



This PDF is generated from authoritative online content, and is provided for convenience only. This PDF cannot be used for legal purposes. For authoritative understanding of what is and is not supported, always use the online content. To copy code samples, always use the online content.

## Content Analyzer Plug-in for GAX

eServices 8.5.2

# Table of Contents

<b>Plugiciel Content Analyzer pour GAX</b>	<b>3</b>
<b>Objets de données d'entraînement</b>	<b>6</b>
<b>Modification d'objets de données d'entraînement</b>	<b>10</b>
<b>Programmer un entraînement</b>	<b>12</b>
<b>Prétraitement du texte</b>	<b>14</b>
<b>Essai des modèles</b>	<b>16</b>
<b>Validation croisée</b>	<b>20</b>
<b>Échantillons pour détecter la langue, l'expression de sentiments et la nécessité d'action</b>	<b>21</b>
<b>Installation du plugiciel Content Analyzer pour GAX</b>	<b>24</b>

# Plugiciel Content Analyzer pour GAX

Genesys Content Analyzer est une option de Genesys Knowledge Management, nécessitant une licence supplémentaire. Elle ajoute un complément de technologie de traitement du langage naturel à Genesys Knowledge Management.

Le plugiciel Content Analyzer pour GAX est un outil administratif qui permet la gestion de Genesys Content Analyzer. Le reste du Knowledge Management est géré par le **plugiciel eServices Manager pour GAX**.

## Conseil

- Content Analyzer est conçu pour un affichage dont la résolution d'écran minimale requise est de 1280x1024 ou une résolution HD intégrale de 1920x1080.
- Pour que Content Analyzer fonctionne avec Microsoft Internet Explorer 11, il est possible que deviez ajuster les paramètres de votre navigateur. Assurez-vous que les options de sécurité suivantes d'Internet Explorer sont activées :
  - Téléchargements/Téléchargement de police
  - Écriture de scripts/Écriture de scripts ASP

## Qu'est-ce que Content Analyzer fait et comment y parvient-il?

Content Analyzer analyse les interactions textuelles et vous indique à quel endroit sur votre arborescence de catégories elles sont susceptibles de se situer.

- Pour ce faire, il utilise un modèle de classification, qui est en fait une représentation statistique d'une **arborescence de catégories**.
- Il applique le modèle aux interactions texte entrantes (courriel, SMS, clavardage, etc.) et produit une liste de catégories. Pour chaque catégorie, il vous indique quelle est la probabilité que l'interaction y appartienne.
- L'analyse en soi est effectuée par le Classification Server lorsque celui-ci détecte un objet **Classify** ou **Analyze** dans une stratégie de routage.

## Conseil

La classification, tout comme le **filtrage**, peut fonctionner sur toute interaction ayant du texte qui lui est associé, de quelque manière que ce soit, qu'il s'agisse du corps de l'interaction (courriel, clavardage) ou d'un élément d'une autre nature (des données d'utilisateur, par exemple). Ce document nomme tous les éléments du même genre

des *interactions d'entraînement* ou des *messages*.

## Alors, d'où viennent les modèles?

Tout d'abord, vous avez besoin d'un *objet de données d'entraînement*. Il s'agit d'une arborescence de catégories et d'un jeu d'objets texte (comme des courriels, des messages SMS, des interactions Twitter, etc.), dont chaque objet est assigné à une catégorie de l'arborescence. Ces objets texte se nomment *messages d'entraînement* ou *interactions d'entraînement*.

Ensuite, exécutez l'entraînement, et celui-ci produira le modèle.

De façon un peu plus détaillée,

1. Créez un **objet de données d'entraînement**.
2. **Modifiez** l'objet de données d'entraînement au besoin :
  - Créez de nouveaux messages d'entraînement.
  - Si vous avez des messages d'entraînement non catégorisés, déplacez-les dans des catégories.
  - Utilisez couper ou copier, puis coller pour déplacer les messages d'entraînement d'une catégorie à l'autre.
3. **Préparez-vous** à exécuter l'entraînement.
  - Choisissez quand et de quelle façon l'exécuter.
  - Si vous le désirez, vous pouvez lui faire **filtrer** les chaînes de texte qui portent à confusion ou qui sont non pertinentes avant son exécution.
4. **Tester le modèle**.

## Importation et exportation

Vous pouvez importer et exporter des modèles et des objets d'entraînement dans le cadre du processus **d'importation et d'exportation d'objets du eServices Manager**. C'est aussi un moyen de copier des objets d'entraînement.

### Avertissement

Une fois que vous avez importé une archive qui comprend un objet de données d'entraînement, il ne vous est plus possible d'importer l'archive de nouveau dans une autre langue. Si vous tentez de le faire, vous recevrez un message vous demandant de changer le nom de la catégorie racine. Toutefois, en procédant ainsi, vous recevrez un message d'erreur vous informant qu'un objet de données de formation portant ce

nom existe déjà.

Pour contourner le problème, vous pouvez renommer l'objet de données d'entraînement avant de procéder à l'importation la deuxième fois.

# Objets de données d'entraînement

Un objet de données d'entraînement correspond à une arborescence de catégories ainsi qu'à un jeu d'objets texte (comme des courriels), appelés *messages d'entraînement* ou *interactions d'entraînement*, dont chacun est attribué à une catégorie de l'arborescence.

Il existe plusieurs façons pour qu'un objet texte se retrouve classé dans une catégorie.

