

LA LUCHA CONTRA EL TERRORISMO Y LA DELINCUENCIA ORGANIZADA: UNA VISIÓN DESDE LA LINGÜÍSTICA Y LA INGENIERÍA DEL CONOCIMIENTO

PEDRO UREÑA GÓMEZ-MORENO

Universidad de Granada
pedrou@ugr.es

107

1. Introducción

El procesamiento del lenguaje natural (PLN) es un campo de estudio multidisciplinar que ha atraído el interés de organismos públicos y de entidades privadas debido al gran número de aplicaciones que tiene tanto en la vida cotidiana como en el ámbito profesional. El PLN, que se encuentra en la base de las conocidas como “tecnologías del lenguaje”, está dirigido a la creación de modelos computacionales que las máquinas puedan utilizar para producir o interpretar el lenguaje humano, no solo a nivel morfosintáctico, sino fundamentalmente a nivel semántico y discursivo (Cole et al. 1997). Son precisamente estos dos últimos niveles los que resultan fundamentales para una comprensión profunda y no meramente formal del lenguaje. De ahí que una de las prioridades del PLN sea crear herramientas de análisis semántico con las que, entre otros, recuperar y procesar la cuantiosa información que diariamente se produce en páginas web, bases de datos y redes sociales. Si bien los contenidos digitales aparecen con frecuencia “estructurados” y su procesamiento resulta a priori sencillo, la mayor parte son de tipo “no estructurado”, en cuyo caso el reconocimiento semántico requiere un procesamiento más complejo. Un ámbito de estudio en el que es especialmente relevante el potencial del PLN es la detección de escenarios de riesgo para la seguridad, entre ellos, los relacionados con el cibercrimen, el terrorismo, o el crimen organizado (Vander Beken 2004; Argamon y Howard 2009; Subrahmanian 2009; Choi et al. 2014).

El presente artículo trata de establecer parámetros para el procesamiento semántico de unidades conceptuales del ámbito de la delincuencia y abrir así nuevas vías que permitan analizar tanto los datos estructurados como los no estructurados.

Son muchos los estudios que, de un modo u otro, combinan las tecnologías del lenguaje con el análisis de datos para combatir la delincuencia y el terrorismo. Tseng et al. (2012), por ejemplo, proponen una herramienta que está basada en redes terminológicas y permite detectar y visualizar conexiones entre entidades del ámbito delictivo. De forma similar, Choi et al. (2014) establecen un marco metodológico para el análisis textual y la recuperación de documentos relacionados con el terrorismo en el que combinan medidas estadísticas con el análisis de redes léxicas de WordNet. Otras propuestas subrayan la importancia de las ontologías especializadas para la recuperación de información. En este sentido, son de especial relevancia para este artículo los estudios sobre el desarrollo ontológico en los campos del derecho (Breuker et al. 2005; Casellas 2011), del terrorismo (Mannes y Golbeck 2005; Golbeck et al. 2006; Chmielewski et al. 2009; Galjano y Popovich 2009; Jiang y Tan 2010) y de la delincuencia (Dzemydiene y Kazemikaitiene 2005; Bak y Jedrzejek 2010; Bak et al. 2013; González-Conejero et al. 2014; Spranger y Labudde 2014). Finalmente, cabe destacar el papel que juegan aquellos consorcios formados por entidades públicas o empresas privadas que ponen a disposición de los investigadores, en particular, y de los usuarios de internet, en general, las bases de datos sobre hechos delictivos y terroristas acaecidos en el ámbito internacional. Por ejemplo, *Global Terrorism Database* (GTB) (LaFree 2011; LaFree et al. 2015) (<http://www.start.umd.edu/gtd>) es un proyecto que reúne información sobre unos 140.000 atentados terroristas ocurridos desde la década de 1970 hasta hoy en distintos países. Cada suceso registrado en la GTB incluye información sobre quiénes cometieron el atentado, dónde lo hicieron, qué armas emplearon, cuál fue el número de víctimas y rehenes o cómo se llevó a cabo, entre otras variables. Los proyectos MOSAIC (*Multi-Modal Situation Assessment and Analytics Platform*) (<http://www.mosaic-fp7.eu>) (Seidler, Adderley, Badii & Raffaelli, 2014) y CAPER (*Collaborative Information Acquisition Processing Exploitation and Reporting for the Prevention of Organised Crime*) (<http://www.fp7-caper.eu>) (Aliprandi et al. 2014) ofrecen soluciones para descubrir organizaciones criminales a través de los contenidos visuales, auditivos o de texto incluidos en los repositorios documentales y en las redes sociales. Finalmente, el proyecto e-POOLICE (*Early Pursuit against Organized Crime Using Environmental Scanning, The Law and Intelligence Systems*) (<https://www.epoolice.eu>) (Brewster et al. 2014) tiene como misión prevenir la delincuencia organizada mediante el análisis de información obtenida de diversas fuentes, entre ellas repositorios de organismos internacionales, documentación gubernamental, páginas web o foros académicos.

Este artículo se enmarca dentro de la corriente ontológica arriba mencionada y, más concretamente, en el contexto de colaboración entre el PLN y la ingeniería del conocimiento, rama ésta última que se ocupa de la creación de modelos de representación semántica tanto del conocimiento común como del especializado (Gruber 1995; Guarino et al. 2009). El artículo muestra cómo se elaboran los modelos conceptuales de varias unidades léxicas pertenecientes a los ámbitos del terrorismo y del crimen organizado, y aborda su integración en FunGramKB (*Functional Grammar Knowledge Base*), una base de conocimiento para el desarrollo de aplicaciones de comprensión y generación lingüística que ha sido diseñada para establecer relaciones lógicas de herencia e inferencia entre unidades conceptuales a partir de sus rasgos definitorios (<http://www.fungramkb.com/>). Para ello, se analiza a modo de caso práctico la conceptualización de las unidades *cartel* (“cártel”), *oriented cluster* (“grupo con fines propios”) y *terrorist cell* (“célula terrorista”). Tanto la información conceptual como la léxica constituyen dos elementos fundamentales de FunGramKB y se incorporan a esta mediante la plataforma online *FunGramKB Suite*. Las unidades seleccionadas representan una pequeña muestra de un repertorio mayor que ha sido elaborado por el proyecto de creación ontológica denominado GLOBALCRIMETERM (Felices y Ureña, 2012; Arcas-Túnez y Periñán-Pascual 2014). El proyecto se ha completado con la elaboración de una taxonomía jerarquizada de conceptos relacionados con los ámbitos de la delincuencia organizada y del terrorismo, e incluye información sobre tipos de armas, tipos de delitos, o tipos de sustancias explosivas y estupefacientes, entre otros. La propuesta que se defiende en este artículo se basa en la necesidad de abordar con nuevos enfoques el problema de la delincuencia del siglo XXI, uno de los cuales puede ser el procesamiento lingüístico. La importancia de FunGramKB radica en que, a diferencia de otros modelos, su motor de razonamiento está diseñado a partir de la comprensión real de las interacciones conceptuales que se producen entre las acciones y las entidades delictivas, y no a partir de parámetros estadísticos o probabilísticos. Si bien estos últimos optimizan el tiempo que se necesita para detectar posibles amenazas, FunGramKB ofrece una comprensión profunda (esto es, semántica) del lenguaje, lo que contribuiría de forma decisiva a mejorar la precisión con la que se puede recuperar información relacionada con las redes de delincuencia.

