

EL CENTRO ELECTRONICO DE DOCUMENTACION DEL TRIBUNAL SUPREMO DE CASACION Y EL USUARIO EXTRANJERO

Por **ETTORE GIANNANTONIO**

CED. Tribunal Supremo de Casación

Un usuario extranjero que quiera consultar los archivos del centro tropieza con tres dificultades principales: la diversidad de los sistemas de consulta, la diversidad de los idiomas de los documentos y la diversidad de los sistemas jurídicos.

Para superar, por los menos en parte, las dificultades originadas por la diversidad del idioma de los documentos es necesario elaborar una serie de instrumentos lingüísticos que permitan al investigador que no conozca el idioma italiano formular las consultas utilizando el propio idioma y obtener una traducción, al menos parcial, de los documentos que le permita orientarse sobre su contenido.

Para poder alcanzar este objetivo es necesario ante todo memorizar un vocabulario que contenga la traducción de las palabras incluidas en los documentos en el idioma del investigador: en la práctica es necesario memorizar tantos vocabularios como idiomas de la Comunidad distintos del italiano.

De esta manera, por ejemplo, el usuario inglés que quiera buscar las leyes o los principios jurídicos o los artículos doctrinales que contengan la palabra incendio deberá teclear: *F ARSON* o bien *F FIRE*.

La memorización de un vocabulario de traducción de las palabras de los documentos de un idioma de la Comunidad no resulta, por otra parte, suficiente para permitir una consulta eficaz de los documentos por parte de un investigador extranjero con palabras del propio idioma. Se sabe, en efecto, que entre lenguaje y significado no existe una relación unívoca.

Características del lenguaje —de cualquier lenguaje natural— son la variedad, la ambigüedad y la indeterminación.

El lenguaje es variado en el sentido de que un mismo concepto puede expresarse de diversas formas, es ambiguo porque una misma palabra puede tener diversidad de significados, es indeterminado en cuanto que no siempre determina con rigurosa precisión su objeto.

La variedad y la ambigüedad constituyen una importante dificultad para la selección de documentos mediante la indicación de las palabras del texto. La variedad, en efecto, implica que el lenguaje del documento buscado puede ser diferente del que usa el investigador debido a la recurrencia de las variaciones morfológicas, la utilización de sinónimos o de perífrasis, de palabras compuestas o de términos más específicos, y puede determinar, en consecuencia, la no selección de un documento pertinente (el denominado silencio o falta de *recall*, plenitud). La ambigüedad del lenguaje, por otra parte, puede determinar la selección de un documento inadecuado (el denominado rumor, *noise*).

Estos dos inconvenientes, propios de toda búsqueda de documentos mediante palabras del lenguaje natural, resultan mucho mayores cuando se trata de la traducción de la pregunta de un idioma a otro. Y en verdad, los inconvenientes originados por la existencia de sinónimos, perífrasis, palabras compuestas o términos más específicos en el lenguaje original se suman a los inconvenientes originados por los sinónimos, perífrasis, palabras compuestas o términos más específicos del idioma de los documentos; y así las dificultades originadas por las polisemias del idioma original se suman a las producidas por las polisemias del idioma de los documentos.

Por ejemplo, en inglés, el significado del término *arson* es similar al de los términos *fire* o *combustion*; así, en italiano, en lugar de incendio (*incendio*) podemos encontrar fuego (*fuoco*), combustión (*combustione*), autocombustión (*autocombustione*).

Para evitar los inconvenientes relacionados con la búsqueda mediante palabras coloquiales en el sistema Italgire-Find se ha realizado un tesoro con características completamente originales. El mismo se basa en un análisis previo del lenguaje para proporcionar al ordenador algunas informaciones fundamentales de carácter gramatical, morfológico y semántico sobre cada palabra. Este «tratamiento» del lenguaje consiste, en concreto, en las siguientes operaciones:

a) La codificación de los lemas. A cada lema se le atribuye una clave, es decir, un símbolo constituido por una letra del alfabeto, que indica al ordenador qué regla de la gramática italiana tiene que aplicar para obtener automáticamente todas las variaciones morfológicas susceptibles de aplicarse a dicho lema.

b) La descomposición de las palabras en sememas fundamentales del lenguaje o, como también hacen los lingüistas, categorías semánticas o factores primos semánticos.

c) La conexión de los términos más específicos y no susceptibles de descomposición en elementos fundamentales con aquellas más generales.

d) La conexión de todas las palabras de la misma raíz con un solo lema denominado cabeza de grupo; con este último se ponen en relación también las palabras que, incluso teniendo raíz distinta, son sinónimos del mismo. De esta manera no es necesaria ni la clasificación manual de cada documento ni la consulta del tesoro por parte del investigador antes de cada operación; es suficiente con que el investigador indique al ordenador si desea la selección de los documentos con la palabra indicada textualmente o bien en toda su extensión conceptual.

La descomposición de las palabras en sememas fundamentales del lenguaje resulta el medio más rápido y conveniente también para poder buscar un documento redactado en un idioma extranjero mediante «palabras-cebo» expresadas en la propia lengua.

En efecto, mientras que la traducción de un idioma a otro cuyas palabras tienen significados particulares, es compleja y larga (por la gran cantidad de vocablos de los principales idiomas) y difícil (al no encontrarse siempre en el otro idioma el término correspondiente capaz de expresar todos los matices, y particularidades del significado), la traducción de los elementos del lenguaje resultará rápida (porque los elementos fundamentales son solamente unos tres mil) y fácil (por estar estos elementos representados por las palabras de significado más elemental e inequívoco).

