

❑ DOCUMENTS ASSOCIES

Liana Ermakova Short Text Contextualization in Information Retrieval: Application to Tweet Contextualization and Automatic Query Expansion. Thèse de doctorat, Université de Toulouse-le-Mirail, 2016. https://www.irit.fr/publis/SIG/my_thesis.pdf

Liana Ermakova. A Method for Short Message Contextualization: Experiments at CLEF/INEX. Conference on Multilingual and Multimodal Information Access Evaluation (CLEF 2015), Toulouse, [Springer](https://www.springer.com), p. 352-363, 2015 (Best paper award). https://www.irit.fr/publis/SIG/2015_CLEF_LE.pdf

Liana Ermakova, Josiane Mothe. IRIT at INEX 2014: Tweet Contextualization Track. Conference on Multilingual and Multimodal Information Access Evaluation (CLEF 2014), 2014. <http://ceur-ws.org/Vol-1180/CLEF2014wn-Inex-ErmakovaEt2014.pdf>

Liana Ermakova, Josiane Mothe. IRIT at INEX 2013: Tweet Contextualization Track. Initiative for the Evaluation of XML Retrieval (INEX 2013), Valencia, Spain, Polytechnic University of Valencia, 2013. <http://www.clef-initiative.eu/documents/71612/58a64b0a-cf0c-4751-a91f-9c8aba4312e1>

❑ CONTACTS

J. MOTHE
Institut de Recherche en Informatique de Toulouse
Equipe Systèmes d'Informations Généralisées
Campus Université Paul Sabatier
118, Route de Narbonne – 31062 Toulouse Cedex 4
Email : Josiane.Mothe@irit.fr
Tel : 05 61 55 64 44



EQUIPE SIG
SYSTEMES D'INFORMATIONS GENERALISEES
<http://www.irit.fr>

r

CONTEXTUALISATION DE TWEETS

Skill

Prototype

Product

UMR 5505 CNRS-INP-UPS

118 ROUTE DE NARBONNE 31062 TOULOUSE CEDEX 4

❑ DESCRIPTION

La contextualisation de tweets vise à générer un résumé permettant de remettre en contexte un texte qui, de par sa taille, ne contient pas l'ensemble des éléments permettant à un lecteur de comprendre tout ou partie de son contenu. Nous visons plus particulièrement les tweets mais les méthodes développées peuvent s'appliquer à tout texte court.

L'objectif de la contextualisation est donc de fournir à son lecteur des informations qui rendent le message compréhensible.

❑ ARCHITECTURE

Fournir un tel résumé implique de s'appuyer sur des ressources disponibles pour le constituer. Dans nos développements, nous nous sommes appuyés sur la ressource Wikipédia qui possède l'avantage d'être libre d'utilisation. Les résumés sont construits comme des extraits de plusieurs documents.

❑ EVALUATION

Nous avons proposé des mesures d'évaluation des contenus des résumés proposés et de leur lisibilité.

$$\text{LogSim}(T, S) = \sum_{t \in T} P(t|T) \times \frac{\min(R(t, T), R(t, S))}{\max(R(t, T), R(t, S))}$$

Ces mesures ont l'avantage d'être plus stables que les mesures de la littérature comme ROUGE.

$$R(t, X) = \log(1 + P(t|X) \times |T|)$$

Whitney Elizabeth Houston (August 9, 1963 February 11, 2012) was an American recording artist, actress, producer, and model. Houston was one of the world's best-selling music artists, having sold over 170 million albums, singles and videos worldwide. Robert Barisford "Bobby" Brown (born February 5, 1969) is an American R&B singer-songwriter, occasional rapper, and dancer. After a three-year courtship, the two were married on July 18, 1992. On March 4, 1993, Houston gave birth to their daughter Bobbi Kristina Houston Brown, her only child, and his fourth. With the missed performances and weight loss, rumors about Houston using drugs with her husband circulated. Following fourteen years of marriage, Brown and Houston filed for legal separation in September 2006. Their divorce was finalized on April 24, 2007, with Houston receiving custody of their then-14-year-old daughter. On February 11, 2012, Houston was found unresponsive in suite 434 at the Beverly Hilton Hotel, submerged in the bathtub.

Figure 1. Exemple d'un résumé contextualisant le tweet "Bobby Brown – Fighting #WhitneyHouston's Family to See Bobbi Kristina". Les phrases utilisées proviennent de Wikipedia.



[Home](#) | [About](#) | [2011](#) | [2012](#) | [2013](#)

INEX 2013 Tweet Contextualization Track

LAST NEWS

2014/05/23: Release of informativeness evaluation postponed to Monday 25 May

2014/05/15: Late runs will be accepted until 20 May

2014/05/09: Run submission deadline extension 16 May

❑ CAMPAGNE D'ÉVALUATION DANS LE CADRE DE CLEF

Nous avons mis en place une campagne d'évaluation dans le cadre de INEX (inclus à partir de 2013 dans la campagne internationale CLEF). Lors de la campagne 2011, 11 équipes avaient participé et soumis 23 runs. La campagne 2012 a quant à elle mobilisé 13 équipes de 10 pays différents et 33 runs ont été soumis. <https://inex.mmci.uni-saarland.de/tracks/qa/>

❑ COLLABORATIONS

Laboratoire d'Informatique d'Avignon
LIMSI-CNRS, Université de Paris-Sud
LSIS – Université d'Aix-Marseille