El resto del artículo se articula como sigue: la Sección 1 presenta una visión general de la estructura de FunGramKB, centrándose para ello en la relación entre dicha estructura y las áreas terminológicas. La Sección 2 aborda el análisis conceptual de las unidades terminológicas relacionadas con los dominios del terrorismo y de la delincuencia organizada. La Sección 3 incluye las conclusiones. Es necesario apuntar que tanto los términos analizados como su definición provienen del inglés, dado que la mayoría de la documentación pertenece a órganos internacionales que utilizan prioritariamente esta lengua como medio de difusión. Asimismo, se proporcionan las correspondientes traducciones al español.

2. FunGramKB

FunGramKB es una base de conocimiento para el procesamiento del lenguaje natural, es decir, es una plataforma diseñada para albergar contenido lingüístico-conceptual, y está dirigida a la realización de tareas relacionadas con el lenguaje y el razonamiento automático (Periñán-Pascual y Arcas-Túnez 2007; Periñán-Pascual y Mairal-Usón 2010). FunGramKB nace como una propuesta teórica y metodológica para el desarrollo o la mejora de aplicaciones tales como la desambiguación semántica, la traducción automática, la recuperación de información o las interfaces de comunicación persona-máquina. Para tener una visión más completa del funcionamiento de la base de conocimiento, es necesario establecer una relación entre esta y el Modelo Léxico Construccional (MLC), esto es, una teoría sobre la formación e interpretación del significado que propone un sistema multinivel para la descripción de procesos semánticos y pragmáticos (Periñán-Pascual 2013; Ruiz de Mendoza 2013, 2014). Entre las principales virtudes del MLC destacan la formalización rigurosa de plantillas léxicas y construccionales, que se combinan mediante mecanismos de inferencia y subsunción, así como una explicación global de procesos metafóricos y metonímicos. Si bien el MLC y FunGramKB tienen un origen y un desarrollo independientes, la colaboración entre ambos ha sido estrecha. Por un lado, el MLC juega un papel fundamental en la implementación del *Gramaticón* de la base de conocimiento. Por otro lado, FunGramKB puede servir como solución computacional para la aplicación del MLC a diversas tareas relacionadas con el PLN.

110

FunGramKB adopta un enfoque en tres niveles (Figura 1) (Periñán-Pascual y Arcas-Túnez 2010). El primer nivel es el conceptual: la base de conocimiento alberga la “Ontología Nuclear”, esto es, un repertorio jerarquizado de unidades representativas del sentido común, ya sean entidades, como por ejemplo “humano”, acciones como “saltar,” o características como “fácil” (Periñán-Pascual y Arcas-Túnez 2007). La Ontología está conectada, por un lado, al “Cognición”, que se define como un repositorio de guiones relacionados con procedimientos como por ejemplo “ver una película”, “comer en un restaurante” o “pagar con tarjeta de crédito”, y, por otro, al “Onomasticón”, módulo encargado de almacenar entidades únicas, tales como *Taj Mahal* o *Napoleón* (Periñán-Pascual y Carrión-Varela 2011).

Las unidades contenidas en los tres módulos (Ontología, Cognición y Onomasticón) aparecen expresadas mediante COREL (*Conceptual Representation Language*), un lenguaje de interfaz basado en la lengua inglesa, que se caracteriza por su sencillez sintáctica y por su expresividad semántica¹. El nivel conceptual formado por la Ontología, el Cognición y el Onomasticón constituye el motor de la base de conocimiento y es el responsable de que ésta pueda *comprender* conceptos propios de la cognición

humana. El segundo nivel es el lingüístico: el componente conceptual se halla conectado al módulo denominado “Lexicón”, que contiene las plantillas léxicas de cada uno de los lexemas que materializan el contenido conceptual en distintas lenguas, entre ellas el español, el inglés, el italiano o el alemán (Mairal-Usón y Perrián-Pascual 2009). El tercer nivel es el gramatical: FunGramKB incluye un “Gramaticón” o módulo que contiene los esquemas construccionales —como la construcción ditransitiva o la estructura de “movimiento causado”— que se utilizan en las distintas lenguas representadas en el Lexicón (Ruiz de Mendoza y Mairal-Usón 2008; Luzondo-Oyón 2014). Para un análisis más detallado de la estructura y funcionamiento de FunGramKB, véase Perrián-Pascual y Mairal-Usón (2009, 2010).