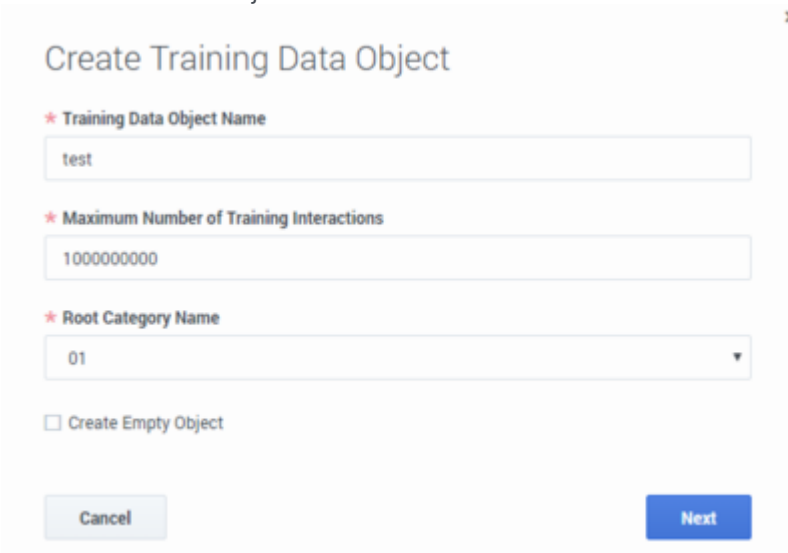
- Chaque fois qu'un agent utilise une réponse standard lors du traitement d'une interaction, cette dernière est classée dans la même catégorie que la réponse standard en question.

Vous constaterez que certains agents sont plus susceptibles que d'autres de choisir la réponse standard appropriée. Vous pouvez choisir de n'utiliser que les attributions de catégorie effectuées par ces agents au moment de créer un nouvel objet de données d'entraînement. —Pour ce faire, consultez la section "[Qui catégorise?](#)" ci-dessous.

- Par définition, les réponses standards sont associées à des catégories.
- Vous pouvez [créer une nouvelle interaction d'entraînement](#) et l'ajouter à un objet de données d'entraînement existant.

## Nouvel objet de données d'entraînement

1. Dans l'onglet **Entraînement** du plugiciel Content Analyzer, cliquez sur l'icône du signe + pour afficher la fenêtre Nouvel objet de données d'entraînement.



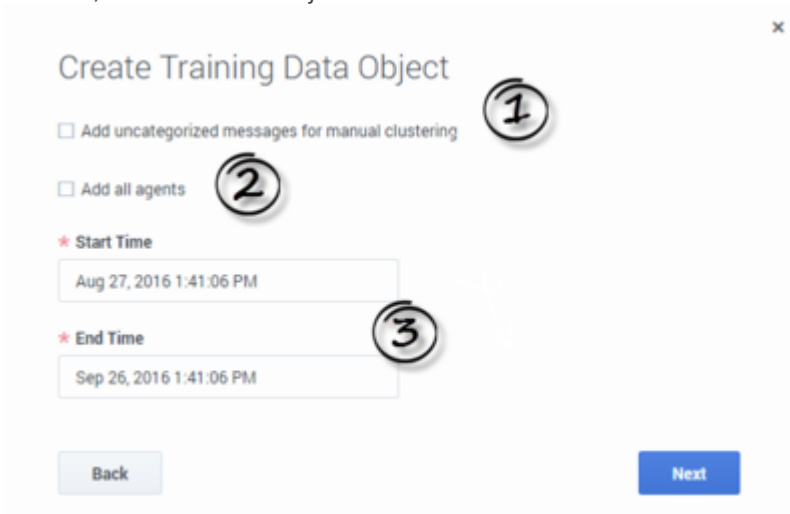
Nouvel objet de données d'entraînement

- Assurez-vous d'utiliser un nom qui respecte les exigences relatives aux caractères : les caractères

autorisés sont les mêmes que pour les **objets du eServices Manager**, et le maximum de caractères permis est 64.

- Si vous n'avez pas de jeu d'objets texte catégorisés à votre disposition, vous pouvez sélectionner **Créer un objet vide** et y **ajouter des interactions d'entraînement** plus tard.
- La figure illustre le nombre maximal d'interactions d'entraînement qui est permis. Le nombre par défaut est 50 000, et le minimum est 1.

2. Ensuite, choisissez les objets texte à inclure.



Nouvel objet de données d'entraînement 2

1. Si vous avez sélectionné **Ajouter des messages non catégorisés**, vous serez en mesure de les **déplacer vers des catégories** plus tard.
2. Utiliser les catégories affectées par tous les agents? (sélectionnez **Ajouter tous les agents**) ou sélectionner les agents les plus fiables? (étape suivante.)
3. Ce laps de temps définit les dates des messages utilisés pour l'entraînement.

### Conseil

En définissant l'intervalle de temps, prenez soin d'éviter ces erreurs courantes :

- Des dates de début et de fin identiques, entraînant l'absence de messages dans l'objet de données d'entraînement
- Un laps de temps trop grand, ce qui peut se traduire par un nombre trop important de messages dans l'ensemble d'entraînement et faire en sorte que les derniers courriels aient moins d'effet sur l'entraînement du modèle.

3. Utiliser la catégorisation faite par quels groupes d'agents? (Si vous avez sélectionné **Ajouter tous les agents** cette boîte de dialogue ne s'affichera pas.) Consultez la section "**Qui catégorise?**" ci-dessous pour des explications plus détaillées sur le sujet.

## Create Training Data Object

### ★ Agent Group

<input type="checkbox"/> ▾	Name	⬇
<input type="checkbox"/>	Chat distribution for processing	
<input type="checkbox"/>	E-mail distribution for processing	
<input checked="" type="checkbox"/>	E-mail QA review group	
<input type="checkbox"/>	SMS or MMS group	
<input checked="" type="checkbox"/>	Supervisors	

Nouvel objet de données d'entraînement 3

4. Choisissez les types de médias.

<input type="checkbox"/> ▾	Name
<input checked="" type="checkbox"/>	email
<input type="checkbox"/>	facebook
<input checked="" type="checkbox"/>	facebookprivatemessage
<input type="checkbox"/>	facebooksession
<input type="checkbox"/>	fax
<input type="checkbox"/>	imchat
<input checked="" type="checkbox"/>	mms
<input type="checkbox"/>	session

Nouvel objet de données d'entraînement 4

5. Le dernier filtre montre combien de messages d'entraînement sont compris dans le nouvel objet de données d'entraînement.



