

Implementación de reglas para el control de modelado con SNOMED CT

Navas H.^a, Lopez Osornio A.^b, Gambarte L.^a, Elías Leguizamón G.^a, Wasserman S.^a, Orrego N.^a, Luna D.^a, Gonzalez B. de Quirós F.^a

^a Área de Informática Médica, Hospital Italiano de Buenos Aires, Argentina

^b Termmed, Buenos Aires, Argentina

Resumen

La utilización de SNOMED CT como terminología de referencia estándar permite lograr interoperabilidad entre sistemas clínicos. SNOMED CT contempla la creación de términos post-coordinados por parte de los usuarios de acuerdo a cada necesidad local. Si bien la creación de estos términos es libre, existen una serie de reglas definidas en el Manual de Usuario de SNOMED CT que deben ser cumplidas.

El Hospital Italiano de Buenos Aires posee un Servidor de Terminología que codifica los términos médicos utilizando SNOMED CT como vocabulario de referencia.

Luego que el Hospital Italiano de Buenos Aires realizara una prueba de interoperabilidad con el Centro Médico Nebraska y detectara una alta tasa de error en los términos post-coordinados (26%) se decidió implementar un sistema de reglas automático dentro del propio Servidor de Terminología acorde a lo definido en el Manual de Usuario de SNOMED CT.

Luego de la implementación de las reglas, la tasa de error disminuyó del 26% al 2% aproximadamente.

Palabras Claves:

Snomed-CT, reglas modelado, control modelado, servidor terminología.

Introducción

SNOMED CT se ha propuesto como terminología estándar en muchos países, con una apropiada cobertura del vocabulario médico [1]. Actualmente es mantenido por una organización multinacional de desarrollo de estándares. Las implementaciones locales de SNOMED CT a menudo requieren ampliar su contenido para satisfacer necesidades particulares. Es por ello que existen mecanismos de post-coordinación que son definidos y promovidos por SNOMED Internacional.

El Área de Terminología Clínica del Hospital Italiano de Buenos Aires migró de un sistema de lista de problemas en texto libre en la Historia Clínica Electrónica a la creación de una terminología local de interfase donde el médico debe elegir la descripción apropiada para el problema o procedimiento [2].

Esta terminología de interfase fue creada a partir de textos ingresados por los propios médicos como problemas o procedimientos durante 4 años mediante texto libre. El vocabulario de interfase terminó creando 21.000 conceptos que se relacionaron con SNOMED CT como conceptos equivalentes o como expresiones post-coordinadas en los casos donde no se pudieron encontrar equivalencias directas. En 16.000 conceptos fue necesario crear una expresión post-coordinada, se trata de una muy alta proporción de post-coordinación, y se relaciona con una premisa de implementación de nuestro vocabulario de permitir la inclusión todos los conceptos que se registraron al menos 10 veces en el repositorio de la lista de problemas. Éste incluye conceptos muy específicos con una gran cantidad de modificadores, como la gravedad, la lateralidad, curso clínico, etc.

No existen herramientas específicas para modelar vocabularios dependientes de la lógica de SNOMED CT. En el pasado se utilizaron programas de modelado de estándares en ontología como Apelon o Protege, pero las reglas específicas de modelado definidas por SNOMED internacional y publicadas en el manual de SNOMED CT no fueron chequeadas o reforzadas por el software. En el Hospital Italiano la herramienta de modelado fue designada "in house" y no incluyó el control de la lógica del modelado.

En el año 2006 se desarrolló una prueba entre el Hospital Italiano de Buenos Aires y el Centro Médico Nebraska, en Estados Unidos [3]. Esta prueba mostró que conceptos post-coordinados con el mismo significado fueron representados de manera diferente, la mayoría debido al no cumplimiento de las guías de post-coordinación de SNOMED Internacional. El 28,8% de los términos post-coordinados no seguía las normas de post-coordinación.

Existen además, diferencias de criterios entre los usuarios que realizan el modelado de las expresiones post-coordinadas [4] [5] [6].

Para mejorar la calidad de las expresiones locales post-coordinadas, se agregó al software la capacidad de verificar, en línea, la correcta post-coordinación de los conceptos, representando las directivas de la guía de SNOMED, en un modelo de base de datos relacional. Estas directivas fueron creadas como reglas de modelado.