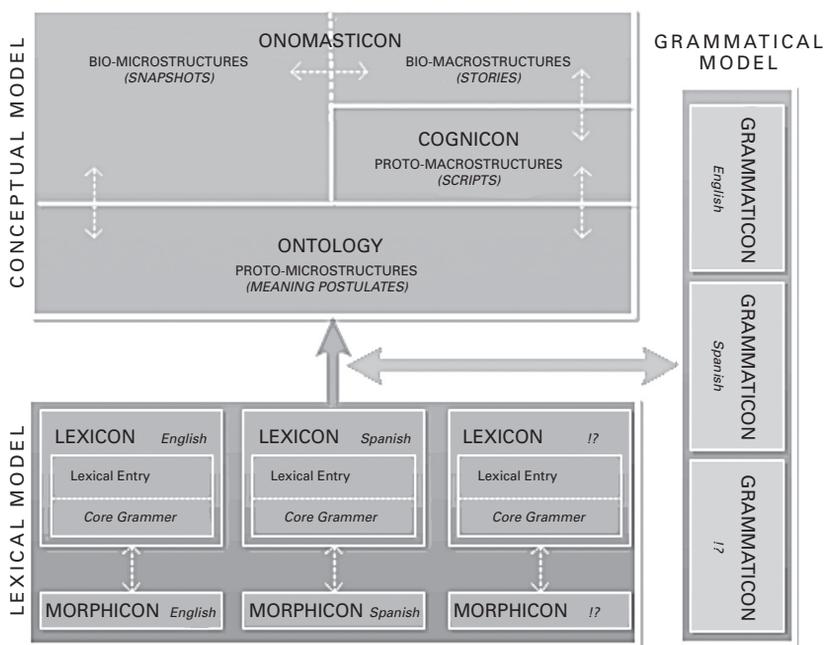


Figura 1. Estructura de FunGramKB.

Si, como se ha mencionado arriba, la Ontología Nuclear incluye entidades, acciones y características del conocimiento que las personas poseemos comúnmente, las “Ontologías Satélite” están formadas por unidades conceptuales relacionadas con ámbitos científico-técnicos como la medicina o el derecho. A pesar de coexistir

como repositorios separados, ambos tipos de ontologías se encuentran conectados mediante una misma estructura jerárquica y un mismo lenguaje de formalización, esto es, COREL. De este modo, ambos pueden intervenir simultánea o alternativamente en una misma secuencia de procesamiento lingüístico o razonamiento artificial. El módulo satélite se encuentra actualmente en desarrollo, tras haberse efectuado un estudio piloto que se ha completado con la creación de la subontología sobre terrorismo y delincuencia organizada, GLOBALCRIMETERM, en la que se incluyen equivalencias léxicas en inglés, español e italiano para cada unidad. La construcción de las Ontologías Satélite como GLOBALCRIMETERM está basada en una adaptación de la metodología de creación ontológica propuesta por Perrián-Pascual y Mairal-Usón (2011), y desarrollada posteriormente de forma específica por Ureña et al. (2011), Felices-Lago y Ureña (2012), y Arcas-Túnez y Perrián-Pascual (2014) para el modelado *ad hoc* de unidades especializadas. La metodología *ad hoc* consta de cuatro fases: (a) elaboración de un corpus representativo del ámbito objeto de estudio; (b) extracción automática de términos a partir del corpus; (c) descripción en lenguaje natural de los términos recuperados durante la extracción y (d) definición en lenguaje COREL de las unidades especializadas. Las fases (a) y (b) aparecen detalladas en varios estudios, principalmente Felices-Lago y Ureña (2012), y Felices-Lago y Ureña (2014), mientras que las fases (c) y (d) aparecen ilustradas en la Sección 2 del presente artículo.

112

3. Conceptualización de las unidades del terrorismo y la delincuencia organizada

A los efectos de la investigación que nos ocupa y siguiendo la definición de Enders y Sandler (2002), el terrorismo se definirá de forma genérica en este artículo como el uso de la violencia con fines políticos, religiosos o ideológicos que se ejerce contra grupos numerosos de personas no relacionadas con la toma de decisiones en dichos ámbitos (Enders y Sandler 2002). Esta acción de extrema gravedad la perpetran colectivos organizados, individuos aislados, conocidos como “lobos solitarios” (*lone wolves*), o la denominada “resistencia acéfala” (*leaderless resistance*) (Carson et al. 2012: 129; Spink et al. 2013: 4). La delincuencia organizada, por otra parte, se refiere a la actividad ilícita que lleva a cabo cualquier red mafiosa cuyos objetivos son el comercio ilegal y el control de un territorio mediante el uso de la violencia (Ashby y Ramos, 2013). Tanto la comisión de atentados terroristas como el tráfico ilegal de mercancías y personas constituyen dos de las amenazas más serias a la seguridad ciudadana a nivel global, de ahí que el análisis de la estructura y modus operandi de la delincuencia organizada sea fundamental para comprender ambos dominios. Poder contar con un modelo tanto de los agentes como

de los medios materiales que estos utilizan supone un paso importante en la creación de herramientas para la detección de la actividad delictiva, ya proceda de foros, páginas webs o documentación incautada. Por ello, esta sección se centra en el análisis de las entidades colectivas *cartel* (“cártel”), *oriented cluster* (“grupo con fines propios”) y *cell* (“célula”); las dos primeras están relacionadas con la delincuencia organizada, mientras que la tercera pertenece al ámbito del terrorismo. Por “entidades colectivas” se entenderá aquí una asociación de personas o entidades que colaboran en la planificación y comisión de actos terroristas u operaciones complejas de tráfico de personas, armamento o droga. La noción de colectividad —en su vertiente no delictiva— ya se encuentra en la Ontología Nuclear de FunGramKB bajo la categoría +COLLECTION_00, que incluye conceptos básicos y terminales¹ como +CATTLE_00, +FLOCK_00 o +FLORA_00. Entre los conceptos que se agrupan bajo el agente criminal colectivo se encuentran entidades como +CARTEL_00, +MAFIA_00, +ORIENTED_CLUSTER_00, +GANG_00, +GUERRILLA_00, +CELL_00 y +DEATH_SQUAD.

3.1. Elaboración de un modelo conceptual de cartel

Cartel (en español, “cartel” o “cártel”) es un término que, a los efectos del presente artículo, se aplica a una organización ilícita dedicada a la producción y comercialización de sustancias estupefacientes ilegales, y cuya área de acción se centra principalmente en México y Colombia (McCarthy 2011). La actividad de los cárteles, encabezada por los capos de la droga, se ha venido desarrollando y diversificando durante décadas, de ahí que resulte complejo determinar con precisión lo que constituye un cártel (McCarthy 2011). Esta complejidad conceptual es relevante para el contenido de la ontología. Para ilustrar el análisis conceptual, esta sección se basa en la definición propuesta por el NDIC (*National Drug Intelligence Center*), que incluye información sobre los objetivos y la estructura de este tipo de organización al tiempo que solventa la diversificación que se acaba de mencionar²:

Drug cartels are large, highly sophisticated organizations composed of multiple DTOs [Drug Trafficking Organisations] and cells with specific assignments such as drug transportation, security/enforcement, or money laundering. Drug cartel command-and-control structures are based outside the United States; however, they produce, transport, and distribute illicit drugs domestically with the assistance of DTOs that are either a part of or in an alliance with the cartel. (NDTA 2010: 8)

Para la versión en COREL de una unidad léxica, es necesario, en primer lugar, identificar las proposiciones de significado más relevantes de la descripción en lenguaje natural para, en segundo lugar, proceder a su conceptualización. En este caso, la unidad “cártel” está asociada al concepto +CARTEL_00. Abajo se

muestran las proposiciones en lenguaje natural y a continuación su correspondiente enunciado COREL:

1. *A cartel is a large (criminal) organisation.*
 +(e1: +BE_00 (x1: +CARTEL_00)Theme (x2: +ORGANIZATION_00)Referent)
 * (e2: +BE_01 (x1)Theme (x3: +BIG_00)Attribute)
 * (e3: +DO_00 (x1)Theme (x4: +CRIME_00)Referent)
2. *A cartel is composed of multiple groups.*
 * (e4: +COMPRISE_00 (x1)Theme (x5: i +GROUP_00)Referent)
3. *A cartel produces, transports and distribute illicit drugs.*
 * (e5: +CREATE_00 (x1)Theme (x6: +DRUG_00)Referent) (f1: +SALE_00 Purpose)
 * (e6: +TAKE_01 (x1)Agent (x6)Theme (x7)Location (x8)Origin (x9)Goal)
 * (e7: +SELL_00 (x1)Agent (x6)Theme (x1)Origin (x10)Goal)

Tal y como muestra esta definición, el desarrollo semántico-conceptual se compone de varias proposiciones identificadas mediante la letra *e*. Cada proposición está encabezada por un evento principal, por ejemplo +BE_00 (“ser”) o +COMPRISE_00 (“contener”), que a su vez está conectado a otros conceptos que denotan entidades, como por ejemplo +ORGANIZATION_00 (“organización”), o cualidades como +BIG_00 (“grande”). Tanto las entidades como las cualidades constituyen los participantes necesarios u opcionales del núcleo eventual. La primera proposición, que es irrefutable (indicada por el signo “+”), tiene como función explicitar taxonómicamente el concepto que el hiperordinario inmediato define. El resto de proposiciones son rebatibles, es decir, pueden considerarse verdaderas o falsas dependiendo del contexto, o dicho de otro modo, podrían no cumplirse en caso de aparecer nuevas proposiciones que nieguen o contradigan el significado de las precedentes.

No existe una única conceptualización para una unidad ontológica concreta, sino que la apariencia final de una definición dependerá del grado de precisión semántica que el ingeniero de conocimiento requiera para lograr un objetivo de procesamiento específico. Por tanto, los enunciados en COREL se ajustarán siempre al significado original de una unidad pero, al mismo tiempo, podrán matizarla o ser más generales que ésta. Por ejemplo, mientras que la descripción en lenguaje natural arriba indicada establece que las organizaciones que trafican con drogas (*Drug Trafficking Organisations*) y que forman parte del cártel pueden realizar labores especializadas dentro de la organización, la definición COREL establece que la producción, distribución y venta de drogas son propias del cártel de forma general y no de un subgrupo dentro de la organización. Otro ejemplo, aunque a la inversa,

es que el NDIC no explicita, por tratarse de información obvia dentro de este contexto, que un cártel es una organización ilegal, mientras que en el caso de COREL este dato sí que aparece explícitamente en la segunda proposición (e2).

Las proposiciones de +CARTEL_00 que se acaban de mostrar, así como las incluidas en los dos apartados siguientes, permiten a la base de conocimiento saber, de manera genérica, lo que constituye una entidad ilegal colectiva. Una segunda fase de implementación consistiría en añadir un rango mayor de especificidad conceptual que le permita a una herramienta de prevención de riesgos detectar movimientos para cometer delitos o formar grupos violentos en lugares de potencial actividad delictiva. El procesamiento lingüístico de los focos de riesgo debe integrar los enunciados lógicos que aquí se proponen para la Ontología Nuclear, con la información contenida en el Cognición y el Onomasticón, ya que el primero está concebido para incluir tipificaciones de entidades delictivas concretas, mientras que el segundo permite recopilar los datos relacionados con los lugares, las fechas o los integrantes de las bandas armadas. La conexión entre los tres módulos, en lo que se refiere a materia terrorista y delictiva, es un aspecto complejo que está fuera del alcance de este artículo, centrado en la Ontología.

3.2. Elaboración de un modelo conceptual de oriented cluster

El término *cluster* (“clúster”), aplicado al ámbito de la delincuencia organizada, tiene su origen en un documento de las Naciones Unidas sobre la delincuencia organizada (UNODC, 2002). En este aparecen por primera vez las expresiones *clustered hierarchy* y *cluster of criminal groups* (“jerarquía agrupada” y “asociación de grupos delictivos”, respectivamente) para denotar un conglomerado de colectivos que colaboran puntualmente para lograr un objetivo delictivo común. El término, a su vez, forma parte de una tipología de grupos organizados junto a *standard hierarchy* (“jerarquía estándar”), *regional hierarchy* (“jerarquía local”) y *core group* (“grupo nuclear”), tipología que todavía sigue proporcionando una base definitoria a un gran número de estudios sobre redes de narcotráfico y criminología (Klerks 2007; Tupman 2009; Dean et al. 2010). Posteriormente, y basándose en la definición de la UNODC (Oficina de las Naciones Unidas contra la Droga y el Delito), la EUROPOL, organismo policial europeo especializado en antiterrorismo y delincuencia organizada (<https://www.europol.europa.eu>), establece los rasgos principales de este tipo de agente colectivo al que incorpora el adjetivo *oriented*:

Single criminal groups with common objectives no longer operate in isolation and this creates a powerful convergence of criminal intentions and resources. Strategic direction for their activities can be determined by policies decided by the leaders of the most dominant criminal group or by regular meetings of the most influential

representatives of the individual groups. The presence of such “oriented clusters” that are led or, at least, coordinated by a common centre of influence is assessed as a major threat. These clusters may combine the strengths of both hierarchies and networks so that very high levels of effectiveness, diversification and specialisation can be achieved. (OCTA 2007: 9)