## Create Training Data Object

Created Training Data Object: SpringfieldOffice

Training Messages created: 121

Back

Finish

Nouvel objet de données d'entraînement - fin

## Remarques

### Qui catégorise?

Le processus de création d'un nouvel objet de données d'entraînement peut prendre de quelques secondes à de nombreuses heures, tout dépendant du nombre d'interactions et d'agents qui sont compris. Pour cette raison, Genesys recommande que 200-300 agents, tout au plus, participent au processus de catégorisation. Vous pouvez créer un groupe d'agents qui servira exclusivement à la création d'objets de données d'entraînement, composé d'agents dont la catégorisation est fiable et qui sont conscients que leurs décisions en matière de catégorisation auront un effet direct sur le modèle de classification. Vous pouvez ensuite sélectionner ce groupe à l'étape 3 ci-dessus.

Vous ne pouvez pas sélectionner des agents séparément.

### Suppression

Vous ne pouvez supprimer aucun objet de données d'entraînement (ni aucune catégorie racine) auquel au moins un modèle est attribué.

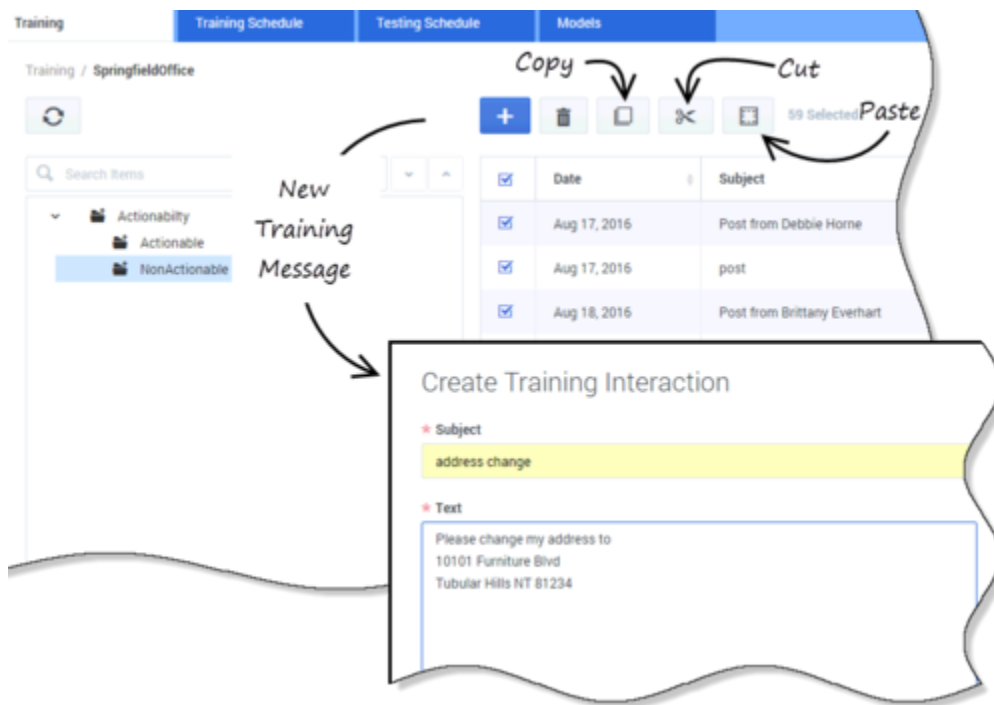
## Étapes suivantes

Vous pouvez :

- Assigner des messages non catégorisés à une catégorie
- Créer de nouvelles interactions d'entraînement

## Modification d'objets de données d'entraînement

Pour modifier un objet de données d'entraînement existant, sélectionnez-le dans l'onglet **Entraînement**, puis cliquez sur l'icône Modifier (engrenage). La figure suivante illustre certaines des choses que vous pouvez faire :



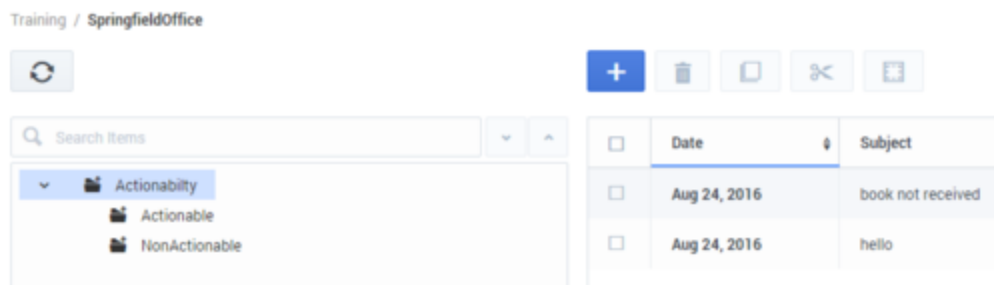
Modifier un objet de données d'entraînement

## Ajouter de nouvelles Interactions d'entraînement

Sélectionner une catégorie, puis cliquer sur l'icône du signe +, comme le montre la figure **Opérations relatives aux objets de données d'entraînement**. Ajouter du texte aux champs **Objet** et **Texte**.

## Classer des messages non catégorisés dans une catégorie

Les interactions d'entraînement non catégorisées sont affichées dans la catégorie racine. Dans la figure suivante, les interactions d'entraînement non catégorisées **livre non reçu** et **bonjour** apparaissent dans la catégorie racine **Nécessité d'action**.