Objetivo

Testear la eficacia de las reglas de control de modelado para la post-coordinación de términos.

Materiales

La definición de las expresiones post-coordinadas sigue el mismo formalismo de cualquier concepto de SNOMED CT. Los conceptos incluidos en la distribución estándar fueron definidos por las autoridades de SNOMED CT y se llaman pre-coordinados. Nuevos conceptos definidos por los usuarios de la terminología son llamados post-coordinados. La Guía de Usuarios de SNOMED CT especifica las reglas para post-coordinar nuevos conceptos, este es el Modelo Semántico de SNOMED CT [7].

El objetivo de la post-coordinación es agregar nuevos conceptos a SNOMED CT, construidos a partir de existentes siguiendo el modelo semántico especificado. Crear expresiones post-coordinadas para representar nuevos conceptos requiere 2 pasos, el primero es definir los supertipos (uno o más) del nuevo concepto. Los supertipos son, en este caso, conceptos más genéricos de SNOMED CT de la distribución estándar. La correcta asignación de los supertipos es lo que determina la jerarquía del concepto. El segundo paso es agregar los atributos para diferenciar el nuevo y más específico concepto post-coordinado de su supertipo incluido en SNOMED CT. Cada atributo es definido como una relación con otro concepto de SNOMED.

Un concepto cuyo significado es completamente explicado por sus atributos y los atributos heredados de su supertipos, es llamado "Fully defined" (Completamente definido); por otra parte, si por alguna razón no es posible explicar el significado del concepto, o tiene algunos atributos diferentes, es llamado "Primitive" (Primitivo).

Según el modelo semántico, diferentes tipos de relaciones son válidas para diferentes jerarquías, pero solo una porción pre-definida de conceptos de SNOMED son válidos como blanco de esas relaciones. Por ejemplo "Severidad" es un atributo válido para conceptos de la jerarquía "Hallazgos", pero no para conceptos ubicados en la jerarquía "Procedimientos". Lo opuesto ocurre para el atributo "Acceso", que es válido para "Procedimientos" pero no para "Hallazgos".

Las reglas del modelo semántico fueron codificadas en un modelo relacional de bases de datos, utilizando una tabla de tipos válidos de relación, una tabla para las jerarquías señaladas por sus conceptos de alto nivel y un cuadro de objetivos, definidos una vez más por el concepto de nivel superior del grupo de conceptos objetivos válidos

El vocabulario del Hospital Italiano se almacena también en una base de datos relacional, incluyendo sus conceptos, descripciones y relaciones [8]. Un procedimiento desarrollado utilizando PL/SQL comprueba si es cada expresión post-coordinada está acorde a las restricciones del modelo semántico de SNOMED.

Las reglas que se ingresaron al Servidor de Terminología fueron:

- Sólo puede existir una y solo una relación "Es un mapeo".
- Si la relación utilizada es "Es un mapeo" el concepto no puede ser "Primitivo".
- Debe existir al menos una relación "es un" o "es un mapeo".
- El concepto debe pertenecer sólo a una jerarquía.
- La jerarquía del concepto no puede ser diferente a la jerarquía del dominio.
- El tipo de relación debe poder utilizarse para esa jerarquía.
- Los conceptos destino de relación deben ser válidos.
- Si una relación depende de la existencia de otra, esta última debe encontrarse presente.
- Existe además una advertencia en el caso que se utilice una "estructura corporal entera".

Métodos

Con el fin de comprobar la eficacia del sistema de control de modelado se tomaron 2 muestras de conceptos post-coordinados del vocabulario del Hospital Italiano.

La primera se realizó en el año 2006 luego de realizar el testeo con el Centro Médico Nebraska y previo a la implementación del sistema de reglas. La muestra contiene 34.253 términos post-coordinados.

La segunda muestra se realizó en el año 2008 luego de la implementación del sistema de reglas (agosto de 2007) y contiene 9.015 términos.

Los conceptos post-coordinados incluyen conceptos en todos los dominios del Sistema de Información Clínica, incluyendo de la lista de problemas, procedimientos, fármacos o dispositivos.