La elaboración de un modelo conceptual de “grupo con fines propios” que se presenta a continuación está basada en la propuesta de la EUROPOL, al tratarse esta de una definición que sintetiza las principales características del término. Al igual que en el caso de ‘cártel’, abajo se muestran los enunciados en lenguaje natural de +ORIENTED_CLUSTER_00 junto con su conceptualización COREL:

1. *An oriented cluster is a criminal organisation.*

+(e1: +BE_00 (x1: +ORIENTED_CLUSTER_00)Theme (x2: +ORGANIZATION_00)Referent)

*(e2: +DO_00 (x1)Theme (x3: +CRIME_00)Referent)

2. *An oriented cluster has a common coordinator.*

*(e3: +HAVE_00 (x1)Theme (x4: +LEADER_00)Referent)

3. *An oriented cluster is made up of independent criminal groups. Each member organisation has its own leader.*

*((e4: +COMPRISE_00 (x1)Theme (x5: i +GROUP_00)Referent) (e5: +BE_01 (x5)Theme (x6: +INDEPENDENT_00)Attribute) (e6: +HAVE_00 (x5)Theme (x7: 1 +LEADER_00)Referent))

La principal diferencia entre un cártel y un grupo con fines propios consiste en que, si bien ambas entidades se componen de varias organizaciones subordinadas, la relación organizativa que hay entre ellas es muy distinta: en la primera existe una estructura jerárquica con un líder, mientras que en la segunda, al tratarse de grupos originalmente independientes, coexisten varios líderes y un coordinador, este último es el encargado de gestionar el clúster en su conjunto. Esta diferencia constituye un aspecto distintivo del postulado de significado de +ORIENTED_CLUSTER_00 y, para especificarlo mediante COREL, se han utilizado dos predicaciones; en concreto, la segunda proposición diferencia al coordinador de los líderes independientes especificados en la tercera proposición.

2.3. Elaboración de un modelo conceptual de cell

Antes de abordar la conceptualización de *cell* (“célula”) es necesario hacer varias aclaraciones respecto al uso y significado de este término. La palabra “célula”, referida a la delincuencia, es una transferencia metafórica desde el dominio de la

biología al dominio judicial. Mientras que en el primer dominio el término se refiere a una forma básica de vida, en el segundo el elemento denotado es una agrupación de personas que persiguen objetivos relacionados principalmente con el terrorismo. El proceso metafórico se mantiene a un nivel sintagmático, en el que ambos dominios comparten colocaciones con *cell*, como por ejemplo *formation of the cell* (“formación de la célula”) (UNODC 2010: 31) o *nucleus of the cell* (“núcleo de la célula”) (UNODC 2010: 32). Por otro lado, el término también participa en algunas formaciones que son específicas del ámbito delincencial, como *sleeping cell* (“célula durmiente”) (UNODC 2012: 114) o *hybrid cell* (“célula híbrida”) (Hreckovski y Dobson 2012: 338). Finalmente, hay que subrayar que, si bien “célula” se aplica preferentemente a grupos terroristas, no es infrecuente que se use para referirse a grupos ilegales no terroristas, como se observa por ejemplo en *cell-like criminal groups* (“grupos delincuentes de tipo celular”), *cell-like structure* (“estructuras [delictivas] de tipo celular”) (UNODC 2010: 117). En este artículo, una célula terrorista se considerará un hipónimo de la colocación más general *terrorist group* (“grupo terrorista”) (UNODC 2010: 15) o *terrorist organisation* (“organización terrorista”) (UNODC 2010: 14) y, por tanto, las definiciones que se analizan en esta sección están referidas en exclusiva a este sentido original de célula como grupo terrorista.

En cuanto su definición, Hreckovski y Dobson (2012: 338) proponen la siguiente descripción de una célula prototípica:

Terrorist acts can be performed by individuals, organized groups, or a state. The main type of terrorist organization is a secretive cell that consists of members who are highly motivated to perform dangerous and deadly operations. Usually, a terrorist cell consists of two to five members with their own logistics. Each terrorist cell does not have information about the others, which is the main reason why it is hard to penetrate the whole organization. The commander of the cell is the only person who communicates and coordinates with higher levels and other cells. Another fact is that, by using technical tools like the Internet, they do not have to know each other or train together. Terrorist organizations may form only one cell to operate on a tactical level, or they may form many cells that operate locally or internationally. (Hreckovski y Dobson 2012: 338)

A continuación se muestra la traducción en COREL de los rasgos más importantes de la descripción anterior:

1. *A cell is an independent group inside a terrorist organisation.*

+ (e1: +BE_00 (x1: +CELL_00) Theme (x2: +ORGANIZATION_00) Referent)

* (e2: +KILL_00 (x1) Theme (x3: m +PEOPLE_00) Referent)

2. *A cell is clandestine.*

*(e3: +HIDE_00 (x1)Theme (x1)Referent (f1: (e4: +BE_01 (x1)Theme (x4: +SECRET_00)Attribute))Purpose)

3. *A cell consists of two to five members.*

*(e4: +COMPRISE_00 (x1)Theme (x5: 1^2^3^4^5 +MEMBER_01)Referent)

4. *A cell resists penetration.*

*(e5: +ENTER_00 (x6: +HUMAN_00)Agent (x2)Theme (x7)Location (x8) Origin (x1)Goal (x9: +DIFFICULT_00)Manner)

Hreckovski y Dobson aportan rasgos sobre cómo percibe la sociedad una célula terrorista, es decir, un grupo armado secreto (*secretive*), motivado (*motivated*) y que a menudo utiliza el ciberespacio para comunicarse (*using technical tools like the Internet*). Sin embargo, estos rasgos resultan secundarios desde el punto de vista de su diferenciación como entidad, ya que pueden ser compartidos por muchos otros tipos de organizaciones ilícitas; de ahí que no se incluyan en la descripción conceptual de +CELL_00. Los mismos autores, por otro lado, sugieren otras características que sí permiten distinguir las células como entidades independientes. En concreto, la característica principal de la célula es la comisión de atentados en sí misma, es decir, mientras que la violencia para los cárteles y las organizaciones de corte mafioso es un instrumento con el que ejecutar o extorsionar a sus enemigos o a los miembros acusados de revelar secretos, para las células terroristas es un fin táctico en sí mismo. Tanto los cárteles y las organizaciones de corte mafioso como las células recurren al blanqueo de capitales y al tráfico de drogas para financiarse, pero en tanto que los cárteles tienen como objetivo el enriquecimiento, las células lo utilizan para perseguir objetivos políticos y/o religiosos. Finalmente, existe una proposición en +CELL_00 sobre el número concreto de miembros que conforman comúnmente la célula y que, si bien refleja el número exacto de la definición en lenguaje natural, es también una proposición rebatible y, por tanto, flexible.

