



ProtAnt (Windows)

Version 1.0.1 (édité le 21 octobre 2015)

Laurence Anthony, Ph.D.

Centre d'Enseignement de la Langue Anglaise de la Science et de l'Ingénierie, École de Science et d'Ingénierie, Université de Waseda, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japon.

Version du manuel d'utilisation : 003 (8 avril 2016)

File	KeyTypes	KeyTokens	NormKeyTypes	
1	5.txt	152	464	284.644
2	7.txt	114	392	256.757
3	6.txt	93	409	254.098
4	4.txt	118	419	243.802

File	Key 1	Key 2	Key 3	Key 4	Key 5	
1	5.txt	islam	blair	muslim	brotherhood	islamic
2	7.txt	islam	blair	muslim	brotherhood	assad
3	6.txt	islam	muslims	muslim	islamic	pastor
4	4.txt	islam	blair	muslim	brotherhood	islamic

Keyword	Freq	Keyness	
1	islam	41	239.602
2	muslims	37	210.93
3	batten	17	147.137
4	football	32	135.558
5	blair	35	132.355

Introduction

ProtAnt est un logiciel gratuit de détection de textes prototypiques développé en collaboration avec Paul Baker de l'université de Lancaster, GB. *ProtAnt* permet d'identifier les mots-clés représentant les traits caractéristiques de l'ensemble d'un corpus de textes (**encodé en UTF-8**) en le comparant à un corpus de référence (**encodé en UTF-8**). Puis, *ProtAnt* analyse les fichiers individuellement et y compte le nombre de mots-clés présents. Les fichiers du corpus sont ensuite classés selon le nombre de mots-clés (normalisé selon différentes méthodes) qu'ils contiennent. *ProtAnt* fonctionne sur tout ordinateur utilisant un système d'exploitation Microsoft Windows (testé sur Win 98/Me/2000/NT, Vista, 7, 8) ou Macintosh OS X (testé jusqu'à OS X 10.9 Mavericks). *ProtAnt* est développé en langage Python et Qt et utilise le compilateur *PyInstaller* pour générer des exécutables pour les différents systèmes d'exploitation.

Mise en route (aucune installation requise)

Windows

Sur Windows, double-cliquez simplement sur l'icône *ProtAnt* pour lancer le programme.

Macintosh OS X

Sur Macintosh, double-cliquez simplement sur le fichier zip *ProtAnt*. Le fichier zip va extraire l'application *ProtAnt*. Il vous suffit ensuite de déplacer l'application *ProtAnt* dans votre dossier application, sur votre bureau ou à tout autre endroit de votre choix. Supprimez le fichier zip lorsque vous avez fini.

Identifier des textes prototypiques

1^{re} étape : Sélectionnez les fichiers cibles du corpus que vous souhaitez analyser. Quatre méthodes à votre disposition :

- a) Dans le menu Fichier, cliquez sur Ouvrir Fichier(s) Cible(s)... et sélectionnez les fichiers que vous souhaitez analyser :
- b) Dans le menu Fichier, cliquez sur Ouvrir Dossier Cible... Et sélectionnez un dossier contenant des fichiers que vous souhaitez analyser.
- c) Cliquez sur le bouton « Ajouter Fichier Corpus » sous l'étiquette « Corpus Cible » et sélectionnez les fichiers à analyser :
- d) Glissez-déposez directement les fichiers dans l'application *ProtAnt*.

Note 1 : Pour retirer les fichiers du corpus de la liste, cliquez sur "Fermer Fichiers Cibles" dans le menu "Fichier", ou sur le bouton "Effacer" sous l'étiquette "Corpus Cible".

2^e étape : Sélectionnez les fichiers du corpus de référence que vous souhaitez utiliser. Quatre méthodes à votre disposition :

- a) Dans le menu Fichier, cliquez sur Ouvrir Fichier(s) Référence(s)...et sélectionnez les fichiers que vous souhaitez analyser :
- b) Dans le menu Fichier, cliquez sur Ouvrir Dossier Référence... et sélectionnez un dossier contenant des fichiers que vous souhaitez analyser.
- c) Cliquez sur le bouton « Ajouter Fichier Corpus » sous l'étiquette « Corpus de Référence » et sélectionnez les fichiers à analyser.
- d) Glissez-déposez directement les fichiers dans l'application *ProtAnt*.

Note 1 : Pour retirer les fichiers du corpus de la liste, cliquez sur "Fermer Fichiers Références" dans le menu "Fichier", ou sur le bouton "Effacer" sous l'étiquette "Corpus de Référence".