Messages non catégorisés

Pour les catégoriser, utilisez **couper** (ou **copier**), puis **coller** pour les déplacer vers des catégories non-racine.

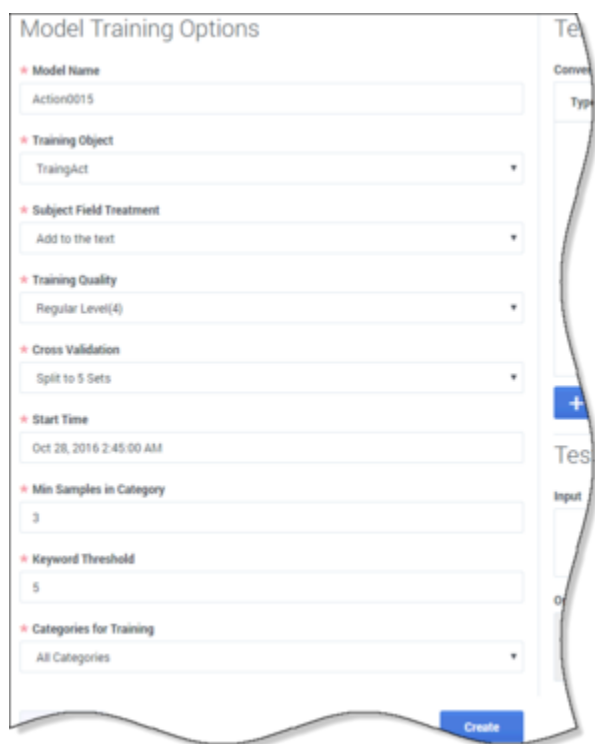
Vous pouvez aussi utiliser couper ou copier, puis coller pour déplacer des interactions d'entraînement d'une catégorie à l'autre.

# Programmer un entraînement

Programmez une séance d'entraînement dans l'onglet **Programmation d'entraînement**.

- Pour remplir les champs avec les valeurs d'une séance d'entraînement existante, cliquez sur l'icône de clone.  C'est une façon pratique de modifier l'heure à laquelle une séance d'entraînement existante sera exécutée.
- Pour remplir les champs avec des valeurs par défaut, cliquez sur l'icône du signe +.

Dans les deux cas, la boîte de dialogue **Options d'entraînement de modèles** s'affichera.



Options d'entraînement de modèles

1. Les **noms de modèles** doivent être composés d'un maximum de 21 caractères et doivent uniquement utiliser les **caractères autorisés**.
2. **Traitement du champ Objet**
  - **Ignorer**—L'entraînement ne prend pas en considération le contenu du champ **Objet**.
  - **Ajouter au texte**—L'entraînement prend en considération le contenu du champ **Objet**.
  - **Ajouter avec importance double**—L'entraînement donne deux fois plus d'importance au contenu du champ **Objet** qu'à celui du corps du courriel.
3. **Qualité de l'entraînement**—Si vous savez que l'objet de données d'entraînement contient de

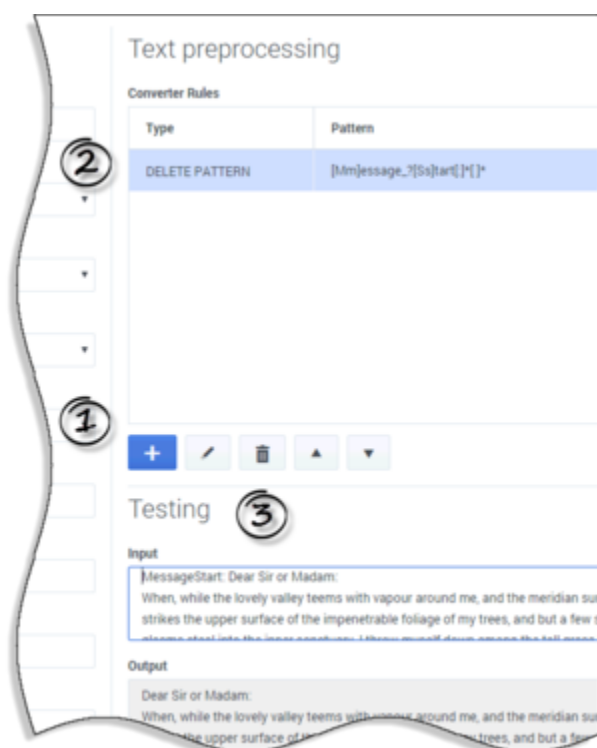
nombreux objets texte classés dans les mauvaises catégories, utilisez les Niveaux de non-fiabilité 10-12. Sinon, utilisez Brouillon ou les Niveaux réguliers 1-6. Veuillez noter ce qui suit :

