

FLaReNet: una red para fomentar los recursos lingüísticos

Fostering Language Resources Network: FLaReNet

Núria Bel

Universitat Pompeu Fabra
Roc Boronat, 138
08018 Barcelona, España
Telf. 0034 935422307
nuria.bel@upf.edu

Nicoletta Calzolari

CNR Istituto de Linguistica Computazionale
Via Giuseppe Moruzzi, 1
56124Pisa, Italia
nicoletta.calzolari@ilc.cnr.it

Resumen: FLaReNet es una red temática cuyo objetivo es la elaboración de estrategias y recomendaciones para la promoción y el desarrollo de las tecnologías lingüísticas y los recursos lingüísticos asociados por su importancia para minimizar el impacto de la diversidad lingüística en una Europa digital multilingüe. Los resultados de este proceso de reflexión conjunta de investigadores y profesionales de todo el mundo serán la base para políticas europeas consensuadas de financiación y promoción del sector.

Palabras clave: Recursos lingüísticos y Tecnologías Lingüísticas

Abstract: FLaReNet is a thematic network whose objective is the preparation of strategies and recommendations for the promotion and development of language technologies and the associated language resources because of their importance for minimizing the impact of the linguistic diversity in a digital and multilingual Europe. The results of this joint process of reflection by researchers and professionals of all around the world will be the basis of European agreed policies for funding and promoting this sector.

Keywords: Language Resources and Technology

1 La Red FLaReNet

Europa es multilingüe y esta realidad tiene que ser respetada y apoyada por todos, ciudadanos e instituciones. Pero para conseguir que la diversidad lingüística no se convierta en una barrera en el mundo de la información digital, es un factor crítico el desarrollo de tecnologías lingüísticas para todas las lenguas de la Unión y para la transferencia entre ellas. Las tecnologías lingüísticas basadas, de una forma u otra, en grandes cantidades de información de una lengua particular (en forma de textos, corpus etiquetados o diccionarios y gramáticas), son los llamados "Recursos Lingüísticos" que son una condición necesaria para el desarrollo de aplicaciones tan importantes como la traducción automática, pero también para la extracción de información o el resumen automático. Todas estas aplicaciones son las que han de ayudar a los ciudadanos a diluir las barreras lingüísticas

y conseguir una Europa digital verdaderamente multilingüe.

La disponibilidad de estos recursos lingüísticos actualmente está muy desequilibrada según la lengua. Esta desigualdad en cuanto calidad y cantidad tiene que ver tanto con motivos comerciales y de mercado, como con políticas lingüísticas promovidas localmente o con el desarrollo de esta área de investigación en cada país. En su mayor parte, los recursos existentes para la mayoría de las lenguas europeas son todavía pequeños e insuficientes para garantizar el buen funcionamiento de las aplicaciones multilingües de las que hemos hablado.

Para garantizar la existencia de aplicaciones en todas las lenguas es necesario generar una política común de apoyo al ámbito y presentar una visión consensuada de qué acciones hay que apoyar y cómo. Estas estrategias han de ser fruto de una reflexión colectiva entre expertos y profesionales para asegurar su efectividad. Las estrategias así elaboradas han de darse a conocer también entre los círculos de las agencias

financiadoras e incluso los políticos, para que el apoyo al sector sea consistente y consiga los frutos esperados: disponibilidad de recursos, tecnología y aplicaciones para todas las lenguas, independientemente de su posible mercado o dimensión demográfica.

El proyecto FlaReNet, *Fostering Language Resources Network*, financiado por el programa econtentplus de la Comisión de la Unión Europea (ECP-2007-LANG-617001), tiene como misión activar y promover la participación de todos los actores europeos relacionados con el ámbito de los recursos lingüísticos y las tecnologías lingüísticas en la reflexión sobre la situación actual y la elaboración de consenso en torno a directrices y recomendaciones para poder satisfacer las demandas del sector industrial que emplea las tecnologías lingüísticas. Es misión de la red FlaReNet describir el estado actual del desarrollo de recursos y tecnologías y proponer estrategias, modelos y maneras de producir y garantizar el suministro de recursos lingüísticos de todas las lenguas de la Unión, identificar prioridades a corto, medio y largo plazo así como proponer una serie de objetivos estratégicos que contribuyan a establecer qué condiciones han de cumplirse para que las tecnologías lingüísticas contribuyan efectivamente a diluir las fronteras lingüísticas en una Europa Multilingüe.

La red cuenta con 38 miembros entre Europa, América y Asia y está gestionada por un Comité Directivo formado por el coordinador de la red, el Istituto de Linguistica Computazionale “Antonio Zampolli” del Consiglio Nazionale de la Ricerca, Pisa, Italia, y los directores de cada uno de los Grupos de Trabajo: Evaluation and Language Resources Distribution Agency, ELDA, Paris; Institute for Language and Speech Processing, Athena RC, Atenas, Institute for Multilingual and Multimedia Information, Centre National de la Recherche Scientifique, Paris, Universitat Pompeu Fabra, Barcelona, Universidad de Viena y la Universidad de Utrecht.

2 Grupos de Trabajo

Las actividades de la red se han estructurado en torno a 5 grupos de trabajo: 1) Elaboración de un mapa de los recursos y las tecnologías lingüísticas, 2) Nuevos modelos para la construcción, reutilización, interconexión y distribución de recursos lingüísticos. 3) Armonización de formatos y estandarización de recursos y

tecnologías. 4) Definición de protocolos de evaluación y validación y 5) Métodos para la construcción y el procesamiento automático de LRS

Los grupos de trabajo están abiertos y pueden participar todas aquellas personas interesadas registrándose en la página web de la red.

3 Participación

Por sus objetivos, la participación en las actividades de la red está abierta a todos los investigadores y profesionales del área de los recursos lingüísticos y las tecnologías lingüísticas.

Además de la organización de talleres y reuniones, FlaReNet se vale de las nuevas modalidades de comunicación y colaboración y ha publicado en su página web www.flarenet.eu diferentes foros en una plataforma wiki que permitirá poder recabar información opiniones de todos aquellos que quieran participar. Únicamente hay que registrarse en la página web y suscribirse a cualquiera de los grupos de trabajo para participar o sugerir actividades.

También está prevista la ampliación de la red mediante la incorporación de nuevas organizaciones públicas y privadas, de investigación y empresas que quieran formalizar su interés y participación en los objetivos de la red. El formulario de solicitud se encuentra también en la página web de FLAReNet.

4 Resultados de FLAReNet

Los resultados de los diferentes grupos de trabajo se integrarán en un único documento de recomendaciones que ha de servir a la Comisión Europea y a los Gobiernos de los diferentes Estados miembros como una hoja de ruta para definir políticas de innovación y desarrollo en torno a las tecnologías lingüísticas. Antes, no obstante, de este informe final FLAReNet irá produciendo informes parciales, resultado de las consultas a expertos, que serán distribuidos a las agencias que lo soliciten y hechos públicos, igual que todos los materiales de los grupos de trabajo, en la página web de la red.

Ya ha sido elaborado el primer informe resultado de la reunión *The European Language Resources and Technologies Forum: Shaping the Future of the Multilingual Digital Europe*, celebrado en Viena en 12 y 13 de febrero de 2009, en el que participaron más de 100 expertos, y cuya documentación se puede encontrar también en www.flarenet.eu.