118

4. Conclusiones

El artículo ha presentado un caso práctico de colaboración entre el análisis lingüístico y la ingeniería del conocimiento para el procesamiento computacional del lenguaje natural. En concreto, se ha descrito la elaboración de un modelo conceptual de unidades provenientes de los dominios del terrorismo y de la delincuencia organizada a partir de un lenguaje artificial de creación ontológica. Para ello, se ha propuesto una descripción de las entidades colectivas “cártel”, “grupo con fines propios” y “célula”, que hacen referencia a los tipos de asociaciones ilegales que

son responsables de la mayor parte de los actos delictivos y de terrorismo ocurridos a nivel internacional. Estas entidades forman parte de una ontología especializada de FunGramKB, una base de conocimiento cuyo propósito es servir como plataforma para el desarrollo de aplicaciones relacionadas con el procesamiento del lenguaje y el razonamiento artificial. Las dos principales aportaciones del presente artículo son: definir ontológicamente las unidades criminológicas seleccionadas, y explorar la conexión, a través de los enunciados lógicos de la base de conocimiento, entre el conocimiento especializado y el conocimiento de sentido común presente en la Ontología Nuclear de FunGramKB. Tal y como se ha sugerido, el contenido semántico de los términos constituye el inicio del aparato conceptual que, en posteriores desarrollos, permitiría, por un lado, mejorar el proceso de recuperación de información en escenarios de riesgo y, por otro, prevenir la delincuencia y el terrorismo.

*Agradecimientos

Este artículo forma parte del proyecto de investigación denominado “Elaboración de una subontología terminológica en un contexto multilingüe (español, inglés e italiano) a partir de la base de conocimiento FunGramKB en el ámbito de la cooperación internacional en materia penal: Terrorismo y crimen organizado”, financiado por el Ministerio de Ciencia e Innovación. Código: FFI2010-15983.

119

Notas

¹ La Ontología incluye tres tipos de unidades: metaconceptos, conceptos básicos y conceptos terminales. Los metaconceptos son elementos de jerarquización general y se expresan mediante el símbolo “#” (por ejemplo, #OBJECT representa el concepto genérico “objeto”). Los conceptos básicos constituyen el segundo nivel —semánticamente más específico— dentro de la Ontología. Se expresan mediante el símbolo “+” (por ejemplo, +HOUSE_00, que denota la entidad “casa”). Finalmente, los conceptos terminales son las unidades de mayor concreción semántica y conforman el nivel más básico de la estructura ontológica, ya que no pueden dividirse en unidades conceptuales menores. Se expresan mediante el signo “\$” (por ejemplo, \$DICTIO-

NARY_00, que se refiere a la entidad “diccionario”). El índice numérico que aparece en los conceptos básicos y terminales permite distinguir entre varios sentidos en casos de polisemia y homonimia (por ejemplo, +FISH_00 denota “pez”, mientras que +FISH_01 denota “pescar”) (Periñán-Pascual & Mairal-Usón 2010).

² NDIC: National Drug Intelligence Center. <http://www.justice.gov/archive/ndic>. Este organismo, dependiente del Ministerio de Justicia estadounidense, cesó su actividad en junio de 2012. A pesar de ello, su descripción de cártel sigue siendo válida para el propósito de este artículo, de ahí que se proponga como base conceptual.

Obras Citadas

ARCAS-TÚNEZ, Francisco y Carlos PERIÑÁN-PASCUAL. 2014. "La Ingeniería del Conocimiento en el Dominio Legal: La Construcción de una Ontología Satélite en FunGramKB". *Signos* 47 (84): 113-139.

ALIPRANDI, Carlo, Juan ARRAIZA-IRUJO, Montse CUADROS, Sebastian MAIER, Felipe MELERO y Matteo RAFFAELLI. 2014. "CAPER: Collaborative information, acquisition, processing, exploitation and reporting for the prevention of organised crime". En Stephanidis, Constantine (ed.) *HCI International 2014-Posters' Extended Abstracts*. Berlín: Springer: 147-152.

ARGAMON, Shlomo y Newton HOWARD (eds.). 2009. *Computational Methods for Counterterrorism*. Berlín: Springer.

ASHBY, Nathan y Miguel Ángel RAMOS. 2013. "Foreign Direct Investment and Industry Response to Organized Crime: The Mexican Case". *European Journal of Political Economy* 30: 80-91.

BAK, Jaroslaw y Czeslaw JEDRZEJEK. 2010. "Application of an Ontology-based Model to a Selected Fraudulent Disbursement Economic Crime". En Casanovas, Pompeu, Ugo Pagallo, Giovanni Sartor y Gianmaria Ajani (eds.) *AI Approaches to the Complexity of Legal Systems: Complex Systems, the Semantic Web, Ontologies, Argumentation, and Dialogue*. Berlín: Springer: 113-132.

BAK, Jaroslaw, Jolanta CYBULKA y Czeslaw JEDRZEJEK. 2013. "Ontological Modeling of a Class of Linked Economic Crimes". En Ngnoc, Thanh Nguyen (ed.) *Transactions on Computational Collective Intelligence IX*. Berlín: Springer: 98-123.

BREUKER, Joost, André VALENTE y Radboud WINKELS. 2005. "Use and Reuse of Legal Ontologies in Knowledge Engineering and Information Management: Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications". En Benjamins, Richard, Pompeu Casanovas, Joost Breuker y Aldo Gangemi (eds.) *Law and the Semantic Web*. Berlín: Springer: 36-64.

BREWSTER, Ben, Simon ANDREWS, Simon POLOVINA, Laurence HIRSCH y Babak AKHGAR. 2014. "Environmental Scanning and Knowledge Representation for the Detection of Organised Crime Threats". En Hernández, Nathalie, Robert Jäschke y Madalina Croitoru (eds.) *Graph-based Representation and Reasoning*. Berlín: Springer: 275-280.

CARSON, Jennifer, Gary LAFREE y Laura DUGAN. 2012. "Terrorist and Non-terrorist Criminal Attacks by Radical Environmental and Animal Rights Groups in the United States, 1970-2007". *Terrorism and Political Violence* 24 (2): 295-319.