La traducción del tesoro a los distintos idiomas de la Comunidad constituirá un importante progreso para la búsqueda de los documentos contenidos en el archivo del CED por medio de las palabras de un idioma extranjero. El investigador, en efecto, podrá pedir mediante un canal específico (canal PK o palabras conceptuales) no sólo los documentos que contengan la palabra utilizada en el vocabulario para traducir el término usado por él, sino también todos los sinónimos, perífrasis o palabras compuestas.

Por ejemplo, el uso del término *fire* en toda su extensión conceptual por parte del investigador permitirá no sólo la obtención de los documentos que tengan la palabra fuego, sino también la de aquellos que contengan la palabra incendio, combustión, ustión o piromanías; la formulación de una banda como esta:

F PK = SALE AND BOVINE AND SICKNESS

permitirá la selección no sólo de los documentos que contengan las palabras venta, bovino o enfermedad, sino también de aquellos que hablan de compraventa de vacas y toros, o de enfermedades específicas como la tuberculosis y el afta epizoótica; el uso del término:

F PK = SWITZERLAND

permitirá la selección no sólo de los documentos que contengan la palabra Suiza; sino también la de aquellos que contengan el adjetivo suizo o la expresión Confederación Helvética o que hablen del Cantón del Tesino o de Zurich.

La posibilidad para el usuario extranjero de utilizar la propia lengua en la búsqueda de los documentos mediante palabras coloquiales no excluye, por otra parte, la otra dificultad consistente en el hecho de que los documentos, una vez seleccionados, se facilitan al usuario en lengua italiana, con la excepción de aquellos archivos (Rebis, BID, Enlex) en los que los documentos están memorizados también en idiomas extranjeros.

Para obviar este inconveniente es necesario —al menos por lo que se refiere a los archivos de jurisprudencia— realizar la traducción del esquema de clasificación utilizado específicamente para este tipo de documentos. En efecto, las voces del esquema de clasificación aparecen encabezando todos los documentos y constituyen, junto con una parte variable propia de cada documento, el título del mismo. La traducción del esquema de clasificación permitirá al usuario extranjero no sólo la utilización de las voces del esquema como canales de selección (canales

de selección SC SC3), sino también tener el título de la máxima traducido al propio idioma y poderse orientar, por tanto, con la suficiente precisión en la documentación seleccionada por el ordenador.

Por ejemplo, si el investigador está interesado en saber si un ciudadano británico divorciado puede casarse libremente en Italia y si es necesario que la sentencia de divorcio pronunciada en Gran Bretaña sea aceptada por el juez italiano, basta con teclear según el *common command language*:

F PK: Gran Bretaña o bien Reino Unido
o incluso, para un jurista inglés
F PK: Great Britain or United Kingdom

La respuesta será:

R 1 72 documentos 0, 2 sec. 140291 * 72
el investigador podrá entonces teclear:

F Marriage or Divorce

la respuesta será:

R 2 2 documentos 0, 2 sec. 72 * 1944

si el investigador quiere leer el documento deberá teclear SHOW.

He aquí como aparecerá el documento a un usuario inglés:

División I Sentence 823 of 14/3/68 * ANNOTATION * Chairman Pece Rapporteur Spagnoletti.

Plaintiff Rossi di Monti Defendant Forbes De Grana Family-Marriage-Contract in Italy by foreign parties - Free state of British or Irish citizen - Divorce Sentence passed abroad acknowledgment - need - exclusion - conditions.

Para comprobar la soltería de un extranjero que pretende contraer matrimonio en Italia, debe tenerse presente que tanto en Inglaterra como en Irlanda una sentencia de divorcio pronunciada en el extranjero tiene validez sin necesidad de ningún reconocimiento, siempre que haya sido pronunciada por la autoridad judicial competente en la jurisdicción territorial en la que las partes tenían el domicilio conyugal.

El usuario obtiene, de esta forma, una primera indicación sobre el contenido del documento.

Otra dificultad con la que tropieza el usuario extranjero al interrogar los bancos de datos jurídicos del centro está constituida por las diferencias entre el propio sistema jurídico y el sistema jurídico italiano. La diversidad, que es particularmente relevante en el caso de los sistemas anglosajones, puede implicar que en el documento selecciona-

do se haga referencia a una institución jurídica que no tiene analogía en el sistema del usuario y le es, por tanto, desconocida o bien a una institución que, incluso teniendo una correspondencia formal, tiene una estructura o una finalidad esencialmente distinta.

En estos casos no puede, en absoluto, traducirse el término extranjero con un término no italiano correspondiente o viceversa; sería, por ejemplo, un grave error traducir al italiano el término *Consideration* o al inglés la expresión «interés legítimo». Es necesario en estos casos dejar el término en su idioma original y disponer un archivo específico en el cual el usuario que lo desee pueda consultar el significado específico del término encontrado.

Se constituye así un vocabulario jurídico electrónico: vocabulario en el que para cada término se dispone de una descripción de la institución correspondiente. El vocabulario, por otra parte, se orientará en sentido comparativo: o sea, la explicación se efectuará en el lenguaje del usuario extranjero y de manera que quedan explicadas las características y diferencias respecto a las instituciones correspondientes en el sistema del usuario extranjero. Esto llevará aparejadas tantas definiciones distintas —y por ello tantos vocabularios diferentes— como sistemas jurídicos extranjeros se considerán.