- Les niveaux réguliers et les niveaux de non-fiabilité représentent deux échelles indépendantes qui sont difficilement comparables. Dans chacun de ces types de niveau, un nombre plus grand se traduit par une meilleure qualité. Le seul moyen de confirmer si, par exemple, un niveau de non-fiabilité de 11 produira de meilleurs ou de moins bons résultats qu'un niveau régulier de 4 est de créer un modèle pour chacun des paramètres et de les tester.
  - En fait, ces niveaux permettent de déterminer le nombre de mots que le système prend en considération et le nombre d'itérations que le processus de traitement exécute. Augmenter chacun de ces niveaux devrait améliorer la qualité du modèle résultant, mais ce ne sera pas le cas à des niveaux plus élevés. Encore une fois, la seule façon de s'en assurer est de tester les modèles résultants, préférablement par **validation croisée**.
  - La durée de l'entraînement augmente à mesure que vous passez de la qualité Brouillon à celle d'un Niveau régulier de 3. Toutefois, après avoir dépassé une qualité de niveau 3, la différence en matière de durée de l'entraînement est minime. Genesys vous recommande de seulement utiliser le plus bas niveau de qualité lorsque vous souhaitez obtenir une lecture préliminaire de l'estimation de la qualité du modèle. Pour la production, utilisez un niveau de qualité de 2-6.
4. La **validation croisée** est expliquée sur une autre page. Choisissez entre aucune validation croisée ou une validation croisée qui séparera les données en 3, en 6 ou en 10 jeux. Si vous sélectionnez la validation croisée, l'entraînement produira une mesure de précision pour le modèle ainsi que le modèle lui-même. Cela a l'avantage de ne nécessiter aucune autre étape d'essai, mais la durée de l'entraînement sera plus longue.
  5. Étant donné que l'entraînement peut utiliser une grande partie des ressources du système, il est préférable de programmer **l'heure de début** à une heure non achalandée. Assurez-vous de programmer une heure de début à un moment postérieur au moment présent.
  6. Le **nombre minimal d'échantillons dans la catégorie** correspond au nombre minimal d'objets texte qu'une catégorie doit avoir pour qu'elle soit comprise dans l'entraînement. Les catégories qui ne contiennent pas ou peu d'objets texte ne sont pas pertinentes pour l'entraînement.
  7. Le **seuil de mots-clés** correspond au nombre minimal d'objets texte dans lesquels un mot-clé doit apparaître afin que celui-ci soit pris en considération dans l'entraînement. Une valeur relativement élevée de ce paramètre peut réduire la durée de l'entraînement, mais aussi la qualité. Ce qui est considéré comme étant de valeur élevée ou faible pour ce paramètre dépend de la taille totale de l'objet de données d'entraînement. Par exemple, si un objet de données d'entraînement contient de 5 à 10 objets texte par catégorie, un seuil de mots-clés élevé correspondrait à **2** ou à **3**. En revanche, si un objet de données d'entraînement contient de 30 à 50 objets texte par catégorie, un seuil de mots-clés élevé correspondrait à **20**.
  8. Les **catégories pour l'entraînement** sont **Toutes les catégories** ou **Catégories terminales seulement**. Le terme "catégorie terminale" se dit d'une catégorie qui ne contient pas de sous-catégorie. Il est possible qu'une arborescence de catégories utilise des catégories non terminales principalement pour organiser celles qui sont terminales. Si c'est le cas, peu ou pas d'objets texte sont associés aux catégories non terminales, donc il y a peu à gagner à ajouter ces dernières à l'entraînement.
  9. Vous pouvez nettoyer votre objet de données d'entraînement en utilisant la **sous-fenêtre de prétraitement du texte** (à la droite de **Options d'entraînement de modèles**) pour éliminer le texte superflu.

# Prétraitement du texte

Dans le cadre de la programmation de l'entraînement, vous pouvez utiliser le volet **Prétraitement du texte** (à la droite de la sous-fenêtre **Options d'entraînement de modèles**) pour éliminer le texte superflu des messages d'entraînement d'un objet de données d'entraînement. Vous créez des filtres (modèles) qui effectuent une recherche pour du texte et procèdent à diverses opérations de suppression. Cela peut s'avérer utile lorsque les courriels que vous souhaitez utiliser pour l'entraînement contiennent une proportion considérable de texte qui est :

- Impertinent ou équivoque à des fins de classification et
- Identifiable par une **expression régulière**.



Prétraitement du texte

1. Cliquez sur l'icône du signe + pour créer une nouvelle règle.
2. Types de règles :
  - **SUPPRIMER APRÈS**—Cherchez une correspondance avec le modèle, puis supprimez tout le texte qui vient après et qui contient le texte correspondant.
  - **SUPPRIMER AVANT**—Cherchez une correspondance avec le modèle, puis supprimez tout le texte qui vient avant et qui contient le texte correspondant.

- **SUPPRIMER TOUT EN CAS DE CORRESPONDANCE**—Cherchez une correspondance avec le modèle, puis supprimez la totalité du message d’entraînement qui contient le texte correspondant.
  - **SUPPRIMER TOUT EN CAS DE NON-CORRESPONDANCE**—Cherchez une correspondance avec le modèle, puis supprimez la totalité du message d’entraînement s’il ne contient pas le texte correspondant.
  - **SUPPRIMER LE MODÈLE**—Cherchez une correspondance avec le modèle, puis supprimez seulement le texte qui correspond au modèle.
3. Testez le modèle. Entrez du texte, et le résultat s’affichera. Si vous apportez des modifications à la règle, vous devrez entrer le texte de nouveau afin de voir le résultat de la règle modifiée. Pour les deux types de SUPPRIMER TOUT, si vous testez le modèle et que celui-ci trouve une correspondance, la fenêtre **Sortie** sera vide. Dans des conditions réelles d’utilisation, si une correspondance est trouvée, le message d’entraînement sera entièrement supprimé de l’objet de données d’entraînement.

# Essai des modèles

Une fois que vous avez créé un modèle, vous pouvez le tester. Le processus de test commande à un modèle d'analyser un objet d'entraînement. Un instant de réflexion vous permettra de déterminer si l'objet d'entraînement :

- Doit faire partie de la même catégorie racine que le modèle
- Ne doit pas être celui qui a été utilisé pour créer le modèle

Programmez le test sur l'onglet Programmation de test : il vous suffit de sélectionner le modèle, l'objet de données d'entraînement sur lequel il sera testé et l'heure de début du test.

