mot ou d'une phrase

Manque de compréhension du contexte: la compréhension du contexte dans le langage naturel est complexe. Les modèles d'exploration de texte peuvent avoir des difficultés avec le sarcasme, l'ironie ou les nuances culturelles, ce qui conduit à des interprétations erronées.

Text Mining et Analyse de Sentiments

Défis

Traitement de grands ensembles de données: le traitement et l'analyse efficaces de grands volumes de données textuelles nécessitent de puissantes ressources informatiques. La mise à l'échelle des techniques de fouille de textes pour traiter les données volumineuses peut s'avérer difficile.

Exploration de textes multilingues: L'exploration de textes dans plusieurs langues introduit des complexités supplémentaires en raison des variations dans la syntaxe, la grammaire et l'expression des sentiments.

Biais dans les données et les modèles: Les systèmes de text mining peuvent, par inadvertance, perpétuer ou amplifier les préjugés sociétaux existants.

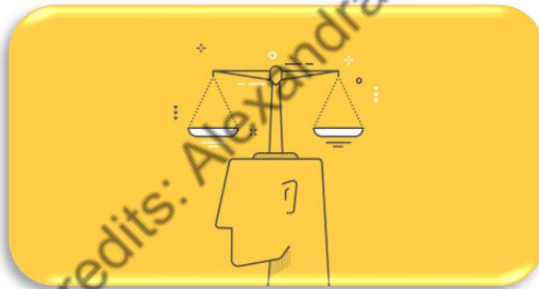
Nature dynamique du langage: la langue évolue au fil du temps avec de nouveaux mots, de nouvelles phrases et de nouvelles significations. Les modèles de text mining peuvent avoir du mal à suivre les changements linguistiques.

Text Mining et Analyse de Sentiments

Considérations éthiques

Préoccupations en matière de protection de la vie privée

L'exploration de textes implique souvent l'analyse de données personnelles contenues dans des textes. Garantir la protection de la vie privée et la confidentialité des personnes figurant dans les données est une considération éthique essentielle.



Consentement éclairé

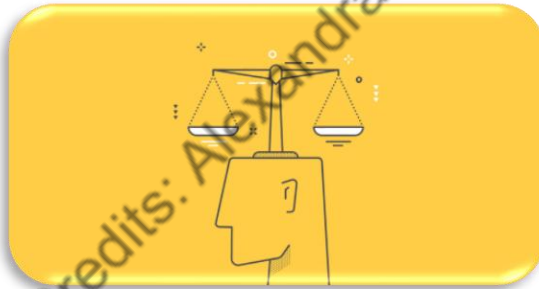
Il est essentiel d'obtenir le consentement éclairé des personnes dont les données sont utilisées dans le cadre du text mining. Les utilisateurs doivent savoir comment leurs données sont analysées et dans quel but.

Text Mining et Analyse de Sentiments

Considérations éthiques

Transparence et explicabilité

Le manque de transparence des modèles de fouille de textes peut soulever des problèmes éthiques. Les utilisateurs et les parties prenantes doivent pouvoir comprendre comment les modèles prennent leurs décisions.



Équité et biais

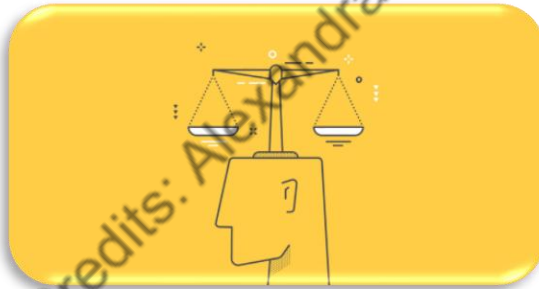
Il est essentiel de garantir l'équité des modèles de text mining et de s'attaquer aux biais. Les biais dans les modèles peuvent aboutir à des résultats discriminatoires, renforçant les stéréotypes ou excluant certains groupes.

Text Mining et Analyse de Sentiments

Considérations éthiques

Responsabilité

Il est important d'établir la responsabilité des décisions prises par les modèles de text mining. Pour une mise en œuvre éthique, il est essentiel de savoir qui est responsable de la résolution des problèmes et des biais.