Cada lista de conceptos fue revisada por un experto que los categorizó de "correcto" o "incorrecto" de acuerdo con el cumplimiento de las normas de post-coordinación de SNOMED del manual de usuario. En el caso de los conceptos incorrectos, se le asignó una descripción más detallada del error.

La clasificación de los errores fue la siguiente:

- Error en la definición de la jerarquía: El concepto se encuentra en una jerarquía no apropiada; por ejemplo: un procedimiento en la jerarquía de problemas (apendicectomía como problema), o un descartable en procedimientos (stent como procedimiento), etc.
- Error en el tipo de relación: Tipo de relación no permitido para la jerarquía; por ejemplo: localización del hallazgo en procedimientos.

- Error en el objetivo de la relación: Concepto objetivo fuera del rango válido para el tipo de relación. La relación está bien, pero el blanco no es correcto; por ejemplo: lateralidad con rodilla derecha, cuando debería ser solamente derecha o se debería utilizar: localización del hallazgo.
- Uso inapropiado de una estructura anatómica como un todo: El concepto objetivo es un hijo del concepto Estructura anatómica como un todo (estructura corporal); por ejemplo cuando se utiliza el concepto "pierna como un todo" en lugar de usar "estructura de la pierna."
- Otros errores: errores fuera de cualquier categoría se agruparon en esta.

El experto también evaluó la precisión de la representación; las expresiones post-coordinadas pueden cumplir con las normas de SNOMED CT, pero no describir con exactitud el significado del término. Utilizando una escala de Likert de 5 puntos, expresó si la representación del significado es la mejor representación posible o si es completamente incorrecta. Esta medición será útil para interpretar que porción del control de calidad de la interfase de terminología puede ser automatizado y cual deberá requerir siempre la interpretación de un ser humano.

Se realizó un análisis estadístico de ambas muestras con prueba de hipótesis nula y se analizó su potencia. Para encontrar una diferencia estadísticamente significativa (0,050), era necesario analizar 50 términos de cada muestra, seleccionados aleatoriamente, ya que tendrían un poder del 80,9% de producir un resultado estadísticamente significativo. Este cálculo asume que la diferencia de proporciones es 0,20 (en concreto, 0,25 frente a 0,05).

El tamaño muestral también permite informar la diferencia existente entre ambas con una precisión de 0,13 puntos aproximadamente (nivel de confianza de 95%).

En concreto, una diferencia observada de 0,20 tendría un intervalo de confianza de 0,07 a 0,33.

Resultados

En la primera muestra de 34.253 se analizaron 300 términos. De ellos el 74% no presentaba errores de ningún tipo. El 26% restante está formado por 0,67% de error en la definición de la jerarquía, 3% de error en el tipo de relación, 6,67% en error en el blanco de la relación, 5% de uso indebido de parte entera, y 10,67% de otros errores.

En la segunda muestra de 9.015 se analizaron 300 términos. De ellos el 97,3% no presentaba errores de ningún tipo. El 2,7% restante está conformado por 0,3% de error en el blanco de la relación y el 2,4% restante por errores de otro tipo.

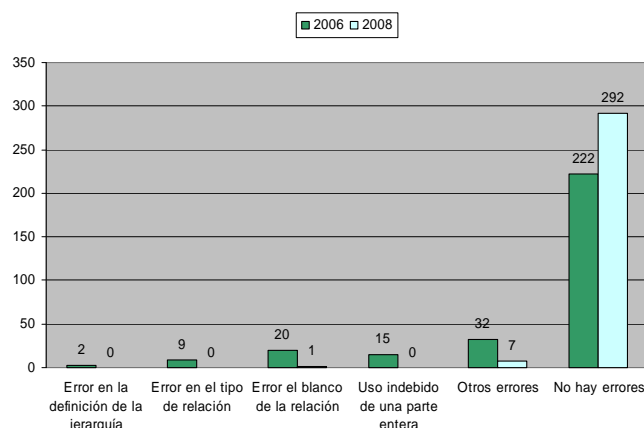


Figura 1: Tipo de error

Del análisis de la precisión de la representación surgió que de la primera muestra el 77% de los términos poseían una representación adecuada (puntos 1 y 2 de la escala), el 10,67% se encontraban mal representados (puntos 4 y 5) y el 12,3% no podían determinarse.