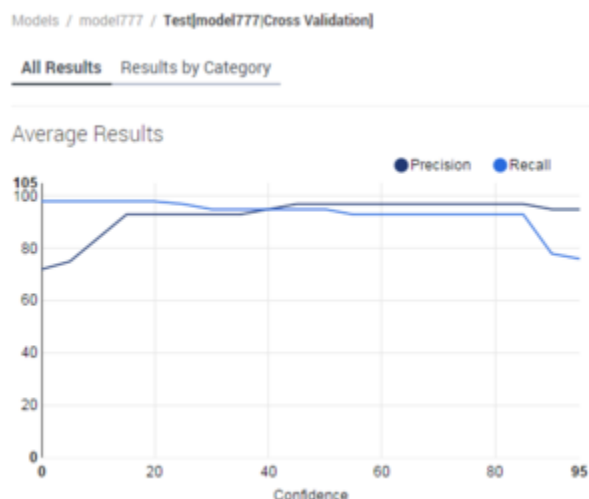
Pour afficher les résultat de test d'un modèle donné, sélectionnez le modèle, sur l'un ou l'autre des onglets Programmation de test ou Modèles, puis cliquez sur l'icône de l'œil. L'icône de l'œil pourra être activée seulement si le modèle que vous sélectionnez a déjà été testé.

Passons aux choses intéressantes... la compréhension des résultats.

## Onglet Tous les résultats

Cet onglet affiche les graphiques Résultats moyens et Présent dans les meilleures catégories N ainsi que le tableau Confusion de catégorisation.

### Résultats moyens



Précision, rappel et confiance

Ce graphique illustre les évaluations (axe vertical) de la précision (noir) et du rappel (bleu) selon un niveau de confiance donné (axe horizontal). Mais que signifient ces termes? Poursuivez votre lecture



pour en savoir plus :

### Confiance

La confiance est un résultat numérique, de 1 à 100, qui indique le pourcentage de probabilité qu'un objet texte appartienne à une certaine catégorie, en fonction du modèle sélectionné.

(À titre de comparaison, *l'exactitude* est une évaluation issue de tests qui porte sur la justesse de l'attribution d'objets texte par un modèle à des catégories. En d'autres mots, la confiance exprime l'estimation d'un modèle en matière de catégorisation, et l'exactitude évalue la justesse de cette estimation.)

### Précision et rappel

Pour comprendre la précision et le rappel, il faut prendre en considération plusieurs façons différentes de voir la performance d'un modèle. Si votre modèle tente d'assigner un certain nombre d'éléments à une catégorie X, vous pouvez faire les déductions suivantes :

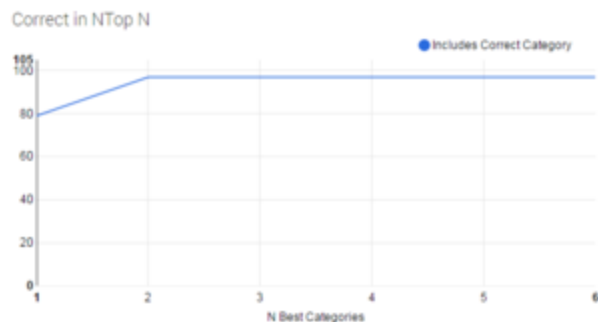
- $a$  = le nombre d'éléments que le modèle assigne à X avec justesse
- $b$  = le nombre d'éléments que le modèle fait l'erreur d'assigner à X
- $c$  = le nombre d'éléments que le modèle fait l'erreur de rejeter de X (à savoir les éléments que le modèle aurait dû assigner à X, mais qui n'ont pas été attribués)

À partir de ces quantités, vous pouvez calculer les mesures de performance suivantes :

- Précision =  $a / (a + b)$
- Rappel =  $a / (a + c)$

Généralement, pour améliorer la précision, vous devez assumer les conséquences d'une diminution du taux de rappel. En d'autres mots, le modèle n'assignera un élément à une catégorie que s'il est persuadé de la justesse de cette attribution. Toutefois, en insistant sur le fait d'être persuadé de l'exactitude de l'opération, le modèle court le risque de rejeter des éléments qui appartiennent réellement à la catégorie. À la figure **Précision, rappel et confiance**, vous pouvez constater cet effet au-delà du niveau de confiance de 85 pour cent.

## Présent dans les meilleures catégories N



Présent dans les meilleures catégories N

Lorsqu'un modèle classe un objet texte, il fournit une liste de catégories et la probabilité (niveau de confiance) que l'objet y appartienne. Étant donné que le modèle classe les catégories fournies avec la plus forte probabilité en premier, quelles sont les chances que la bonne catégorie apparaisse parmi les deux premières, les trois premières, etc.?

- L'axe horizontal, **Meilleures catégories N**, montre les meilleures catégories (c'est-à-dire, les mieux classées) : les deux premières, trois premières, etc.
- L'axe vertical, **Comprend la bonne catégorie**, correspond au pourcentage de probabilité que la bonne catégorie se trouve parmi les catégories les mieux classées suivantes.

Supposons que la catégorie obtenue dont la probabilité est la plus forte corresponde à la première estimation du modèle. Si l'on se rend jusqu'à la catégorie classée quatrième, c'est comme si l'on donnait quatre tentatives d'estimation au modèle. Plus le modèle a de tentatives, plus il est susceptible de trouver la bonne réponse. Moins le modèle a de tentatives, meilleur il sera en classification.

Le modèle présenté dans la figure est plutôt efficace : sa première estimation (à savoir la catégorie obtenue la mieux classée) se situe tout juste en-dessous de 80 pour cent du temps, et nous ne lui avons donné qu'une seule autre tentative pour atteindre près de 100 pour cent d'exactitude.

Une façon d'utiliser cette évaluation est d'aviser les agents du nombre de catégories à analyser quand vient le temps de choisir une réponse standard. S'il y a une probabilité de 95 pour cent que la bonne catégorie fasse partie des trois mieux classées, vous pouvez informer les agents de ne prendre en considération que les trois premières catégories.