Risques pour la sécurité

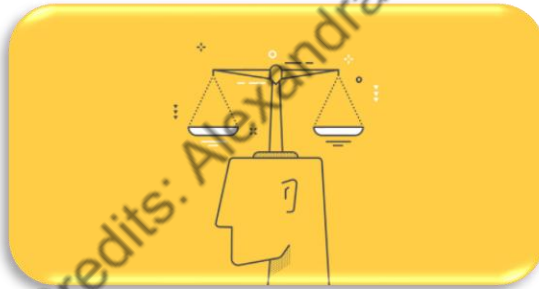
La protection de la sécurité des données textuelles est une considération éthique. L'accès non autorisé à des informations sensibles pourrait entraîner des violations et des dommages potentiels.

Text Mining et Analyse de Sentiments

Considérations éthiques

Impact sur l'emploi

L'automatisation par le text mining peut avoir un impact sur l'emploi dans certaines industries. Les considérations éthiques comprennent la prise en compte des déplacements d'emplois potentiels et le soutien aux travailleurs touchés par ces changements.



Pratiques trompeuses

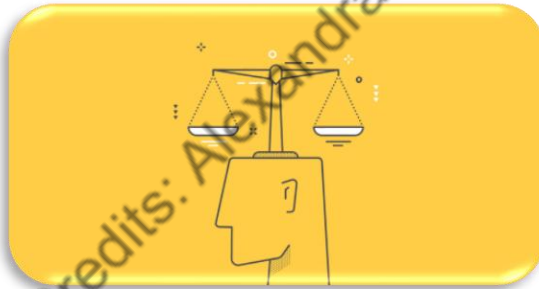
L'utilisation du text mining pour des pratiques trompeuses, telles que la diffusion de fausses informations ou la manipulation de l'opinion publique, soulève des questions éthiques et peut avoir des conséquences considérables.

Text Mining et Analyse de Sentiments

Considérations éthiques

Implications en matière de droits de l'homme

Les activités de text mining doivent s'aligner sur les principes des droits de l'homme, en évitant les activités susceptibles de porter atteinte aux droits individuels, à la liberté d'expression ou à la liberté d'association.



Impact sur l'environnement

Les opérations de text mining à grande échelle peuvent avoir des répercussions sur l'environnement en raison de la consommation d'énergie des centres de données. La mise en œuvre de pratiques durables est une considération éthique.

Text Mining et Analyse de Sentiments

Outils et bibliothèques

1. **ggplot2 (R)**: un paquetage R permettant de créer des visualisations statiques, animées et interactives.

Caractéristiques principales:

- Syntaxe de haut niveau pour créer des visualisations complexes.
- Forte communauté et documentation.
- Tracés facilement personnalisables.

2. **Altair**: une bibliothèque déclarative de visualisation statistique en Python qui permet des visualisations concises et élégantes. Elle est particulièrement adaptée aux visualisations interactives.

Caractéristiques principales:

- Syntaxe déclarative pour spécifier les visualisations.
- Rapide et facile à utiliser pour les graphiques courants.
- Intégration avec Pandas.

Text Mining et Analyse de Sentiments

Outils et bibliothèques

3. **Bokeh**: une bibliothèque Python pour créer des visualisations interactives dans le navigateur web. Elle est conçue pour les navigateurs modernes et supporte les tableaux de bord interactifs.

Caractéristiques principales:

- Tracés interactifs avec zoom, panoramique et infobulles.
- Prise en charge des données en continu et en temps réel.
- Intégration avec Jupyter Notebooks.

4. **Holoviews**: une bibliothèque Python permettant de construire facilement des visualisations complexes. Elle est construite sur Bokeh et Matplotlib.

Caractéristiques principales:

- API de haut niveau pour créer des visualisations.
- Prise en charge d'ensembles de données volumineux et en continu.
- Facilité de composition de plusieurs visualisations.

Bibliographie

- [1] Subramaniam, A. (2020). What Is Big Data Analytics| Big Data Analytics Tools and Trends| Edureka.
- [2] Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.
- [3] <https://www.microsoft.com/en-us/microsoft-viva/glint>