3^e étape : Définissez la méthode de calcul de la statistique *keyness*, sa valeur seuil (p-value) et la fréquence minimum de mot-clé que vous souhaitez appliquer pour générer les mots-clés du corpus cible.

4^e étape : Déterminez la fréquence normalisée qui sera affichée dans la fenêtre de résultats.

5^e étape : Vous pouvez choisir d'ignorer la casse dans le corpus cible (tous les mots s'écrivent alors en minuscule), il est également possible de personnaliser la définition des tokens en indiquant l'expression régulière (regex) adéquate.

6^e étape : Cliquer sur "Démarrer" pour commencer le processus de segmentation.

Note 1 : Les résultats de la détection des textes prototypiques s'affichent dans la fenêtre supérieure droite. Les mots-clés apparaissant dans chacun des fichiers (classés) du corpus s'affichent dans la fenêtre centrale droite. La liste complète des mots-clés s'affiche dans la fenêtre inférieure droite. Toute colonne peut être triée par ordre croissant ou décroissant en cliquant sur son titre.

Note 2 : Le processus de détection des textes prototypiques peut être arrêté à tout moment en cliquant sur le bouton "Arrêter".

Caractéristiques supplémentaires

Les résultats peuvent être sélectionnés, copiés et collés en utilisant les raccourcis connus selon le système d'exploitation :

Windows :	CTRL-A ⇒ Sélectionner tout	CTRL-C ⇒ Copier	CTRL-V ⇒ Coller
Macintosh :	CMD-A ⇒ Sélectionner tout	CMD-C ⇒ Copier	CMD-V ⇒ Coller

NOTES

Commentaires/Suggestions/Correction de bugs

Toutes les nouvelles versions et corrections des bugs sont listées dans l'historique de révisions présent en fin de document. Toutefois, si vous trouviez un bug, ou si vous aviez la moindre suggestion permettant d'améliorer le programme, n'hésitez pas à me le faire savoir et je les prendrai en considération dans les futures versions.

Ce logiciel est disponible en tant que 'logiciel libre' (gratuitiel), entendu selon le sens de la licence ci-dessous. Il est important pour le financement de mes recherches d'avoir des retours de tout succès rencontré grâce à ce logiciel. Par conséquent, si vous trouvez le logiciel utile, envoyez-moi, s'il vous plait, un email en décrivant brièvement la manière dont vous l'utilisez.

CITER/FAIRE RÉFÉRENCE À *ProtAnt*

Pour citer ou faire référence à *ProtAnt*, merci d'utiliser la méthode suivante, élaborée selon les normes de l'APA [Ndt : l'APA ne propose pas de version française de ses modèles, la traduction s'inspire ici des propositions émises par la bibliothèque de l'université de Laval, pour plus de détails, voir <http://www.bibl.ulaval.ca/services/citation-de-sources/style-apa>] :

Anthony, L. et Baker, P. (Année d'édition). *ProtAnt* (Version NUMÉRO DE LA VERSION) [Logiciel]. Tokyo, Japon : Université de Waseda. Récupéré de <http://www.laurenceanthony.net/>

Par exemple, si vous téléchargez *ProtAnt 1.0.0*, qui est sorti en 2015, vous citerez ou ferez référence selon ce qui suit :

Anthony, L. and Baker, P. (2015). *ProtAnt* (Version 1.0.0) [Logiciel]. Tokyo, Japon : Université de Waseda. Récupéré de <http://www.laurenceanthony.net/>

Veuillez noter que les instructions de l'APA ne sont pas totalement claires en ce qui concerne la citation de logiciels, ainsi la nécessité de la déclaration "Récupéré de ..." est discutable. Pour plus de détails : <http://owl.english.purdue.edu/owl/resource/560/10/>

LICENCE pour ProtAnt

ProtAnt 1.0 et toute mise à jour mineure diffusée par AntLab Solutions (dénommés conjointement : le « Logiciel »).

CONDITIONS RÉGISSANT L'UTILISATION DU LOGICIEL

Le Logiciel est protégé par un copyright (droit d'auteur) et ne doit pas être utilisé, diffusé, modifié, adapté, distribué, transmis, transféré ou publié ou autrement reproduit sous quelque forme et par quelque moyen que ce soit, autres que ceux expressément prévus par les conditions exposées ci-dessous. En installant Le Logiciel, vous acceptez d'être lié par les conditions de la licence. Cette licence ProtAnt (« Licence ») est conclue entre AntLab Solutions, Tokyo, Japon en tant que concédant, et vous-même, en tant que titulaire de la Licence, à la date de votre utilisation du Logiciel. Le Logiciel est dit en utilisation lorsqu'il est chargé dans la mémoire vive (RAM) ou installé dans la mémoire permanente de l'ordinateur, par exemple sur le disque dur ou un dispositif de stockage.