## Confusion de catégorisation

Le tableau Confusion de catégorisation répertorie jusqu'à 10 paires de catégories que le modèle est susceptible de confondre.

Category Confusion

Category One	Path One	Category Two	Path Two	Confusion
Tigers	/Animal Planet	Wolves	/Animal Planet	4
Apes	/Animal Planet	Bats	/Animal Planet	7

Confusion de catégorisation

La colonne Confusion présente le pourcentage de probabilité que le modèle classe par erreur un

élément de la Catégorie 1 dans la Catégorie 2. Par exemple, la figure **Confusion de catégorisation** montre que ce modèle classe les tigres comme étant des loups à raison de 4 pour cent du temps.

Une probabilité de 50 pour cent représenterait une confusion totale; le modèle n'est pas capable de différencier les loups des tigres. Une probabilité de 100 pour cent voudrait dire que le modèle identifie toujours les loups comme étant des tigres et les tigres comme étant des loups—une inversion totale.

Si une paire de catégories obtient une probabilité de plus de 20 pour cent et que ces deux catégories ont plus de trois ou quatre membres, vous devriez envisager de les modifier. Vous pouvez les modifier des deux façons suivantes :

- Les fusionner, c'est-à-dire en venir à la conclusion qu'elles sont semblables au point de n'en faire qu'une seule catégorie.
- Les différencier davantage en ajoutant plus d'interactions d'entraînement fortement contrastées dans l'objet de données d'entraînement.

## Onglet Résultats par catégorie

Cet onglet affiche les mêmes évaluations que l'onglet **Tous les résultats**, mais pour une seule catégorie.



Résultats par catégorie

La **Confusion de catégorisation** montre la ou les catégories qui sont susceptibles d'être confondues avec la catégorie sélectionnée à gauche.

# Validation croisée

Dans la validation croisée, le Training Server suit les étapes suivantes :

1. Il crée un seul modèle en utilisant toutes les données.
2. Il divise les données en  $x$  partitions,  $x$  = correspondant à 3, à 5 ou à 10.
3. Il crée ensuite un certain nombre de modèles partiels, soit un pour chaque partition, et chacun d'entre eux utilise une combinaison différente de  $x - 1$  partitions.  
Par exemple, si les données sont divisées en trois partitions (A, B et C), le Training Server créera le modèle X en utilisant les partitions A et B, le modèle Y en utilisant les partitions A et C ainsi que le modèle Z en utilisant les partitions B et C.
4. Il teste chacun de ces modèles partiels en fonction de la partition qui a été omise lors de sa création.  
Dans l'exemple précédent, il teste le modèle X avec la partition C, le modèle Y avec la partition B et le modèle Z avec la partition A.
5. Il regroupe les résultats de tous ces tests et les présente comme l'évaluation de l'ensemble du modèle.

Les idées suivantes sous-tendent le concept de la validation croisée :

- La meilleure façon de tester un modèle est de l'appliquer à des données qui n'ont pas été utilisées lors de la création du modèle en question.
- Un modèle créé à l'aide de la plupart des données est d'une utilité semblable à un modèle créé à l'aide de toutes les données. Ainsi, tester tous les modèles possibles de 90 pour cent, par exemple, sera une bonne indication de la qualité du modèle de 100 pour cent.

Étant donné que la validation croisée prolonge la durée nécessaire pour créer un modèle, il est préférable de ne pas la sélectionner pour les objets d'entraînement très volumineux ou ceux pour lesquels vous avez sélectionné une qualité d'entraînement d'un niveau régulier de 6.

# Échantillons pour détecter la langue, l'expression de sentiments et la nécessité d'action

L'installation du **eServices Manager** comprend des échantillons qui détectent

- La langue
- L'expression de sentiments
- La nécessité d'action

## Règles de filtrage

Genesys fournit des **échantillons de règles de filtrage** qui analysent les interactions pour détecter l'expression de sentiments et la nécessité d'action

## Modèles

### Détection de la langue

Dans le cadre de l'installation du **eServices Manager**, Genesys fournit un modèle qui classe les courriels en anglais, en français, en allemand, en italien, en portugais, en russe ou en espagnol.

Pour importer ce modèle et son objet d'entraînement,

1. Ouvrez eServices Manager, puis sélectionnez **inconnu** comme langue. Si cette option n'est pas disponible, vous devez en créer une dans le Genesys Administrator ou le Configuration Manager (sous **Attributs métier > Langues**).
2. Ensuite, importez le fichier lang.kme, qui se trouve dans le répertoire <eServicesManagerHome>\LanguageModel.

### Conseil

Si vous acceptez le chemin par défaut proposé lors de l'installation de GAX, <eServicesManagerHome> est normalement

- Windows : C:\genesys\GCTI\

- Linux : /home/genesys/GCTI

L'objet d'entraînement est composé de sept catégories, soit une pour chaque langue. Chacune des catégories contient un certain nombre d'objets texte dans la langue donnée. Vous pouvez également **ajouter plus d'objets texte** à ces catégories. Cela peut s'avérer particulièrement pratique si vous avez une collection d'objets texte (comme des courriels) dont le sujet est pertinent pour votre entreprise. Après avoir ajouté des objets texte, vous devez entraîner un nouveau modèle pour pouvoir tirer parti des données ajoutées.

Vous pouvez aussi ajouter d'autres langues au modèle de la façon suivante :

1. eServices Manager : Dans l'onglet **Réponses standards** (toujours avec **inconnu** sélectionné comme langue), **ajoutez une catégorie** pour la nouvelle langue à l'arborescence de catégories **LanguageDetection**.
2. Content Analyzer :
  1. Dans l'onglet **Entraînement**, sélectionnez l'objet d'entraînement **LanguageDetection**, sélectionnez la catégorie de la nouvelle langue dans l'objet d'entraînement, puis **ajoutez des objets texte** dans la langue en question.
  2. **Entraînez un nouveau modèle** qui comprend la nouvelle langue.