CASELLAS, Núria. 2011. *Legal Ontology Engineering: Methodologies, Modelling Trends, and the Ontology of Professional Judicial Knowledge*. Berlín: Springer.

CHMIELEWSKI, Mariusz, Andrzej GALKA, Piotr JAREMA, Kamil KRASOWSKI y Artur KOSIŃSKI. 2009. "Semantic Knowledge Representation in Terrorist Threat Analysis for Crisis Management Systems". En Nguyen, Ngoc Thanh, Ryszard Kowalczyk y Shyi-Ming Chen (eds.) *Computational Collective Intelligence: Semantic Web, Social Networks and Multiagent Systems*. Berlín: Springer: 460-471.

COLE, Ronald, Joseph MARIANI, Hans USZKOREIT, Annie ZAENEN y Victor ZUE (eds.). 1997. *Survey of the State of the Art in Human Language Technology*. Cambridge University Press and Giardin. Disponible en: <http://www.dfki.de/~hansu/HLT-Survey.pdf>

DEAN, Geoff, Ivar FAHSING y Petter GOTTSCHALK. 2010. *Organized Crime: Policing Illegal Business Entrepreneurialism*. Oxford: Oxford University Press.

CHOI, Dongjin, Byeongkyu, Ko, Heesun Kim Pankoo Kim. 2014. "Text Analysis for Detecting Terrorism-related Articles on the Web". *Journal of Network and Computer Applications* 38: 16-21.

DZEMYDIENE, Dale y Egle KAZEMIKAITIENE. 2005. "Ontology-based Decision Support System for Crime Investigation Processes". En Vasile-

- cas, Olegas, Wita Wojtkowski, Jože Zupančič, Albertas Caplinskas, Gregory Wojtkowski y Stanisław Wrycza (eds.) *Information Systems Development*. Berlin: Springer: 427-438.
- ENDERS, Walter y Todd SANDLER. 2002. "Patterns of Transnational Terrorism, 1970-1999: Alternative Time-series Estimates". *International Studies Quarterly* 46 (2): 145-165.
- FELICES-LAGO, Ángel y Pedro UREÑA GÓMEZ-MORENO. 2012. "Fundamentos Metodológicos de la Creación Subontológica en FunGramKB". *Onomázein* 26 (2): 49-67.
- FELICES-LAGO, Ángel y Pedro UREÑA GÓMEZ-MORENO. 2014. "FunGramKB Term Extractor: A Tool for Building Terminological Ontologies from Specialised Corpora". En Nolan, Brian y Carlos Perrián-Pascual (eds.) *Language Processing and Grammars: The Role of Functionally Oriented Computational Models*. Amsterdam: John Benjamins: 251-269.
- GALJANO, Phillip y Vasily POPOVICH. 2009. "Theoretical Investigation of Terrorism". En Popovich, Vasily, Christophe Claramunt, Thomas Devogele, Manfred Schrenk y Kyryll Korolenko (eds.) *Information Fusion and Geographic Information Systems: Proceedings of the Fourth International Workshop, 17-20 May 2009*. Berlin: Springer: 227-239.
- GOLBECK, Jennifer, Aaron MANNES y James HENDLER. 2006. "Semantic Web Technologies for Terrorist Network Analysis". En Popp, Robert y John Yen (eds.) *Emergent Information Technologies and Enabling Policies for Counter-terrorism*. Nueva Jersey: John Wiley: 125-137.
- GONZÁLEZ-CONEJERO, Jorge, Rebeca VARELA-FIGUEROA, Juan MUÑOZ-GÓMEZ y Emma TEODORO. 2014. "Organized Crime Structure Modelling for European Law Enforcement Agencies Interoperability through Ontologies". En Casanovas, Pompeu, Ugo Pagallo, Monica Palmirani y Giovanni Sartor (eds.) *AI Approaches to the Complexity of Legal Systems*. Berlin: Springer: 217-231.
- GRUBER, Thomas. 1995. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing?". *International Journal of Human-Computer Studies* 43 (5): 907-928.
- GUARINO, Nicola, Daniel OBERLE y Steffen STAAB. 2009. "What is an Ontology?". En Staab, Steffen y Rudi Studer (eds.) *Handbook on ontologies*. Berlin: Springer: 1-17.
- HRECKOVSKI, Boris y Robert DOBSON. 2012. "Terrorist Attacks on the Civilian Community". En Lennquist, Sten (ed.) *Medical Response to Major Incidents and Disasters: A Practical Guide for all Medical Staff*. Berlin: Springer: 337-351.
- JIANG, Xing y Ah-Hwee TAN. 2010. "CRCTOL: A Semantic-based Domain Ontology Learning System". *Journal of the American Society for Information Science and Technology* 61 (1): 150-168.
- KLERKS, Peter. 2007. "Methodological Aspects of the Dutch National Threat Assessment". *Trends in Organized Crime* 10 (4): 91-101.
- LAFREE, Gary. 2011. "Using Open Source Data to Counter Common Myths about Terrorism". En Forst, Brian, Jack Greene y James Lynch (eds.) *Criminologists on Terrorism and Homeland Security*. Cambridge: Cambridge University Press: 411-442.
- LAFREE, Gary, Laura DUGAN y Erin MILLER. 2015. *Putting Terrorism in Context: Lessons from the Global Terrorism Database*. Londres: Routledge.
- LUZONDO-OYÓN, Alba. 2014. "Constraining Factors on the Family of Resultative Constructions". *Review of Cognitive Linguistics* 12 (1): 30-63.
- LUZONDO-OYÓN, Alba y Francisco RUIZ DE MENDOZA IBÁÑEZ. 2015. "Argument Structure Constructions in a Natural Language Processing Environment". *Language Sciences* 48, 70-89.
- MAIRAL-USÓN, Ricardo y Carlos PERIÁN-PASCUAL. 2009. "The Anatomy of the Lexicon Component within the Framework of a Conceptual Knowledge Base". *Revista Española de Lingüística Aplicada* 22 (1): 217-244.
- MANNES, Aaron y GOLBECK, Jennifer. 2005. Building a terrorism ontology. Presentado en ISWC Workshop on Ontology Patterns for the Semantic Web.
- McCARTHY, Dennis. 2011. *An Economic History of Organized Crime: A National and Transnational Approach*. Londres: Routledge.
- NDTA: "National Drug Threat Assessment" 2010. U. S. Department Of Justice, National

Drug Intelligence Center (NDIC). Disponible en: <http://www.justice.gov/archive/ndic/pubs38/38661/38661p.pdf>.