1. Matériel sous licence

Les présentes conditions régissent votre utilisation du Logiciel mais n'incluent pas les versions ultérieures (par exemple ProtAnt 2.0).

2. Concession de licence

AntLab Solutions vous accorde une licence personnelle, non-exclusive et non-transférable (la « Licence ») pour l'utilisation du Logiciel selon les conditions spécifiques énumérées ci-dessous.

a) Utilisation à but non-commercial (gratuitiel) :

Vous pouvez utiliser Le Logiciel dans un but non-lucratif sur un ou plusieurs ordinateurs ou sur un réseau tant que vous demeurez l'unique utilisateur du Logiciel. (On entend par "réseau" toute combinaison de deux ou plusieurs ordinateurs qui sont électroniquement reliés et capables de partager l'utilisation d'un même programme logiciel). Vous n'êtes en aucun cas autorisé à vendre, louer, distribuer, transférer, sous-licencier le Logiciel ou à en disposer de toute autre façon, en tout ou partie, en contrepartie de toute valeur commerciale ou rétribution réelle ou potentielle.

b) Utilisation à titre d'évaluation commerciale (Essai) :

Vous pouvez évaluer (essayer) le Logiciel à des fins commerciales pendant une période n'excédant pas quatorze (14) jours à partir de la date de téléchargement, sur un ou plusieurs ordinateurs ou sur un réseau, tant que vous demeurez l'unique utilisateur du Logiciel.

c) Utilisation commerciale

En contrepartie du paiement d'un droit de licence commerciale fixé par AntLab Solutions, vous pouvez utiliser le Logiciel à des fins commerciales ou non lucratives, sur un ou plusieurs ordinateurs ou sur un réseau, tant que vous demeurez l'unique utilisateur du Logiciel. (On entend par "réseau" toute combinaison de deux ou plusieurs ordinateurs qui sont électroniquement reliés et capables de partager l'utilisation d'un même programme logiciel). Vous obtiendrez une licence distincte pour chaque utilisateur supplémentaire du Logiciel (que ces utilisateurs soient connectés sur un réseau ou non). Vous n'êtes en aucun cas autorisé à vendre, louer, distribuer, transférer, sous-licencier le Logiciel ou à en disposer de toute autre façon, en tout ou partie, en contrepartie de toute valeur commerciale ou rétribution réelle ou potentielle.

3. Résiliation

Vous pouvez résilier cette Licence à tout moment en désinstallant Le Logiciel et en le supprimant. La Licence sera également résiliée si vous enfreignez l'une des conditions de la Licence.

4. Droits de propriété

Le Logiciel est concédé sous licence et ne vous est aucunement vendu. Antlab Solutions se réserve tous les droits qui ne vous sont pas expressément accordés. La propriété du Logiciel et ses droits propriétaires associés, y compris mais sans s'y limiter les brevets et demandes de brevets, sont conservés par Antlab Solutions. Le Logiciel est protégé par les lois japonaises sur les droits d'auteur (copyright) et par des traités internationaux. Par conséquent, vous êtes tenu de vous conformer à ces lois et traités lors de l'utilisation du Logiciel. Vous acceptez de ne supprimer du Logiciel aucun droit d'auteur, aucune marque commerciale ou autre indication de propriété liés à Antlab Solutions.

5. Distribution

À l'exception des cas spécifiquement définis dans la Section 2, ou dans le cas d'accord écrit signé par Antlab Solutions et vous-même, vous n'êtes pas autorisé à distribuer Le Logiciel, en tout ou partie, sous quelque forme et par quelque moyen que ce soit.

6. Transfert et restrictions d'utilisation

Vous n'êtes en aucun cas autorisé à vendre, licencier, sous-licencier, prêter, louer, partager, attribuer, transmettre, télétransmettre, exporter, distribuer ou autrement transférer le Logiciel à des tiers, à l'exception des cas spécifiquement prévus dans le présent Contrat de Licence ou dans tout autre contrat conclu avec Antlab Solutions. Vous n'êtes pas en droit de modifier, procéder à une ingénierie inverse, décompiler, décrypter, extraire ou autrement désassembler le Logiciel.