Vous pouvez faire de même pour toute langue prise en charge par E-mail Server (E-mail Server prend en charge les mêmes langues que la version de JRE fournie avec Genesys eServices). Toutefois, Genesys n'a testé aucune langue autre que celles énumérées ci-dessus.

## La nécessité d'action

Pour utiliser l'échantillon de nécessité d'action, importez le fichier `Actionability.kme`, qui se trouve dans le répertoire `<eServicesManagerHome>\ActionabilityModel`.

Cela vous donne :

- Un modèle **Actionability** pour analyser la nécessité d'action.
- L'objet d'entraînement **Actionability** qui a créé ce modèle.
- Une arborescence de catégories **Actionability** qui contient les catégories à attribuer à des interactions en réponse à l'analyse.

## L'expression de sentiments

Pour déployer l'échantillon d'expression de sentiments, utilisez la procédure suivante :

1. Dans le Configuration Manager ou le Genesys Administrator, créez une langue nommée **English\_Sentiment**.
2. En réglant eServices Manager à cette langue, importez le fichier `EnglishSentiment.kme`, qui se trouve dans le répertoire `<eServicesManagerHome>\SentimentModel`.

Cela vous donne :

- Un modèle **SentimentSampleModel** pour analyser l'expression de sentiments.
- L'objet d'entraînement **Sentiment** qui a créé ce modèle.
- Un arborescence de catégories **SentimentDetection** qui contient les catégories à attribuer à des interactions en réponse à l'analyse.
- FirstName
- FullName
- LastName
- Signature

# Installation du plugiciel Content Analyzer pour GAX

## Conditions préalables

- Genesys Administrator Extension (GAX) 8.5.230.03 ou version plus récente
- UCS 8.5.200.17 ou version plus récente
- Classification Server 8.5.200.05 ou version plus récente
- Si une version antérieure du plugiciel Content Analyzer pour GAX est installée sur l'hôte actuel, **désinstallez cette version.**

## Créer un fichier .zip

1. Créez un dossier temporaire sur votre bureau.
2. Copiez le dossier contenant le progiciel d'installation à partir de son emplacement d'origine (généralement le CD du produit) dans le dossier temporaire.
3. Copiez le dossier de modèles à partir de son emplacement d'origine dans le dossier temporaire.
4. Compressez le dossier temporaire.

Vous pouvez aussi communiquer avec le service client de Genesys pour obtenir le fichier .zip qui contient le progiciel d'installation et les modèles qui y sont associés.

## Installer avec l'assistant d'installation GAX

1. Ajoutez l'option suivante à l'onglet **Options d'application** de l'objet d'application Genesys Administrator Extension (GAX) :
  - Nom de la section : **asd**
  - Nom de l'option : **plugin\_ip\_list**
  - Valeur de l'option : CntAnlzPlgnAdm64 (hôte Windows) ou CntAnlzPlgnAdm (hôte Linux)

### Conseil

Cette option est aussi utilisée par eServices Manager et le **plugiciel Privacy Manager pour Administrator**. Si



vous utilisez à la fois le Privacy Manager et le eServices Manager, la valeur peut correspondre à une liste séparée par des virgules; par exemple eSMngrPLgnAdm,PrivacyMng

2. Redémarrez GAX.
3. **Téléversez le progiciel d'installation de Content Analyzer.** Si vous avez déjà téléversé le progiciel d'installation, vous devez maintenant le faire de nouveau.
4. Effectuez le **processus d'installation du plugiciel.**
5. Redémarrez GAX.

## Configurer

1. Créez une application de type Application Cluster pour Content Analyzer.
2. Ajoutez des locataires à l'application de Content Analyzer.
3. Précisez l'hôte et le port. L'application GAX et l'objet d'application de type Application Cluster doivent être configurés avec le même hôte afin de permettre des connexions TLS communes.
4. Connectez l'application de Content Analyzer au UCS et au Classification Server.
5. Configurez ces connexions : TLS, PADD, etc.
6. Accédez à l'application GAX et indiquez le nom de l'application de Content Analyzer comme étant la valeur de l'option **app-name** dans la section **[gax-content-analyzer]**.
7. Si vous avez besoin d'assistance multilingue, configurez le paramètre UTF-8 JVM pour tous les composants Java (UCS, GAX Server, navigateur) dans le fichier .ini correspondant (ou le fichier .bat si vous avez démarré le composant à partir de la ligne de commande) :  
Fichier .ini : [JavaArgs] -Dfile.encoding=UTF-8 ...  
Fichier .bat : set JAVA\_OPTS=%JAVA\_OPTS% -Dfile.encoding=UTF-8 ...
8. Redémarrez GAX.

### Conseil

À propos de l'assistance multilingue : Dans la fenêtre **Préférences d'utilisateur** de GAX (sous le menu Préférences), **Utiliser les paramètres système** fait référence à l'utilisation des paramètres qui se trouvent dans la fenêtre **Préférences système**. Cela ne correspond pas au paramètre Région et langue de l'ordinateur hôte.

Si **Utiliser les paramètres système** est sélectionné dans le champ **Langue** des **Paramètres de l'utilisateur pour l'emplacement** de GAX, cela fait référence à la langue qui est définie dans le champ **Langue** des **Préférences système**. Veuillez noter que si vous souhaitez pouvoir choisir toute autre langue que l'anglais dans ce champ, vous devez installer le plugiciel de module linguistique de la langue en question.

## Accès basé sur le rôle

Pour donner l'accès au menu GAX et au logiciel Content Analyzer pour GAX, accordez les privilèges suivants aux utilisateurs dans GAX :

- **COM/Access Configmanager**
- **content-analyzer/Content Analyzer Plug-in for GAX Access**