En la muestra del año 2008 se encontró que el 96,33% de los términos se encontraban correctamente representados, el 2% se encontraron mal representados y el 1,67% no podían determinarse.

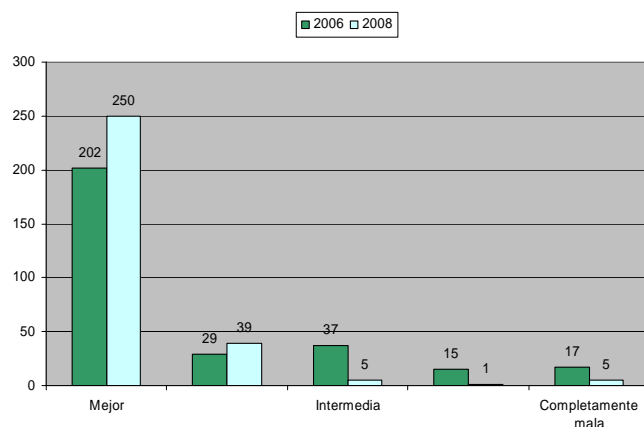


Figura 2: Calidad de la representación

Conclusión

La implementación de un sistema de reglas automático para la post-coordinación de términos mejora la representación de los mismos, ya sea por la correcta utilización de las relaciones de SNOMED CT, como también por la adecuada representación de los conceptos médicos.

Claramente un sistema de reglas permitirá mejorar la interoperabilidad con otros Centros de Salud que la lograda en el año 2006 con el Centro Médico de Nebraska.

Tiene también efecto positivo en el ámbito educativo al mejorar la capacitación de los usuarios encargados de modelar los

términos (codificadores), al obligarlos a seguir las reglas buscando opciones adecuadas para la representación de los términos.

Este sistema de reglas podría disminuir la variabilidad entre usuarios al momento de interpretar y generar las expresiones post-coordinadas, aumentando el consenso.

Agradecimientos

A todos aquellos que colaboraron en el trabajo, en especial a Guada y Santi.

Referencias

- [1] Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. AMIA Annual Symposium proceedings / AMIA Symposium. 2003:699-703.
- [2] Osornio AL, Luna D, Gambarte ML, Gomez A, Reynoso G, de Quiros FG. Creation of a local interface terminology to SNOMED CT. Studies in health technology and informatics. 2007;129(Pt 1):765-9.
- [3] Campbell JRLO, Alejandro; de Quiros, Fernan; Luna, Daniel and Reynoso, Guillermo. Semantic Interoperability and SNOMED CT: A Case Study in Clinical Problem Lists. In: Kuhn KAW, James R; Leong, Tze-Yun, ed. *Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems 2007*: [2401]-[2].
- [4] Andrews JE, Patrick TB, Richesson RL, Brown H, Krischer JP. Comparing heterogeneous SNOMED CT coding of clinical research concepts by examining normalized expressions. Journal of biomedical informatics. 2008 Feb 5.
- [5] Andrews JE, Richesson RL, Krischer J. Variation of SNOMED CT coding of clinical research concepts among coding experts. J Am Med Inform Assoc. 2007 Jul-Aug;14(4):497-506.
- [6] Chiang MF, Hwang JC, Yu AC, Casper DS, Cimino JJ, Starren JB. Reliability of SNOMED-CT coding by three physicians using two terminology browsers. AMIA Annual Symposium proceedings / AMIA Symposium. 2006:131-5.
- [7] Organisation TIHTSD. SNOMED Clinical Terms User Guide. In: Organisation TIHTSD, ed. 2008.
- [8] Gambarte ML, Osornio AL, Martinez M, Reynoso G, Luna D, de Quiros FG. A practical approach to advanced terminology services in health information systems. Studies in health technology and informatics. 2007;129(Pt 1):621-5.

Dirección para correspondencia

Hernán Navas

Departamento de Informática Médica

Hospital Italiano de Buenos Aires

Gascón 450 (C1181ACH) Buenos Aires – Argentina

e-mail: hernan.navas@hospitalitaliano.org.ar