7. Garanties

ANTLAB SOLUTIONS NE FOURNIT AUCUNE GARANTIE CONCERNANT LE LOGICIEL ET EN PARTICULIER, NE GARANTIT PAS QUE LE LOGICIEL FONCTIONNERA CONFORMÉMENT A LA DOCUMENTATION JOINTE POUR TOUTE COMBINAISON DE PLATEFORME MATÉRIELLE, D'ENVIRONNEMENT LOGICIEL OU DE CONFIGURATION, OU QU'IL SERA COMPATIBLE AVEC TOUT SYSTÈME INFORMATIQUE. SI LE LOGICIEL VENAIT À ÊTRE DÉFECTUEUX POUR QUELQUE RAISON QUE CE SOIT, VOUS ASSUMEREZ LA TOTALITÉ DES FRAIS DE TOUTE RÉPARATION OU DE TOUT REMPLACEMENT NÉCESSAIRES.

8. Clause de non-responsabilité

ANTLAB SOLUTIONS NE GARANTIT AUCUNEMENT QUE LE LOGICIEL OU SERVICE SOIT EXEMPT DE BUGS, DÉFAUTS, ERREURS OU OMISSIONS. LE LOGICIEL OU SERVICE EST FOURNI EN L'ÉTAT ET ANTLAB SOLUTIONS N'ASSUMERA AUCUNE AUTRE GARANTIE OU CONDITION, EXPLICITE OU IMPLICITE, CONCERNANT LE LOGICIEL, Y COMPRIS MAIS SANS S'Y LIMITER, LES GARANTIES IMPLICITES OU CONDITIONS DE QUALITÉ MARCHANDE OU DE CONFORMITÉ À UNE UTILISATION PARTICULIÈRE.

9. Limitation de responsabilité

ANTLAB SOLUTIONS NE SERA TENU D'ASSUMER AUCUNE RESPONSABILITÉ NI OBLIGATION DE VERSER DES DOMMAGES ET INTÉRÊTS OU TOUTE AUTRE COMPENSATION CONCERNANT, MAIS SANS S'Y LIMITER, LES COÛTS LIÉS AU REMPLACEMENT DE MATÉRIEL OU À LA PERTE DE DONNÉES, TOUT MANQUE À GAGNER, TOUTE PERTE DE REVENUS, NI NE SERA TENU D'ASSUMER AUCUN AUTRE DOMMAGE DIRECT, INDIRECT, ANNEXE, SPÉCIAL, GÉNÉRAL, PUNITIF OU CONSÉCUTIF, RÉSULTANT DE CETTE LICENCE OU DE L'UTILISATION OU DE L'INCAPACITÉ D'UTILISATION DU LOGICIEL OU SERVICE. LE CUMUL GLOBAL DES RESPONSABILITÉS ASSUMÉES PAR ANTLAB SOLUTIONS (EN VERTU DU CONTRAT (Y COMPRIS LES VIOLATIONS FONDAMENTALES), DE LA GARANTIE, DES RESPONSABILITÉS DÉLICITUELLES (Y COMPRIS DÉLITS DE NÉGLIGENCE), DE LA RESPONSABILITÉ DU FAIT DES PRODUITS, DE LA LÉGISLATION SUR LA PROPRIÉTÉ INTELLECTUELLE, OU TOUTE AUTRE DOCTRINE JURIDIQUE) EN RAPPORT AVEC LE LOGICIEL ET/OU LA PRÉSENTE LICENCE NE PEUT EN AUCUN CAS EXCÉDER LE MONTANT DU DROIT DE LICENCE RÉGLÉ PAR VOUS-MÊME À ANTLAB SOLUTIONS. EN OUTRE, ANTLAB SOLUTIONS NE SAURAIT EN AUCUN CAS ÊTRE TENU POUR RESPONSABLE DE L'EXÉCUTION TARDIVE OU DE L'INEXÉCUTION DE SES OBLIGATIONS LIÉES À LA PRÉSENTE LICENCE FAISANT SUITE À TOUTE CAUSE OU CONDITION ÉCHAPPANT À SON CONTRÔLE RAISONNABLE.

10. Juridiction compétente

Les présentes conditions sont régies par la législation japonaise et les tribunaux japonais seront seuls compétents.

PROBLÈMES CONNUS

Aucun à ce jour.

HISTORIQUE DES RÉVISIONS

1.0.1

Une modification mineure a été apportée.

Bug corrigé : Lorsqu'une classe générale de caractères Unicode (par exemple `\p{L}` pour les lettres) était utilisée pour la définition des tokens, les textes écrits dans une autre langue que l'anglais n'étaient pas traités correctement. Le problème est maintenant résolu.

1.0.0

Cette version est la première du programme.

Copyright 2015 Laurence Anthony. Tous droits réservés.

Traduit de l'anglais par Jean-Yves Préault (jeanyvespreault@gmail.com).