OCTA: "EU Organised Crime Threat Assessment" 2007. Disponible en: <https://www.europol.europa.eu/sites/default/files/publications/octa2007.pdf>.

PERIÑÁN-PASCUAL, Carlos. 2013. "Towards a Model of Constructional Meaning for Natural Language Understanding". En Nolan, Brian y Elke Diederichsen (eds.) *Linking Constructions into Functional Linguistics: The Role of Constructions in Grammar*. Amsterdam: John Benjamins: 205-230.

PERIÑÁN-PASCUAL, Carlos y Francisco ARCAS-TÚNEZ. 2007. "Cognitive Modules of an NLP Knowledge Base for Language Understanding". *Procesamiento del Lenguaje Natural* 39 (1): 197-204.

PERIÑÁN-PASCUAL, Carlos y Francisco ARCAS-TÚNEZ. 2010. "Ontological Commitments in FunGramKB". *Procesamiento del Lenguaje Natural* 44 (1): 27-34.

PERIÑÁN-PASCUAL, Carlos y Francisco ARCAS-TÚNEZ. 2014. "The Implementation of the FunGramKB CLS Constructor". En Nolan, Brian y Carlos Periñán-Pascual (eds.) *Language Processing and Grammars: The Role of Functionally Oriented Computational Models*. Amsterdam: John Benjamins: 165-196.

PERIÑÁN-PASCUAL, Carlos y María de los Llanos CARRIÓN-VARELA. 2011. "FunGramKB y el Conocimiento Cultural". *Anglogermánica Online* 1 (1): 87-105.

PERIÑÁN-PASCUAL, Carlos y Ricardo MAIRAL-USÓN. 2009. "Bringing Role and Reference Grammar to Natural Language Understanding". *Procesamiento del Lenguaje Natural* 43 (1): 265-273.

PERIÑÁN-PASCUAL, Carlos y Ricardo MAIRAL-USÓN. 2010. "La Gramática de COREL: Un Lenguaje de Representación Conceptual". *Onomázein* 21 (1): 11-45.

PERIÑÁN-PASCUAL, Carlos y Ricardo MAIRAL-USÓN. 2011. "The COHERENT Methodology in FunGramKB". *Onomázein* 24 (1): 13-33.

RUIZ DE MENDOZA IBÁÑEZ, Francisco. 2013. "Meaning Construction, Meaning Interpretation

and Formal Expression in the Lexical Constructional Model". En Nolan, Brian y Elke Diederichsen (eds.) *Linking Constructions into Functional Linguistics: The Role of Constructions in Grammar*. Amsterdam: John Benjamins: 231-270.

RUIZ DE MENDOZA IBÁÑEZ, Francisco. 2014. "Low-level Situational Cognitive Models within the Lexical Constructional Model and their Computational Implementation in FunGramKB". En Nolan, Brian y Carlos Periñán-Pascual (eds.) *Language Processing and Grammars: The Role of Functionally Oriented Computational Models*. Amsterdam: John Benjamins: 367-390.

RUIZ DE MENDOZA IBÁÑEZ, Francisco y Ricardo MAIRAL-USÓN. 2008. "Levels of Description and Constraining Factors in Meaning Construction: An Introduction to the Lexical Constructional Model". *Folia Linguistica* 42 (3-4): 355-400.

SEIDLER, Patrick, Richard ADDERLEY, Atta BADI y Matteo RAFFAELLI. 2014. "MOSAIC: Criminal Network Analysis for Multi-modal Surveillance and Decision Support". En *Advances in Social Networks Analysis and Mining (ASONAM)*, 2014 IEEE/ACM: IEEE: 257-260.

SPINK, John, Douglas MOYER, Hyeonho PARK y Justin HEINONEN. 2013. "Defining the Types of Counterfeiters, Counterfeiting, and Offender organizations". *Crime Science* 2 (8): 1-10.

SPRANGER, Michael y Dirk LABUDE. 2014. "Towards Establishing an Expert System for Forensic Text Analysis". *International Journal on Advances in Intelligent Systems* 7 (1-2): 247-256.

SUBRAHMANIAN, V. S. (ed.). 2009. *Handbook of Computational Approaches to Counterterrorism*. Berlín: Springer.

TSENG, Yuen-Hsien, Zih-Ping Ho, Kai YANG y Chun-Cheng CHEN. 2012. "Mining Term Networks from Text Collections for Crime Investigation". *Expert Systems with Applications* 39 (11): 10082-10090.

TUPMAN, W. A. 2009. "Ten Myths about Terrorist Financing". *Journal of Money Laundering Control* 12 (2): 189-205.

UNODC (United Nations Office on Drugs and

Crime). 2002. "Results of a Pilot survey of Forty Selected Organised Criminal Groups in 16 Countries". Disponible en: https://www.unodc.org/pdf/crime/publications/Pilot_survey.pdf.

UNODC (United Nations Office on Drugs and Crime). 2010. "Digest of Terrorist Cases". Disponible en: https://www.unodc.org/documents/terrorism/Publications/Digest_of_Terrorist_Cases/English.pdf.

UNODC (United Nations Office on Drugs and Crime). 2012. "The Use of the Internet for Terrorist Purposes". Disponible en: https://www.unodc.org/documents/frontpage/Use_of_Internet_for_Terrorist_Purposes.pdf.

UREÑA GÓMEZ-MORENO, Pedro, Ángela ALAMEDA-HERNÁNDEZ y Ángel FELICES-LAGO. 2011. "Towards a Specialised Corpus of Organized Crime and Terrorism". En Carrió, María Luisa, Josefa Contreras, Françoise Olmo, Hanna Skorczynska, Inmaculada Tamarit y Debra Westwall (eds.) *La Investigación y la Enseñanza Aplicadas a las Lenguas de Especialidad y a la Tecnología*. Valencia: Universitat Politècnica de Valencia: 301-306.

VANDER BEKEN, Tom. 2004. "Risky Business: A Risk-based Methodology to Measure Organized Crime." *Crime, Law & Social Change* 41 (1): 471-516.